JGR Atmospheres

RESEARCH ARTICLE
10.1029/2024JD042578

Surface Energy Balance Responses to Radiative Forcing in the Central Arctic From MOSAiC and Models

Anne Sledd[1,2], **Matthew D. Shupe**[1,2], **Amy Solomon**[1,2], **and Christopher J. Cox**[2]

[1]Cooperative Institute for Research in Environmental Sciences, University of Colorado Boulder, Boulder, CO, USA, [2]National Oceanic and Atmospheric Administration Physical Sciences Laboratory, Boulder, CO, USA

Key Points:

- The responses of surface fluxes to radiative forcing are calculated using observations over the Central Arctic sea ice pack for a full year
- Broadly there are two regimes of how surface fluxes respond, based on if the skin temperature can vary or if it is constrained by melt
- Flux responses are a useful diagnostic to evaluate compensating biases in surface fluxes in forecast systems and climate models

Supporting Information:

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

Correspondence to:

A. Sledd,
anne.sledd@colorado.edu

Citation:

Sledd, A., Shupe, M. D., Solomon, A., & Cox, C. J. (2025). Surface energy balance responses to radiative forcing in the Central Arctic from MOSAiC and models. *Journal of Geophysical Research: Atmospheres*, *130*, e2024JD042578. https://doi.org/10.1029/2024JD042578

Received 26 SEP 2024
Accepted 21 FEB 2025

Abstract The Arctic surface energy budget (SEB) couples the atmosphere with the sea ice, making it useful for both studying surface processes as well as evaluating models. Improved understanding of atmosphere-ice interactions is required to improve models, requiring year-round observations to address seasonally dependent biases. This work uses novel observations from the MOSAiC expedition to quantify the responses of surface fluxes to radiative forcing over sea ice throughout a complete annual cycle. We identify two primary regimes of flux response: an ice growth regime in winter and an ice melt regime in summer. In the growth regime, changes in radiative forcing impact upwelling longwave, sensible heat, and subsurface heat fluxes, whereas in the melt regime changes in radiative forcing primarily alter the amount of melt and subsurface transmission because the surface temperature is fixed. These observed responses of surface fluxes to radiative forcing are used to evaluate seven weather forecast models during the ice growth regime. In most models, the responses of surface fluxes to radiative forcing do not match observations. Many models also have biased downwelling longwave. One model (the Coupled Arctic Forecast System; CAFS) adequately captures both the mean radiative forcing and the flux responses in winter. CAFS is further evaluated against observations spanning the full MOSAiC year, demonstrating sufficient agreement to provide a more generalized understanding of these SEB process relationships across the Arctic.

Plain Language Summary Arctic sea ice grows or melts based on whether there is more energy going into it or leaving it. The amount of energy going into the sea ice depends on the atmosphere, and this energy exchange couples the surface below to the atmosphere above. Understanding this coupled relationship is important for accurate weather forecasts in the Arctic. We quantify how extra energy from the atmosphere, such as increased thermal heating from a cloud, is balanced at the surface using observations. During winter, additional energy from the atmosphere is balanced by the surface warming, sea ice growth decreasing, and mixing of the air above the surface. In summer additional energy causes the surface to melt or it moves deeper into the sea ice. Broadly, the different ways that energy is distributed depends on whether or not the sea ice is growing or melting. We use these relationships to test how well seven weather forecast models represent the balance of energy in winter. Most models disagree with the observations, either with how much additional energy is coming from the atmosphere or where the extra energy goes. Recognizing how models diverge from observations can help focus where the models need to be improved.

1. Introduction

The surface energy budget (SEB) is a useful framework for studying the coupling between the atmosphere and sea ice in the Arctic. The exchange of energy at the surface determines if sea ice grows or melts, that is, if there is more energy going to the ice or leaving it. However, due to a relative lack of observations over sea ice, some basic understanding of the atmosphere-ice coupling is still lacking (e.g., Jung et al., 2016). Better understanding is required for improving models to have reliable predictions in the Arctic, especially in light of the changing sea ice cover and subsequent feedbacks (e.g., Maslowski et al., 2012).

The balance of surface fluxes is seasonally dependent. In winter when there is no incoming solar radiation (i.e., downwelling shortwave radiative flux, SWD), the key surface fluxes are all interdependent, connected through the surface temperature (e.g., Sterk et al., 2013), either directly, as in upwelling longwave radiative flux (LWU), or through temperature gradients in the atmosphere (sensible heat flux; SH) and snow-ice column (conduction). The relative magnitude of these fluxes largely depends on whether or not clouds are present (e.g., Persson et al., 2002; Persson et al., 2017). As the year progresses, the balance of surface fluxes changes. SWD adds energy

to snow and sea ice, with the magnitude of SW absorbed by the surface dependent on solar zenith angle and surface albedo (Shupe & Intrieri, 2004). This extra energy warms the surface, eventually leading to surface melt. Cloud cover is more frequent in summer (Shupe et al., 2011), and the net radiative impact of clouds changes. Although clouds still increase downwelling longwave radiation (LWD) they also decrease SWD, sometimes having a net cooling effect on the sea ice surface depending on the surface albedo (Intrieri et al., 2002; Shupe & Intrieri, 2004).

Many of the aforementioned processes important to the SEB are difficult for models to realistically represent in the Arctic (Jung et al., 2016). Key among model deficiencies is accurately simulating clouds in the Arctic (e.g., Curry et al., 1996; Inoue et al., 2021). Most weather forecast models underestimate the occurrence and water content of liquid-containing clouds in winter (Solomon et al., 2023), whereas in summer, simulated clouds can have too much liquid (Sedlar et al., 2020; Tjernström et al., 2021). These cloud biases cause further problems in modeled radiative fluxes that vary by season (e.g., Rinke et al., 2006). Surface representation, for example, snow cover, and coupling to the atmosphere are also important for accurately modeling the surface temperature and near surface stability (e.g., Batrak & Müller, 2019; Pithan et al., 2016; Steeneveld et al., 2006). In turn, surface temperature biases impact the presence of surface-based temperature inversions and turbulent heat transfer (Pithan et al., 2014). Complicating matters further, because these processes are all interdependent (e.g., Walden et al., 2017), an error in one flux or process must be compensated elsewhere in the SEB (e.g., Sotiropoulou et al., 2016). All of these issues impact the SEB, making it a useful tool for model evaluation (e.g., Miller et al., 2018). However, because the processes that determine these biases are seasonally dependent, observations are needed throughout the year for proper evaluation.

The Multidisciplinary drifting Observatory for the Study of Arctic Climate (MOSAiC) expedition provides a unique opportunity to study the SEB over sea ice for a full year in the Arctic. The MOSAiC expedition took place from 2019 to 2020 in the central Arctic where the icebreaker *Polarstern* drifted with the sea ice for a full year taking measurements of the coupled atmosphere-ice-ocean system (Nicolaus et al., 2022; Shupe et al., 2022). Such comprehensive year-round observations have not been made in the Arctic since the SHEBA experiment in 1997–1998 (Uttal et al., 2002), and since then the Arctic sea ice has become younger and thinner on average (Meier & Stroeve, 2022). MOSAiC observations provide a unique opportunity to better understand the Arctic SEB as well as evaluate models using energy budget based process relationships.

This paper uses MOSAiC observations to answer the following questions: How do surface fluxes respond to radiative forcing over sea ice throughout the year? What drives seasonal differences in the flux responses? Are weather forecast models able to represent these process relationships? To answer these questions, observational results are presented first. Then the observed relationships are used to evaluate several weather forecast models. Based on the evaluation, additional data from one model is used to further explore the variability of surface flux responses beyond the MOSAiC year and across the Arctic. We end this work with a discussion of the results and a summary.

2. Data

2.1. Observations

The German icebreaker *Polarstern* (Knust, 2017) was frozen into the sea ice in the central Arctic Ocean on 4 October 2019. Instruments were installed in the Central Observatory (CO) within 2 km of *Polarstern*, including a 10-m tower that measured turbulent fluxes and meteorological variables with a nearby station for measuring surface radiative fluxes. The Distributed Network (DN; Rabe et al., 2024) consisted of additional instruments installed within tens of kilometers of the CO. Included in the DN were three atmospheric surface flux stations (ASFS), with similar measurements as the CO, and seasonal ice mass balance buoys (IMBs). These installations operated in various configurations until 31 July 2020. On 21 August 2020 instruments were installed on a new ice floe near the North Pole, and measurements were taken until 20 September 2020. The number of stations contributing data in a given month are provided in Table 1.

For the present analysis, we use broadband radiative fluxes and turbulent sensible and latent heat (LH) fluxes from the CO and ASFS whenever available. These measurements are described in detail by C. J. Cox et al. (2023), and a brief summary of some important details is provided here. Radiative fluxes were measured at 1 m above the surface for downward fluxes and 3 m above the surface for upward fluxes at the CO, and 2 m above the surface at

Table 1
Number of Stations Reporting Data in Each Month for Radiative (SW, LW), Turbulent (SH, LH), and Subsurface (Conduction, Storage) Fluxes

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Radiative fluxes	3	3	4	4	3	2	3	3	3	3	2	3
Turbulent fluxes	3	3	4	4	3	2	2	2	2	3	2	3
Subsurface fluxes	2	7	8	8	6	4	2	2	2	1	0	0

ASFS for both up- and downward fluxes. LW uncertainties are estimated to be 2.6 Wm^{-2} (LWD) and 1 Wm^{-2} (LWU), and SW uncertainties are 3.1–4.5 W m^{-2} depending on the specific radiometer. Turbulent fluxes were calculated using 20-Hz sonic anemometer data, resampled to 10-Hz, and the eddy-covariance technique. SH data have an estimated uncertainty of 4.8 W m^{-2}, and LH have an uncertainty of 50%. Surface skin temperature is calculated from LW fluxes assuming a surface emissivity of 0.985. The default time resolution is 10 min for all observational data.

To characterize the thermal conductivity of the sea ice and snow, we use temperature profiles from two types of IMBs. Both buoys have vertical resolution of 2 cm. Snow and Ice Mass Balance Arrays (SIMBA) are thermistor string-type IMBs (Jackson et al., 2013; Lei, Cheng et al., 2022) with measurements taken every 6 hours. SIMBAs also measure temperature differences after pulses of heat over 120 s, which can be used to distinguish interfaces between the atmosphere, snow, ice, and ocean (Zuo et al., 2018). We use five SIMBAs that were installed at the CO and remote sites. Seasonal ice mass balance (SIMB) buoys are spar buoys that include acoustic rangefinders to measure the top and bottom of the snow-ice system (Perovich et al., 2023; Planck et al., 2019; Polashenski et al., 2011). Data from SIMBs were taken every 4 hours. We use three SIMBs that were installed at each of the remote sites. To calculate sea ice and snow conductive heat fluxes in winter from these temperature profiles, we use estimates of thermal conductivity and density from Sledd et al. (2024). Estimates are based on inverting the one-dimensional heat equation assuming energy is only transferred through vertical conduction. These estimates were only made through 14 March 2020, when there was no SWD. We expand on these estimates to include solar absorption, described in the Methods section.

2.2. Models

Seven weather forecast models are evaluated against observations. These models vary in their representation of snow, sea ice, surface-atmosphere coupling, clouds, and turbulence parametrizations (Table 2). For six of the models, 1 hour grid cell averages are used from the grid cell closest to the location of *Polarstern* over winter (15 October 2019–15 March 2020) for the first 24 hr of lead time from daily initialization. The output from these models is described in more detail in Solomon et al. (2023). Output from one additional model, an experimental

Table 2
Models Used in This Study

Model	Horizontal resolution	Snow and sea ice treatment	Reference
Météo-France ARPEGE-GELATO forecast system (ARPEGE)	7.5 km	GELATO w/1 snow layer and 10 sea ice layers	Bazile et al. (2020)
NOAA-PSL-Coupled Arctic Forecast System (CAFS)	10 km	CICE w/1 snow layer and 7 sea ice layers	Solomon et al. (2024)
German Weather Service forecast system (DWD)	13 km	Sea ice model with 1 sea ice layer and snow representation through an empirical temperature relationship	Zängl et al. (2015)
HARMONIE-AROME forecast system, experimental configuration (H-AROME)	2.5 km	1D sea ice model with 4 ice layers and 12 snow layers	Bengtsson et al. (2017)
ECMWF Integrated Forecast System (IFS)	18 km	Constant 1.5 m sea ice thickness. No snow on sea ice.	Haiden et al. (2018)
U.S. Navy-ESPC forecast system (NAVY)	37 km	CICE w/1 snow layer and 7 sea ice layers.	Barton et al. (2021)
NOAA Unified Forecast System, experimental configuration (UFS)	13 km	CICE6	Jacobs (2021)

Note. Further details available in Solomon et al. (2023).

prototype of the NOAA Unified Forecast System (UFS), is only available for 3 hr instantaneous values with 72 hr lead times and initialization over 3 December 2019–14 March 2020; thus it is treated distinctly in the analysis. After an initial assessment of all models, further analysis is done using one of the models, the Coupled Arctic Forecast System (CAFS; Solomon et al., 2024). CAFS is a fully coupled regional forecast model with the following components: CICE version 5 with 1 snow layer and 7 ice layers, WRF ARW version 3.6.1, POP ocean model version 2, and the NCAR CLM land model version 4.5. Global 0.5° GFS forecast fields are used for lateral forcing. The further analysis uses 6 hour averages taken from the sea ice model output of CAFS over the full MOSAiC year from grid cells closest to the CO and ASFS as well as output from the full Arctic basin from additional years. This additional data from CAFS uses the same model physics as the earlier data set but has slightly different initial conditions.

3. Methods

3.1. Defining the SEB

For this work we define the SEB as follows:

$$LWD - LWU + SWD - SWU - SWT - SH - LH + G = M, \tag{1}$$

where shortwave (SW) fluxes are upwelling (SWU), downwelling (SWD), and transmitted below the surface interface layer (SWT). Turbulent fluxes are defined as positive when directed away from the surface. M is calculated as a residual of the other terms, which can all be directly measured or estimated. This can be considered the energy that is available for sea ice and snow melt, although, as a residual, it also accumulates the uncertainty in other measurements. Moreover, M is considered to only include melt but not freezing because sea ice grows at its bottom surface and not at its top surface. G represents the subsurface heat flux, which we define two ways. For studying the SEB from observations alone we consider the energy budget of a 6 cm slab so that G is the sum of conduction, C, into the slab from below and the change in temperature of the slab, that is, its storage, S:

$$G = C + S, \tag{2}$$

where C is calculated using effective thermal conductivity, k, and the temperature gradient at 6 cm below the surface:

$$C = -k * \Delta T/\Delta z. \tag{3}$$

S is calculated using density, ρ, heat capacity, c_p, and temperatures at the surface, T_{skin}, and 6 cm below the surface, T_{-6cm} at time, t, n:

$$S = -\frac{\rho * c_p * \Delta T}{\Delta t} = -\frac{\rho * c_p * \left(T_{-6cm}^{n+1} - T_{-6cm}^{n-1} + T_{skin}^{n+1} - T_{skin}^{n-1}\right)}{2(t_{n+1} - t_{n-1})}. \tag{4}$$

Density is averaged over the 6 cm slab (three IMB levels) when density estimates are available in winter and early spring. As the surface melts, lower levels from the IMB profiles are used to calculate density, typically increasing in magnitude. We limit the density to a maximum of 497 kg m^{-3} to represent the surface scattering layer that forms on the top of bare sea ice (Macfarlane, Dadic, et al., 2023). The 6 cm slab is chosen because determining the precise location of the snow-atmosphere interface is difficult from the IMBs, and that impacts the conductive flux estimate needed to close the budget. A 6 cm slab from the top of the surface can also be reasonably assumed to have the properties of a surface scattering layer during summer when density estimates are no longer available from IMBs.

Temperature measurements from IMBs are only available at 4 or 6 hr increments, so we combine them with skin temperatures from ASFS and the CO to estimate subsurface heat fluxes at 10 min resolution to match the time resolution of atmospheric fluxes. For a given IMB, skin temperature is taken from the nearest ASFS or the CO. Thielke et al. (2023) found that skin temperatures on level ice did not have large spatial variability across the MOSAiC floe, typically less than 1 K. For each 4 or 6 hr interval for the IMB, the ASFS skin temperature is

linearly adjusted so that the endpoints of the interval are equal to the IMB surface temperature estimates. We expect the temperature variability beneath the surface to be damped compared to the surface, so the temperature 6 cm below the surface is linearly interpolated to 10 min increments. Density and the conductive flux are also linearly interpolated to 10 min time steps.

For IMB measurements in winter, the surface is defined as where the effective thermal conductivity remains below 0.6 W m^{-1} K^{-1} based on Sledd et al. (2024). When these estimates are no longer available we use surface estimates from the acoustic sounder from SIMB measurements and the algorithm based on differences in heating rates for SIMBA measurements (Lei, Cheng et al., 2022).

Although using a slab definition for G works well to account for observational uncertainty, it is not easily comparable with model output. Therefore, to evaluate the models with observations, we define the energy budget at the surface-atmosphere interface, assuming an infinitely thin layer. With this definition of the energy budget there is no slab or mass that changes temperature, so G is equal to the conductive flux at the surface. Because of the high uncertainty of observed conduction at the surface, we use the sum of radiative and turbulent fluxes to calculate the total subsurface flux, the middle equation of:

$$LWD - LWU + SWD - SWU - SH - LH = M - SWT - G = NA, \tag{5}$$

where NA is the "net atmospheric flux" equal to the total subsurface flux ($M - SWT - G$). In winter, when there is no SW radiation or melt, the right-hand side of Equation 5 is equal to the conduction at the surface.

To calculate SWT, we assume exponential decay (i.e., Beer's law) of SW flux in the ice and snow:

$$SWT = SWN e^{-\kappa_s D_s}\left(f_v I_{0v} e^{-\kappa_{iv} D_i} + f_n I_{0n} e^{-\kappa_{in} D_i}\right), \tag{6}$$

where $SWN = SWD - SWU$, $e^{-\kappa_s D_s}$ represents SWT through snow, and the term in parantheses represents the SWT through ice. $D_{s,i}$ is the slab thickness of snow or ice, totaling 6 cm. κ_s is the snow extinction coefficient, assumed to be 16 m^{-1} until 5 May 2020, after which it linearly declines to 3 m^{-1} on 17 June 2020 when the snow had significantly melted on level ice surfaces. Sea ice optical properties are taken from Persson (2012). I_{0v} is the ice surface transmission for visible radiation (0.95), and I_{0n} is the ice surface transmission for near-IR radiation (0.37). Solar radiation is weighted as 60% visible to 40% near-IR ($f_v = 0.6$; $f_n = 0.4$). Extinction coefficients for visible and near-IR radiation are $\kappa_{iv} = 0.72$ m^{-1} and $\kappa_{in} = 0.37$ m^{-1}. For comparison, Castellani et al. (2022) found up to 70% of incident light is extinguished in the surface scattering layer. The response of SWT to LWD + SWN is about 10% smaller when assuming no near-IR is transmitted.

The same constants are used to calculate SW absorption in the snow and ice column to estimate thermal conductivity and density after 14 March 2020. This means including a source term in the 1D heat equation used by Sledd et al. (2024):

$$\frac{\partial(\rho c_p T)}{\partial t} = \frac{\partial(k \partial T/\partial z)}{\partial z} + \kappa * SWN * e^{-\kappa z}. \tag{7}$$

SWN is used from the ASFS closest to each IMB. Varying the snow extinction coefficient has a negligible impact on the calculated storage and conductive flux in spring and summer. This method returns reasonable thermal conductivities and densities in early spring, but it is no longer viable once the snowpack warms and experiences sign changes in the vertical temperature gradient. Mean profiles of thermal conductivity and density are used after this point. The surface at MOSAiC had considerable spatial heterogeneity that led to large variability in SW absorption and transmission (Tao et al., 2024), so it is not necessarily surprising that measurements from one location for SW properties do not work well with the IMB measurements. It is also possible for the IMBs to absorb SW and bias the temperature measurements (e.g., Huwald et al., 2009). These facts mean our subsurface heat flux estimates in spring and summer are highly uncertain. However, in late spring and summer the snow and sea ice near the surface is approximately isothermal as it melts (e.g., Perovich et al., 2023), so we expect the subsurface heat flux to be small.

3.2. Quantifying Flux Responses to Atmospheric Forcing

To study the impact of atmospheric forcing on the SEB we use the framework of Miller et al. (2017), which assumes that variability in atmospheric radiative forcing, defined as LWD + SWN (the forcers), drives changes in the remaining surface fluxes (the responders). We use regressions to quantify the responses of other flux terms to atmospheric forcing. The response is calculated as the slope from linear least squares regression between two parameters, and effectively represents the sensitivity of a given flux term to changes in radiative forcing. An example to illustrate this approach is given as the introduction in the Results section. Variability in the radiative forcing is primarily due to the presence or absence of clouds, the solar zenith angle, and the surface albedo. In winter with no solar contribution, the atmospheric forcing reduces to the typical bimodal behavior of LW radiation, whereas in summer the forcing is somewhat more complicated due to the additional solar contributions. Through this approach we can learn about the coupling between the SEB and atmospheric variability throughout the year.

This framework is also useful to evaluate models and determine if they can produce the right fluxes for the right reasons (e.g., Miller et al., 2018). For example, too few liquid containing Arctic clouds is a longstanding issue with models (e.g., Curry et al., 1996; Inoue et al., 2021), but by studying the response of surface fluxes to radiative forcing we can still understand if the coupling between the surface and atmosphere is reasonable despite incorrect atmospheric forcing (e.g., Day et al., 2020). To further assess models we also calculate biases of the regressions as follows:

$$response\ bias = \frac{\sum(m_{mod} - m_{obs}) * x + (b_{mod} - b_{obs})}{n}, \tag{8}$$

where m is the slope from the model or observations, b is the y-intercept, x is the observed LWD + SWN, and n is the number of points. Finally, we calculate the difference in means of atmospheric radiative forcers.

If the uncertainty in the measurements were zero, the sum of slopes relating the radiative forcing to all responding terms would equal one:

$$\frac{\partial(LWD + SWN)}{\partial(LWD + SWN)} = \frac{\partial(LWU + SH + LH + M - SWT - G)}{\partial(LWD + SWN)}, \tag{9}$$

meaning that all incoming energy was balanced by outgoing energy. In reality there is both measurement uncertainty in the vertical fluxes as well as different measurement footprints and unaccounted horizontal fluxes that complicate closing the SEB (e.g., Foken, 2008). If the measurement uncertainty is randomly distributed, it would have no impact on the calculated slope; however, if there is a bias in when the measurements are more uncertain (e.g., only during clear time periods) the derived slopes would be incorrect. In winter, testing the maximum possible error in LW fluxes during only radiatively clear or cloudy times has a limited impact on slopes (<0.02). Including the maximum possible error for G and SH could lead to an error in slope <0.08. Although not negligible, this worst-case difference is similar to or less than the spatial variability. In summer there is no significant impact. Turbulent fluxes may also transiently become forcing terms, for example, in the presence of leads or during the passage of fronts. Under such conditions, our assumption of a radiatively driven flux system breaks down.

4. Results

In this section we first explain the approach of estimating the response of SEB terms with monthly regressions from observations for one example month, then quantify monthly responses of fluxes over the full MOSAiC year. This annual perspective is used to evaluate the CAFS model along with several other weather forecast models. It will be shown that CAFS represents the energy partitioning sufficiently to warrant further analysis. The final section uses spatial output from CAFS to explore the variability of flux responses to atmospheric forcing beyond the MOSAiC drift.

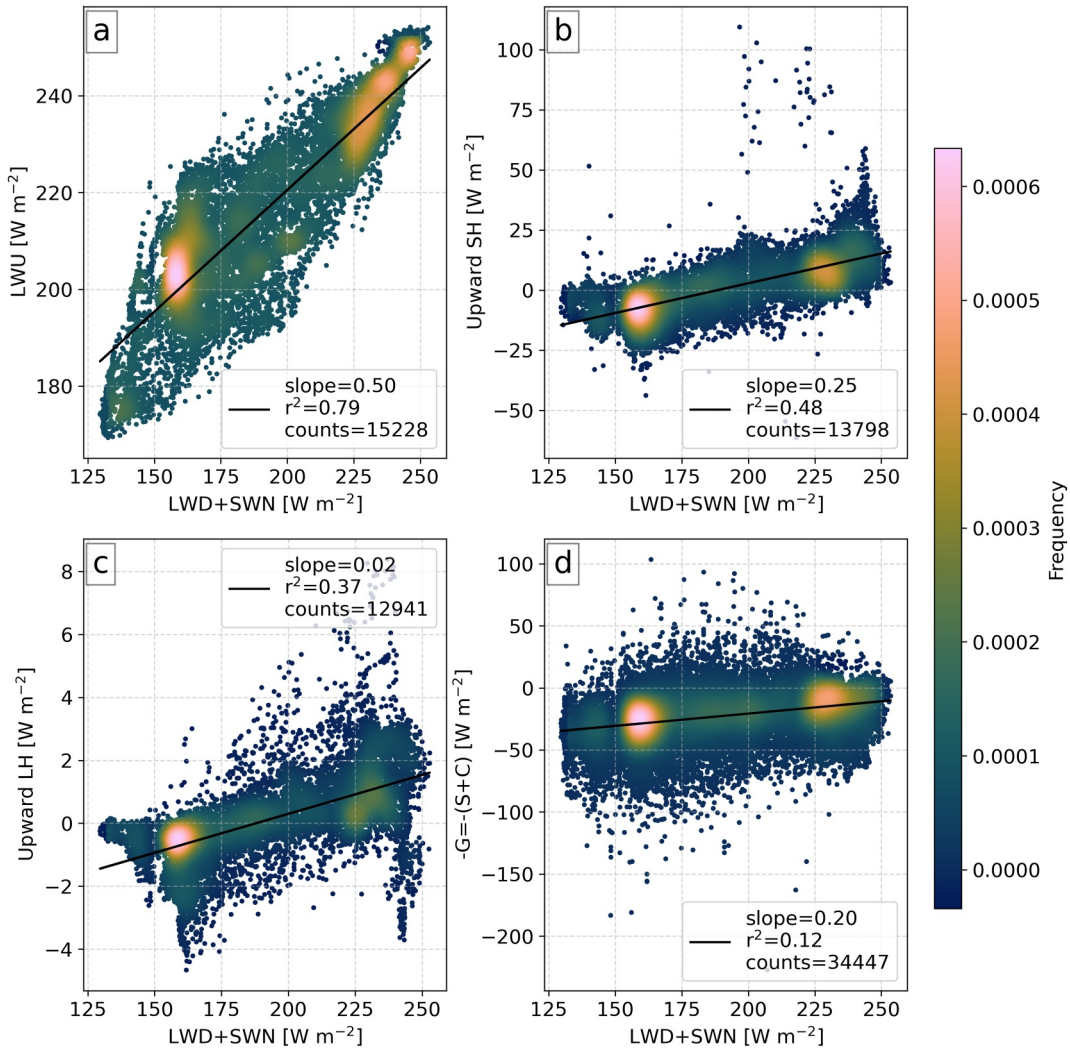

Figure 1. Observed regressions between radiative forcing and individual responding fluxes in December 2019. Shading represents frequency of occurrence, and best fit lines are found using linear least squares regression. Data from all sites are included. Each marker represents one observation at 10 min frequency.

4.1. Quantifying Flux Relationships

Observations from the month of December are used to illustrate the general process relationship approach employed in this work, that is, how changes in radiative forcing alter the responding fluxes. In December, when SW radiation is negligible, an increase of LWD into the surface increases LWU as a result of increasing the surface skin temperature. The two clusters in Figure 1a are the clear and cloudy states of the Arctic, consistent with past findings that the winter Arctic SEB is approximately bimodal (e.g., Shupe & Intrieri, 2004; Stramler et al., 2011). It is the changes in atmospheric opacity between these two states that drive the LWU response to LWD. The LWU response in December 2019 is 0.5, so 50% of incoming LWD goes to changing the surface temperature and thereby the surface emission. The r^2 value indicates the assumed linear relationship between LWU and LWD explains nearly 80% of the variance in the data.

Changes in surface temperature impact the remaining surface fluxes in December, which explain how the rest of an increase in LWD is balanced. Increases in LWD also increase the upwards SH and LH and decrease the upward conduction at the surface (Figures 1b–1d). Overall, 25% of an increase in LWD is mitigated through turbulent mixing, mainly the SH response, in December. The r^2 value is lower for SH than LWU, only 0.48, possibly because SH may not always be radiatively driven and instead may act independently of radiative forcing under certain

conditions, such as during frontal passages or from lead plumes (e.g., Grachev et al., 2013; Jozef et al., 2023). In December, 20% of an increase in LWD goes to a combination of warming the snowpack (storage) and decreasing conduction to the snow surface layer from below (Figure 1d). Changes in storage occur quickly over the 6 cm slab, whereas changes in conduction are slower as changes in the temperature take some time to propagate down into the snow and ice (e.g., Lei et al., 2014). The subsurface heat flux has an even lower r^2 value in December, 0.12, but still shows some bimodal behavior. In December, and in general throughout the MOSAiC year, LH is small at the surface over snow, and so its response to radiative forcing is also small, only 0.02 (Figure 1c).

4.2. Flux Relationships From the MOSAiC Year

We can learn details about how the atmosphere and sea ice interact over the course of the annual cycle from Figure 2, with marked differences between the winter and summer seasons. The annual variability of monthly flux relationships is given for each of the responding terms (LWU, SH, LH, G, and SWT) as well as for the total of all of these terms. In winter, when solar processes are minimal, the surface temperature and fluxes respond to changes in LWD. In summer, with substantial solar input, the surface is constrained to be near its melting point causing diminished response in some surface fluxes and an increase in melt and sub-surface energy transfer. These two regimes, and the transitions between them, are described in the following section by examining the contributions of each individual term. For context, the response of the NA flux (NA = SWN + LWN-SH-LH; Equation 5) is also provided, which must be equivalent to all energy that is transferred into and below the surface (i.e., the sum of G, SWT, and any energy going to melt snow and ice). This parameter can be compared to the responses of SWT and G alone to infer the relationship for the energy that contributes to surface melt.

Flux responses to radiative forcing are fairly consistent from October through February (Figure 2a). As was true in December (Figure 1), over months without SWD about half of any additional energy at the surface results in changes to LWU (red circles), and the remaining energy is approximately split between changes in SH (green squares) and G (orange right triangles). Mean responses of these fluxes have similar month-to-month variability, on the order of 0.10, but the subsurface heat flux has the greatest spatial variability. This fact makes sense given that subsurface heat fluxes are point measurements sampling snow and ice conditions that can vary significantly in depth, whereas the SH and LWU footprints integrate across larger areas. Mean LH responses (blue X) are low throughout winter months. In winter, when there is no SWD or melt, the NA flux (purple upward triangle) should be equal to the subsurface heat flux, or at least within measurement uncertainties, which it is.

Although the responses of some terms start to shift in the spring, the major change in flux responses occurs once the surface temperature approaches the melting point in summer and becomes relatively invariant. This dependence on temperature variability, as opposed to SWD, explains why the mean LWU response remains larger than 0.50 through May since melt onset did not occur until May 25. Apparently, the LWU response is not strongly dependent on the mean temperature state since the monthly mean response is consistent, whereas the monthly mean temperature changes from October through May (Figure 2b). In contrast, in June through August when the skin temperature was limited by surface melt, the LWU response is 0.13 or less, depending on the number of refreeze events. Once the skin temperature is consistently below freezing in September, the LWU response approaches the previous winter values.

Similar to LWU, the subsurface heat flux response begins declining in May and reaches zero in June. In order for melt to occur, the snow and ice interior must warm, which leads to smaller temperature gradients in the subsurface. In turn, smaller temperature gradients mean smaller conductive fluxes with greater radiative forcing, which results in less responsiveness of the subsurface heat flux to atmospheric forcing. Although we do not have reliable estimates of effective thermal conductivity and density coincident with the IMBs after May, the vertical and temporal temperature gradients are small near the surface after melt onset, when the surface temperature is close to the ocean temperature, so the subsurface heat flux is near zero regardless of subsurface properties.

The SH response also decreases to near zero, but this change occurs earlier in April. In March the mean SH response peaks, and the near-surface atmosphere is more frequently stable than earlier months (not shown). After March, the monthly mean SH response declines to less than 0.04 from April through July. Although SH still varied in these months, it no longer appears to be radiatively driven. In summer this is in large part because the surface temperature is constrained to near 0°C and the near-surface stability is driven by temperature advection. Interestingly, in August the turbulent fluxes have the opposite response to radiative forcing as the rest of the year, meaning that increased radiative forcing leads to enhanced downward SH. Tjernström et al. (2019) found that

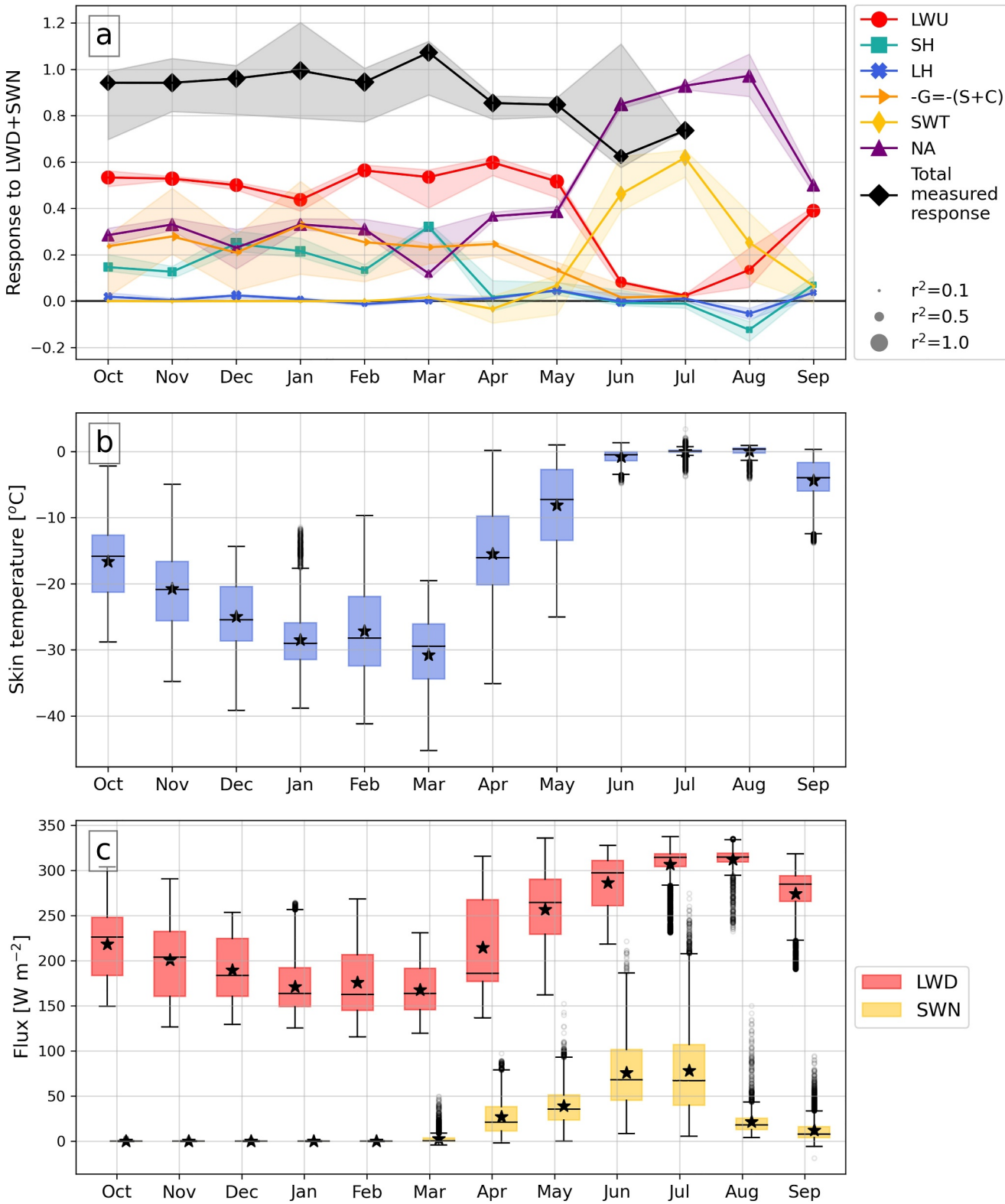

Figure 2.

warm air advection from open ocean over melting sea ice in late summer led to a shallow surface inversion because the surface temperature could not respond, and therefore SH increased into the surface. This process could explain the negative SH response in August, although we caution that measurements are only available from 10 days in this month.

After polar sunrise (15 March 2020 at *Polarstern*) SWN was initially small, and the SWT responses remain close to zero through May. In the early spring there is little SW absorbed since the albedo is relatively high before the surface starts melting (Figure 2c). In April the SWT response to radiative forcing is slightly negative due to the timing of available data and the relative variability of LWD and SWN. Data from two of the remote sites do not include the entire month. From the available data at these sites, there is less SWT when the atmosphere is cloudy and LWD + SWN is large because of the LWD. However, there is more SWT when the sky is clear but LWD + SWN is low because SWN is still small early in spring. After melt onset the snowpack albedo lowers, leading to less SW being reflected by the surface and consequently more SW being transmitted below the surface layer. In June and July about half of the net incoming energy goes to the SWT response, that is, is transmitted below the surface layer to be absorbed in the ice interior or ocean. As incoming SW declines in August and September the SWT response also decreases.

After melt onset, nearly all additional energy goes toward melt or is transmitted below the surface. Only about 60% of radiative forcing is transmitted below the surface layer, so the remaining 40% of radiative forcing results in melt at the surface. The NA response (purple upward triangle) should be balanced by the subsurface flux (G and SWT) plus melt. During June and July when we expect the subsurface heat flux response to be near zero, the difference between the NA response and SWT response should be equal to the melt response, approximately 0.3–0.4. The melt response, that is, the difference between NA and SWT response, increases to roughly 0.8 in August when there is less SWT. Overall, the balance of additional energy when the surface is melting is primarily split between melting the surface and being transmitted to the snow, ice and ocean below.

One way to test how well the SEB closes is to regress the radiative forcing against all measured responding terms summed together (Equation 9). If all fluxes were perfectly measured and perfectly representative of the same closed system, the slope would be one and the radiative forcing would be balanced by responding fluxes. Assessing this closure in winter is relatively straightforward since all terms are measured. For October through March, the mean slope from the sum of responding fluxes (Figure 2a black diamond) are within 0.1 of the expected 1. The spatial variability is largely due to variability of the subsurface heat flux due to the IMBs sampling different sea ice and snow conditions across the floe. This result indicates that the flux measurements are relatively self consistent in winter, with a tendency for the subsurface heat flux to be slightly smaller than the atmospheric residual. In spring the subsurface measurements have greater uncertainty than in winter, and the response of all measured fluxes to radiative forcing does not balance as well. The mean total measured flux response drops to 0.85 in April and May. This imbalance is likely due to the subsurface flux, either the increased uncertainty in spring and/or the fact that only two IMBs are used that do not appear to be as representative of the atmospheric measurements. Additionally, the impact of leads on SH in spring can be significant at times, but this behavior does not fit well within this framework, which assumes the system to be radiatively driven. Leads can directly influence the measurements of SH while not influencing the radiative forcing of the surface. The SWT estimates are also uncertain, although the amount of SWD that is absorbed early in spring is still rather low. Direct observations of melt were not made, so once the melt season began, the total measured response is correspondingly less than one.

4.3. Model Evaluations Against Observed Winter Responses

In the winter months, the relationships among observed surface flux forcers and responders is quite consistent and not complicated by the temporally variable contributions from solar energy transfer or seasonal variations in local

Figure 2. Annual cycles of (a) fluxes responses to radiative forcing (downwelling longwave and net shortwave; LWD + SWN), (b) observed skin temperature, and (c) individual forcing fluxes. In panel (a), markers are the mean of slopes calculated across available sites in a given month weighted by the number of points from each station, and the shading represents the total range across the sites. The size of each marker is proportional to the mean r^2 value of the best fit line for the month. The total response is calculated by regressing the sum of measured fluxes against the radiative forcing. Fluxes include upwelling longwave (LWU), sensible and latent heat (SH and LH), subsurface flux (G) that depends on storage (S) and conduction (C), transmitted shortwave (SWT), and the net atmospheric flux (NA = SWN + LWN-SH-LH). Box and whisker plots in panels (b, c) are calculated from all available observations in each month. Black horizontal lines represent medians, and black stars indicate means. Outliers are points outside of the range (Q1−1.5 × IQR, Q3 + 1.5 × IQR), where Q1 is the 25th percentile, Q3 is the 75th percentile, and IQR = Q3−Q1.

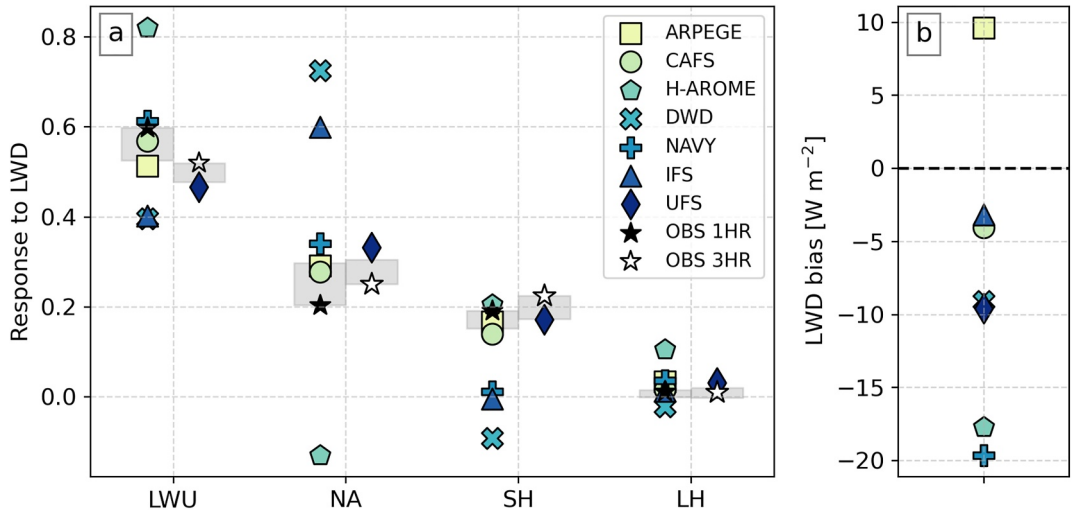

Figure 3. (a) Response of surface fluxes to radiative forcing during months without incoming solar radiation from observations (OBS) and weather forecast models listed in Table 1. Observations from the CO are plotted as stars for two different time averages and periods to match output from different models. From each model, hourly averages are taken from the grid cell nearest the *Polarstern* which was within 0.6 km of the CO. The range of slopes from all observational sites are represented as the gray rectangles for each flux. During winter the net atmospheric flux (NA) is equal to conduction at the surface. (b) The bias of LWD over winter for each model compared to observations. Regressions of individual fluxes for each model are available in Figure S1–S2 in Supporting Information S1.

sea-ice fraction. Thus, this time of year offers a prime opportunity to evaluate SEB processes in numerical models. In this section we briefly evaluate seven state of the art weather forecast models against observed flux relationships in winter. Since the observations primarily represent processes over sea ice, the comparison is limited from 15 October 2019 to 14 March 2020 when model grid cells are dominated by sea ice coverage with minimal contribution from open ocean or leads (SIC > 0.9). In nonwinter months, when model grid cells can contain a significant fraction of open ocean, the grid-cell average relationships do not represent the same processes that are represented by the observations. Thus, a proper model assessment at this time of year is much more challenging and not attempted here for all models. Instead, following the intermodel comparison during winter, we focus on one model for which we can distinguish the sub-grid processes over the sea ice from those over the open ocean surfaces to evaluate against observations for the full annual cycle.

Models represent the subsurface in a variety of ways (Table 2) that are not easily comparable to observations. For the previous analysis we used a finite interface definition of the SEB to account for observational uncertainty and to simplify the representation of specific terms. However, using a 6-cm surface slab would be complicated when comparing observations to models where one model may have a single snow layer, whereas another may have no snow layer at all. Therefore, in this section we define the SEB for the surface-atmosphere interface only and calculate the subsurface fluxes as a residual from the radiative and turbulent fluxes, LWN-SH-LH. In winter the NA flux should be equal to the subsurface heat flux, which is approximately true (within ∼10%) in Figure 2.

Historically models have struggled to represent the Arctic SEB (e.g., Rinke et al., 2006; Solomon et al., 2023), and this fact remains true for most of the models included in Figure 3. Despite the relatively simple balance of surface fluxes in winter, most of the seven models have responses to radiative forcing that are consistently outside the spatial variability of observed results, and each model displays unique relationships among the different parameters. Integrated Forecast System (IFS) (upward triangle) and DWD (X) both show too little response in LWU and near zero response in SH, balanced by too much NA response (Figure 3a). Both of these models lack snow on top of sea ice. Because there is no insulating layer on top of the sea ice, these models overestimate conduction during clear-skies, leading to a greater change in conduction as a function of changing atmospheric radiative forcing. Neglecting snow causes positive surface temperature biases during clear-skies (e.g., Batrak & Müller, 2019), such that the response in LWU to variability in radiative forcing is reduced. The near zero responses of SH are related to the damped variability of surface temperature as a function of changes in radiative

forcing, but could also point to further issues with the boundary layer schemes and/or surface-atmosphere coupling in these models.

Although the rest of the models explicitly represent snow on sea ice, there are still biases in their surface flux responses. The LWU response in H-AROME (pentagons) is overestimated by over 30%, consistent with a thicker snow pack and colder surface temperatures compared to the rest of the models (Table S1 in Supporting Information S1). Although the SH response is similar to observations, the NA response has the opposite sign as observations. Thus, in this model, conduction weakly increases when LWD increases. This behavior is partly a compensation of the relatively large LH response in H-AROME. The NAVY (+) model appears to have negligible response of SH to radiative forcing, which is compensated by slightly overestimated responses in LWU, LH, and NA. UFS (diamonds) has responses that are generally close to the observed ranges, with small overestimates of the responses in NA and LH compensated by small underestimates of the responses in SH and LWU. All three of these models have notable negative LWD biases (Figure 3b), suggesting they have too few liquid containing clouds.

Two of the models closely match the observed flux responses (CAFS and ARPEGE), but they differ in their LWD representation. The flux responses from these models are all close to observed spatial variability suggesting that they are able to reasonably represent the coupling between the atmosphere and surface in response to variability in atmospheric forcing. ARPEGE has a +10 W m^{-2} LWD bias compared to observations (Figure 3b) because it compensates for low liquid water paths with unrealistically high ice water paths (Solomon et al., 2023). CAFS has an opposite LWD bias, about −4 W m^{-2}, which is comparable to IFS and smaller than all of the other models.

Although imperfect, the SEB in CAFS has both similar forcing and responses to observations, suggesting that it is a reasonable tool for investigating the coupled Arctic atmosphere-surface system. We use output from CAFS for the remainder of this work, first comparing CAFS to observations for the full MOSAiC year at monthly resolution and then using additional CAFS output to explore variability in flux responses beyond the MOSAiC drift.

4.4. CAFS Annual Cycle Evaluation

To evaluate CAFS over the full annual cycle, we use 6 hour averages from the sea ice model output from grid cells closest to the CO and the remote observing sites. Only times when both observations are available and the grid cell has sea ice present in CAFS are included. Observations are also averaged to 6 hour samples. Averaging the observations in this way increases the LWU response by 0.05 or less throughout the year and decreases the SH response by 0.05 or less because variability in 10 min observations that can increase the slope is usually reduced after averaging. In winter the NA response decreases by 0.05, similar to SH, but its response is increased by 0.10 in June through August, likely because of smoothing out the diurnal cycle.

In addition to reasonably representing the observed all-winter responses (Figure 3), CAFS also captures much of the monthly variability in flux responses seen in observations (Figure 4a). From October through February the responses for all terms in CAFS are within the spatial variability of observations, even matching the relatively small month-to-month changes, for example, the LWU minimum in January. During winter, CAFS has less spatial variability (shading around lines) across sites for SH and NA responses than observations. This minor disagreement is likely due to parametrizations for SH in CAFS that use constant transfer coefficients while they can vary in the real world. Additionally, NA (i.e., conduction) depends on thermal conductivity that is constant in CAFS despite observed spatial variability (e.g., Macfarlane, Löwe, et al., 2023; Sledd et al., 2024).

Although the responses calculated from CAFS are quite close to those observed in winter, there are still some biases (Figure 4b). Small negative biases in LWU and SH responses are compensated by positive biases in NA response. The monthly biases in LWD (Figure 4c) range from 0 to about −20 W m^{-2} in individual winter months, which may be different from Figure 3b due to differences in initialization between the two model data sets. Along with the more frequent times with LWD less than 150 W m^{-2} (e.g., Figure 5a), CAFS also has more occurrences of LWU below 200 W m^{-2}, indicating that CAFS has a slight cold bias compared to observations. The colder surface also occurs with slightly more instances of downward SH, for example, SH < 0 W m^{-2} in Figure 5h, although the slope is still quite close to observations over all of winter and the individual months.

Beginning in spring CAFS deviates more from observations. In March the observed SH response reaches its maximum, but in CAFS the slope is less than half that. SH simulated by CAFS in March is not as variable as in the

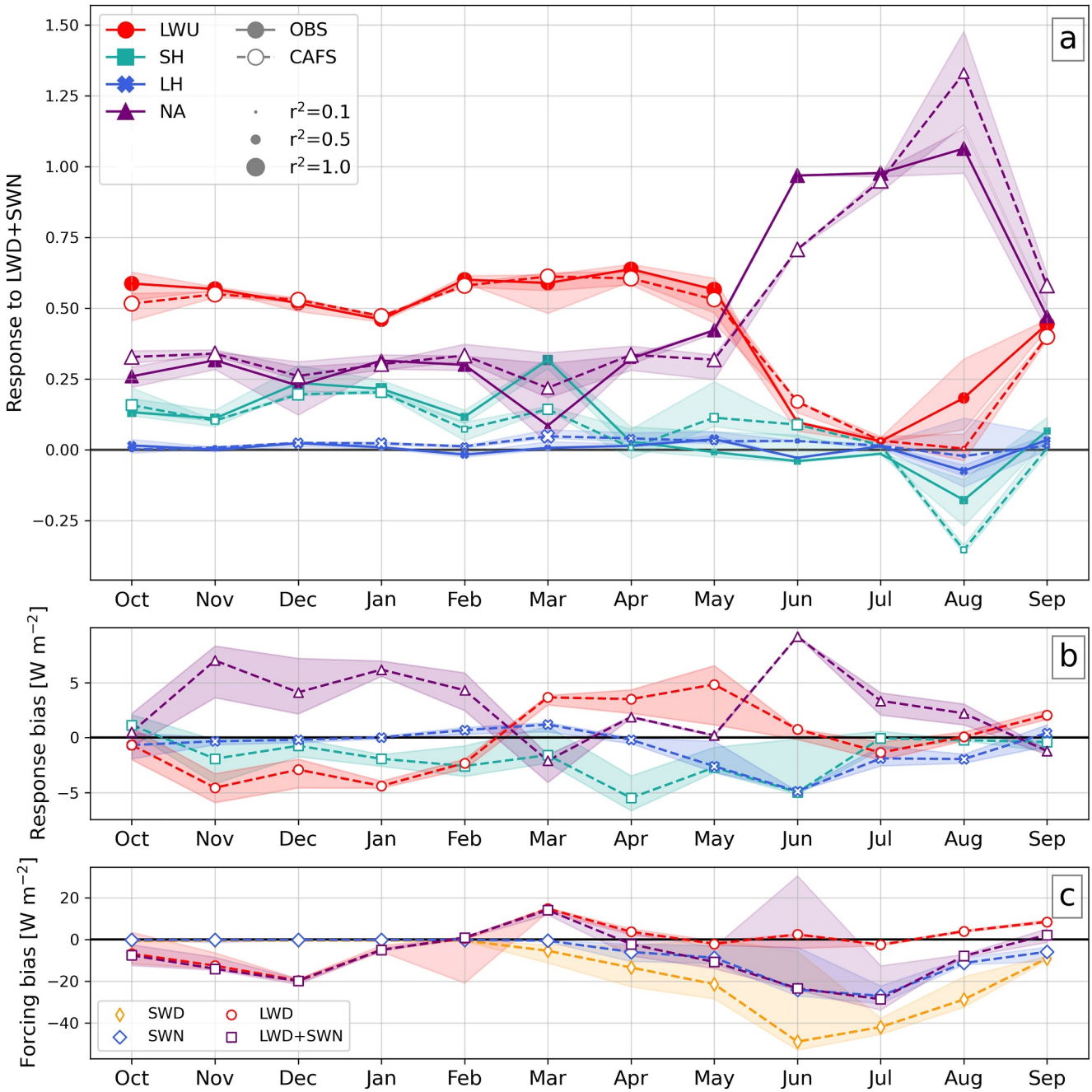

Figure 4. (a) Annual cycle of flux responses to radiative forcing from observations (solid markers and lines) and CAFS (unfilled markers and dashed lines). The net atmospheric response (NA) is equal to conduction at the surface when there is no solar insolation or melt, and otherwise it is the sum of conduction and SWT. (b) Bias in flux responses relative to observed responses (Equation 9). Markers and colors are the same as in panel (a). (c) Bias in incoming radiative fluxes relative to observations. In all plots, markers are the mean of slopes calculated across available sites in a given month, and the shading represents the total range across the sites.

observations; most of the SH flux in CAFS falls within the range of −20 to 10 W m^{-2}, whereas in the observations the range is −30 to 20 W m^{-2}. The narrower range of SH flux in CAFS is despite a larger range of radiative forcing compared to observations in March (Figure 5m). The smaller SH response is compensated by an increased NA response, which is 0.20 in CAFS compared to 0.07 in observations. March is also the only month when CAFS has a notable positive LWD bias, 15 W m^{-2}.

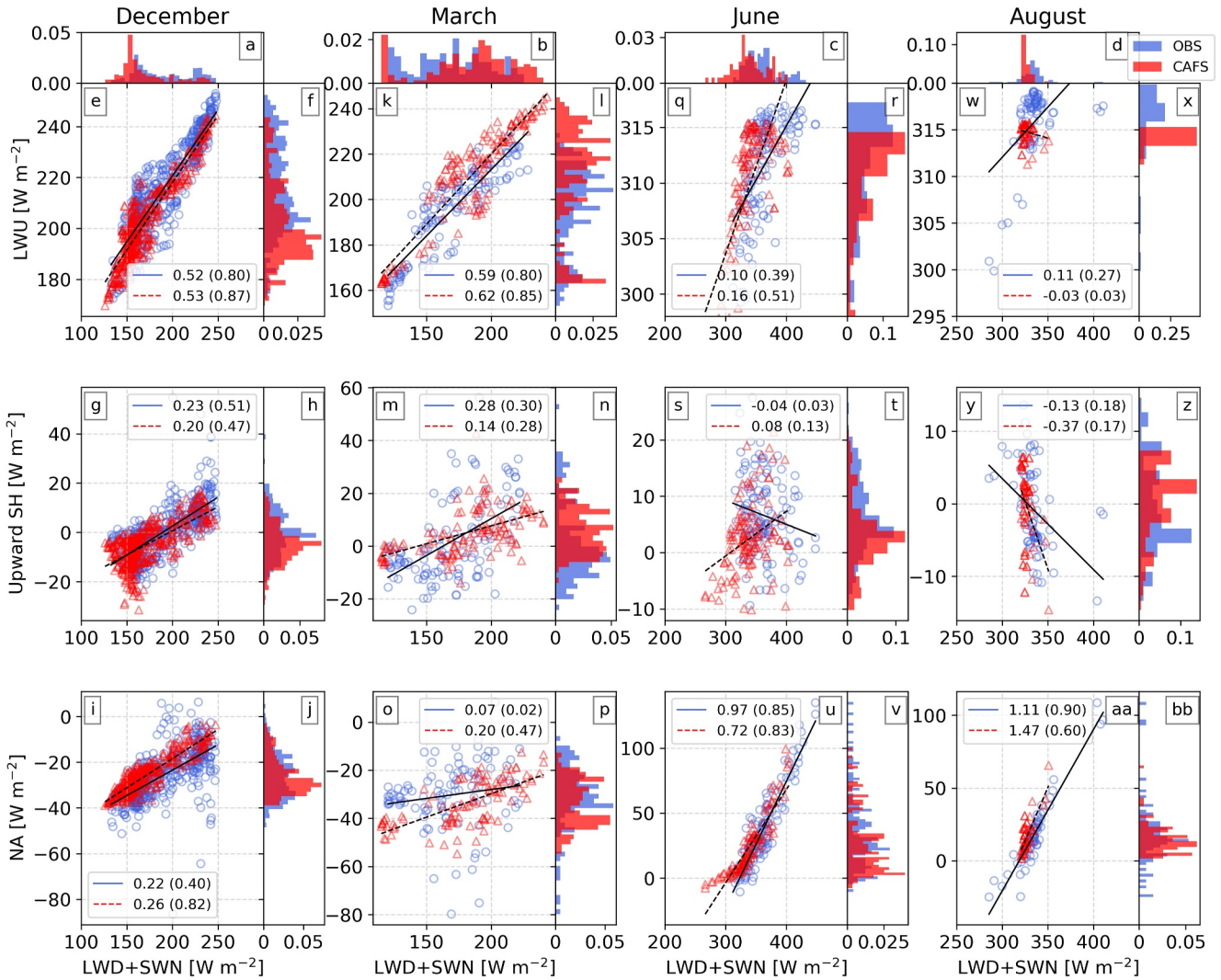

Figure 5. Joint histograms of radiative forcing (LWD + SWN) versus LWU (top row), SH (middle row), and the NA flux (bottom row) for December, March, June, and August of the MOSAiC year for observations (OBS; blue) and CAFS (red). The slopes from the best fit line, found using linear least squares regression, is given in the legend along with the r^2 value in parentheses.

The flux responses later in spring follow the general trends as in observations, albeit with a few larger differences than in winter. The LWU response from CAFS is within 0.03 of observations before declining similarly to observations in June and July. The NA response increases in CAFS after May, similar to observations, but it is initially smaller. Not as much additional energy goes to the subsurface in CAFS because there is still a positive SH response in May and June. June in particular has some periods of radiative forcing that are smaller than observations (Figure 5c). These time periods result in the surface cooling and refreezing more than in the observations, which is consistent with more downward SH that leads to a stronger slope. However, the r^2 value from regressing SH against radiative forcing is low for both observations and CAFS (<0.15), so our framework of radiatively driven fluxes may not be the most appropriate for evaluating SH in these months.

The clouds in CAFS, although well represented compared to most models, have seasonal biases that impact the radiative forcing (Figure 4c). Sedlar et al. (2020) found that clouds in CAFS were too optically thick in summer, leading to reduced SWD compared to field observations, and Figure 4c is consistent with this result. The monthly mean SWD bias in CAFS grows from −5 W m^{-2} in March up to −49 W m^{-2} in June. The negative SWD bias is partly compensated by a higher median albedo in CAFS compared to observations, resulting in SWN biases ranging from 0 to −24 W m^{-2} over the same months. The lack of LWD bias during these months could be due to

the fact that the cloud optical depth (i.e., liquid water path) increases, the sensitivity of LWD to these increases saturates before the sensitivity of SWD does (Shupe & Intrieri, 2004).

Comparing CAFS and observations in August should be done cautiously because only 10 days of observations are available in this month. In Figure 5d, CAFS simulates a smaller range of radiative forcing (~320–360 W m^{-2}) compared to observations (~290–410 W m^{-2}). The surface temperature is also more constrained in CAFS, with most of the data falling within the range of −0.6-0°C, compared to observations with a range of −3 to 0.7°C. This difference causes the CAFS LWU response to be nearly 0, whereas the observed response is 0.11 (Figure 5w). Although CAFS has the same sign of SH response as the observations, the magnitude is about twice as large as in observations but with a smaller r^2 (Figure 5y). SH in CAFS only has a slightly smaller range than the observations but over a much smaller range of radiative forcing. The steeper slope in CAFS is mainly driven by relatively more downward SH for greater LWD + SWN.

When the surface starts to transition back toward its winter state in September, CAFS flux responses are again closer to observations. The LWU and LH responses in CAFS are equal to or barely less than those from observations. In observations the SH response is once again negative in September, −0.06, whereas it is zero in CAFS. SH flux is most often slightly upward or close to zero in both observations and CAFS. However, CAFS has some occurrences of downward SH greater than observations when radiative forcing is large.

4.5. Variability of Flux Responses Across the Arctic From CAFS

In the prior section, the flux responses in CAFS were shown to be mostly within the variability of the observations for all variables in most months. This finding provides interpretive context and justification for using CAFS to expand our analysis of flux partitioning to the pan-Arctic sea ice cover, which is the focus of this section. We use CAFS forecasts from February 2018–December 2023 and separate the spatial analysis into two periods, the growth regime and the melt regime. These regimes are determined by a threshold of 1 W m^{-2} for the energy available for surface melt reported by the model. These two regimes were identified in the prior section as being the principal characteristic of the annual cycle in the analyzed process relationships observed at MOSAiC. For the following analysis, the slope between parameters within each grid cell is calculated across the CAFS domain. Different from the previous sections, we use the surface conductive flux variable, FC, from CAFS. We also subtract FC from the NA flux to calculate the response of melt and SWT together. We do not separate these subsurface terms because their representation in CAFS was not verified in the previous sections, and doing so is beyond the scope of this work. To explain some of the spatial variability in flux responses from CAFS, we use effective depth, h_{eff}:

$$h_{eff} = h_{ice} + h_{snow} * \frac{k_{ref}}{k_{snow}}, \tag{10}$$

where h_{ice} is the sea ice thickness, h_{snow} is the snow depth, k_{ref} is the reference thermal conductivity, 2 W m^{-1} K^{-1}, and k_{snow} is the snow thermal conductivity in CAFS, 0.3 W m^{-1} K^{-1}. Snow depth scaled this way can be thought of as the equivalent thickness of sea ice that would conduct the same energy as the snow. For example, snow with a thermal conductivity of 0.3 W m^{-1} K^{-1} would only conduct 15% of the same energy as sea ice.

Across the Arctic Ocean domain LWU and FC responses are to first order dictated by whether or not there is surface melt (Figures 6a–6d). Nearly 90% of LWU responses during the growth regime fall within 0.4–0.7, similar to the monthly variability from MOSAiC before melt onset when the responses fall to approximately zero. About 90% of FC responses are within 0.1–0.4 during the growth regime and nearly all responses have magnitudes less than 0.1 during the melt regime. Contours of effective depth roughly correspond to LWU and FC response intervals, where thicker effective depths correspond to greater LWU response and lower FC response. When melt occurs, the surface and near-surface ice temperatures are invariant in response to radiative forcing, such that LWU and FC responses are near zero.

Unlike LWU and FC, weak spatial variability in SH is found in CAFS during both melt and growth regimes in the Central Arctic. During the growth regime, SH has responses on the order of 0–0.10 across the Arctic (Figure 6e) as a result of typically stronger downward SH under weak atmospheric radiative forcing conditions. The SH responses during the melt regime are generally of similar magnitude to those during the growth regime (Figure 6f). Although the surface temperature cannot adjust while the surface is melting, the air temperature aloft,

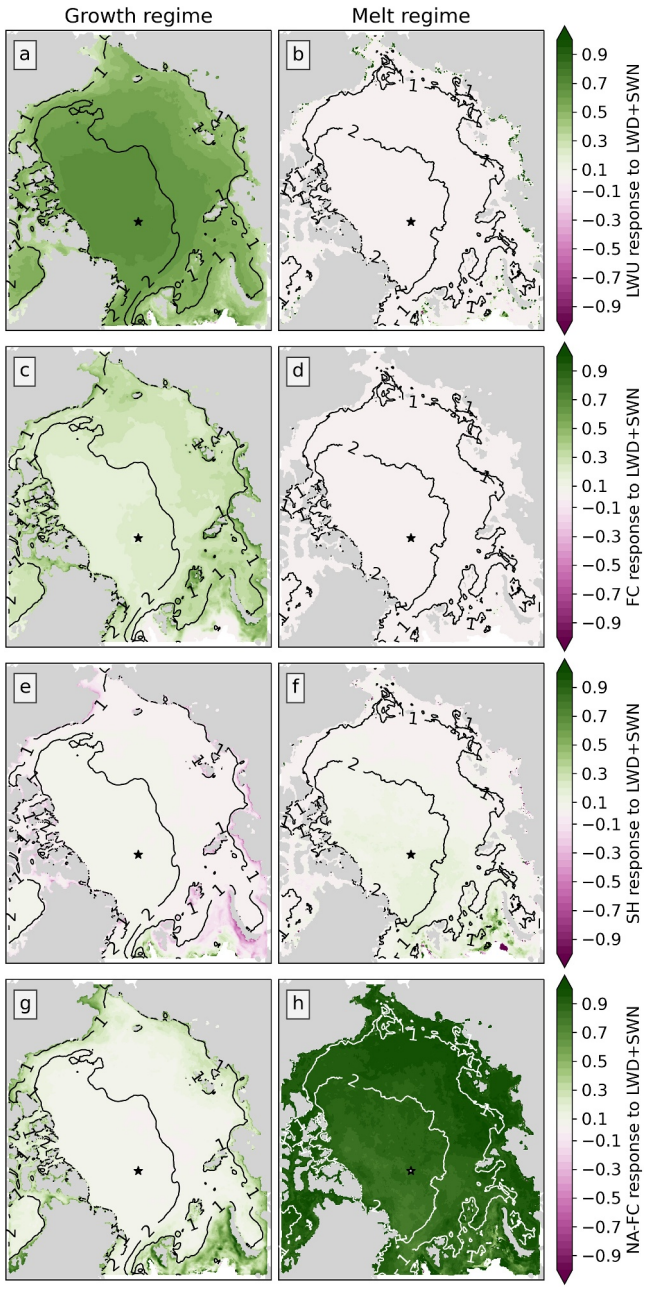

Figure 6. Flux responses to radiative forcing using CAFS data from February 2018–December 2023. Slopes in each grid cell are calculated for times when the surface is in the growth regime (left column) and melt regime (right column) based on a threshold of 1 W m^{-2} for the energy available for melt from the CAFS sea ice model. Only time steps where sea ice is present for the entire 6 hr are included. Contour lines show mean effective depth in meters. The black star indicates the grid cell used in Figure 7.

and thus ice-air temperature gradient, can still change as a result of advection. For example, warm air advected from open ocean over melting sea ice can form surface-based temperature inversions that produce additional downward SH (e.g., Tjernström et al., 2019). Cold air advected from colder sea ice over melting sea ice can have the opposite response, where there is increased upward SH flux (e.g., Vihma et al., 2005). These changes in SH coincide with different amounts of radiative forcing but are not driven by those differences, as is also the case in observations when the surface is melting. The largest SH responses during melt are in the North Atlantic sector of the Arctic (Figure 6f), where poleward advection from the warm ocean over the sea ice is common (e.g., Woods et al., 2013). Looking at regressions from an individual grid cell (87°N, 115°E; black star in Figure 6) we see that most of the data during the melt regime are around 0°W m^{-2} and insensitive to radiative forcing (Figure 7b). However, there are enough cases of more downward SH with more LWD + SWN (warm advection) and more upward SH with less LWD + SWN (cold advection) to influence the slope. Although the r^2 value is not large during the melt regime, it is actually similar to the r^2 value during the growth regime for this grid cell. Although in observations radiation may be the dominant driver of SH in winter, in CAFS advection may create the scatter in Figure 7 and drives the variability during melt.

As in the MOSAiC year, during melting conditions variability in radiative forcing is mostly partitioned into the subsurface, either contributing to melt at the surface or absorbed in the interior ice or transmitted to the ocean. When examining the response of the NA flux minus FC during the growth regime, we expect the response to be small because there is no melt and limited solar input for much of the time. The majority of data at this time falls between 0 and 0.2, with the larger values representing areas along the coast and poleward of the North Atlantic that have average effective depths equal to or less than 1 m (Figure 6g). These regions have some SW absorbed or transmitted below the surface because the growth regime can include times with solar insolation. Unsurprisingly, during the melt regime the largest response to variability in atmospheric radiative forcing is in the subsurface and melt, with over 90% of grid cells having responses greater than or equal to 0.75 (Figure 6h). The smaller responses are in the central Arctic where the effective depth is thicker and SH comprises the rest of the response.

The spatial variability of flux responses during the growth regime is influenced by snow and sea ice cover. For the flux responses shown in Figure 8 scatter increases for effective depths thinner than about a meter, and the sensitivity to effective depth saturates for effective depths thicker than about 2 m. Effective depth has a fairly constrained relationship with the LWU response (Figure 8a). Thicker snow and ice are associated with colder surface temperatures when the atmosphere is optically thin in winter, so the range of LWU is larger for thicker snow than thinner, assuming similar surface temperatures during cloud cover. Effective depth has less influence on SH and FC responses (Figures 8b and 8c). SH depends on the surface temperature because of the near surface temperature gradient, but as previously discussed, SH can also be driven by advective forcing, leading to scatter. This noise needs to be balanced in another flux, which is evident in the scatter in FC responses (Figure 8c). We would expect thicker effective depths to decrease FC and its response to radiative forcing, which is broadly true for the more frequent values (brighter colors in Figure 8c), but it is apparent that for thinner effective depths there is compensation from the SH response. The NA-FC response represents SW penetration below the surface during the growth regime. Because SW transmission through snow is quite small, the NA-FC response is most sensitive to effective depth when there is little to no snow cover.

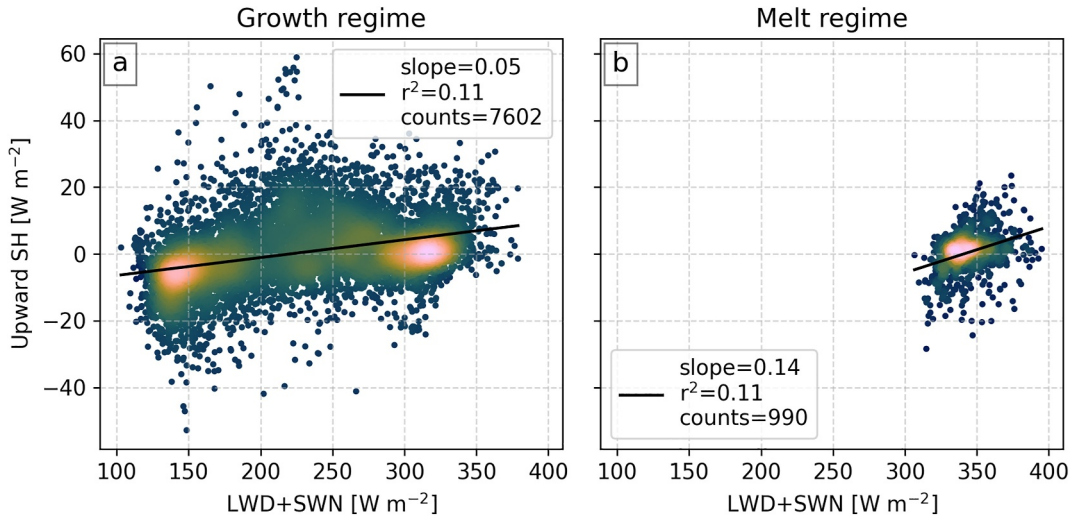

Figure 7. Regressions of radiative forcing versus SH from one grid cell (87N, 115E; black star in Figure 6) in CAFS during melt and growth regimes using data from February 2018–December 2023.

5. Discussion and Conclusions

When studying the responses of surface fluxes over sea ice to radiative forcing, we have found that broadly there are two primary regimes, an ice growth regime and an ice melt regime. The ice growth regime occurs during polar night and transitions to the ice melt regime as increasing solar insolation pushes the surface toward the

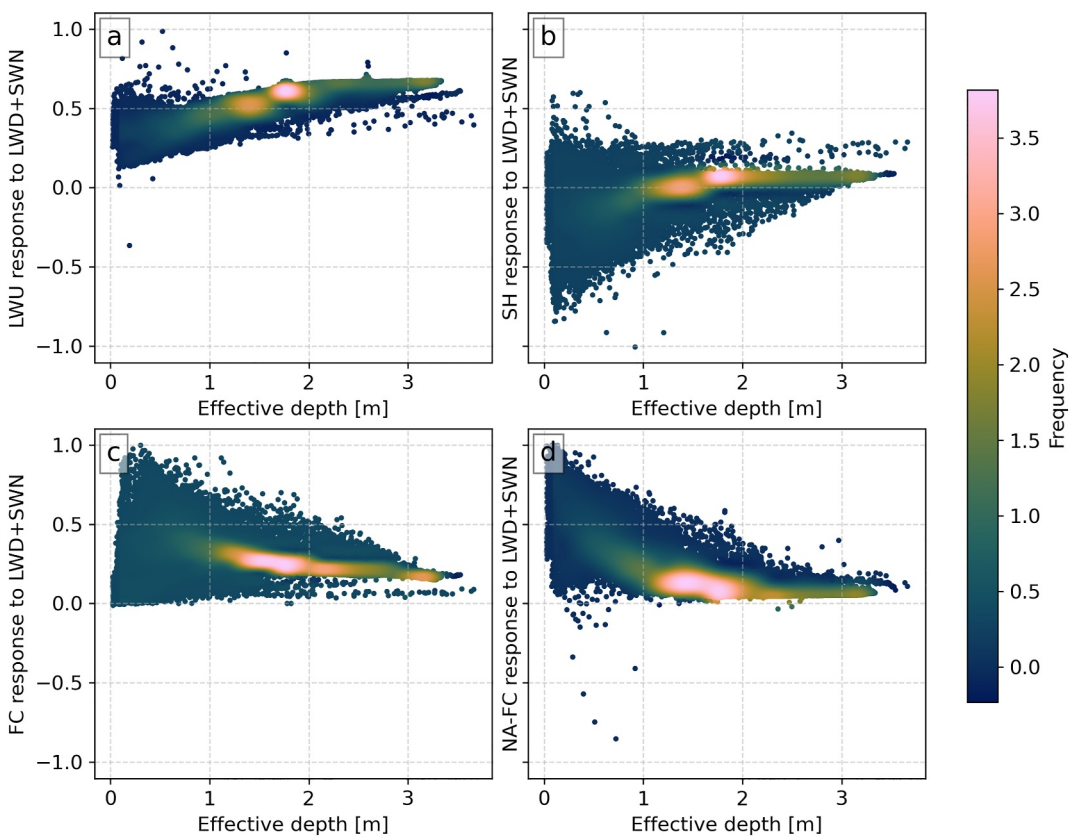

Figure 8. Regressions of flux responses to radiative forcing from the growth regime in CAFS (Figures 6a–6c, 6e and 6g) versus mean effective depth. Grid cells with mean sea ice concentration less than 0.05 are excluded.

unresponsive temperature of melt. The system transitions from the melt regime back to the growth regime in fall when the surface stops melting and the surface temperature can again respond. The growth regime is defined by the interdependence of surface fluxes, consistent with past work (Miller et al., 2017; Sterk et al., 2013). From MOSAiC observations, roughly half of any additional radiative forcing is partitioned into a surface temperature response (LWU), whereas the rest is partitioned into SH and FC. In contrast, during the melt regime most additional energy contributes to melt or is transmitted below the surface. Although the precise magnitude of the responses may be sensitive to measurement uncertainty, the relative magnitudes are robust. Persson et al. (2002) found that whether or not the surface temperature could respond to radiative forcing was key to melt onset at SHEBA, although they did not quantify this relationship. The seasonal change in responses contrasts with Miller et al. (2017) who used this framework with observations from the central Greenland ice sheet where melt is negligible.

Although we can learn about the coupling between surface fluxes from the driver-responder framework, it has limitations when considering the annual cycle of SH responses. In winter over sea ice, SH and the near surface stability are driven by radiation. However, SH can also respond to large-scale advection. This finding is consistent with Jozef et al. (2023) who found that in summer stable boundary layers were connected with advective processes rather than radiative cooling during MOSAiC. Tjernström et al. (2019) also observed greater SH into the surface associated with warm advection over melting sea ice that coincided with a change in radiation without being driven by it, similar to the negative SH response in August observed at MOSAiC. Although this work has increased understanding of how SH is related to radiative forcing, the SH response to advection does not fit in this framework. Furthermore, it means that additional diagnostics are needed to evaluate and improve summertime turbulent fluxes in models.

The regime dependence on surface temperature is confirmed by calculating slopes from CAFS when there is and is not surface melt occurring. We believe CAFS is a reasonable model for helping provide a more generalized view of these SEB process relationships across the Arctic. The additional annual analysis of CAFS is possible because output is available from its sea ice model, allowing for a more comparable evaluation between it and the observations. Because the surface flux observations from MOSAiC represent snow and sea ice, comparing them to model grid cell averages with open ocean can lead to incorrect conclusions of model performance. Therefore in order to evaluate models against MOSAiC observations it is recommended to use output from the sea ice model, including surface fluxes, near surface wind, and skin temperature. The annual model evaluation presented here could be repeated for additional models if this output were more often available. One goal of this work is to provide process-orientated diagnostics for model evaluation and development. Such diagnostics are an important tool to assess if models are simulating the right responses for the right reasons.

Mean biases in fluxes and biases in flux responses appear, to first order, to be independent in models: that is, being unbiased in the forcing does not guarantee accuracy in the partitioning of the response (e.g., IFS) and being biased in the forcing does not alone prevent accuracy in reflecting the responses (e.g., ARPEGE). We therefore assert that the framework presented here can help isolate errors in the coupling between ice and atmospheric components of coupled systems independent from the mean climate state of the model. For example, one key area for model improvement is explicitly modeling snow on sea ice. Past work has identified snow as important for accurately modeling skin temperature and has pointed toward the benefit of more complex snow schemes (e.g., Arduini et al., 2022; Batrak & Müller, 2019). Models without snow (IFS and DWD) have unrealistic flux responses, but so does the model with a sophisticated snow scheme and 12 snow layers (H-AROME). Correct snow depth is also important for the SEB, which requires accurate estimates of precipitation and snow redistribution in coupled models.

MOSAiC only represents one drift track during 1 year. The Arctic-wide results from CAFS improve confidence in the relationships from MOSAiC being robust, even if there is some spatial variability in the responses for a given flux in a particular month. However, CAFS is not without its own biases. Future work should investigate the flux response to radiative forcing from other field campaigns, for example, SHEBA. Additional observations in regions not sampled by MOSAiC, such as marginal seas, would also be worthwhile. These regions have different responses than MOSAiC observations in CAFS but it is unclear if they are due to model deficiencies or unique physical processes.

Data Availability Statement

Winter-only model data are available from https://yopp.met.no/, except for UFS data that is available from Sledd and Cox (2024). Additional CAFS output is available from Solomon et al. (2025). ASFS and Met City measurements are available from C. Cox et al. (2023), C. Cox et al. (2023a, 2023b, 2023c). SIMB data are available from Perovich et al. (2022). SIMBA data are available from Lei, Cheng, Hoppmann, et al. (2022a, 2022b), Lei, Cheng, Zuo, and Hoppmann (2022a, 2022b), Lei et al. (2021). Derived snow and ice thermal properties are available from Sledd et al. (2023).

Acknowledgments
Data sets used in this article were produced as part of the international Multidisciplinary drifting Observatory for the Study of the Arctic Climate (MOSAiC) with the tag MOSAiC20192020 and the Project_ID: AWI_PS122_00. Some data used here were collected by the Atmospheric Radiation Measurement User Facility, a DOE Office of Science user facility operated by the Biological and Environmental Research program. We thank all people involved in the expedition of the Research Vessel *Polarstern* (Knust, 2017) during MOSAiC in 2019–2020 as listed in Nixdorf et al. (2021). This work was supported by DOE Atmospheric System Research Program (DE-SC0021341), the National Science Foundation (OPP-1724551), NOAA's Physical Sciences Laboratory (PSL) (NOAA Cooperative Agreement NA22OAR4320151) and NOAA's Global Ocean Monitoring and Observing Program (GOMO) (FundRef https://doi.org/10.13039/100018302). The authors appreciate assistance with the UFS experimental simulation data from partners at NOAA/EMC and Lynker, in particular Nick Szapiro. The authors additionally thank David Clemens-Sewall and the anonymous reviewers for feedback on the manuscript. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of NOAA or the U.S. Department of Commerce.

References

Arduini, G., Keeley, S., Day, J. J., Sandu, I., Zampieri, L., & Balsamo, G. (2022). On the importance of representing snow over sea-ice for simulating the arctic boundary layer. *Journal of Advances in Modeling Earth Systems*, *14*(7), e2021MS002777. https://doi.org/10.1029/2021ms002777

Barton, N., Metzger, E. J., Reynolds, C. A., Ruston, B., Rowley, C., Smedstad, O. M., et al. (2021). The navy's earth system prediction capability: A new global coupled atmosphere-ocean-sea ice prediction system designed for daily to subseasonal forecasting. *Earth and Space Science*, *8*(4), e2020EA001199. https://doi.org/10.1029/2020ea001199

Batrak, Y., & Müller, M. (2019). On the warm bias in atmospheric reanalyses induced by the missing snow over arctic sea-ice. *Nature Communications*, *10*(1), 4170. https://doi.org/10.1038/s41467-019-11975-3

Bazile, E., Azouz, N., Napoly, A., & Loo, C. (2020). Impact of the 1d sea-ice model gelato in the global model arpege (Tech. Rep.). *Technical Reports Series*.

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., et al. (2017). The harmonie–arome model configuration in the aladin–hirlam nwp system. *Monthly Weather Review*, *145*(5), 1919–1935. https://doi.org/10.1175/mwr-d-16-0417.1

Castellani, G., Veyssière, G., Karcher, M., Stroeve, J., Banas, S. N., Bouman, A. H., et al. (2022). Shine a light: Under-ice light and its ecological implications in a changing arctic ocean. *Ambio*, *51*(2), 307–317. https://doi.org/10.1007/s13280-021-01662-3

Cox, C., Gallagher, M., Shupe, M., Persson, O., Blomquist, B., Grachev, A., et al. (2023). Met city meteorological and surface flux measurements (level 3 final), multidisciplinary drifting observatory for the study of arctic climate (MOSAiC), central arctic, october 2019 - September 2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A2PV6B83F

Cox, C., Gallagher, M., Shupe, M., Persson, O., Grachev, A., Solomon, A., et al. (2023a). Atmospheric surface flux station #40 measurements (level 3 final), multidisciplinary drifting observatory for the study of arctic climate (MOSAiC), central arctic, october 2019 - September 2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A25X25F0P

Cox, C., Gallagher, M., Shupe, M., Persson, O., Grachev, A., Solomon, A., et al. (2023b). Atmospheric surface flux station #50 measurements (level 3 final), multidisciplinary drifting observatory for the study of arctic climate (MOSAiC), central arctic, october 2019 - September 2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A2XD0R00S

Cox, C., Gallagher, M., Shupe, M., Persson, O., Grachev, A., Solomon, A., et al. (2023c). Atmospheric surface flux station #30 measurements (level 3 final), multidisciplinary drifting observatory for the study of arctic climate (MOSAiC), central arctic, october 2019 - September 2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A2FF3M18K

Cox, C. J., Gallagher, M. R., Shupe, M. D., Persson, P. O. G., Solomon, A., Fairall, C. W., et al. (2023). Continuous observations of the surface energy budget and meteorology over the arctic sea ice during MOSAiC. *Scientific Data*, *10*(1), 519. https://doi.org/10.1038/s41597-023-02415-5

Curry, J. A., Schramm, J. L., Rossow, W. B., & Randall, D. (1996). Overview of arctic cloud and radiation characteristics. *Journal of Climate*, *9*(8), 1731–1764. https://doi.org/10.1175/1520-0442(1996)009<1731:ooacar>2.0.co;2

Day, J. J., Arduini, G., Sandu, I., Magnusson, L., Beljaars, A., Balsamo, G., et al. (2020). Measuring the impact of a new snow model using surface energy budget process relationships. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2020MS002144. https://doi.org/10.1029/2020ms002144

Foken, T. (2008). The energy balance closure problem: An overview. *Ecological Applications*, *18*(6), 1351–1367. https://doi.org/10.1890/06-0922.1

Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., & Persson, P. O. G. (2013). The critical richardson number and limits of applicability of local similarity theory in the stable boundary layer. *Boundary-Layer Meteorology*, *147*(1), 51–82. https://doi.org/10.1007/s10546-012-9771-0

Haiden, T., Janousek, M., Vitart, F., Bouallègue, Z. B., Ferranti, L., Prates, F., & Richardson, D. (2018). *Evaluation of ecmwf forecasts, including the 2018 upgrade*. European Centre for Medium Range Weather Forecasts.

Huwald, H., Higgins, C. W., Boldi, M.-O., Bou-Zeid, E., Lehning, M., & Parlange, M. B. (2009). Albedo effect on radiative errors in air temperature measurements. *Water Resources Research*, *45*(8). https://doi.org/10.1029/2008wr007600

Inoue, J., Sato, K., Rinke, A., Cassano, J. J., Fettweis, X., Heinemann, G., et al. (2021). Clouds and radiation processes in regional climate models evaluated using observations over the ice-free arctic ocean. *Journal of Geophysical Research: Atmospheres*, *126*(1), e2020JD033904. https://doi.org/10.1029/2020jd033904

Intrieri, J., Fairall, C., Shupe, M., Persson, P., Andreas, E., Guest, P., & Moritz, R. (2002). An annual cycle of arctic surface cloud forcing at sheba. *Journal of Geophysical Research*, *107*(C10), SHE–13. https://doi.org/10.1029/2000jc000439

Jackson, K., Wilkinson, J., Maksym, T., Meldrum, D., Beckers, J., Haas, C., & Mackenzie, D. (2013). A novel and low-cost sea ice mass balance buoy. *Journal of Atmospheric and Oceanic Technology*, *30*(11), 2676–2688. https://doi.org/10.1175/jtech-d-13-00058.1

Jacobs, N. A. (2021). Open innovation and the case for community model development. *Bulletin of the American Meteorological Society*, *102*(10), E2002–E2011. https://doi.org/10.1175/bams-d-21-0030.1

Jozef, G. C., Cassano, J. J., Dahlke, S., Dice, M., Cox, C. J., & de Boer, G. (2023). Thermodynamic and kinematic drivers of atmospheric boundary layer stability in the central arctic during the multidisciplinary drifting observatory for the study of arctic climate (mosaic). *Atmospheric Chemistry and Physics*, *23*(20), 13087–13106. https://doi.org/10.5194/acp-23-13087-2023

Jung, T., Gordon, N. D., Bauer, P., Bromwich, D. H., Chevallier, M., Day, J. J., et al. (2016). Advancing polar prediction capabilities on daily to seasonal time scales. *Bulletin of the American Meteorological Society*, *97*(9), 1631–1647. https://doi.org/10.1175/bams-d-14-00246.1

Knust, R. (2017). Polar research and supply vessel polarstern operated by the alfred-wegener-institute. *Journal of large-scale research facilities JLSRF*, *3*, A119. https://doi.org/10.17815/jlsrf-3-163

Lei, R., Cheng, B., Hoppmann, M., Zhang, F., Zuo, G., Hutchings, J. K., et al. (2022). Seasonality and timing of sea ice mass balance and heat fluxes in the arctic transpolar drift during 2019–2020. *Elementa Science of the Anthropocene*, *10*(1), 000089. https://doi.org/10.1525/elementa.2021.000089

Lei, R., Cheng, B., Hoppmann, M., & Zuo, G. (2021). Temperature and heating induced temperature difference measurements from simba-type sea ice mass balance buoy 2019t67, deployed during MOSAiC 2019/20 [Dataset]. *PANGEA*. https://doi.org/10.1594/PANGAEA.938128

Lei, R., Cheng, B., Hoppmann, M., Zuo, G., & Lan, M. (2022a). Temperature and heating induced temperature difference measurements from simba-type sea ice mass balance buoy 2019t62, deployed during MOSAiC 2019/20 [Dataset]. *PANGEA*. https://doi.org/10.1594/PANGAEA.940231

Lei, R., Cheng, B., Hoppmann, M., Zuo, G., & Lan, M. (2022b). Temperature and heating induced temperature difference measurements from simba-type sea ice mass balance buoy 2019t70, deployed during MOSAiC 2019/20 [Dataset]. *PANGEA*. https://doi.org/10.1594/PANGAEA.940659

Lei, R., Cheng, B., Zuo, G., & Hoppmann, M. (2022a). Temperature and heating induced temperature difference measurements from simba-type sea ice mass balance buoy 2019t63, deployed during MOSAiC 2019/20 [Dataset]. *PANGEA*. https://doi.org/10.1594/PANGAEA.940593

Lei, R., Cheng, B., Zuo, G., & Hoppmann, M. (2022b). Temperature and heating induced temperature difference measurements from simba-type sea ice mass balance buoy 2019t65, deployed during MOSAiC 2019/20 [Dataset]. *PANGEA*. https://doi.org/10.1594/PANGAEA.940634

Lei, R., Li, N., Heil, P., Cheng, B., Zhang, Z., & Sun, B. (2014). Multiyear sea ice thermal regimes and oceanic heat flux derived from an ice mass balance buoy in the arctic ocean. *Journal of Geophysical Research: Oceans*, *119*(1), 537–547. https://doi.org/10.1002/2012jc008731

Macfarlane, A. R., Dadic, R., Smith, M. M., Light, B., Nicolaus, M., Henna-Reetta, H., et al. (2023b). Evolution of the microstructure and reflectance of the surface scattering layer on melting, level arctic sea ice. *Elementa Science of the Anthropocene*, *11*(1), 00103. https://doi.org/10.1525/elementa.2022.00103

Macfarlane, A. R., Löwe, H., Gimenes, L., Wagner, D. N., Dadic, R., Ottersberg, R., et al. (2023). Temporospatial variability of snow's thermal conductivity on arctic sea ice. *The Cryosphere*, *17*(12), 5417–5434. https://doi.org/10.5194/tc-17-5417-2023

Maslowski, W., Clement Kinney, J., Higgins, M., & Roberts, A. (2012). The future of arctic sea ice. *Annual Review of Earth and Planetary Sciences*, *40*(1), 625–654. https://doi.org/10.1146/annurev-earth-042711-105345

Meier, W. N., & Stroeve, J. (2022). An updated assessment of the changing arctic sea ice cover. *Oceanography*, *35*(3/4), 10–19. https://doi.org/10.5670/oceanog.2022.114

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., Lenaerts, J. T., 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*(10), 4777–4796. https://doi.org/10.1029/2017jd027377

Nicolaus, M., Perovich, D. K., Spreen, G., Granskog, M. A., von Albedyll, L., Angelopoulos, M., et al. (2022). Overview of the MOSAiC expedition: Snow and sea ice. *Elementa Science of the Anthropocene*, *10*(1), 000046. https://doi.org/10.1525/elementa.2021.000046

Nixdorf, U., Dethloff, K., Rex, M., Shupe, M., Sommerfeld, A., & Perovich, D. K. (2021). *Mosaic extended acknowledgement*. Zenodo.

Perovich, D., Raphael, I., Moore, R., Clemens-Sewall, D., Lei, R., Sledd, A., & Polashenski, C. (2023). Sea ice temperature and mass balance measurements during the MOSAiC drift campaign. *Elementa: Science of the Anthropocene*, *11*(1). https://doi.org/10.1525/elementa.2023.00017

Perovich, D., Raphael, I., Moore, R., Polashenski, D. C.-S. C., & Planck, C. (2022). Measurements of ice mass balance and temperature from autonomous seasonal ice mass balance buoys in the arctic ocean, 2019-2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A2SB3X05G

Persson, P. O. G. (2012). Onset and end of the summer melt season over sea ice: Thermal structure and surface energy perspective from SHEBA. *Climate Dynamics*, *39*(6), 1349–1371. https://doi.org/10.1007/s00382-011-1196-9

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), SHE–21. https://doi.org/10.1029/2000jc000705

Persson, P. O. G., Shupe, M. D., Perovich, D., & Solomon, A. (2017). Linking atmospheric synoptic transport, cloud phase, surface energy fluxes, and sea-ice growth: Observations of midwinter SHEBA conditions. *Climate Dynamics*, *49*(4), 1341–1364. https://doi.org/10.1007/s00382-016-3383-1

Pithan, F., Ackerman, A., Angevine, W. M., Hartung, K., Ickes, L., Kelley, M., et al. (2016). Select strengths and biases of models in representing the arctic winter boundary layer over sea ice: The larcform 1 single column model intercomparison. *Journal of Advances in Modeling Earth Systems*, *8*(3), 1345–1357. https://doi.org/10.1002/2016ms000630

Pithan, F., Medeiros, B., & Mauritsen, T. (2014). Mixed-phase clouds cause climate model biases in arctic wintertime temperature inversions. *Climate Dynamics*, *43*(1–2), 289–303. https://doi.org/10.1007/s00382-013-1964-9

Planck, C. J., Whitlock, J., Polashenski, C., & Perovich, D. (2019). The evolution of the seasonal ice mass balance buoy. *Cold Regions Science and Technology*, *165*, 102792. https://doi.org/10.1016/j.coldregions.2019.102792

Polashenski, C., Perovich, D., Richter-Menge, J., & Elder, B. (2011). Seasonal ice mass-balance buoys: Adapting tools to the changing arctic. *Annals of Glaciology*, *52*(57), 18–26. https://doi.org/10.3189/172756411795931516

Rabe, B., Cox, C. J., Fang, Y.-C., Goessling, H., Granskog, M. A., Hoppmann, M., et al. (2024). The MOSAiC distributed network: Observing the coupled arctic system with multidisciplinary, coordinated platforms. Elementa. *Science of the Anthropocene*, *12*(1). https://doi.org/10.1525/elementa.2023.00103

Rinke, A., Dethloff, K., Cassano, J. J., Christensen, J., Curry, J. A., Du, P., et al. (2006). Evaluation of an ensemble of arctic regional climate models: Spatiotemporal fields during the SHEBA year. *Climate Dynamics*, *26*(5), 459–472. https://doi.org/10.1007/s00382-005-0095-3

Sedlar, J., Tjernström, M., Rinke, A., Orr, A., Cassano, J., Fettweis, X., et al. (2020). Confronting arctic troposphere, clouds, and surface energy budget representations in regional climate models with observations. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD031783. https://doi.org/10.1029/2019jd031783

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*(3), 616–628. https://doi.org/10.1175/1520-0442(2004)017<0616:crfota>2.0.co;2

Shupe, M. D., Rex, M., Blomquist, B., Persson, P. O. G., Schmale, J., Uttal, T., et al. (2022). Overview of the MOSAiC expedition: Atmosphere. *Elementa Science of the Anthropocene*, *10*(1), 00060. https://doi.org/10.1525/elementa.2021.00060

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

Sledd, A., & Cox, C. J. (2024). NOAA UFS data for surface energy balance responses to radiative forcing in the central arctic from MOSAiC and models [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.13840705

Sledd, A., Shupe, M., Solomon, A., Perovich, C. C. D., & Lei, R. (2023). Winter snow and sea ice thermal conductivity, density, and conductive heat flux profiles from the multidisciplinary drifting observatory for the study of arctic climate (MOSAiC) expedition, central arctic, october 2019 - March 2020 [Dataset]. *Arctic Data Center*. https://doi.org/10.18739/A2NK3669K

Sledd, A., Shupe, M. D., Solomon, A., Cox, C. J., Perovich, D., & Lei, R. (2024). Snow thermal conductivity and conductive flux in the central arctic: Estimates from observations and implications for models. *Elementa: Science of the Anthropocene*, *12*(1). https://doi.org/10.1525/elementa.2023.00086

Solomon, A., Bazile, E., Day, J., Frank, H. P., Remes, T., Tolstykh, M., & Doyle, J. D. (2025). Hourly forecast data of near-surface and surface fields for the wintertime period 15 October 2019 to 15 March 2020 at the location of the Multidisciplinary drifting Observatory for the Study of Arctic Climate expedition from 7 operational and quasi-operat [dataset]. *PANGAEA*. https://doi.org/10.1594/PANGAEA.974438

Solomon, A., Cox, C. J., Intrieri, J. M., Persson, P., de Boer, G., Shupe, M. D., et al. (2024). Description and evaluation of the NOAA experimental coupled arctic forecast system (cafs) model. *NOAA Technical Memorandum, OAR-PSL 318*.

Solomon, A., Shupe, M. D., Svensson, G., Barton, N. P., Batrak, Y., Bazile, E., et al. (2023). The winter central arctic surface energy budget: A model evaluation using observations from the MOSAiC campaign. *Elementa Science of the Anthropocene*, *11*(1), 00104. https://doi.org/10.1525/elementa.2022.00104

Sotiropoulou, G., Sedlar, J., Forbes, R., & Tjernström, M. (2016). Summer arctic clouds in the ecmwf forecast model: An evaluation of cloud parametrization schemes. *Quarterly Journal of the Royal Meteorological Society*, *142*(694), 387–400. https://doi.org/10.1002/qj.2658

Steeneveld, G., Van de Wiel, B., & Holtslag, A. (2006). Modelling the arctic stable boundary layer and its coupling to the surface. *Boundary-Layer Meteorology*, *118*(2), 357–378. https://doi.org/10.1007/s10546-005-7771-z

Sterk, H., Steeneveld, G., & Holtslag, A. (2013). The role of snow-surface coupling, radiation, and turbulent mixing in modeling a stable boundary layer over arctic sea ice. *Journal of Geophysical Research: Atmospheres*, *118*(3), 1199–1217. https://doi.org/10.1002/jgrd.50158

Stramler, K., Del Genio, A. D., & Rossow, W. B. (2011). Synoptically driven arctic winter states. *Journal of Climate*, *24*(6), 1747–1762. https://doi.org/10.1175/2010jcli3817.1

Tao, R., Nicolaus, M., Katlein, C., Anhaus, P., Hoppmann, M., Spreen, G., et al. (2024). Seasonality of spectral radiative fluxes and optical properties of arctic sea ice during the spring–summer transition. *Elementa: Science of the Anthropocene*, *12*(1). https://doi.org/10.1525/elementa.2023.00130

Thielke, L., Fuchs, N., Spreen, G., Tremblay, B., Birnbaum, G., Huntemann, M., et al. (2023). Preconditioning of summer melt ponds from winter sea ice surface temperature. *Geophysical Research Letters*, *50*(4), e2022GL101493. https://doi.org/10.1029/2022gl101493

Tjernström, M., Shupe, M. D., Brooks, I. M., Achtert, P., Prytherch, J., & Sedlar, J. (2019). Arctic summer airmass transformation, surface inversions, and the surface energy budget. *Journal of Climate*, *32*(3), 769–789. https://doi.org/10.1175/jcli-d-18-0216.1

Tjernström, M., Svensson, G., Magnusson, L., Brooks, I. M., Prytherch, J., Vüllers, J., & Young, G. (2021). Central arctic weather forecasting: Confronting the ecmwf ifs with observations from the arctic ocean 2018 expedition. *Quarterly Journal of the Royal Meteorological Society*, *147*(735), 1278–1299. https://doi.org/10.1002/qj.3971

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), 255–276. https://doi.org/10.1175/1520-0477(2002)083<0255:shbota>2.3.co;2

Vihma, T., Lüpkes, C., Hartmann, J., & Savijärvi, H. (2005). Observations and modelling of cold-air advection over arctic sea ice. *Boundary-Layer Meteorology*, *117*(2), 275–300. https://doi.org/10.1007/s10546-004-6005-0

Walden, V. P., Hudson, S. R., Cohen, L., Murphy, S. Y., & Granskog, M. A. (2017). Atmospheric components of the surface energy budget over young sea ice: Results from the N-ICE2015 campaign. *Journal of Geophysical Research: Atmospheres*, *122*(16), 8427–8446. https://doi.org/10.1002/2016jd026091

Woods, C., Caballero, R., & Svensson, G. (2013). Large-scale circulation associated with moisture intrusions into the arctic during winter. *Geophysical Research Letters*, *40*(17), 4717–4721. https://doi.org/10.1002/grl.50912

Zängl, G., Reinert, D., Rípodas, P., & Baldauf, M. (2015). The icon (icosahedral non-hydrostatic) modelling framework of dwd and mpi-m: Description of the non-hydrostatic dynamical core. *Quarterly Journal of the Royal Meteorological Society*, *141*(687), 563–579. https://doi.org/10.1002/qj.2378

Zuo, G., Dou, Y., & Lei, R. (2018). Discrimination algorithm and procedure of snow depth and sea ice thickness determination using measurements of the vertical ice temperature profile by the ice-tethered buoys. *Sensors*, *18*(12), 4162. https://doi.org/10.3390/s18124162