



What surface radiative fluxes reveal about Arctic cloud modelling accuracy

Yaël Le Gars¹, Jean-Christophe Raut¹, and Louis Marelle¹

¹LATMOS/IPSL, Sorbonne Université, UVSQ, CNRS, Paris, France

Correspondence: Yaël Le Gars (yael.le-gars@latmos.ipsl.fr)

Abstract. Low-level clouds exert a strong control on the Arctic surface energy budget, yet their representation in regional atmospheric models remains a major source of uncertainty. We evaluate the Weather Research and Forecasting (WRF) model against observations from the Norwegian Young Sea Ice Experiment (N-ICE2015), conducted north of Svalbard from polar night to polar day. The analysis focuses on downward surface shortwave (SW^{\downarrow}) and longwave (LW^{\downarrow}) radiation under synchronous

5 cloudy conditions to diagnose cloud-related radiative biases. While near-surface meteorology is generally well reproduced, pronounced seasonal radiative errors emerge. A dominance analysis based on a simplified two-layer emission framework shows cloud emissivity, primarily controlled by liquid water path (LWP), is the leading contributor to LW^{\downarrow} errors. During the spring transition, the model underestimates cloud occurrence and simulates optically too thin clouds, causing excessive SW transmission and insufficient LW trapping. During polar day, a marked negative SW^{\downarrow} bias develops at high LWP. Radiative

10 errors are largest for low LWP, where cloud optical properties are highly sensitive to variations in liquid water content. Sensitivity experiments demonstrate that improved representations of sea ice cover and surface albedo reduce polar day SW^{\downarrow} biases, while modifying prescribed cloud droplet number concentration alters optical thickness but introduces compensating errors. Clouds diagnosed as surface-decoupled exhibit lower LWP and larger radiative biases, and this regime is overrepresented in the model. These results highlight the need for consistent representation of surface properties, boundary-layer structure, and

15 mixed-phase microphysics to improve simulations of Arctic surface radiation.

1 Introduction

Low-level clouds play a pivotal role in the Arctic climate system by modulating the surface energy budget (SEB) through a balance of competing radiative effects. While these clouds reflect incoming shortwave (SW) solar radiation, leading to surface cooling, they simultaneously absorb and re-emit longwave (LW) terrestrial radiation, producing a warming effect, thereby

20 regulating surface temperature, snowmelt onset, and sea ice evolution (Persson and Vihma, 2017). Unlike the global average, in the Arctic, where solar radiation does not reach the surface for half of the year, clouds exert a net positive effect for most of the year via their sole LW interaction. Hereby, they reduce wintertime surface cooling by $40\text{--}50\text{ W}\cdot\text{m}^{-2}$, punctuated by a short period of surface cooling in summer when the cloud albedo effect prevail (Curry et al., 1996; Intrieri et al., 2002; Nakanishi and Michibata, 2025). Cloud radiative properties drive the magnitude of these interactions and depend on multiple parameters.

25 LW cloud forcing is determined by cloud temperature, height and emissivity, which relies on both hydrometeors concentration



and phase, while SW forcing is a function of cloud transmittance, solar zenith angle and surface albedo (Shupe and Intrieri, 2004; Griesche et al., 2024).

The Atlantic sector of the Arctic region exemplifies this complex environment with persistent and extensive cloud cover. Ground-based observations from the multi-year Ice-Atmosphere arctic Ocean Observing System (IAOOS) datasets (Maillard et al., 2021), with buoys mostly drifting within the Atlantic sector, reveal low-cloud covers of $\sim 85\%$ between May and October with lower coverage during the other months of measurements (April, November, December), suggesting a seasonality of those clouds. Satellite observations show cloud fractions exceeding 80% annually and low-level cloud cover above 60% , with a notable presence (55% annually) of mixed-phase clouds (MPCs) that contain both supercooled liquid droplets and ice crystals (Wang and Key, 2005; Liu et al., 2012; Jiang et al., 2024; Mioche et al., 2015). These MPCs exert dominant control over the Arctic surface cloud radiative effect (CRE) due to the disproportionate influence of minimal liquid water amounts on cloud optical depth and radiative transfer. Supercooled droplets, being smaller and more numerous than ice crystals, enhance cloud opacity and thus modulate SW and LW surface radiation (Rogers and Yau, 1989; Shupe and Intrieri, 2004). Even subtle variation in liquid water path (LWP) or phase partitioning can drastically alter CRE, precipitation efficiency and cloud lifetime (Tao et al., 2014).

Despite their importance, the accurate modelling of Arctic low-level clouds remains notoriously challenging. General Circulation Models (GCMs) often show systematic biases, including underestimation of cloud fraction (Kay et al., 2016; Boeke and Taylor, 2016; Taylor et al., 2019; Nakanishi and Michibata, 2025) and liquid water content particularly in winter and spring, while some overestimate liquid-rich clouds in summer (Shaw et al., 2022; Nakanishi and Michibata, 2025), together with substantial inter-model variability. Regional and mesoscale models, though benefiting from finer resolution and physically constrained dynamics, continue to struggle with realistic representations of cloud fraction, macrophysical properties and phase partitioning (Sedlar et al., 2020; Inoue et al., 2021), heavily depending on microphysical parameterizations that vary widely in performance (Klein et al., 2009).

This study presents a detailed evaluation of the Weather Research and Forecasting (WRF) model against comprehensive surface-based observations collected during the N-ICE2015 campaign in the Arctic Ocean north of Svalbard from January to June 2015, encompassing measurements from the polar night to the polar day. Focusing on downwelling SW and LW surface radiation under cloudy conditions, the study aims to derive model cloud evaluation. Section 2 describes the model setup and the observations used for model evaluation. Section 3 presents the results along with their discussion, and Sect. 4 provides a broader discussion and conclusion.

2 Methodology and datasets

Model biases on surface radiation, as well as the role of cloud properties are investigated using the WRF model (Skamarock et al., 2019). Model outputs are compared with meteorological variables and radiative fluxes obtained during the N-ICE2015 campaign near Svalbard (15 January - 22 June 2015). Simulations cover the whole period, starting on 1 January 2015, allowing

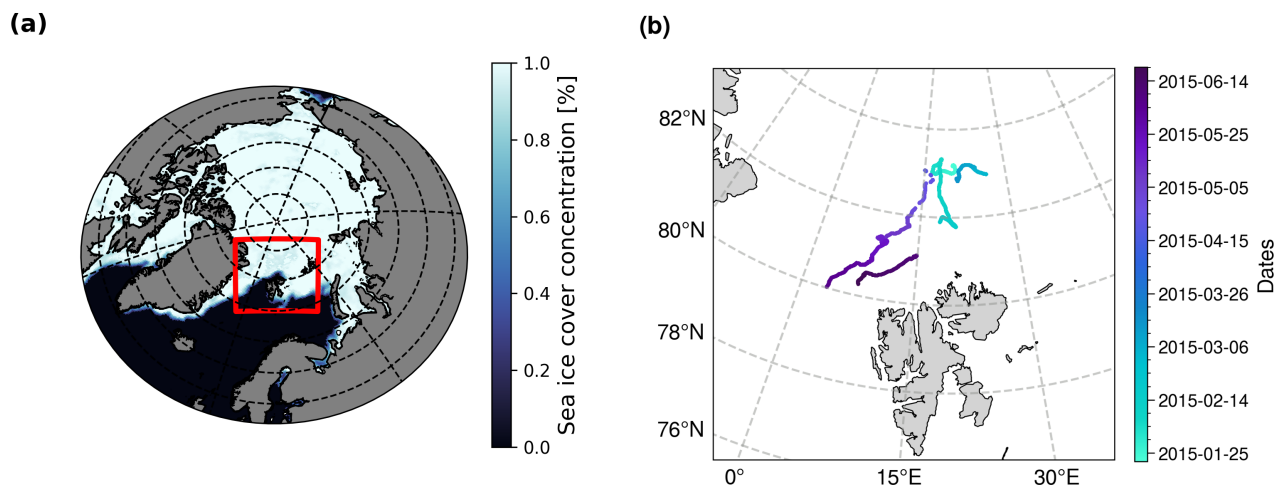


Figure 1. (a) WRF domain over the Arctic. (b) WRF domain with the drift trajectory of the Lance vessel during N-ICE2015. Breaks during the drift indicate the breakup of ice floes and transport to a new ice floe, without measurements.

for two weeks of initial spin up before the beginning of field observations used for evaluation, and cover the European Arctic region.

60 2.1 N-ICE2015 observations

The N-ICE2015 campaign was conducted from January to June 2015. The main objective was to improve the understanding of Arctic sea ice and the effects of its rapid shift to a younger and thinner one (Granskog, 2016). Through mid-January to late-June, the Norwegian Polar Institute's research vessel *Lance* drifted in the European Arctic sea ice, tethered in successive ice floes. Its drift can be seen in Fig. 1b. Besides observations on the ocean, ice dynamics, snow and sea ice, a comprehensive set of atmospheric observations was carried out. In this study, we use measurements of surface meteorology (Hudson et al., 2015), surface radiative fluxes and surface albedo (Hudson et al., 2016) and observations from radiosondes (Hudson et al., 2017). More details on the drift and on instrumental uncertainties are well described by (Walden et al., 2017). Surface meteorology and radiative fluxes are measured every minute. Balloon launches are performed twice a day, at approximately 11:00 and 23:00 UTC. Only data of radiative fluxes associated with a quality control flag of 0, corresponding to unambiguously valid data, are selected.

70 2.2 WRF model setup

The WRF model version 4.6.1, including developments for polar regions described in Marelle et al. (2025) is used to perform simulations from 1 January to 23 June 2015, over a domain of 1500 km × 1350 km spatial extent, with a 15 × 15 km² horizontal resolution and 72 vertical levels from the surface up to about 50 hPa. Figure 1a depicts the simulation domain, centered over the observations used for model evaluation. The control simulation (hereafter referred to as WRF-CTRL) is driven by the NCEP



FNL analysis ($1^\circ \times 1^\circ$, 27 vertical levels) (National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce, 2000), providing the initial and boundary meteorological conditions, along with sea surface temperature (SST) and sea ice concentrations (SIC). Boundary conditions, SST and SIC are being updated every 6 hours, and WRF simulations are conducted without any nudging to FNL meteorological fields within the simulation domain. Grid-scale cloud microphysics is modelled using the Thompson scheme (Thompson et al., 2008), accounting for 2-moment ice prediction and 1-moment liquid, rain and snow predictions and interacts with radiation through the RRTMG scheme (Iacono et al., 2008). Cloud liquid mass is distributed assuming a prescribed and fixed cloud droplet number concentration (CDNC), for which the default value of 100 cm^{-3} is used. The Kain-Fritsch cumulus parameterisation is used to represent sub-grid clouds (Kain, 2004). The full setup of the physical options used in this study is summarized in Table A1. The cloud fraction is calculated using Xu and Randall (1996), following Marelle et al. (2025). Model outputs are saved at hourly intervals.

A second simulation (hereafter WRF-SEAICE) is used in section 3.3.3 and section 3.3.5. It uses the same set up as WRF-CTRL but replaces the default SIC values with the daily 6.25 km grid resolution sea ice product from the University of Bremen, derived from AMSR2 (Advanced Microwave Scanning Radiometer 2) satellite observations (Melsheimer and Spreen, 2019).

Two sensitivity simulations to CDNC are performed and discussed in section 3.3.4. These simulation use the same setup as WRF-SEAICE, but prescribe lower CDNC values: 50 cm^{-3} (WRF-50CDNC) and 10 cm^{-3} (WRF-10CDNC).

Observed hourly-averaged radiative fluxes used in section 3 are calculated by averaging the 1-minute measurements over the preceding hour. Simulated hourly radiative fluxes are computed directly from the accumulated radiative fluxes, available every hour. The WRF model outputs are spatially co-located with the N-ICE2015 observations, taking the mean latitude and longitude of the observational platform over each hour and using bi-linear interpolation on the model grid, allowing direct comparison between the model and observations.

3 Results

We first present a general evaluation of the modeled surface meteorology and surface radiative fluxes in the WRF-CTL simulation, compared to observations obtained during the N-ICE2015 campaign (section 3.1). In section 3.2 we identify the model radiative biases in cloudy conditions specifically. We investigate the main processes driving these biases in the model in section 3.3.

3.1 Near-surface meteorology and surface radiative fluxes

Fig. 2 displays the time series of the 2-meter temperature (T_{2m}), 2-meter relative humidity (RH_{2m}), downward SW (SW^\downarrow) and downward LW (LW^\downarrow) over the 5-months period of the N-ICE2015 campaign. As meteorological conditions strongly vary during the drift, we define 3 periods based on the sunlight cycle. The first period (P1) is the polar night, when no sunlight reaches the top of the atmosphere (TOA). P1 lasts until 19 March. During the second period (P2), incoming solar radiation at the TOA increases and is associated with a rapid change in surface temperature. P2 extends from 18 April to 25 May. The third period (P3) is the polar day, when incoming SW radiation becomes nearly constant. Surface temperature also becomes



near-constant at 0°C , after the onset of snow melt (Granskog et al., 2018). P3 extends until the end of measurements on 20 June. Hereafter, all results and analyses are presented and examined separately according to these three periods. The statistical performance of the WRF model across all periods and key variables is summarized in Table 1.

The first two months of the campaign were marked by six intense winter storms driven by large-scale atmospheric circulation, bringing bursts of heat and moisture leading to heavy snowfall (Walden et al., 2017; Cohen et al., 2017). The hourly observed and simulated values of $T_{2\text{m}}$ and $RH_{2\text{m}}$ are shown respectively in Fig. 2a and b. During those stormy events, the temperature rises from -40° or -30°C to near 0°C in about a day. These events are associated with quasi-saturated air and are well captured both in time ($r = 0.96$) and magnitude by the model. Outside these events, the model exhibits an overall cold bias during the P1 period (Normalized Mean Bias (NMB) of -0.6% , Mean Bias (MB) of -1.47 K). This bias persists during P2, during which the model accurately follows the observed seasonal warming trend, but systematically underestimates the absolute temperatures (NMB = -0.6% , MB = -1.57 K). During P3, $T_{2\text{m}}$ cold biases diminish, with the model showing good agreement to observations (RMSE (Root Mean Square Error) = 1.9 K , NMB = -0.1%) under the milder conditions of the polar day.

Despite very low absolute moisture contents within the Arctic atmospheric boundary layer (ABL), the cold temperatures maintain a persistently high $RH_{2\text{m}}$. Although the model reproduces the general temporal evolution of $RH_{2\text{m}}$, the agreement is weaker. Correlation coefficients are moderate to low (ranging from 0.60 for P1 to ~ 0.35 for P2 and P3) and RMSE generally exceeds 10%. Notably, the model shows a positive NMB ($+7\%$) during the storm-influenced P1 and transitions to a negative bias (-6% and -7% , respectively) during P2 and P3. Such biases are expected to affect the representation of low-level cloud formation. The 10-meter wind speed ($WS_{10\text{m}}$, not shown) is simulated with good skill across the entire campaign. Correlations are strong ($r \geq 0.82$) and RMSE values are low ($\leq 2.2\text{ m}\cdot\text{s}^{-1}$). The SW^{\downarrow} and LW^{\downarrow} radiative fluxes (irradiances) averaged over 24 hours are displayed on Fig. 2c and d respectively. The observed daily mean downwelling fluxes are used to examine the overall model behavior, smoothing out sub-daily variability for better readability. The statistical evaluation (Table 1) reveals a strong seasonal dependence in model performance, with notable compensations of errors between SW^{\downarrow} and LW^{\downarrow} during the transition period P2. The model overpredicts SW^{\downarrow} during P2 (daily NMB = 12.8% , hourly NMB = 13.5%) together with good temporal correlation ($r_{\text{daily}} = 0.76$ and $r_{\text{hourly}} = 0.87$). On contrast, the bias becomes negative during P3, underlying a substantial underestimation (daily NMB = -15% , hourly NMB = -12.2%) alongside moderate temporal correlation ($r_{\text{daily}} = 0.41$ and $r_{\text{hourly}} = 0.70$).

The model demonstrates high skill to predict LW^{\downarrow} during the polar night (P1), with excellent correlation ($r_{\text{daily}} = 0.95$ and $r_{\text{hourly}} = 0.85$) and small bias (daily NMB = 0.5% , hourly NMB = 0.6%). This suggests a faithful representation of clear-sky near-surface atmospheric emission and the radiative impact of warm and moist intrusions during winter storms episodes. Model performance markedly degrades during P2, with the emergence of a pronounced negative bias (daily NMB = -8.9% , hourly NMB = -9.5%) and correlation dropping ($r_{\text{daily}} = 0.53$ and $r_{\text{hourly}} = 0.35$). This is associated with the SW^{\downarrow} positive bias discussed above. This opposite behaviour is typically associated with an atmosphere that is too transparent, resulting from an underestimation of the cloud optical depth (COD). This degradation coincides with a marked seasonal shift in boundary-layer structure during N-ICE2015, with the frequency of surface-based inversions decreasing from 58% in Winter to 7%

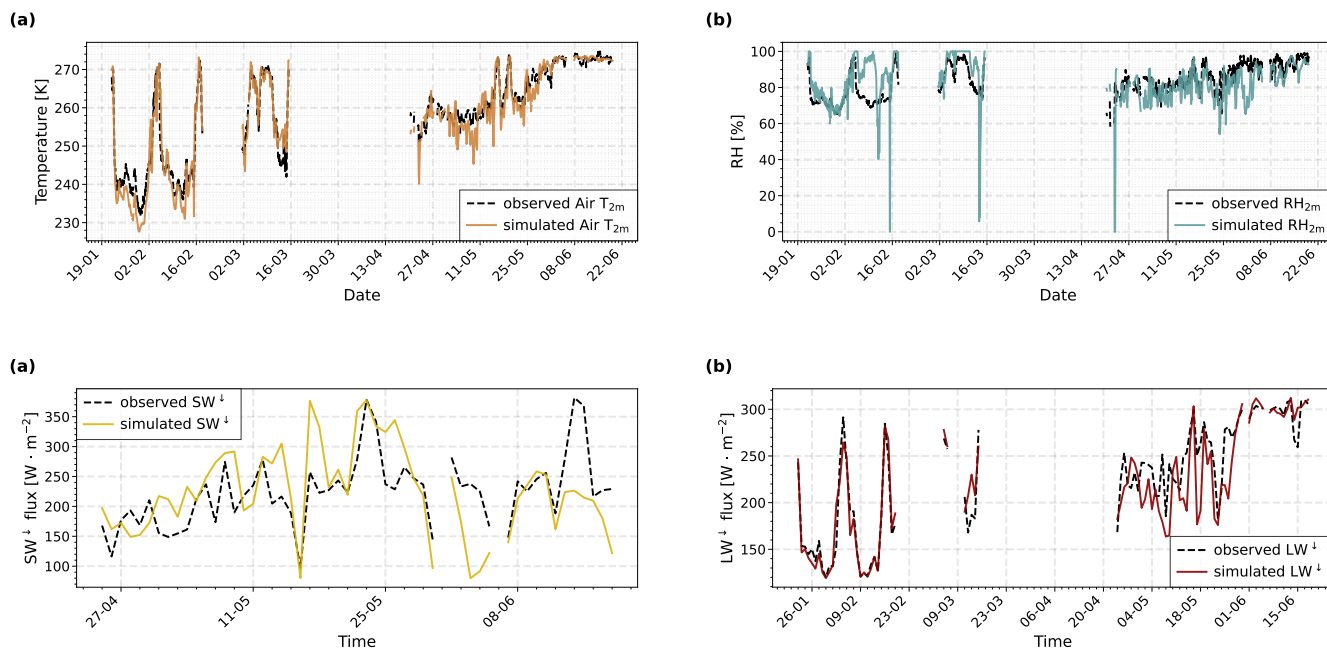


Figure 2. Time series of observed (black) and simulated (colored) near-surface meteorology and surface radiative fluxes along the drift: hourly instantaneous 2-m air temperature, relative humidity, and daily averaged downward SW and LW fluxes.

in Spring and early-Summer (Kayser et al., 2017). The predominance of observed unstable stratification reflects turbulent mixing initiated by surface fluxes resulting from radiative heating of the surface. Surface snowmelt in both observations and simulations starts only at the turn of May-June, suggesting a model limitation in simulating unstable conditions over frozen and snow-covered surfaces. An initial cold bias in near-surface air temperature, such as seen during P2, would reduce turbulent heat fluxes and thereby act as a positive feedback that would reinforce the cold bias, as shown by Hong and Jiang (2024), and promote stronger surface-based inversions. The model accuracy in simulating LW[↓] improves during P3, with the negative NMB reduced to -1.6% for daily averages and -2.2% for hourly averages.

150 3.2 Radiative fluxes under cloudy skies

Radiative biases can arise from various sources, as detailed in section 1. In this study we focus on Arctic clouds and their effect on the surface radiative budget. For this reason, we focus hereafter on cloudy conditions only. This requires a robust method to identify coincident cloudy cases in both observations and simulations, as detailed below.

3.2.1 Cloud cases selection

155 Surface net longwave radiation (LW_{net} , i.e. the downward minus the upward longwave components) is a good proxy for distinguishing cloudy from clear-sky conditions from radiative observations (Maillard et al., 2021; Silber and Shupe, 2022;



Table 1. Statistical evaluation of WRF simulations against N-ICE2015 observations for key variables, over the full period and seasonal subsets. Correlation coefficients (r), root mean-square errors (RMSE) and normalised mean bias (NMB) are shown.

Variable	Period	Correlation (r)	RMSE	NMB [%]	MB
Hourly T_{2m}	Full period	0.97	3.6 K	-0.4	-0.99 K
	P1	0.96	4.2 K	-0.4	-1.02 K
	P2	0.83	3.6 K	-0.6	-1.57 K
	P3	0.82	1.9 K	-0.1	-0.17 K
Hourly RH_{2m}	Full period	0.45	11.3 %	-0.8	-0.69 %
	P1	0.60	12.6 %	7.1	5.70 %
	P2	0.34	10.5 %	-5.9	-4.85 %
	P3	0.36	9.8 %	-7.1	-6.54 %
Hourly WS_{10m}	Full period	0.85	1.8 m.s ⁻¹	1.4	0.09 m.s ⁻¹
	P1	0.84	2.2 m.s ⁻¹	4.4	0.30 m.s ⁻¹
	P2	0.82	1.6 m.s ⁻¹	0.1	0.00 m.s ⁻¹
	P3	0.92	1.3 m.s ⁻¹	-3.5	-0.23 m.s ⁻¹
Daily averaged SW^{\downarrow}	Full period	0.52	65.2 W.m ⁻²	0.1	0.24 W.m ⁻²
	P1	/	/	/	/
	P2	0.76	53.6 W.m ⁻²	12.8	27.16 W.m ⁻²
	P3	0.41	78.1 W.m ⁻²	-15.0	-36.03 W.m ⁻²
Daily averaged LW^{\downarrow}	Full period	0.91	27.0 W.m ⁻²	-3.5	-7.99 W.m ⁻²
	P1	0.95	16.6 W.m ⁻²	0.5	0.88 W.m ⁻²
	P2	0.53	37.2 W.m ⁻²	-8.9	-20.83 W.m ⁻²
	P3	0.57	23.6 W.m ⁻²	-1.6	-4.58 W.m ⁻²
Hourly averaged SW^{\downarrow}	Full period	0.77	80.8 W.m ⁻²	2.4	5.50 W.m ⁻²
	P1	/	/	/	/
	P2	0.87	68.9 W.m ⁻²	13.5	29.07 W.m ⁻²
	P3	0.70	95.9 W.m ⁻²	-12.2	-29.97 W.m ⁻²
Hourly averaged LW^{\downarrow}	Full period	0.85	34.9 W.m ⁻²	-4.0	-8.95 W.m ⁻²
	P1	0.89	25.0 W.m ⁻²	0.6	0.97 W.m ⁻²
	P2	0.35	45.8 W.m ⁻²	-9.5	-22.32 W.m ⁻²
	P3	0.44	30.7 W.m ⁻²	-2.2	-647 W.m ⁻²



Stramler et al., 2011). In the absence of clouds, there is strong surface radiative cooling ($LW_{net} \ll 0$), whereas it is close to zero ($LW_{net} \approx 0$) in the presence of clouds that enhance the downwelling LW flux, compensating surface cooling. The probability density function (PDF) of the observed LW_{net} during N-ICE2015 (Fig. 3) shows clear bimodality, with one clear sky mode (centered around -45 W.m^{-2} in P1 and close to -70 W.m^{-2} in P2 and P3) and a second mode corresponding to the presence of opaque clouds (centered near -5 W.m^{-2} in P1, between -10 and -15 W.m^{-2} in P2 and P3).

Based on these results, we consider conditions as cloudy when LW_{net} exceeds -10 W.m^{-2} in P1 and -25 W.m^{-2} in P2 and P3. While this method reliably identifies optically thick clouds, thinner clouds may be missed, particularly when the corresponding LW_{net} is found between the two modes. Stramler et al. (2011) found that, during winter, 100 % of net LW fluxes in the $\sim 0 \text{ W.m}^{-2}$ mode indeed correspond to cloudy cases, but that 51 % of values within the $\sim -40 \text{ W.m}^{-2}$ mode were also associated with the presence of clouds. Nevertheless, although optically thin clouds are likely missed, cases identified as cloudy can be considered, with a high level of confidence, as genuinely cloudy.

The simulated LW_{net} distribution (Fig. 3) displays period-dependent biases. During P1, the simulated distribution is narrower than the observed one with a less pronounced opaque cloud mode and a positive bias in the clear-sky mode (centered over -35 W.m^{-2}). The latter results from an underestimation of the upwelling component. During P2, cloud and clear modes align well with observations, with cloud modes peaks differing by only a $\sim 5 \text{ W.m}^{-2}$ gap. Yet, the magnitude is not well represented, with the cloudy being mode largely underestimated. This is evidenced by an overestimation by nearly a factor of two of the frequency of values exceeding -25 W.m^{-2} . Conversely, the clear-sky mode is largely overestimated as well as the intermodal region. During P3, the simulated and observed distributions agree well, with better representation of the cloudy mode frequency.

We define a second criterion, based on the model diagnostic of total cloudiness (CLT). The CLT is diagnosed using a maximum-random overlap method, whose formula is given in eq. (B1), to remain consistent with the model RRTMG scheme. An arbitrary threshold of $CLT \geq 0.95$ is adopted to ensure that most of the downward radiation has indeed interacted with clouds. Applying the first criterion, we identify, for each period, the set of hourly cases classified as cloudy in the observations (i.e. when observed LW_{net} exceeds the prescribed threshold). Model cloudy cases are defined by applying the second criteria on total cases. Observed and simulated cloud occurrence frequencies are reported in Table B1.

Both observed and simulated clouds frequencies are lowest during P1 and highest during P3, although notable discrepancies exists in magnitude and temporal overlap (defined here as the proportion of simulated cloudy cases among observed cloudy hours). As expected, using LW_{net} leads to a substantially lower observed cloud occurrence during P1 (19 %) than reported in Arctic climatologies (Shupe et al., 2011; Huschke, 1969). This is primarily due to the frequent omission of optically thin clouds such as thin ice clouds, which weakly impacts surface longwave radiation during polar night conditions. The simulated cloud occurrence during P1 is also lower than climatological estimates (38 %) but is sensitive to the choice of the CLT threshold (increases to 51 % using a CLT threshold of 0.75). During P2, the model underestimates cloud occurrence (48 %, rising to 53 % using a CLT threshold value of 0.75) and shows weak temporal overlap (50 %), indicating that opaque clouds are rarely simulated. During P3, observations and model presents quite similar cloud occurrence frequencies (87 % from observations,

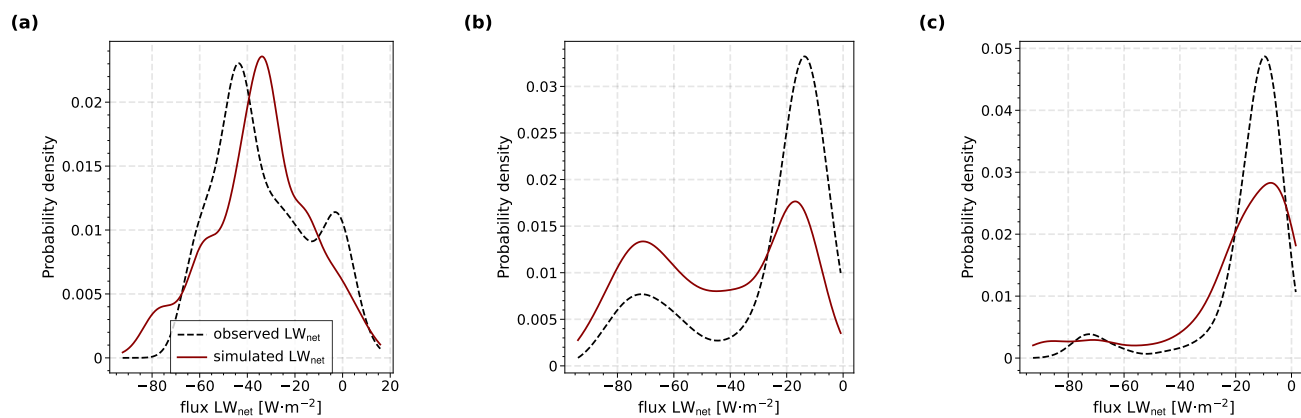


Figure 3. Probability density functions of observed (dashed black) and simulated (solid red) hourly surface net longwave radiation (LW_{net}) for (a) P1, (b) P2, and (c) P3. Kernel density estimation using Scott’s bandwidth.

78 % from simulations) together with good temporal overlap (80 %), indicating better model performance in representing cloud occurrence during this period. A weak sensitivity of simulated cloud frequency to the CLT threshold value is found during P3.

Overall, while the CLT threshold has a slight to moderate impact on the model cloud frequency, it does not affect the result presented in the following sections nor their interpretation.

195 To enable a like-for-like comparison of cloud scenes, we define a *synchronous cloudy set* comprising hours that satisfy both the observed LW_{net} criterion and the model CLT criterion.

3.2.2 Downward radiative fluxes distributions

The PDF of surface SW^{\downarrow} and LW^{\downarrow} are shown in Fig. 4 for each period. Only the LW^{\downarrow} is depicted for P1 as SW^{\downarrow} is zero during the polar night. During P1 (Fig. 4a), the distribution shape is retrieved, but underestimating the peak of distribution of about
200 $10 \text{ W}\cdot\text{m}^{-2}$, i.e. the effective emission of the warm and opaque clouds occurring during winter storms events.

During P2, the model captures the peak of the LW^{\downarrow} distribution (Fig. 4c), but underestimates its magnitude. This is due to an overestimated number of cases with low LW^{\downarrow} values ($< 220 \text{ W}\cdot\text{m}^{-2}$). Meanwhile, the SW^{\downarrow} distribution during this same period (Fig. 4b) is shifted toward higher fluxes in the model compared to observations. This pattern is consistent with a model representation of clouds that are optically too thin, allowing excessive shortwave transmission while insufficiently trapping
205 longwave radiation.

Fig. 4d and Fig. 4e present respectively the distribution in SW^{\downarrow} and LW^{\downarrow} during P3. The LW^{\downarrow} distribution is broadly reproduced, though with the model again overestimating the lower values ($< 275 \text{ W}\cdot\text{m}^{-2}$). The SW^{\downarrow} distribution exhibits larger discrepancies. Notably, the lowest SW^{\downarrow} values ($< 100 \text{ W}\cdot\text{m}^{-2}$) are strongly overestimated, while those above $> 275 \text{ W}\cdot\text{m}^{-2}$ are underestimated. These biases likely reflect errors in COD. Still, CRE values are also influenced by uncertainties in temper-

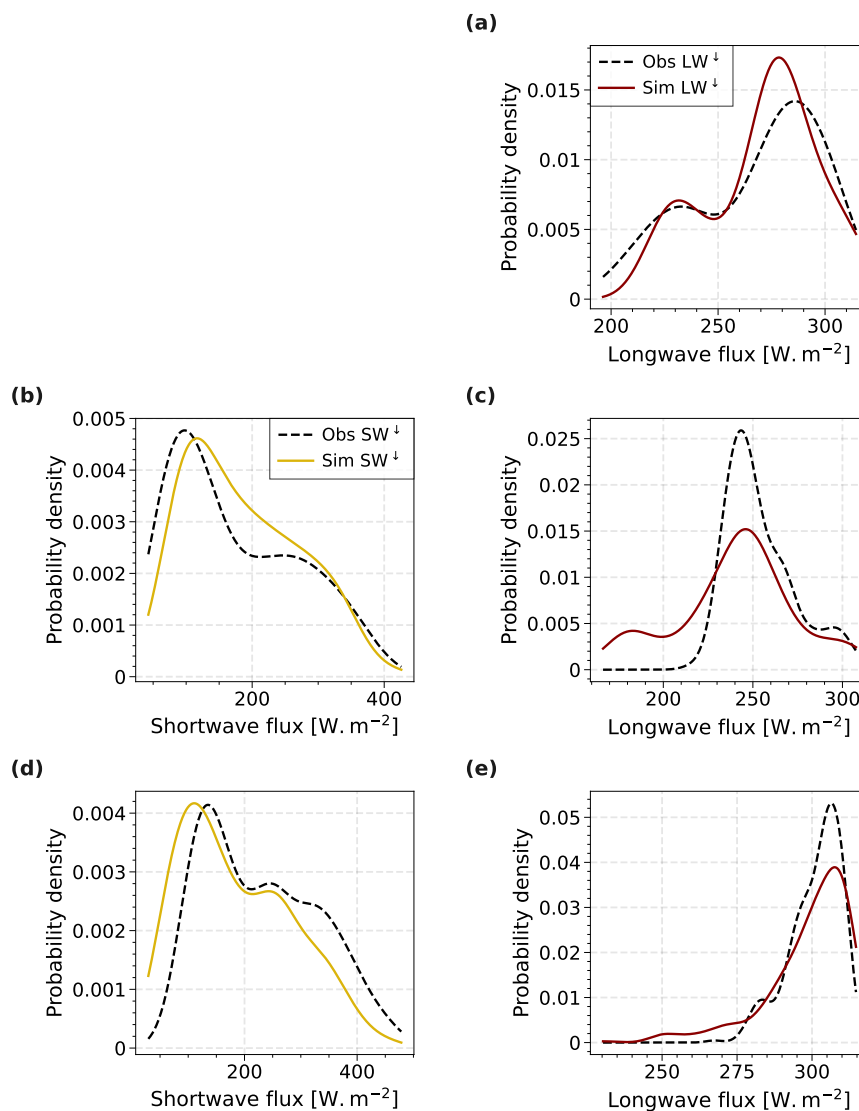


Figure 4. Probability density functions of observed (dashed black) and simulated (solid coloured) downward radiative fluxes for cloudy-sky: (a) LW^{\downarrow} during P1 period, (b) SW^{\downarrow} during P2 period, (c) LW^{\downarrow} during P2 period, (d) SW^{\downarrow} during P3 period and (e) LW^{\downarrow} during P3 period. Kernel density estimation using Scott’s bandwidth.

210 ature for LW and in surface albedo for SW. The role of those physical variables in driving the uncertainties is evaluated and discussed in Sect. 3.3.



3.3 Cloud radiative biases and drivers

3.3.1 Temperature biases

Under cloudy sky conditions, the magnitude of LW^\downarrow is driven by temperature profile, cloud height, and by the combined emis-
215 sivity of the clear-sky near-surface atmosphere and the overlying cloud layer. We can relate the amount of LW^\downarrow to these factors using a simple two-layer atmospheric emission model, in which a cloudless near-surface layer emits at its own temperature (T_{2m}) and emissivity (ϵ_0), and an overlying cloud layer emitting at its own temperature (the cloud base temperature T_{cb}) and emissivity (ϵ_c).

$$LW^\downarrow = \epsilon_0 \cdot \sigma \cdot T_{2m}^4 + \epsilon_c \cdot (1 - \epsilon_0) \cdot \sigma \cdot T_{cb}^4 \quad (1)$$

220 ϵ_0 is the near-surface cloudless emissivity, parametrised here as $\epsilon_0 = 0.83 - 0.18 \times 10^{-0.067 e_0}$, following the formulation of (Niemelä et al., 2001), where e_0 is the near-surface (here, 2 m) vapour pressure in hPa.

During N-ICE2015, three observation buoys from the IAOS program equipped with a Mie backscatter lidar at a wavelength of 808 nm (Mariage et al., 2016) drifted within the Arctic sea ice close to the research vessel. Using the corresponding dataset, Maillard et al. (2021) showed that almost all clouds bases were located below 90 m. We do not have direct measurements of
225 cloud-base temperature at these altitudes, since atmospheric profiles were performed only twice daily. Since the cloud base is usually very low, and there is only a weak temperature gradient in the lowest 100 m (with 90 % of values ranging between 0.6 and 2 °C, Maillard et al. (2021)), here we use T_{2m} as a proxy for the observed cloud base temperature. Therefore, eq. (1) simplifies as :

$$LW^\downarrow = [\epsilon_0 + \epsilon_c(1 - \epsilon_0)] \cdot \sigma \cdot T_{2m}^4 \quad (2)$$

230 We use this equation to calculate ϵ_{cloud} for all cases classified as cloudy according to both the observational and the model criteria from section 3.2.1, using eq. (2) for the observations and eq. (1) for the model, given the observed and simulated values of LW^\downarrow . For the model calculation, T_{cb} is defined as the temperature of the lowest model layer in which the liquid water content (LWC) exceeds 10^{-5} kg.kg⁻¹ or in which the combined ice and snow water content (ISWC) exceeds 10^{-6} kg.kg⁻¹. The derived cloud emissivity values occasionally exceed unity during P1. In observations, $\epsilon_{cloud} > 1$ reflects limitations of
235 the simplified two-layer radiative model under surface-based inversions, where near-surface temperature does not represent the cloud radiating temperature. A similar behavior in the model suggests effective multi-layer cloud emission at temperatures higher than T_{cb} when clouds are embedded within an inversion. ϵ_{cloud} values greater than 1 were therefore excluded from the analysis.

We analyse the contributions of errors in each term of eq. (1) to the LW^\downarrow bias, defined as $\Delta LW^\downarrow = LW^\downarrow_{model} - LW^\downarrow_{obs}$. Due
240 to strong covariances among the four predictors of the two-layer model (ϵ_0 , T_{2m} , T_{cb} , ϵ_c), we apply dominance analysis within a multiple linear regression framework. This yields the relative contribution of each predictor to the total explained variance of ΔLW^\downarrow in an order-independent and robust manner. The resulting variance contributions are reported in Table 2.



Table 2. Relative contribution of the four variables from the two-layer atmospheric emission model to the variance of ΔLW^\downarrow .

Period	T_{2m} [%]	ϵ_0 [%]	T_{cb} [%]	ϵ_c [%]	R^2
P1	33.1	1.7	28.0	36.8	0.996
P2	28.0	2.0	19.9	51.7	0.998
P3	23.4	3.1	14.3	59.0	0.998

Among all predictors, ϵ_c consistently dominates the variance of ΔLW^\downarrow across all seasons. This is particularly the case during P2 (52 %) and P3 (59 %), highlighting that cloud properties misrepresentation (water content, phase, and effective radii) have the strongest impact on LW^\downarrow errors. The contribution of T_{2m} is also substantial, ranging from 23 % (P3) to 33 % (P1), reflecting the role of near-surface temperature in modulating surface LW^\downarrow . Errors in T_{cb} are also found to be substantial. A T_{cb} error can result either from temperature profile misrepresentation or cloud height error. Comparison with co-located IAOOS cloud observations shows that during P1, cloud bases are well captured, with 95 % within the first model level. Thus, $T_{cb} \approx T_{2m}$ during P1 and T_{cb} errors comes from local temperature errors. During P2, cloud bases are less accurately captured, with a broader model distribution (nearly 60 % of the model clouds located within the first layer but a mean of 200 m). During P3, model cloud bases are also higher than observed (mean \approx 300 m, median = 300 m). Thus, misrepresented cloud heights may cause T_{cb} errors during P2 and especially P3. The contribution of ϵ_0 is negligible across all seasons (\leq 3 %), indicating that clear-sky emissivity errors play only a minor role in the variability of ΔLW^\downarrow .

Overall, cloud emissivity emerges as the primary driver of surface LW^\downarrow error, even though consistent contribution of T_{2m} and T_{cb} are found across all periods. Because ϵ_c also modulates SW radiation, its influence on LW^\downarrow errors extends to the SW spectrum. Therefore, this motivates a closer examination of the relationship between radiative errors and the modelled cloud water.

3.3.2 Effect of cloud water content on radiative biases

As we are missing reliable high-resolution observational datasets for cloud water content (stemming from uncertainties in lidar- and radar-derived particulate extinction coefficient, particularly for liquid water content) this analysis relies solely on simulated cloud water content. We therefore investigate the distribution of errors in SW^\downarrow and LW^\downarrow radiation as a function of the modelled cloud water. The statistical performance of the WRF model across all periods in representing surface downward radiative fluxes for the selected synchronous cloudy cases is summarized in Table 3.

Fig. 5 shows that radiative fluxes errors have a strong relationship with modelled LWP during P2 and P3. During P1 (Fig. 5a), no clear monotonic relationship is observed, consistent with the results from Table 2, which shows the weaker control of ϵ_c on LW^\downarrow errors and the strong influence of temperature biases. Though, cases with the lowest LWP values are generally associated to LW^\downarrow underestimation. During P2, a strong overestimation of SW^\downarrow (Fig. 5b) and underestimation of LW^\downarrow (Fig. 5c) is found for low LWP values. These biases reduce as LWP increase, becoming negligible for LWP greater than approximately 30–40 $g \cdot m^{-2}$. Similar relationships are encountered during P3, with a reduction in mirroring SW^\downarrow (Fig. 5d) and LW^\downarrow (Fig. 5e)



Table 3. Statistics of radiative errors in surface downward fluxes during P1, P2, and P3, under synchronous cloudy conditions, showing mean bias (MB), normalised mean bias (NMB), and normalised mean absolute error (NMAE).

Period	Band	MB ($\text{W}\cdot\text{m}^{-2}$)	NMB (%)	NMAE (%)
P1	SW	/	/	/
	LW	3.18	1.19	6.35
P2	SW	11.66	6.68	15.84
	LW	-13.73	-5.40	6.83
P3	SW	-41.71	-18.16	25.55
	LW	-2.30	-0.76	2.44

270 errors up to approximately $30\text{--}40 \text{ g}\cdot\text{m}^{-2}$, albeit concerning fewer data points. The overestimation of SW^\downarrow and underestimation of LW^\downarrow result from clouds that are optically too thin in the model. The cloud scene selection method, based on a threshold in LW_{net} , appears to result in observed scenes that are unlikely to present LWP lower than $\sim 30 \text{ g}\cdot\text{m}^{-2}$. It seems that the model particularly struggles to simulate clouds with sufficient LWP during P2, as reflected by the NMB of $+6.68\%$ for SW^\downarrow and -5.40% for LW^\downarrow . Cases corresponding to both the highest radiative errors and the lowest LWP were found to have a liquid

275 water fraction (LWF) of about 1, suggesting that these maxima of errors — positive for SW^\downarrow and negative for LW^\downarrow — are caused by deficiencies in cloud water content rather than by phase partitioning. In contrast, SW^\downarrow was markedly underestimated during P3 (NMB = -18.16% , Fig. 5d). For LWP greater than $> 40 \text{ g}\cdot\text{m}^{-2}$, SW^\downarrow errors become more negative as LWP increases. Meanwhile, we see weak observed biases in LW^\downarrow radiation (Fig. 5e) along with saturation of the modelled LW^\downarrow (not shown) for LWP greater than $30\text{--}40 \text{ g}\cdot\text{m}^{-2}$, so as for P2. Indeed, high LWP thick clouds behave as black bodies, whereby any additional

280 increase in LWP does not change the LW cloud forcing. As shown in Stephens (1978) and Chen et al. (2006), clouds become nearly fully opaque at LWP values greater than $30\text{--}40 \text{ g}\cdot\text{m}^{-2}$, with negligible sensitivity above $80 \text{ g}\cdot\text{m}^{-2}$ and the highest sensitivity below $10 \text{ g}\cdot\text{m}^{-2}$. Consistent with our findings, Shupe (2011) report that LW cloud forcing is sensitive to LWP below $30 \text{ g}\cdot\text{m}^{-2}$. Consequently, at low LWP, a small variation in LWP will lead to a large change in LW^\downarrow .

The observed pronounced negative SW^\downarrow bias during P3 can arise from multiple sources. In contrast to the LW spectrum,

285 no saturation effect exists in the SW one: an increase in COD will result in a reduced surface SW^\downarrow . An overestimation of the modelled COD may result either from an excessive LWP or from an overestimation of the CDNC, with smaller and more numerous droplets enhancing cloud scattering efficiency and further reduce the SW^\downarrow radiation reaching the surface (Peng et al., 2002). An other source of errors may arise from misrepresentation of surface albedo, which can lead to error in SW^\downarrow calculation through multiple reflections between the surface and an overlying cloud. The latter is investigated in 3.3.3

290 No correlation arises from radiative fluxes errors distribution with respect to ice and snow water path (ISWP) (Fig. S1). This indicates that, to first order, ISWP misrepresentation are not the primary driver of the radiative fluxes errors.

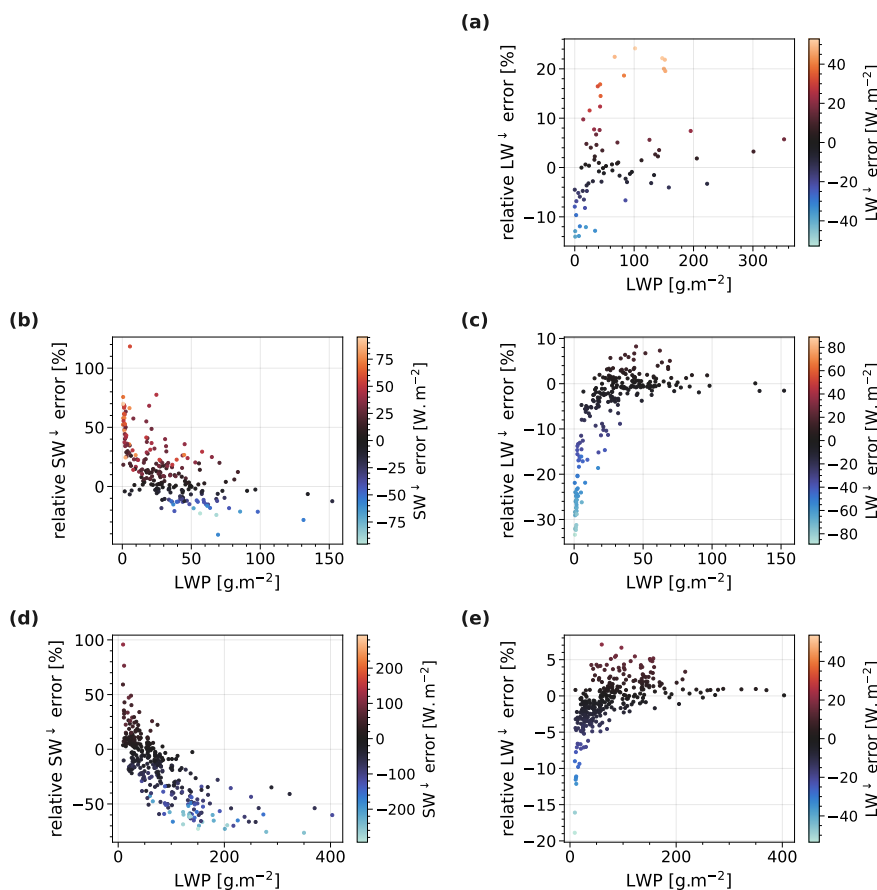


Figure 5. Radiative fluxes biases as a function of modelled liquid water path (LWP), showing relative error (model minus observation) on y-axis and colour-coded by the raw error. **(a)** LW^{\downarrow} errors during P1, **(b)** SW^{\downarrow} errors during P2, **(c)** LW^{\downarrow} errors during P2, **(d)** SW^{\downarrow} errors during P3 and **(e)** LW^{\downarrow} biases during P3.

3.3.3 Effect of surface albedo and surface type on radiative biases

In the WRF-CTRL simulation presented so far, the modelled daily sea ice fraction used for the initial and surface boundary conditions is provided by the NCEP FNL reanalysis. However, the interpolated surface albedo values in FNL were found to be markedly lower than observed, especially during P3, with a bias reaching up to -0.7 in late June. This bias is associated with concurrent underestimation of SW^{\downarrow} and higher modelled LWP (Fig. S4). The default prescribed value for dry snow albedo of 0.82 used in WRF agrees well with observations (Fig. S3), indicating that the grid-scale albedo underestimation instead stems from limitations in the resolution and accuracy of the FNL sea ice product, combined with the bilinear interpolation applied to the $15 \times 15 \text{ km}^2$ simulation grid. This is particularly critical in the marginal ice zone, where sea ice representation is highly inaccurate, close to the area where the *Lance* icebreaker operated.



To improve the representation of surface albedo and surface type, the simulation WRF-SEAICE uses the higher-resolution AMSR2 sea ice satellite observation product (section 2.2). As a result, modelled albedo in WRF-SEAICE agrees much better with observations (Fig S3). This better representation of the sea ice surface in the WRF-SEAICE run limits errors in SW^{\downarrow} associated with multiple cloud-surface reflections, and those associated surface-type-dependent turbulent heat fluxes, which can influence cloud properties (Kay and Gettelman, 2009; Morrison et al., 2012).

The WRF-SEAICE simulation improves agreement with observed T_{2m} . P2 and P3 cloud frequencies are not affected by this change. Cloud scenes only slightly differ between the two simulations (temporal overlap of 88 % in P2 and 96 % in P3). Model cloud frequency during P1 was reduced in WRF-SEAICE (from 38 to 29 %) and differs in the selected cloud scenes set with only two thirds of the cloud scenes from WRF-CTRL. Very limited effects of WRF-SEAICE is found on the distribution of LW^{\downarrow} errors and P2 SW^{\downarrow} errors. During P3, however, the improved surface characterisation in WRF-SEAICE reduces both the prevalence and magnitude of negative SW^{\downarrow} errors (Fig. 6), bringing the NMB close to zero and reducing the NMAE from 25.55 % to 20.11 %. Overall, the surface radiative net balance is substantially improved in WRF-SEAICE and is now in good agreement with observations (Fig. S6). During P3, when the albedo bias was most pronounced, the mean bias in daily mean net radiative fluxes is reduced from $+31.32 \text{ W.m}^{-2}$ in WRF-CTRL to -2.32 W.m^{-2} in WRF-SEAICE.

We found that WRF-SEAICE affects only slightly the distribution of LWP values (Fig. S5), while the distribution of SW^{\downarrow} errors is shifted positively across all LWP ranges, suggesting that the improvements primarily operate through modifications to multiple SW reflections rather than through changes in cloud water content. Despite the reduced SW^{\downarrow} underestimation, a systematic negative bias persists for LWP values greater than about 100 g.m^{-2} .

Two main explanations remain plausible : either LWP values greater than 100 g.m^{-2} are unlikely for the selected cloud scenes in reality, or the model distributes the LWC over an excessive number of smaller droplets. This is further investigated in section 3.3.4.

3.3.4 Effect of CDNC on radiative fluxes

Fig. 7 shows that reducing the prescribed CDNC shifts the distribution of SW^{\downarrow} errors toward more positive values, consistent with the aerosol first indirect (Twomey) effect. A moderate reduction of CDNC from 100 cm^{-3} to 50 cm^{-3} causes no substantial shift of the distribution of LWP values across all periods (Fig. S7). The distribution of LWP over lower CDNC alleviates the SW^{\downarrow} negative biases during P3 but does not fully suppress the occurrence of large negative errors at high LWP (Fig. 7 b). A stronger reduction to 10 cm^{-3} markedly decreases the frequency of large LWP values (e.g. $> 40 \text{ g.m}^{-2}$ from 32 % to 17 % during P2 and $> 70 \text{ g.m}^{-2}$ from 46 % to 16 % during P3), in line with the second indirect (Albrecht) effect. In the WRF-10CDNC simulation, the combined influence of these two aerosol indirect effects effectively removes the negative SW^{\downarrow} biases at high LWP. However, both sensitivity experiments lead to a substantial increase in the frequency and magnitude of positive SW^{\downarrow} biases, evidenced by an increase of SW^{\downarrow} error NMB in P2 from 8 to 15 to 29 % (for WRF-SEAICE, WRF-50CDNC and WRF-10CDNC respectively) and in P3 from ~ 0 to 8 to 28 %.

Overall, model CRE is sensitive to CDNC. Lowering CDNC reduces the representation of optically thick cloud cases and mitigates negative SW^{\downarrow} biases. In the meantime, it degrades the overall radiative balance by increasing and enhancing positive

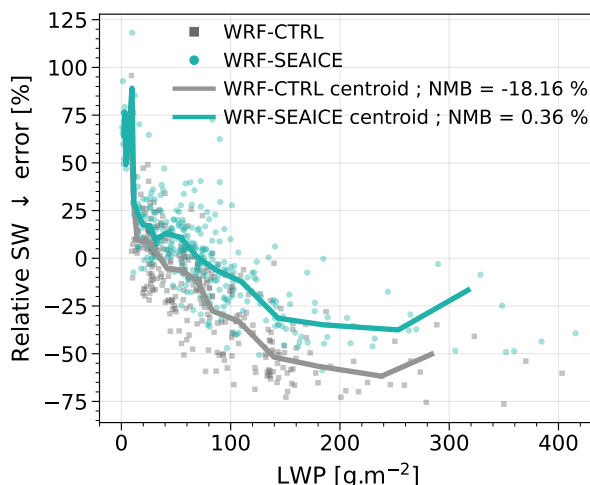


Figure 6. Comparison of WRF-CTRL (gray) and WRF-SEAICE (blue) simulations : SW^{\downarrow} biases as a function of modelled LWP, showing relative error (model minus observation) on y-axis and colour-coded by the raw error.

335 SW^{\downarrow} biases and LW^{\downarrow} deficits. Nevertheless, in-situ aircraft observations shows that the reference value of 100 cm^{-3} is not unrealistic for the spring and early-summer conditions considered here, with the effective CDNC likely lying between 50 and 100 cm^{-3} for Arctic low-level clouds during this season (Ehrlich et al., 2019; Mech et al., 2022; Wendisch et al., 2024).

Negative SW^{\downarrow} biases are primarily associated with cases where the LWF approaches unity, suggesting that they may also reflect deficiencies in cloud phase partitioning. In the simulations, the ice fraction tends to precipitate efficiently as snow. One hypothesis would be that precipitation removes ice-phase condensate too efficiently, thereby limiting mixed-phase processes of ice growth in regards to liquid (Wegener-Bergeron-Findeisen mechanism, riming), that would result in excessive LWP.

340

3.3.5 Effect of ABL coupling state on radiative fluxes

Low-level Arctic clouds are tightly controlled by the thermodynamic structure of the boundary-layer and by surface-atmosphere coupling. Numerous observational and modelling studies have shown that the degree of coupling between the cloud layer and the surface strongly affects their vertical extent, LWP, persistence and CRE (Taylor and Monroe, 2023; Yu et al., 2019; Taylor et al., 2019). In coupled conditions, turbulent mixing efficiently transports heat and moisture from the surface to the cloud layer, favouring the growth of thicker clouds with higher COD.

345

A common metric to characterise the stability of the lower atmosphere is the Lower Tropospheric Stability (LTS), defined as the difference in potential temperature (θ) between the 700 hPa pressure level and the surface. This large-scale metric, designed to link thermodynamic stability to stratiform cloud cover, is inherently integrated over a deep atmospheric layer. In the Arctic, where the boundary-layer is usually much thinner, regional Arctic studies have generally adapted the LTS calculation by using a lower reference level such as 850 hPa or 925 hPa (Taylor and Monroe, 2023; Serreze et al., 2025) to characterise stability

350

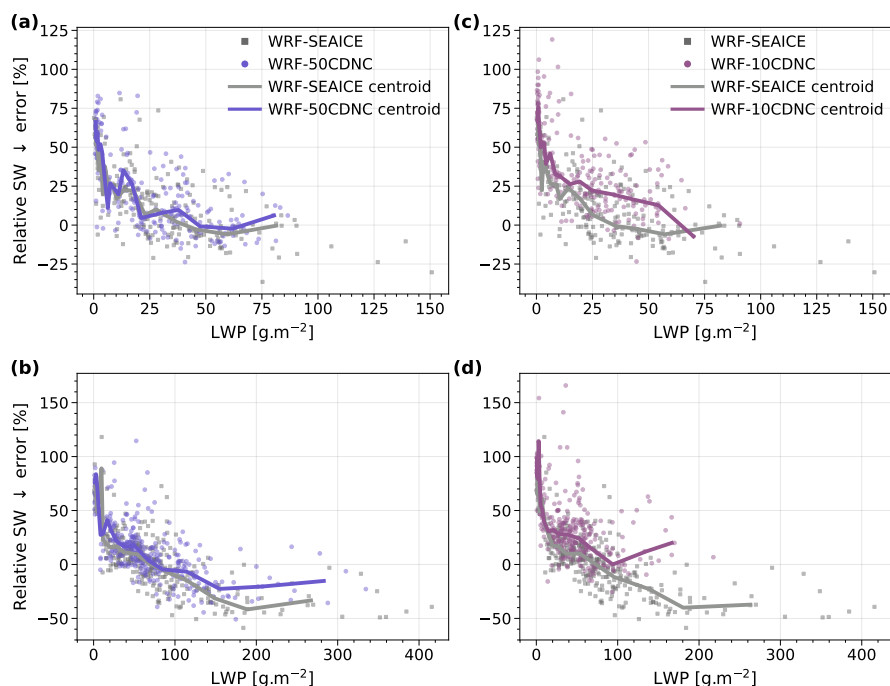


Figure 7. Distribution of SW^{\downarrow} relative error as a function of modelled LWP for different values of CDNC. Results from WRF-50CDNC and WRF-10CDNC are shown relative to the reference simulation WRF-CTRL (100 CDNC, shown in grey). WRF-50CDNC is shown for (a) P2 and (b) P3. WRF-10CDNC is shown for (c) P2 and (d) P3.

within the very shallow, cloud-topped ABLs, which are prevalent in the Arctic. As discussed in section 3.3.1, modelled lower cloud layer was generally found in the first tens to hundreds of meters above the surface. Consequently, any bulk metric based on a fixed pressure level several hundred meters above the surface is inherently disconnected from the key physical process governing surface-cloud interactions: the thermal structure of the *sub-cloud* layer.

To assess cloud-surface coupling and decoupling, we follow the recommendations from Gierens et al. (2020), based on sub-cloud vertical gradients of potential temperature (θ). A cloud is classified as surface-decoupled for any difference of θ that exceeds 0.5 K, and as surface-coupled otherwise. For each cloud scene, the model lowest cloud layer was determined following the LWC threshold described in section 3.3.1. When the cloud base lies within the first model level, this method may systematically diagnose coupled conditions even in the presence of a near-surface inversion. In such cases, the θ difference is instead evaluated between the second and first model levels. Finally, to restrict the analysis to low-level clouds, the cloud-base height is capped at the diagnosed ABL height when it exceeds it.

Fig. 8 shows the distribution of radiative fluxes errors and LWP for surface-coupled and surface-decoupled clouds. The classification reveals distinct radiative error signatures for the two cloud regimes. For decoupled clouds, both SW^{\downarrow} and LW^{\downarrow} errors exhibit a stronger bias compared to their coupled counterparts. Specifically, decoupled clouds in periods P2 and P3



are associated with a substantial positive NMB in SW^\downarrow flux (e.g. 14.62 % and 13.22 %, Fig. 8a and c). Conversely, coupled clouds show a lower SW^\downarrow NMB (e.g. 5.78 % in P2 and -3.86 % in P3). In the LW spectrum, greater underestimations occur for surface-decoupled clouds (e.g. -11.27 % and -4.73 % respectively for P2 and P3) in comparison to the coupled ones (e.g. -3.11 % and -0.67 %).

The distribution of the LWP values further underscores this regime dependency. LWP values are lower when clouds are decoupled (with a median LWP of 14 and 52 $g.m^{-2}$ for P2 and P3, respectively) than for their coupled counterparts (29 and 73 $g.m^{-2}$). This difference in hydrometeor content likely contributes to the observed radiative error patterns. Notably, the most extreme SW^\downarrow and LW^\downarrow errors occur at very low LWP values, a condition predominantly encountered for surface-decoupled clouds. We note that LWP dominates the total cloud water path (CWP), the latter showing a similar distinction between regimes.

From atmospheric radiosounding measurements and results from Maillard et al. (2021) which indicate that cloud bases are quasi-exclusively located within the first 90 meters, we determined the observed coupling or decoupling of cloud scenes temporally coincident with the radiosoundings. Following Gierens et al. (2020) methodology and considering a cloud base height of 100 m, we find that, among the relatively small sample of cloud scenes (26 during P2 and 29 during P3), all but one are identified as surface-coupled. When comparing these observations to the model classification of cloud scenes, we find that 18 out of 25 observed-coupled clouds coincide with the coupling state in the model during P2, and that 25 out of 29 observed-coupled are also coupled according to the model. These results indicate that, over this fractional sample of cloud scenes, the model tends to overestimate the decoupling state, especially during P2.

This suggests that the biases observed in the radiative fluxes may be partly explained by an inaccurate representation of the ABL thermodynamic structure in the model, which can misrepresent the vertical mixing and surface coupling of low-level clouds. Alternatively, or in addition, these biases may reflect limitations in the modelling of cloud properties, such as hydrometeor content and COD, when the cloud layer is thermodynamically disconnected from the surface. Overall, the model appears more robust in capturing cloud radiative properties when the cloud layer is coupled to the surface.

4 Discussion and conclusion

This study provides a comprehensive assessment of the error sources in surface radiative fluxes over Arctic sea ice in the regional WRF model. Using observations from the N-ICE2015 campaign, the model performance was evaluated across three periods : polar night, the springtime radiative transition, and polar day.

Our results show that while the WRF model satisfactorily captures near-surface meteorology, this apparent agreement often masks underlying physical deficiencies. Specifically, accurate surface temperatures partly result from compensating errors between SW and LW radiative fluxes, underscoring the necessity of process-level evaluations to ensure the model's physical consistency.

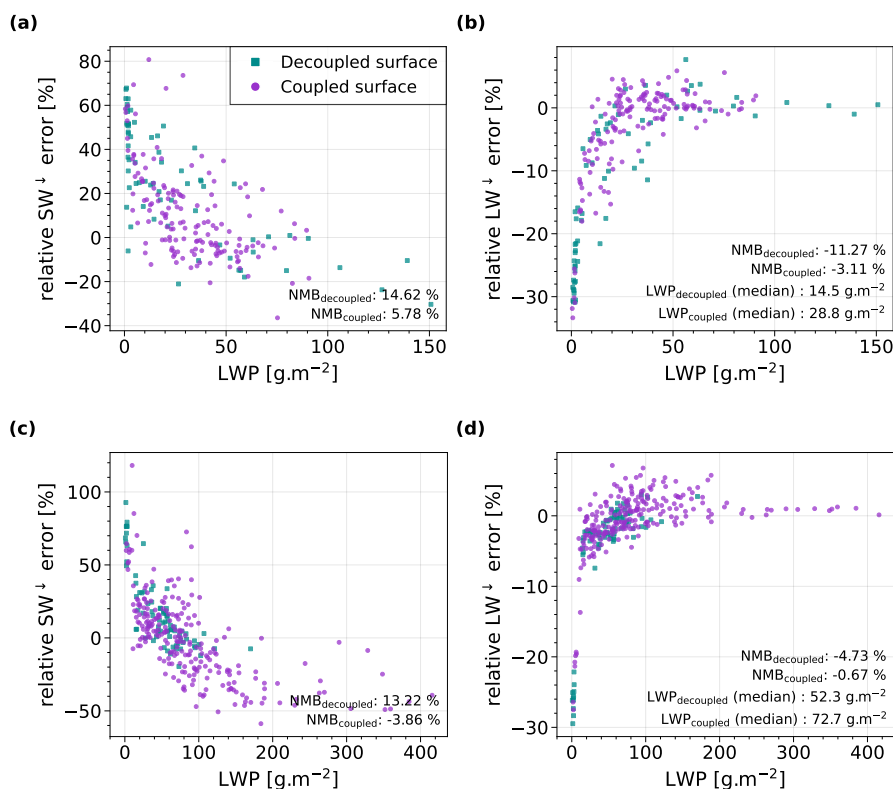


Figure 8. Cloud scenes distributions of radiative fluxes errors and LWP for surface coupled and decoupled states. (a) SW \downarrow and (b) LW \downarrow during P2, (c) SW \downarrow and (d) LW \downarrow during P3. NMB and median LWP are indicated for each regime and period.

Under cloudy conditions, a dominance analysis identifies cloud emissivity, primarily regulated by the LWP, as the leading driver of LW \downarrow errors across all seasons.

400

Taken together, the results indicate that the model struggles to represent the seasonal evolution of cloud optical properties. During the spring transition, clouds are too often underrepresented and, when present, remain optically too thin, leading to excessive SW transmission and insufficient LW trapping. This bias is tightly linked to the model's difficulty in forming and sustaining clouds with sufficient liquid water, leading to an overrepresentation of low-LWP regimes (below 30–40 $\text{g}\cdot\text{m}^{-2}$) where radiative fluxes are most sensitive to small variations in cloud properties. In contrast, during the polar day, this imbalance reverses, with a pronounced underestimation of SW radiation. Sensitivity experiments show that a substantial part of this bias is attributable to misrepresentation of surface albedo linked to inaccurate sea ice cover, as the use of higher-resolution sea ice fields (AMSR2) markedly reduces the negative SW errors. However, a SW deficit persists for high-LWP cases (larger than 100 $\text{g}\cdot\text{m}^{-2}$, consistent with an overestimation of cloud opacity. Ice and snow water paths show no systematic relationship with surface radiative errors, indicating that liquid-containing clouds dominate the radiative signal. Additional sensitivity tests further shows

410



that if modifying prescribed cloud droplet number concentration alters cloud optical thickness, realistic concentration does not remove the SW deficit but instead redistributes errors and introduces compensating effects across radiative components. Together, these results point to a misrepresentation of the seasonal adjustment of cloud optical depth rather than a simple systematic bias.

415 Crucially, we show that cloud radiative performance is strongly conditioned by boundary-layer structure. Surface-decoupled clouds exhibit lower LWP and substantially larger radiative errors than surface-coupled clouds. The apparent overestimation of surface-decoupled regimes in the model likely limits the development of optically thicker, liquid-rich clouds. Conversely, the reduced occurrence of optically thick clouds may also contribute to the maintenance of a more stable and decoupled boundary layer.

420

Reducing Arctic surface radiative flux biases therefore requires coordinated improvements in mixed-phase microphysics, LWP distribution, surface properties, and stable boundary-layer representation. Ensuring physically consistent cloud–radiation interactions is essential for robust simulation of the Arctic surface energy budget under ongoing climate change. Our results demonstrate that improving Arctic radiative flux simulations cannot rely on isolated parameter tuning but requires a physically
425 consistent representation of the coupled boundary layer–cloud–surface system. In particular, accurate simulation of low-level liquid-containing clouds in both low-LWP and high-LWP regimes, realistic sea ice albedo fields, and faithful boundary-layer coupling are jointly necessary to avoid compensating shortwave–longwave errors. As Arctic amplification accelerates and cloud regimes evolve, reducing structural biases in cloud optical properties becomes critical for credible projections of sea ice retreat and regional climate feedbacks. Process-constrained evaluation of cloud–radiation coupling, such as presented here,
430 should therefore be a cornerstone of future Arctic regional model development.



Appendix A: WRF model setup

Table A1. WRF model setup

Land-Surface	NOAH-LSM (Tewari et al., 2004)
Surface Layer	MYNN (Olson et al., 2021)
PBL	MYNN2 (NAKANISHI and NIINO, 2009)
Microphysics	Thompson (Thompson et al., 2008) Single-moment scheme + double-moment cloud ice
Convection	Kain-Fritsch (Kain, 2004)
Radiation	RRTMG (Iacono et al., 2008)

Appendix B: Cloud diagnostics

B1 Cloud overlap formulation

The total cloudiness of a column (CLT) can be calculated using a maximum-random overlap assumption, as it is used in WRF
435 RRTMG scheme :

$$\text{CLT} = 1 - \prod_{g=1}^n (1 - C_g) \quad \text{with} \quad C_g = \max_{k \in G_g} C_k \quad (\text{B1})$$

Here, C_g is the maximum cloud fraction within group g , and G_g is the set of contiguous vertical layers where the cloud fraction $C_k > 0$, meaning that each group G_g contains adjacent cloudy layers (i.e., no clear-sky layers in between), and n is the total number of such groups in the vertical profile.

440 B2 Cloud occurrence statistics



Table B1. Cloudiness occurrence according to the criteria detailed in section 3.2.1. The frequency of observed clouds is calculated relative to the total number of cases, whereas the frequency of simulated clouds is calculated relative to the same total. The temporal overlap is defined as the proportion of simulated cloudy cases within the set of observed cloudy cases.

	P1	P2	P3	Full period
Total cases	808	683	454	1945
Frequency of observed clouds (%)	19	67	87	58
Frequency of simulated clouds (%)	38	48	78	51
Temporal overlap (%)	48	50	80	59



Code and data availability. The modified version of the WRF model optimized for polar regions used in this study can be found at <https://github.com/WRF-Chem-Polar/WRF-Chem-Polar/>. The WRF preprocessing system (WPS) is available at <https://archive.softwareheritage.org/swh:1:dir:21227ff84043afa53bb870245da4061fe7f0c7ab;origin=https://github.com/wrf-model/WPS;visit=swh:1:snp:096256316e752343901abad92a7dd9c2529f48cb;anchor=swh:1:rev:5a2ae63988e632405a4504cfb143ce7f0230a7a0>. The
445 NCEP FNL global analyses are available at <https://doi.org/10.5065/D6M043C6> and AMSR2 sea ice concentration data for the Arctic at <https://doi.org/10.1594/PANGAEA.898399>. The N-ICE2015 observational datasets for surface meteorology, radiation, and radiosondes respectively, can be found at <https://doi.org/10.21334/NPOLAR.2015.056A61D1>, <https://doi.org/10.21334/NPOLAR.2016.A89CB766>, and <https://doi.org/10.21334/NPOLAR.2017.216DF9B3>.

Author contributions. YLG, JCR and LM contributed to project conceptualization, developed the methodology, data curation, and devel-
450 opped software. YLG and LM performed the data validation. JCR and LM supervised this project and provided resources. JCR administered funding acquisition. YLG performed the formal analysis and investigation, prepared the figures, and wrote the initial manuscript draft. All authors contributed to manuscript revisions.

Competing interests. None of the authors has any competing interests.

Acknowledgements. This research has received funding from the French National Research Agency (ANR) via the project MPC2 (n° ANR-
455 22-CEA01-0009-02), and from the European Union's Horizon 2020 research and innovation programme under Grant 101003826 via project CRiceS (Climate Relevant interactions and feedbacks: the key role of sea ice and Snow in the polar and global climate system). Computer analyses benefited from access to IDRIS HPC resources (GENCI allocations A019017141 and A017017141), and from the IPSL mesocenter ESPRI facility which is supported by CNRS, UPMC, Labex L-IPSL, CNES and Ecole Polytechnique.



References

- 460 Boeke, R. C. and Taylor, P. C.: Evaluation of the Arctic surface radiation budget in CMIP5 models, *Journal of Geophysical Research: Atmospheres*, 121, 8525–8548, <https://doi.org/10.1002/2016JD025099>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD025099](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD025099), 2016.
- Chen, Y., Aires, F., Francis, J. A., and Miller, J. R.: Observed Relationships between Arctic Longwave Cloud Forcing and Cloud Parameters Using a Neural Network, *Journal of Climate*, 19, 4087 – 4104, <https://doi.org/10.1175/JCLI3839.1>, 2006.
- 465 Cohen, L., Hudson, S. R., Walden, V. P., Graham, R. M., and Granskog, M. A.: Meteorological conditions in a thinner Arctic sea ice regime from winter to summer during the Norwegian Young Sea Ice expedition (N-ICE2015), *Journal of Geophysical Research: Atmospheres*, 122, 7235–7259, <https://doi.org/10.1002/2016JD026034>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD026034](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD026034), 2017.
- Curry, J. A., Rossow, W. B., Randall, D., and Schramm, J. L.: Overview of Arctic Cloud and Radiation Characteristics, *Journal of Climate*, 9, 1731–1764, <https://www.jstor.org/stable/26201371>, publisher: American Meteorological Society, 1996.
- 470 Ehrlich, A., Wendisch, M., Lüpkes, C., Buschmann, M., Bozem, H., Chechin, D., Clemen, H.-C., Dupuy, R., Eppers, O., Hartmann, J., Herber, A., Jäkel, E., Järvinen, E., Jourdan, O., Kästner, U., Kliesch, L.-L., Köllner, F., Mech, M., Mertes, S., Neuber, R., Ruiz-Donoso, E., Schnaiter, M., Schneider, J., Stapf, J., and Zanatta, M.: A comprehensive in situ and remote sensing data set from the Arctic CLOUD Observations Using airborne measurements during polar Day (ACLOUD) campaign, *Earth System Science Data*, 11, 1853–1881, <https://doi.org/10.5194/essd-11-1853-2019>, publisher: Copernicus GmbH, 2019.
- 475 Gierens, R., Kneifel, S., Shupe, M. D., Ebell, K., Maturilli, M., and Löhnert, U.: Low-level mixed-phase clouds in a complex Arctic environment, *Atmospheric Chemistry and Physics*, 20, 3459–3481, <https://doi.org/10.5194/acp-20-3459-2020>, publisher: Copernicus GmbH, 2020.
- Granskog, M.: Arctic Research on Thin Ice: Consequences of Arctic Sea Ice Loss, <https://eos.org/science-updates/arctic-research-on-thin-ice-consequences-of-arctic-sea-ice-loss>, 2016.
- 480 Granskog, M. A., Fer, I., Rinke, A., and Steen, H.: Atmosphere-Ice-Ocean-Ecosystem Processes in a Thinner Arctic Sea Ice Regime: The Norwegian Young Sea ICE (N-ICE2015) Expedition, *Journal of Geophysical Research: Oceans*, 123, 1586–1594, <https://doi.org/10.1002/2017JC013328>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017JC013328](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2017JC013328), 2018.
- Griesche, H. J., Barrientos-Velasco, C., Deneke, H., Hünnerbein, A., Seifert, P., and Macke, A.: Low-level Arctic clouds: a blind zone in our knowledge of the radiation budget, *Atmospheric Chemistry and Physics*, 24, 597–612, <https://doi.org/10.5194/acp-24-597-2024>, publisher: Copernicus GmbH, 2024.
- 485 Hong, X. and Jiang, Q.: Impact of Land Surface Snow Processes on the Arctic Stable Boundary Layer, *Journal of Hydrometeorology*, 25, 479–494, <https://doi.org/10.1175/JHM-D-23-0040.1>, publisher: American Meteorological Society Section: Journal of Hydrometeorology, 2024.
- 490 Hudson, S. R., Cohen, L., Walden, V., and Institute, N. P.: N-ICE2015 surface meteorology, <https://doi.org/10.21334/NPOLAR.2015.056A61D1>, 2015.
- Hudson, S. R., Cohen, L., Walden, V. P., Institute, N. P., and University, W. S.: N-ICE2015 surface broadband radiation data, <https://doi.org/10.21334/NPOLAR.2016.A89CB766>, 2016.
- Hudson, S. R., Cohen, L., Kayser, M., Maturilli, M., Kim, J.-H., Park, S.-J., Moon, W., and Granskog, M. A.: N-ICE2015 atmospheric profiles from radiosondes, <https://doi.org/10.21334/NPOLAR.2017.216DF9B3>, 2017.
- 495



- Huschke, R. E.: ARCTIC CLOUD STATISTICS FROM 'AIR-CALIBRATED' SURFACE WEATHER OBSERVATIONS., <https://apps.dti.cmil/sti/html/tr/AD0698740/>, number: RM6173PR, 1969.
- Iacono, M. J., Delamere, J. S., Mlawer, E. J., Shephard, M. W., Clough, S. A., and Collins, W. D.: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models, *Journal of Geophysical Research: Atmospheres*, 113, <https://doi.org/10.1029/2008JD009944>, 2008.
- Inoue, J., Sato, K., Rinke, A., Cassano, J. J., Fettweis, X., Heinemann, G., Matthes, H., Orr, A., Phillips, T., Seefeldt, M., Solomon, A., and Webster, S.: Clouds and Radiation Processes in Regional Climate Models Evaluated Using Observations Over the Ice-free Arctic Ocean, *Journal of Geophysical Research: Atmospheres*, 126, e2020JD033904, <https://doi.org/10.1029/2020JD033904>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020JD033904>, 2021.
- 505 Intrieri, J. M., Fairall, C. W., Shupe, M. D., Persson, P. O. G., Andreas, E. L., Guest, P. S., and Moritz, R. E.: An annual cycle of Arctic surface cloud forcing at SHEBA, *Journal of Geophysical Research: Oceans*, 107, SHE 13–1–SHE 13–14, <https://doi.org/10.1029/2000JC000439>, 2002.
- Jiang, Z., Ding, M., Zhong, L., Li, Y., and Hu, X.: Seasonal variations of Arctic cloud in recent 14 years using CALIPSO-GOCCP, *Atmospheric Research*, 309, 107598, <https://doi.org/10.1016/j.atmosres.2024.107598>, 2024.
- 510 Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, *Journal of Applied Meteorology*, 43, 170 – 181, [https://doi.org/10.1175/1520-0450\(2004\)043<0170:TKCPAU>2.0.CO;2](https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2), place: Boston MA, USA Publisher: American Meteorological Society, 2004.
- Kay, J. E. and Gettelman, A.: Cloud influence on and response to seasonal Arctic sea ice loss, *Journal of Geophysical Research: Atmospheres*, 114, <https://doi.org/10.1029/2009JD011773>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD011773](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2009JD011773), 2009.
- 515 Kay, J. E., L'Ecuyer, T., Chepfer, H., Loeb, N., Morrison, A., and Cesana, G.: Recent Advances in Arctic Cloud and Climate Research, *Current Climate Change Reports*, 2, 159–169, <https://doi.org/10.1007/s40641-016-0051-9>, 2016.
- Kayser, M., Maturilli, M., Graham, R. M., Hudson, S. R., Rinke, A., Cohen, L., Kim, J.-H., Park, S.-J., Moon, W., and Granskog, M. A.: Vertical thermodynamic structure of the troposphere during the Norwegian young sea ICE expedition (N-ICE2015), *Journal of Geophysical Research: Atmospheres*, 122, 10,855–10,872, <https://doi.org/10.1002/2016JD026089>, [_eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD026089](https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1002/2016JD026089), 2017.
- 520 Klein, S. A., McCoy, R. B., Morrison, H., Ackerman, A. S., Avramov, A., Boer, G. d., Chen, M., Cole, J. N. S., Del Genio, A. D., Falk, M., Foster, M. J., Fridlind, A., Golaz, J.-C., Hashino, T., Harrington, J. Y., Hoose, C., Khairoutdinov, M. F., Larson, V. E., Liu, X., Luo, Y., McFarquhar, G. M., Menon, S., Neggers, R. A. J., Park, S., Poellot, M. R., Schmidt, J. M., Sednev, I., Shipway, B. J., Shupe, M. D., Spangenberg, D. A., Sud, Y. C., Turner, D. D., Veron, D. E., Salzen, K. v., Walker, G. K., Wang, Z., Wolf, A. B., Xie, S., Xu, K.-M., Yang, F., and Zhang, G.: Intercomparison of model simulations of mixed-phase clouds observed during the ARM Mixed-Phase Arctic Cloud Experiment. I: single-layer cloud, *Quarterly Journal of the Royal Meteorological Society*, 135, 979–1002, <https://doi.org/10.1002/qj.416>, [_eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.416](https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/qj.416), 2009.
- 525 Liu, Y., Key, J. R., Ackerman, S. A., Mace, G. G., and Zhang, Q.: Arctic cloud macrophysical characteristics from CloudSat and CALIPSO, *Remote Sensing of Environment*, 124, 159–173, <https://doi.org/10.1016/j.rse.2012.05.006>, 2012.
- 530 Maillard, J., Ravetta, F., Raut, J.-C., Mariage, V., and Pelon, J.: Characterisation and surface radiative impact of Arctic low clouds from the IAOOS field experiment, *Atmospheric Chemistry and Physics*, 21, 4079–4101, <https://doi.org/10.5194/acp-21-4079-2021>, 2021.



- Marelle, L., Myhre, G., Thomas, J. L., and Raut, J.-C.: Aerosol Background Concentrations Influence Aerosol-Cloud Interactions as Much as the Choice of Aerosol-Cloud Parameterization, *Geophysical Research Letters*, 52, e2024GL111780, <https://doi.org/10.1029/2024GL111780>, _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2024GL111780>, 2025.
- 535 Mariage, V., Pelon, J., Blouzon, F., and Victori, S.: IAOOS Microlidar Development and First Results Obtained During 2014 and 2015 Arctic Drifts, *EPJ Web of Conferences*, 119, 02005, <https://doi.org/10.1051/epjconf/201611902005>, 2016.
- Mech, M., Ehrlich, A., Herber, A., Lüpkes, C., Wendisch, M., Becker, S., Boose, Y., Chechin, D., Crewell, S., Dupuy, R., Gourbeyre, C., Hartmann, J., Jäkel, E., Jourdan, O., Kliesch, L.-L., Klingebiel, M., Kulla, B. S., Mioche, G., Moser, M., Risse, N., Ruiz-Donoso, E., Schäfer, M., Stapf, J., and Voigt, C.: MOSAiC-ACA and AFLUX - Arctic airborne campaigns characterizing the exit area of MOSAiC, *Scientific Data*, 9, 790, <https://doi.org/10.1038/s41597-022-01900-7>, 2022.
- 540 Melsheimer, C. and Spreen, G.: AMSR2 ASI sea ice concentration data, Arctic, version 5.4 (NetCDF) (July 2012 - December 2019), <https://doi.org/10.1594/PANGAEA.898399>, 2019.
- Mioche, G., Jourdan, O., Ceccaldi, M., and Delanoë, J.: Variability of mixed-phase clouds in the Arctic with a focus on the Svalbard region: a study based on spaceborne active remote sensing, *Atmospheric Chemistry and Physics*, 15, 2445–2461, <https://doi.org/10.5194/acp-15-2445-2015>, publisher: Copernicus GmbH, 2015.
- 545 Morrison, H., de Boer, G., Feingold, G., Harrington, J., Shupe, M. D., and Sulia, K.: Resilience of persistent Arctic mixed-phase clouds, *Nature Geoscience*, 5, 11–17, <https://doi.org/10.1038/ngeo1332>, publisher: Nature Publishing Group, 2012.
- Nakanishi, M. and Michibata, T.: How Does Cloud Emissivity Feedback Affect Present and Future Arctic Warming?, *Ocean-Land-Atmosphere Research*, 4, 0089, <https://doi.org/10.34133/olar.0089>, 2025.
- 550 NAKANISHI, M. and NIINO, H.: Development of an Improved Turbulence Closure Model for the Atmospheric Boundary Layer, *Journal of the Meteorological Society of Japan. Ser. II*, 87, 895–912, <https://doi.org/10.2151/jmsj.87.895>, 2009.
- National Centers for Environmental Prediction, National Weather Service, NOAA, U.S. Department of Commerce: NCEP FNL Operational Model Global Tropospheric Analyses, continuing from July 1999, <https://doi.org/10.5065/D6M043C6>, place: Boulder, CO, 2000.
- Niemelä, S., Räisänen, P., and Savijärvi, H.: Comparison of surface radiative flux parameterizations: Part I: Longwave radiation, *Atmospheric Research*, 58, 1–18, [https://doi.org/10.1016/S0169-8095\(01\)00084-9](https://doi.org/10.1016/S0169-8095(01)00084-9), 2001.
- 555 Olson, J. B., Smirnova, T., Kenyon, J. S., Turner, D. D., Brown, J. M., Zheng, W., and Green, B. W.: A Description of the MYNN Surface-Layer Scheme, Tech. rep., <https://repository.library.noaa.gov/view/noaa/30605>, 2021.
- Peng, Y., Lohmann, U., Leaitch, R., Banic, C., and Couture, M.: The cloud albedo-cloud droplet effective radius relationship for clean and polluted clouds from RACE and FIRE.ACE, *Journal of Geophysical Research: Atmospheres*, 107, AAC 1–1–AAC 1–6, <https://doi.org/10.1029/2000JD000281>, _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2000JD000281>, 2002.
- 560 Persson, O. and Vihma, T.: The atmosphere over sea ice, in: *Sea Ice*, pp. 160–196, John Wiley & Sons, Ltd, ISBN 978-1-118-77837-1, <https://doi.org/10.1002/9781118778371.ch6>, section: 6 _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9781118778371.ch6>, 2017.
- Rogers, R. R. and Yau, M. K.: A short course in cloud physics, no. 113 in *International series in natural philosophy*, Pergamon press, Oxford
- 565 New York Beijing [etc.], 3rd ed edn., ISBN 978-0-08-034863-6, 1989.
- Sedlar, J., Tjernström, M., Rinke, A., Orr, A., Cassano, J., Fettweis, X., Heinemann, G., Seefeldt, M., Solomon, A., Matthes, H., Phillips, T., and Webster, S.: Confronting Arctic Troposphere, Clouds, and Surface Energy Budget Representations in Regional Climate Models With Observations, *Journal of Geophysical Research: Atmospheres*, 125, e2019JD031783, <https://doi.org/10.1029/2019JD031783>, publisher: John Wiley & Sons, Ltd, 2020.



- 570 Serreze, M. C., Cassano, E., Crawford, A., Cassano, J. J., and Zhang, C.: The Observed Evolution of Arctic Amplification over the Past 45 Years, <https://doi.org/10.5194/egusphere-2025-3690>, 2025.
- Shaw, J., McGraw, Z., Bruno, O., Storelvmo, T., and Hofer, S.: Using Satellite Observations to Evaluate Model Microphysical Representation of Arctic Mixed-Phase Clouds, *Geophysical Research Letters*, 49, e2021GL096191, <https://doi.org/10.1029/2021GL096191>, 2022.
- Shupe, M. D.: Clouds at Arctic Atmospheric Observatories. Part II: Thermodynamic Phase Characteristics, *Journal of Applied Meteorology and Climatology*, <https://doi.org/10.1175/2010JAMC2468.1>, 2011.
- 575 Shupe, M. D. and Intrieri, J. M.: Cloud Radiative Forcing of the Arctic Surface: The Influence of Cloud Properties, Surface Albedo, and Solar Zenith Angle, *Journal of Climate*, https://journals.ametsoc.org/view/journals/clim/17/3/1520-0442_2004_017_0616_crfota_2.0.co_2.xml, 2004.
- Shupe, M. D., Walden, V. P., Eloranta, E., Uttal, T., Campbell, J. R., Starkweather, S. M., and Shiobara, M.: Clouds at Arctic Atmospheric Observatories. Part I: Occurrence and Macrophysical Properties, *Journal of Applied Meteorology and Climatology*, <https://doi.org/10.1175/2010JAMC2467.1>, 2011.
- 580 Silber, I. and Shupe, M. D.: Insights on sources and formation mechanisms of liquid-bearing clouds over MOSAiC examined from a Lagrangian framework, *Elementa: Science of the Anthropocene*, 10, 000071, <https://doi.org/10.1525/elementa.2021.000071>, 2022.
- Skamarock, C., Klemp, B., Dudhia, J., Gill, O., Liu, Z., Berner, J., Wang, W., Powers, G., Duda, G., Barker, D. M., and Huang, X.: A Description of the Advanced Research WRF Model Version 4, <https://api.semanticscholar.org/CorpusID:196211930>, 2019.
- 585 Stephens, G. L.: Radiation Profiles in Extended Water Clouds. II: Parameterization Schemes, *Journal of the Atmospheric Sciences*, https://journals.ametsoc.org/view/journals/atsc/35/11/1520-0469_1978_035_2123_rpiewc_2_0_co_2.xml, 1978.
- Stramler, K., Genio, A. D. D., and Rossow, W. B.: Synoptically Driven Arctic Winter States, *Journal of Climate*, <https://doi.org/10.1175/2010JCLI3817.1>, 2011.
- 590 Tao, W.-K., Lang, S., Zeng, X., Li, X., Matsui, T., Mohr, K., Posselt, D., Chern, J., Peters-Lidard, C., Norris, P. M., Kang, I.-S., Choi, I., Hou, A., Lau, K. M., and Yang, Y.-M.: The Goddard Cumulus Ensemble model (GCE): Improvements and applications for studying precipitation processes, *Atmospheric Research*, 143, 392–424, <https://doi.org/10.1016/j.atmosres.2014.03.005>, 2014.
- Taylor, P. C. and Monroe, E.: Isolating the Surface Type Influence on Arctic Low-Clouds, *Journal of Geophysical Research: Atmospheres*, 128, e2022JD038098, <https://doi.org/10.1029/2022JD038098>, <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022JD038098>, 2023.
- 595 Taylor, P. C., Boeke, R. C., Li, Y., and Thompson, D. W. J.: Arctic cloud annual cycle biases in climate models, *Atmospheric Chemistry and Physics*, 19, 8759–8782, <https://doi.org/10.5194/acp-19-8759-2019>, publisher: Copernicus GmbH, 2019.
- Tewari, M., Chen, F., Wang, W., Dudhia, J., LeMone, M., Mitchell, K., Ek, M., Gayno, G., Wegiel, J., and Cuenca, R.: Implementation and verification of the unified noah land surface model in the WRF model, *Bulletin of the American Meteorological Society*, pp. 2165–2170, combined Preprints: 84th American Meteorological Society (AMS) Annual Meeting ; Conference date: 11-01-2004 Through 15-01-2004, 2004.
- 600 Thompson, G., Field, P. R., Rasmussen, R. M., and Hall, W. D.: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part II: Implementation of a New Snow Parameterization, *Monthly Weather Review*, 136, 5095–5115, <https://doi.org/10.1175/2008MWR2387.1>, 2008.
- 605 Walden, V. P., Hudson, S. R., Cohen, L., Murphy, S. Y., and Granskog, M. A.: Atmospheric components of the surface energy budget over young sea ice: Results from the N-ICE2015 campaign, *Journal of Geophysical Research: Atmospheres*, 122, 8427–8446, <https://doi.org/10.1002/2016JD026091>, 2017.



- 610 Wang, X. and Key, J. R.: Arctic Surface, Cloud, and Radiation Properties Based on the AVHRR Polar Pathfinder Dataset. Part I: Spatial and Temporal Characteristics, *Journal of Climate*, 18, 2558–2574, <https://doi.org/10.1175/JCLI3438.1>, publisher: American Meteorological Society Section: *Journal of Climate*, 2005.
- 615 Wendisch, M., Crewell, S., Ehrlich, A., Herber, A., Kirbus, B., Lüpkes, C., Mech, M., Abel, S. J., Akansu, E. F., Ament, F., Aubry, C., Becker, S., Borrmann, S., Bozem, H., Brückner, M., Clemen, H.-C., Dahlke, S., Dekoutsidis, G., Delanoë, J., De La Torre Castro, E., Dorff, H., Dupuy, R., Eppers, O., Ewald, F., George, G., Gorodetskaya, I. V., Grawe, S., Groß, S., Hartmann, J., Henning, S., Hirsch, L., Jäkel, E., Joppe, P., Jourdan, O., Jurányi, Z., Karalis, M., Kellermann, M., Klingebiel, M., Lonardi, M., Lucke, J., Luebke, A. E., Maahn, M., 620 Mahernndl, N., Maturilli, M., Mayer, B., Mayer, J., Mertes, S., Michaelis, J., Michalkov, M., Mioche, G., Moser, M., Müller, H., Neggers, R., Ori, D., Paul, D., Paulus, F. M., Pilz, C., Pithan, F., Pöhlker, M., Pörtge, V., Ringel, M., Risse, N., Roberts, G. C., Rosenburg, S., Röttenbacher, J., Rückert, J., Schäfer, M., Schaefer, J., Schemann, V., Schirmacher, I., Schmidt, J., Schmidt, S., Schneider, J., Schnitt, S., Schwarz, A., Siebert, H., Sodemann, H., Sperzel, T., Spreen, G., Stevens, B., Stratmann, F., Svensson, G., Tatzelt, C., Tuch, T., Vihma, T., Voigt, C., Volkmer, L., Walbröl, A., Weber, A., Wehner, B., Wetzel, B., Wirth, M., and Zinner, T.: Overview: quasi-Lagrangian observations 625 of Arctic air mass transformations – introduction and initial results of the HALO–(*A C*)³ aircraft campaign, *Atmospheric Chemistry and Physics*, 24, 8865–8892, <https://doi.org/10.5194/acp-24-8865-2024>, publisher: Copernicus GmbH, 2024.
- Xu, K.-M. and Randall, D. A.: A Semiempirical Cloudiness Parameterization for Use in Climate Models, *Journal of the Atmospheric Sciences*, https://journals.ametsoc.org/view/journals/atsc/53/21/1520-0469_1996_053_3084_ascpfu_2_0_co_2.xml, 1996.
- 625 Yu, Y., Taylor, P. C., and Cai, M.: Seasonal Variations of Arctic Low-Level Clouds and Its Linkage to Sea Ice Seasonal Variations, *Journal of Geophysical Research: Atmospheres*, 124, 12 206–12 226, <https://doi.org/10.1029/2019JD031014>, _eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2019JD031014>, 2019.