Standardized daily high-resolution large-eddy simulations of the Arctic boundary layer and clouds during the complete MOSAiC drift

JAMES | Journal of Advances in Modeling Earth Systems

RESEARCH ARTICLE
10.1029/2024MS004296

Standardized Daily High-Resolution Large-Eddy Simulations of the Arctic Boundary Layer and Clouds During the Complete *MOSAiC* Drift

N. Schnierstein[1], J. Chylik[1,2], M. D. Shupe[3,4], and R. A. J. Neggers[1]

[1]Institute for Geophysics and Meteorology, University of Cologne, Cologne, Germany, [2]Department of Meteorology, University of Reading, Reading, UK, [3]Physical Sciences Laboratory, National Oceanic and Atmospheric Administration, Boulder, CO, USA, [4]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, Boulder, CO, USA

Key Points:

- A standardized LES setup based on campaign data is developed with an aim to supplement the local measurements during the *Multidisciplinary drifting Observatory for the Study of Arctic Climate* drift
- Independent drift-long statistics on key aspects of the surface energy budget, thermodynamic structure, and clouds are reproduced
- Sensitivity tests indicate microphysics, ice-radiation interaction and surface representation are critical for successful daily simulations

Correspondence to:
N. Schnierstein,
nschnier@uni-koeln.de

Citation:
Schnierstein, N., Chylik, J., Shupe, M. D., & Neggers, R. A. J. (2024). Standardized daily high-resolution large-eddy simulations of the Arctic boundary layer and clouds during the complete *MOSAiC* drift. *Journal of Advances in Modeling Earth Systems*, *16*, e2024MS004296. https://doi.org/10.1029/2024MS004296

Received 20 FEB 2024
Accepted 28 OCT 2024

Abstract This study utilizes the wealth of observational data collected during the recent *Multidisciplinary drifting Observatory for the Study of Arctic Climate* (*MOSAiC*) drift experiment to constrain and evaluate close to two-hundred daily Large-Eddy Simulations (LES) of Arctic boundary layers and clouds at high resolutions. A standardized approach is adopted to tightly integrate field measurements into the experimental configuration. Covering the full drift represents a step forward from single-case LES studies, and allows for a robust assessment of model performance against independent data under a range of atmospheric conditions. A homogeneously forced domain is simulated in a Lagrangian frame of reference, initialized with radiosonde and value-added cloud profiles. Prescribed boundary conditions include various measured surface characteristics. Time-constant composite forcing is applied, primarily consisting of subsidence rates sampled from reanalysis data. The simulations run for 3 hours, allowing turbulence and clouds to spin up while still facilitating direct comparison to *MOSAiC* data. Key aspects such as the vertical thermodynamic structure, cloud properties, and surface energy fluxes are well reproduced and maintained. The model captures the bimodal distribution of atmospheric states that is typical of Arctic climate. Selected days are investigated more closely to assess the model's skill in maintaining the observed boundary layer structure. The sensitivity to various aspects of the experimental configuration and model physics is tested. The model input and output are available to the scientific community, supplementing the *MOSAiC* data archive. The close agreement with observed meteorology justifies the use of LES for gaining further insight into Arctic boundary layer processes and their role in Arctic climate change.

Plain Language Summary The Arctic is one of the regions most affected by global climate change, warming up to four times as fast as the rest of the globe. It is also a particularly inaccessible region to conduct measurements. Fortunately, between 2019 and 2020 the *Multidisciplinary drifting Observatory for the Study of Arctic Climate* (*MOSAiC*) campaign collected an unprecedented amount of data in the Arctic. In this study, numerous of these measurements are incorporated into high-resolution computer simulations of the lowest part of the Arctic atmosphere. This simulation data complements and contextualizes the observations and enables insight into complex physical processes, for example, cloud formation, cloud ice production, or turbulent mixing. The Arctic is an extreme place, and models often struggle to represent the atmosphere accurately. Therefore, the main achievement of this study is to successfully simulate 190 atmospheric situations as measured during the campaign. The generated data set performs well when compared to independent observations. Single cases deliver information about individual atmospheric conditions, and the collection gives insight into how key climate variables behaved throughout the *MOSAiC* year.

1. Introduction

The ongoing rapid warming of the Arctic region is a significant contributor to global climate change. Due to amplification processes, the Arctic warms up to four times as fast as the rest of the globe (Rantanen et al., 2022; Serreze & Barry, 2011). Past research suggests various feedback processes cause this Arctic Amplification (AA) (Wendisch et al., 2023), including the albedo (Dai et al., 2019; Jenkins & Dai, 2021; Screen & Simmonds, 2010; Thackeray & Hall, 2019), Planck (Pithan & Mauritsen, 2014), water-vapor, lapse rate (Linke et al., 2023; Stuecker et al., 2018), and cloud feedback (Middlemas et al., 2020; Philipp et al., 2020). Processes in the turbulent boundary layer and associated clouds play an important role in most of these feedbacks, motivating and

prioritizing research on this part of the Arctic atmosphere (Taylor et al., 2018). While measurements at high latitudes play a key role in this research, modeling efforts can expand and deepen this effort, by providing insight into Arctic boundary layer physics and dynamics covering a broad range of scales, from the global scale down to the turbulent scale.

The observational investigation of Arctic boundary layers is challenging for various reasons. The dynamic sea-ice complicates the operation of permanent measuring stations. Reliable meteorological data from ship-based and airborne measurement campaigns is sparse, and often suffers from gaps in the spatial and temporal coverage. The *Multidisciplinary drifting Observatory for the Study of Arctic Climate* (*MOSAiC*, Shupe et al. (2020)) was a major international field campaign that took place in the Arctic from September 2019 to October 2020, intending to fill essential data gaps in in situ observations in the central Arctic. The campaign covered a complete annual cycle and involved the deployment of the German research vessel *Polarstern*, which drifted with the sea ice across the Arctic for over a year. During this time an unprecedented amount of highly diverse data was successfully collected, which is now available to the research communities in atmospheric, cryospheric, and oceanic sciences. For a more detailed overview refer to Shupe et al. (2022); Nicolaus et al. (2022); Rabe et al. (2022).

The MOSAiC data set creates unique opportunities for high-resolution modeling of Arctic boundary layers and clouds, in particular Large-Eddy Simulation (LES) (Fridlind & Ackerman, 2018). The spatial and temporal discretizations applied in LES are fine enough to resolve small-scale boundary layer processes such as cloud dynamics and turbulence, while parameterization techniques represent remaining unresolved subgrid-scale variability, the surface energy exchange, and micro-physical processes. Various model intercomparison studies have provided insight into model skill, identifying strengths but also weaknesses and sensitivities in representing (mixed-phase) cloud properties (Klein et al., 2009; Fridlind et al., 2012; Ovchinnikov et al., 2014; R. G. Stevens et al., 2018). There is strong evidence that a high spatial resolution ($\mathcal{O}$(1 m)), especially in the vertical, is needed to accurately represent the turbulent dynamics in an Arctic atmosphere, characterized by strong inversions and small amounts of energy (Couvreux et al., 2020; Rauterkus & Ansorge, 2020; van der Linden et al., 2020). Several LES studies evaluating model results against data from recent field campaigns in the Arctic have emphasized the importance of a careful setup to achieve an accurate representation of the observed Arctic atmosphere in the virtual domain (Chylik et al., 2023; Morrison et al., 2011; Neggers et al., 2019; Savre & Ekman, 2015), including precise initial and boundary conditions, appropriate model parameterizations and resolution, and improving model-observation comparability. The shear variety and long time coverage of *MOSAiC* data sets is a step forward in this respect, helping to more tightly constrain the experimental setup while simultaneously providing sufficient further independent data to evaluate resolved model results.

The overarching goal of this study is to create a suite of near two-hundred daily high-resolution LES experiments that together represent the conditions of the lower Arctic atmosphere as observed during the full *MOSAiC* year. Days are selected for simulation based on the availability of data required to constrain the numerical experiments. The work, which is a substantial computational effort, has two main science objectives. First, the aim is to generate a library of virtual realizations of the Arctic atmosphere at resolutions high enough to resolve most of the turbulence spectrum, and which are realistic enough to supplement the MOSAiC measurements with information on small-scale variability that is difficult to measure. The second aim is to identify remaining weaknesses in the model in a statistically significant way, by evaluating the simulation results against a selection of *MOSAiC* data sets under a broad range of year-round conditions (Peng et al., 2023).

A standardized model setup is adopted to simulate all cases, which is crucial for various reasons. Consistent representativity for the full year (the first main objective) can only be achieved when the initial and boundary conditions, model settings, and evaluation are all applied without any manual adjustment on a case-to-case basis. This approach thus sets it apart from single-case LES studies, which tend to prioritize optimization against short-term data sets. In addition, discovering systematic biases or parameterization problems for the whole year (the second main objective) is only possible when the same model setup is applied in all daily experiments.

Finally, a standardized setup is needed for effectively conducting internally-consistent year-round impact studies on aspects of the model and experimental setup. A standardized LES setup has recently successfully been applied to conduct long-term simulations at various meteorological sites (Gustafson Jr. et al., 2020; Neggers et al., 2012; Schalkwijk et al., 2015; van Laar et al., 2019). This includes Arctic sites (Kiszler et al., 2023), but not yet in the central Arctic. In this study, the *Polarstern* and the surrounding measurement stations on the sea ice are used to this purpose, acting as a drifting 1-year meteorological site.

The model evaluation in this study is limited to a statistically significant assessment of performance for a set of basic meteorological quantities that were robustly and continuously measured during the drift. This includes terms in the surface energy and radiative budgets, integrated cloud water, and near-surface meteorology. Reproduction of the well-known bimodal nature of the Arctic atmosphere will also be investigated (Sverdrup, 1933). The evaluation of more detailed properties of clouds and turbulence against measurements is beyond the scope of this paper, and will be addressed in future research.

Section 2 gives a detailed description of the method, including an overview of the data sets used, the numerical model framework, and the configuration of the experiments based on the *MOSAiC* data. The results of the simulation and their evaluation against measurements are presented in Section 3. A detailed discussion of the results' meaning, implications, and limitations is found in Section 4, including smaller-grid sensitivity tests on key aspects of the model and experiment setup. Finally, Section 5 summarizes the main conclusions of this study and provides an outlook on future research it might inspire.

2. Data and Method

2.1. *MOSAiC* Data Sets

At the foundation of the drift-covering library of high-resolution Large Eddy Simulations of the Arctic atmospheric boundary layer discussed in this study are the observational data sets collected during the *MOSAiC* drift. Table 1 gives a complete overview of the data sets used to set, constrain and evaluate the LES experiments. It includes the variable, its units and full name, the dimensionality, the associated instrument, and the scientific study presenting the data set. Refer to these publications for a more detailed technical and scientific description of these data sets and the associated instruments.

Vertical profiles of the thermodynamic state, including temperature and water vapor specific humidity, are provided by the daily 11:00 UTC radiosonde launches during the *MOSAiC* campaign (Dahlke et al., 2023; Maturilli et al., 2022) or have been directly calculated from the measured variables. Vertical cloud liquid and ice water content profiles are obtained from the value-added cloud product based on remote sensing data sets as described by Shupe et al. (2015) (*ShupeTurner* data set). Information on cloud condensation nuclei concentrations is based on aerosol measurements (Koontz et al., 2020), while ice nucleating particle concentrations are based on the aerosol measurements by Creamean (2022). The skin temperature of the sea ice is based on the brightness temperature measurements from the *MetCity* location on the sea ice (C. Cox et al., 2023; C. J. Cox et al., 2023). Also based on tower measurements are estimates of the surface aerodynamic roughness length. Finally, surface albedo values are obtained from satellite products (Istomina et al., 2020; Spreen et al., 2008).

A few data sets required for the configuration of the daily numerical control experiments are not mentioned in Table 1 because they are based on ERA5 reanalysis data (Hersbach et al., 2018a, 2018b, 2023). These include vertical profiles of horizontal wind and time series of the sea-ice fraction. One motivation for using reanalysis products for these variables is to make the simulated domain reflect a larger area and avoid introducing ultra-local effects. The reason for using the sea-ice concentration from ERA5 is that (a) assimilates interpolated satellite products (Donlon et al., 2012; ECMWF, 2016) and (b) it is time-continuous as a result. The impact of the sea-ice fraction on the results will be tested in Section 4. A few variables part of the prescribed large-scale forcings are also derived from ERA5 because they were not measured during the *MOSAiC* drift. These include large-scale vertical motion (subsidence) and pressure gradients, adopting the procedure implemented by Neggers et al. (2019) and van Laar et al. (2019).

The resulting simulations are evaluated against additional *MOSAiC* data sets independent from those used in the experimental design. Simulated short- and longwave radiative fluxes are compared to sensor data of the radiation station described by C. Cox et al. (2023), C. J. Cox et al. (2023), and Riihimaki (2021). This installation was part of *MetCity* with sensors placed at 3 m (upwelling) and 1.5 m (downwelling) height. Simulated near-surface turbulent heat fluxes are evaluated against data described by C. Cox et al. (2023) and C. J. Cox et al. (2023), who discuss two different heat flux data sets. One is directly retrieved from sonic-anemometer turbulence measurements, while the other is a bulk calculation based on Monin-Obukhov theory. The latter is used in this study, although C. Cox et al. (2023) and C. J. Cox et al. (2023) do mention potential over-simplifications, such as the assumption of constant roughness lengths. Since the surface flux parameterization in the model is also based on Monin-Obukhov theory, the use of bulk measurements yields the fairest comparison of the model to data.

Journal of Advances in Modeling Earth Systems

10.1029/2024MS004296

Table 1
Overview of Observational Data Sets Utilized to Constrain (C) and Evaluate (E) the Conducted Simulations

Variable	Unit	Description	Dimension	Instrument	Citation	Usage
α	%	Surface albedo	(t)	AMSR-E Satellite	Spreen et al. (2008); Istomina et al. (2020)	C
SIC	%	Sea-ice concentration	(t)	MODIS/AMSR2	Ludwig et al. (2020)	E
RH	%	Relative humidity	(z, t)	Radiosonde	Maturilli et al. (2022) and Dahlke et al. (2023)	C & E
T	K	Temperature	(z, t)	Radiosonde	Maturilli et al. (2022) and Dahlke et al. (2023)	C & E
p	Pa	Pressure	(z, t)	Radiosonde	Maturilli et al. (2022) and Dahlke et al. (2023)	C & E
q_{l}	kg kg^{-1}	Liquid water specific humidity	(z, t)	Multi-sensor retrieval	Shupe et al. (2015)	C & E
q_{i}	kg kg^{-1}	Ice water specific humidity	(z, t)	Multi-sensor retrieval	Shupe et al. (2015)	C & E
$\mathrm{LW_d}$	W m^{-2}	Longwave downward radiative flux	(t)	Radiation station	C. Cox et al. (2023), C. J. Cox, et al. (2023), and Riihimaki (2021)	E
$\mathrm{LW_u}$	W m^{-2}	Longwave upward radiative flux	(t)	Radiation station	C. Cox et al. (2023), C. J. Cox et al. (2023) and Riihimaki (2021)	E
$\mathrm{SW_d}$	W m^{-2}	Shortwave downward radiative flux	(t)	Radiation station	C. Cox et al. (2023), C. J. Cox et al. (2023) and Riihimaki (2021)	E
$\mathrm{SW_u}$	W m^{-2}	Shortwave upward radiative flux	(t)	Radiation station	C. Cox et al. (2023), C. J. Cox et al. (2023) and Riihimaki (2021)	E
T_{skin}	K	Skin temperature	(t)	Radiation station	C. Cox et al. (2023), C. J. Cox et al. (2023) and Riihimaki (2021)	C
$U_{10\,\mathrm{m}}$	m s^{-1}	10 m wind speed	(t)	Meteorological tower	C. Cox et al. (2023) and C. J. Cox et al. (2023)	E
$T_{10\,\mathrm{m}}$	K	10 m temperature	(t)	Meteorological tower	C. Cox et al. (2023) and C. J. Cox et al. (2023)	C & E
$H_{s,10\,\mathrm{m,bulk}}$	W m^{-2}	10 m Bulk sensible heat flux	(t)	Meteorological tower	C. Cox et al. (2023) and C. J. Cox et al. (2023)	E
$H_{s,10\,\mathrm{m,turb}}$	W m^{-2}	10 m Turbulent sensible heat flux	(t)	Meteorological tower	C. Cox et al. (2023) and C. J. Coxet al. (2023)	E
$z_{0,m}$	m	Roughness length for momentum	(t)	Meteorological tower	C. Cox et al. (2023) and C. J. Cox et al. (2023)	C
c_{CCN}	m^{-3}	Cloud condensation nuclei concentration	(t)	ARM optical particle counter	Koontz et al. (2020)	C
c_{INP}	m^{-3}	Ice nucleating particle concentration	(t)	DRUM aerosol sampler	Creamean (2022)	C

2.2. Model Setup

2.2.1. LES Code

The Dutch Large Eddy Simulation framework (*DALES*, Heus et al. (2010)) was utilized in this research. This code has been effectively employed in numerous boundary layer studies (de Roode et al., 2016; Van der Dussen et al., 2013; van Laar et al., 2019), particularly for conducting research in the Arctic region (Chylik et al., 2023; de Roode et al., 2019; Egerer et al., 2021; Neggers et al., 2019). The foundation of the model is the Ogura-Phillips anelastic equations for a set of prognostic variables, including liquid water potential temperature Θ_l, the wind velocity components $\{u, v, w\}$, total water specific humidity q_t, and the mass concentration as well as number concentration of various hydrometeor species. Momentum advection is calculated by a fifth-order central difference scheme and scalar advection by a κ-limiter scheme (Hundsdorfer et al., 1995). A prognostic subfilter-scale turbulent kinetic energy (SFS-TKE) model (Deardorff, 1980) calculates the subgrid-scale transport of heat, moisture, and momentum. The time integration makes use of a third-order Runge-Kutta scheme. Finally, to remove spurious perturbations such as gravity waves in the top of the domain, a sponge layer is applied across the top 25% of model levels, with increasing relaxation times of 6 min at the top of the domain to infinity at sponge layer bottom. To better represent mixed-phase clouds in a wide range of conditions, the cloud microphysics scheme, the radiation scheme, and the surface scheme were enhanced as described in more detail below.

2.2.2. Microphysics

The mixed-phase, double-moment micro-physics scheme by Seifert and Beheng (2006a) considers the mass and number concentration of five hydro-meteors; cloud droplets, cloud ice crystals, and the precipitating hydrometeors snow, graupel, and rain. The cloud condensation nuclei concentration is single-species and prognostic (number concentration changes due to microphysical processes and turbulent transport), with activation dependent on supersaturation with respect to water. The ice nucleating particle concentration is also single-species but prescribed, and its activation depends on temperature and supersaturation with respect to ice. Freezing modes represented in the Seifert and Beheng (2006a) scheme include deposition nucleation, immersion freezing and contact freezing (treated together as heterogeneous freezing), and homogeneous freezing at very low temperatures. The microphysics scheme was initially designed for simulating ice clouds in the mid-latitudes (Seifert & Beheng, 2006b) and, therefore, requires evaluation and partial adaptation to suit the extremely cold conditions in the high Arctic. Two key changes to the default implementation in DALES (Chylik et al., 2023) were made. First, limiting the effect of immersion freezing and contact freezing of cloud droplets at low temperatures. The overall heterogeneous freezing rates are calculated using the maximum of the actual atmospheric temperature and $T_{\text{het,lim}} = -15°\,\text{C}$, as temperatures below this threshold lead to the edges of the valid range of the parameterization being approached and exceeded (Pruppacher & Klett, 1996). Second, the maximum number concentration of nucleated ice particles produced by deposition-nucleation is limited to $c_{N,\text{ice}} = 200\ \text{L}^{-1}$ to prevent an unrealistically high concentration of tiny ice particles, which would exceed the estimated concentration of available ice nucleating particles. However, these limits do not apply to homogeneous freezing of cloud droplets (Jeffery & Austin, 1997) nor to the secondary production of ice by the Hallett-Mossop process, which are represented in the Seifert and Beheng (2006a) model. It is worth noting that developing microphysics schemes fit for the Arctic is a highly researched topic (Fridlind et al., 2012; Ong et al., 2022) that cannot be covered within the scope of this study, but by including the mentioned changes, reasonable ice specific humidity and ice water path have been achieved.

2.2.3. Radiation

By default, *DALES* uses a four-stream solver based on (Fu & Liou, 1993; Liou et al., 1988) in combination with Monte Carlo Spectral Integration (Pincus & Stevens, 2009) to calculate the vertical component of the radiative fluxes in the short- and longwave. Previously, the optical properties of ice water content in the radiation calculation were estimated to be identical to a liquid water content of the same mass and a prescribed constant number concentration. While this was expected to introduce an inevitable error in general, it proved to be an unusable assumption for thin ice clouds in the Arctic. For this work, the description of the optical properties of ice crystals followed Fu and Liou (1993). This parameterization was developed for cirrus ice clouds with effective ice

diameter in the range 20 – 120 µm and has been used in past Arctic LES studies (e.g., Savre and Ekman (2015); Dimitrelos et al. (2020)).

Prior to this study the *DALES* code had never been applied to central Arctic conditions. A necessary step was to supply the radiation scheme with realistic information about the effective diameter of ice crystals. These are here based on the physical properties of ice crystals (Baran, 2005; McFarquhar & Heymsfield, 1998), as estimated from the microphysics scheme. How microphysical properties of ice crystals such as size, geometry, and density can be linked to radiative properties is not yet fully understood, and is an active research topic (Ham et al., 2017; Konoshonkin et al., 2017; Mitchell, 2002; Ryan, 2000). For this reason, the mean particle diameter calculated by the microphysics bulk scheme is for simplicity directly used as the radiative effective diameter. This simplification should be kept in mind when interpreting the results presented in this study. Testing more complex models for the effective diameter of ice particles in the LES is for now considered a future research topic.

Finally, two further simplifications in the treatment of radiation should be mentioned. First, in grid boxes with a mean ice crystal size outside of the range defined above, the effective diameter is set to the upper (respectively lower) bound. Second, the solar zenith angle is set constant for the duration of each simulation, and is calculated based on the location and time of the radiosonde used for initialization of the model. The zenith angle affects the solar radiation, but three-dimensional effects are not considered, and radiative transfer works purely in the vertical.

2.2.4. Surface Parameterization

The parameterization of surface processes, particularly flux calculations, becomes necessary when using *DALES* due to the unresolved surface-roughness scale. Assuming that the first model level lies in the atmospheric surface layer, computing the exchange between atmosphere and surface becomes possible based on the Monin-Obukhov theory and the surface layer bulk Richardson number Ri_B (Louis, 1979):

$$Ri_B = \frac{z_1}{L} \frac{\left[\ln \frac{z_1}{z_{0,h}} - \Psi_H\left(\frac{z_1}{L}\right) + \Psi_H\left(\frac{z_{0,h}}{L}\right)\right]}{\left[\ln \frac{z_1}{z_{0,m}} - \Psi_M\left(\frac{z_1}{L}\right) + \Psi_M\left(\frac{z_{0,m}}{L}\right)\right]^2}. \quad (1)$$

here, z_1 is the height of the first model level, L the Obukhov length, $z_{0,h}$ and $z_{0,m}$ the roughness lengths for heat and momentum, respectively, and Ψ_H and Ψ_M the integrated stability functions. For a more detailed description refer to Heus et al. (2010).

The stability functions for the unstable and neutral conditions are unchanged from the initially released code base, while for the stable conditions, the stability functions by Grachev et al. (2007) are newly implemented. They were specifically developed for the Arctic regime based on data from the SHEBA campaign suiting this application. The use of roughness lengths derived from *MOSAiC* measurements is discussed in Section 2.3.4.

Originally, the computation of sensible and latent heat fluxes at the surface relied on a single uniform surface skin temperature. However, a slightly different approach was introduced, which considers two distinct skin temperatures: the ice skin temperature measured and the ocean skin temperature set to $T_{skin,ocean} = -1.8°$ C. By computing fluxes for both temperatures and then taking a weighted average based on the sea-ice fraction, this approach allows for partial inclusion of open ice effects into the simulation. If the ice temperature exceeded −1.8° C, ocean and sea-ice temperature were assumed to be identical.

The ice skin temperature is prescribed and constant throughout the simulation. This one-way surface-atmosphere coupling reduces the complexity of the simulation immensely, and the error it introduces is expected to be minimal, primarily because of relatively short simulation times of 3 h.

2.3. Experimental Configuration

2.3.1. Domain and Grid

The choice of domain and grid size is crucial for the quality of a simulation. The adopted spatial discretization reflects the limits computational resources impose on simulating a full year at turbulence-resolving resolutions.

Table 2
Resolution and Domain Size for Different Setups

Setup	Resolution (h × v)	Domain (h × v)	Usage
Production	20 m × (10 – 185 m)	6.4 km × 12 km	Definitive runs
Test	25 m × (10 – 185 m)	0.8 km × 12 km	Impact studies

Ensuring that turbulent structures are well represented is imperative, which requires a sufficient grid resolution. To avoid the unwanted effects of periodic boundaries in the horizontal directions, the domain must be appropriately sized to accommodate some dynamics that act on scales larger than the mixed layer depth, such as secondary horizontal circulations. In the standardized setup, the horizontal domain and grid utilized are:

$$L_x = L_y = 6400\,\text{m}$$
$$\Delta x = \Delta y = 20\,\text{m}$$

In principle a Lagrangian frame of reference is adopted, with the simulated domain moving with the low level trajectory toward the *Polarstern* location. The domain is initialized with the Polarstern radiosonde profile but only simulated for 3 hours. In this approach, which could be described as "short-range Lagrangian with end-point initialization," the model results are in principle still comparable to stationary measurements at the *MOSAiC* site, with the drift in the mean state being minimized by continuous nudging (described below in Section 2.3.5).

The top of the vertical domain is at $L_z \approx 12\,\text{km}$. The lowest 1200 m of the field is resolved with a grid spacing of $\Delta z_{\text{min}} = 10\,\text{m}$. Above this level, the grid spacing increases exponentially with height until a maximum grid spacing of $\Delta z_{\text{max}} = 185\,\text{m}$. The non-regular vertical grid allows for high resolution in the cloud-containing boundary layer while also accommodating tropospheric features such as higher-level clouds that primarily exert radiative effects on the lower domain. In total, the grid consists of 286 vertical levels. For a more detailed description refer to Appendix A.

In the process of finding a compromise between domain size and resolution with fixed available computational resources, it was decided to prioritize resolution: The 6.4 km × 6.4 km domain is enough to allow for thermodynamic heterogeneities in the virtual area but might lack the evolution of larger mesoscale structures. Since this work is focused on representing realistic conditions and smaller-scale processes in the Arctic atmospheric boundary layer, the resolution that has been selected should be sufficient. The small grid spacing, especially in the lowest layers, improve the representation of turbulence and all turbulence-driven processes, for example, cloud evolution and life cycle, entrainment, shear-layer mixing, detailed distribution of microphysical interactions, surface fluxes, etc. This setup is referred to as Production setup in the following.

A second configuration (Test) was employed: By reducing the horizontal domain to 800 m × 800 m and the horizontal resolution to 25 m × 25 m the computational expense is reduced drastically. The vertical grid is kept unchanged. While this resolution and domain size decrease the simulation quality, the results are still greatly informative. The TEST configuration proved to be crucial in curating and testing the standardized setup and the integration of measurements in an efficient manner. This procedure of using micro-grids to inform the setup of expensive runs was inspired by Neggers et al. (2019). Table 2 summarizes the two setups.

2.3.2. Initial Profiles

The initial conditions of each simulation are mostly derived from data gathered during the *MOSAiC* campaign to constrain the experiments as much as possible with measurements. All initial profiles, most of which concern prognostic variables in the model, are horizontally homogeneous in the simulation domain.

Initial profiles of the two prognostic thermodynamical state variables in *DALES*, total water specific humidity q_t and liquid water potential temperature θ_l, are derived from data from the 11:00 UTC radiosonde launched at the *Polarstern* (Maturilli et al., 2022).

Corrections were applied to account for elevated launch height and initial sensor adjustment after launch by Dahlke et al. (2023). For this purpose, the radiosonde data was combined with measurements from the *MetCity* meteorological tower (C. Cox et al., 2023; C. J. Cox et al., 2023). The prognostic model variables q_t and θ_l in the *DALES* code are derived from the radiosonde temperature, humidity, and pressure measurements. Further, for ice layers, the derived q_v profile was bounded by the saturation specific humidity $q_{\text{sat,i}}$, to limit the impact of measurement uncertainties on the sensitive initialization of ice clouds. The full procedure is detailed in Appendix B.

The two prognostic variables of suspended cloud mass, cloud liquid water content (LWC) and ice water content (IWC), are also initialized with measurement-based profiles. In the *DALES* code, these variables are denoted as q_l and q_i, respectively. Initializing cloud variables with observations reflects our objective of simulating and resolving small-scale turbulent variability around the *MOSAiC* instrumentation, which is often driven by cloud processes. One could initialize with a cloud-free state and let mixed-phase clouds develop by themselves over time. However, previous LES intercomparison studies on mixed-phase clouds have shown that this is difficult to achieve in a short simulation time (Neggers et al., 2019; R. G. Stevens et al., 2018), and the simulations tend to drift significantly away from the initialized atmospheric thermodynamic state as the clouds form. These points motivated adopting a direct initialization of cloud mass. To achieve this, accurate placement and phasing of the initial cloud mass relative to the vertical atmospheric thermodynamic profile is imperative. Because cloud mass is not directly measured by the radiosonde, the *ShupeTurner* data product (Shupe et al., 2015) is used in this study. The multi-sensor cloud retrieval product features vertically resolved LWC and IWC at 1 min resolution. To ensure that extreme values are avoided and the overall thermodynamic state of the boundary layer is captured, the initial conditions used are mean LWC and IWC profiles. This approach focuses on representing the average conditions across the entire domain rather than the potential heterogeneities in certain areas. In each case, the 1 min profiles from the data set are averaged for 15 min after the radiosonde launch time, around 11:00 UTC, depending on the day. This time window is sufficient to average across individual cloud-turbulence elements yet brief enough to prevent alterations in the overall large-scale atmospheric conditions.

The prognostic humidity budget equation in *DALES* is formulated in terms of $q_t = q_v + q_l$ as the total humidity and excludes suspended cloud ice which is treated as a separate prognostic variable. To achieve an internally consistent initialization of all humidity state variables, q_v, q_l, and q_i are combined in a height-dependent way, as follows. In areas, where turbulent mixing of vertical layers increase the overall ice production (below z_{cut}, see Equation B7), the ice and liquid water specific mixing ratio are added and given to the model as only liquid. The expected behavior of the model is to convert the additional liquid into ice via the microphysics scheme and represent realistic IWC and LWC distributions after a spin-up phase. In areas of expected deposition ice production (above z_{cut}), the retrieved IWC is directly included in the model. If ice is measured, the model will place ice clouds directly with an estimated effective radius of 55 µm. This shortens the lengthy deposition spin-up process and produces realistic, weakly turbulent, high-level ice clouds.

The initial profile of the cloud condensation nuclei (CCN) concentration in the *DALES* mixed-phase microphysics scheme (see Section 2.2.2) are estimated based on the observations from the Cloud Condensation Nuclei Particle Counter located aboard *Polarstern* (Koontz et al., 2020). The measurements show high seasonal and day-to-day variability, as well as variability during some days. However, with an emphasis on the consistency of the setup, the mean value of each day is used as the initial value of CCN concentration for the simulation. The value from the preceding recorded day was used for days missing data. In the absence of a consistent CCN profiling of the Arctic air, the near-surface measurements are used as vertically constant initial profiles for the full atmospheric column. This is considered an adequate proxy for the conditions in the boundary layer.

The number concentration of activated ice nucleating particles (INP) is not yet treated prognostically in the current version of the *DALES* microphysics scheme. In the code, the INP concentrations are determined from temperature-dependent activation spectra (Seifert & Beheng, 2006a), prescribed by exponential functions with constant parameters proposed by Reisner et al. (1998). Laboratory measurements of INP activation from in situ samples of Arctic air extracted near *Polarstern* during *MOSAiC* (Ansmann et al., 2023; Creamean, 2022) showed a high time-variation in activation spectra, motivating initialization with daily observed values. The associated observed INP number concentrations were significantly lower than the values proposed by the parameterization mentioned above, often by more than two orders of magnitude. This is not an effect of the ambient temperature on site, since the samples were analyzed at a broader range of temperatures in the laboratory (Creamean et al., 2018). To keep the simulations as representative as possible of the observed atmosphere, the measured activation spectra were used instead. A logarithmic regression was applied to each measured spectra from the laboratory samples, and the resulting slope and intercept parameters were then inserted as new values in the deposition-nucleation scheme instead of the original constant values. Finally, limitations in the sampling of the spectra during the drift necessitate adopting two further simplifications in the initialization of the INP profile in the model. First, due to the 3-day sampling rate, the spectra on sampling-free days are set at the values from the last preceding day of availability. Second, due to the absence of consistent and continuous INP profiling during the drift, the activation spectra are simply assumed constant with height.

2.3.3. Lower Boundary Conditions

Similar to the initial profiles, the lower boundary conditions for model variables are mainly based on *MOSAiC* data. Additional information is taken from the ERA5 reanalysis. The boundary conditions are prescribed, constant in time, and horizontally homogeneous in the simulated domain.

The sea-ice fraction at the location of the *Polarstern* is taken from ERA5 reanalysis data. With the research vessel frozen in solid pack ice for most of the simulated days, the sea-ice fraction at the ship was often close to 100%. Accordingly, short-time reductions in the sea-ice fraction, for example, due to lead events, are not considered in the control experiments for simplicity but could simply be added in future sensitivity experiments.

As described in Section 2.2.4, two surface skin temperatures are considered in the *DALES* bulk surface parameterization; one for sea ice and one for open water. For the ice surface skin temperature T_{seaice} the brightness temperature measurements at the *MetCity* installation are used (C. Cox et al., 2023; C. J. Cox et al., 2023). For the skin temperature of open water, the ocean temperature is used, which is assumed to follow:

$$T_{\text{ocean}} = \max(T_{\text{seaice}}, -1.8°\text{C}), \tag{2}$$

where the numeric value represents the typical freezing temperature of ocean water. T_{seaice} can reach up to 0° C during melting season. Taking the maximum of sea-ice temperature and −1.8° C ensures that the sea-ice temperature is never higher than the temperature of the ocean. In practice, and for consistency with the profile initialization described above, for T_{seaice} the surface value of the combined data product by Dahlke et al. (2023) is used. This value is the skin temperature measured at *MetCity*. Preliminary experiments demonstrated that the skin temperature is a crucial boundary condition to get right, as it directly determines the upward longwave radiation and strongly affects the low-level stability in the simulations. For this reason, no simulations were carried out for days on which *MetCity* data were not available. Using ERA5 values as a replacement was no option for these days because (a) this would introduce inconsistencies in the boundary conditions during the drift, and (b) warm biases are known to exist in the reanalysis regarding the surface skin temperature (Day et al., 2020). As a guiding principle, integrating different data sets in the simulation for the same quantity goes against the approach of a standardized setup.

2.3.4. Surface Properties

The surface roughness length for momentum $z_{0,\text{m}}$ plays a key role in the Monin-Obukhov theory for calculating turbulent fluxes. Accordingly, adopting realistic values is important to get the flux boundary condition right. However, in many model applications, including LES research, the surface roughness length is simply estimated, for example, ECMWF (2021); Michaelis et al. (2020). Fortunately, $z_{0,\text{m}}$ can be derived from daily *MOSAiC* data of the near-surface wind, turbulence, and temperature measurements (C. Cox et al., 2023; C. J. Cox et al., 2023). This data provides an accurate boundary condition for $z_{0,\text{m}}$ for each simulated day. In the literature the surface roughness length for heat $z_{0,\text{h}}$ in the Arctic is often assumed as $z_{0,\text{h}} = 0.1 \cdot z_{0,\text{m}}$ (e.g., IFS Model ECMWF (2021); Michaelis et al. (2020)). For lack of direct measurements of $z_{0,\text{h}}$ during the drift, this value is also adopted here for the control experiments.

The final boundary condition to be considered concerns the surface albedo α. During polar day, the reflectivity of the predominantly frozen surface directly controls the solar contribution to the surface energy budget (SEB). The horizontal scale that α represents is set by the horizontal dimension of the simulated domain. Landscape features such as melt ponds and leads are known to affect the average reflectivity of an area on a meter to kilometer scale (Niehaus et al., 2023), while Arctic weather also influences surface properties on much larger scales. To capture both effects at once, the satellite α product described by Spreen et al. (2008); Istomina et al. (2020) is adopted here. Two main issues affect the accuracy of the satellite estimates of surface albedo; cloud cover and limited satellite overpasses. As a result, satellite albedo products often have significant spatial and temporal gaps. To address this, a rolling average over a week-long period is used to smooth out as many inconsistencies as possible in the data. Second, spatial interpolation in the resulting maps gives a reasonable estimate of the albedo in areas around *Polarstern*. It is possible, that the interpolation in time and space introduces an albedo bias toward cloud-free conditions. A further improvement was not considered for two reasons. Albedo measurements for cloudy

conditions would be derived from a second data set, which clashes with the standardized setup. Second, the uncertainty by this bias is expected to be small when used in the simplified radiation scheme employed in *DALES*.

2.3.5. Large-Scale Forcing and Nudging

The simulated portion of the atmosphere is not isolated from the larger-scale flow in which it is embedded. Various larger-scale processes impact the development of the atmospheric boundary layer, including large-scale advection, vertical motion, and pressure gradients. In *DALES*, these forcings are considered and applied in a horizontally uniform way, while maintaining height-dependency. The associated tendencies are either fully prescribed or partially interactive with the domain-average profile. This forcing method is described in detail by van Laar et al. (2019) and is, in principle, adopted here, with a few notable exceptions as described below.

In the prognostic budget equations for $\varphi \in \{q_t, \theta_l, u, v\}$ the tendency due to large-scale subsidence S_φ^{subs} is constructed using a prescribed large scale vertical motion w_s and the local vertical gradient, calculated from the vertical profile of the domain averaged variable $\overline{\varphi}$,

$$S_\varphi^{\text{subs}} = -w_s \frac{\partial \overline{\varphi}}{\partial z} \tag{3}$$

(Siebesma et al., 2003). This means the significant impact of subsidence on inversions is captured. In the momentum equations the pressure gradient (p) and Coriolis (f) forces are represented in combination, through the departure of the actual wind from the geostrophic wind,

$$S_u^{\text{p+f}} = -f(\overline{v} - \overline{v}_g), \tag{4}$$

$$S_v^{\text{p+f}} = f(\overline{u} - \overline{u}_g), \tag{5}$$

with f the latitude-dependent Coriolis parameter and subscript g indicating the geostrophic state. The latter is calculated from the horizontal pressure fields from ERA5.

In the standardized case generation, the time and location of the *Polarstern* at 11:00 UTC are determined first. Following Neggers et al. (2019), then the six-hour upstream (backward) trajectory of the 950 hPa air mass is estimated from the ERA5 reanalysis data. The quantities required for calculating the forcing terms are then calculated at all trajectory points across an area of 1 × 1 degrees, with horizontal advection corrected for the air mass movement. This means all horizontal advective terms are zero per definition at the height of diagnosis (950 hPa). Finally, the forcing terms are then time-averaged over 6 hours preceding the arrival of the low-level air mass at the *Polarstern*. This yields time-constant composite forcings, which express how the air mass was modulated by larger-scale processes during the period in which any clouds observed at the *Polarstern* formed. To limit the influence of large-scale processes on the evolution of the simulated domain, horizontal advection of temperature and moisture is set to zero. This approach isolates boundary layer processes as drivers of atmospheric changes.

Continuous Newtonian nudging is applied above the boundary layer thermal inversion to prevent excessive drift of the mean thermodynamic and kinematic state in the free troposphere and above (Derbyshire et al., 2004; Neggers et al., 2012; Randall & Cripe, 1999; Sobel & Bretherton, 2000). The nudging tendency S_φ^{n} is formulated in terms of the spatially averaged profiles (Heus et al., 2010; Neggers et al., 2012),

$$S_\varphi^{\text{n}} = -\frac{1}{t^{\text{n}}}(\overline{\varphi} - \varphi^{\text{n}}). \tag{6}$$

Here, $\overline{\varphi}$ is the horizontal mean of an arbitrary scalar, t^{n} is the nudging time scale, and φ^{n} is the profile toward which the model profile is relaxed, which in this case is the initial profile based on the corrected 11:00 UTC radiosonde data. It is only applied above the boundary layer height z_i, which is adaptively calculated based on the maximum liquid water potential temperature gradient. This method is designed to prevent slow drift in the mean state but to simultaneously preserve turbulent perturbations, and is fundamentally different from the sponge layer dampening described earlier (Neggers et al., 2017; Sandu & Stevens, 2011; M. Zhang et al., 2013). The nudging

time scale decreases linearly from infinity at z_i to $t^n = 10800\,\mathrm{s}$ at $z_i + 300\,\mathrm{m}$. Height z_i as used to define the bottom of the nudging layer is defined to be situated in a specified range, as follows:

$$z_i,\ \text{where}\ \left.\frac{\partial \Theta_v}{\partial z}\right|_{z=z_i} = \max\left(\left.\frac{\partial \Theta_v}{\partial z}\right|_z\right),\ z \in [100\,\mathrm{m},\ 5000\,\mathrm{m}] \tag{7}$$

This specification is a practical protection against artificially high gradients near the surface and the tropopause inversion.

The final external forcing to be discussed is the radiative forcing at the top of the simulated domain. The downward radiative fluxes at the model ceiling are calculated by including the part of the full radiosonde (initial) profile in the radiative flux calculations that is situated above the ceiling. This way, impacts of the upper atmosphere on both longwave and shortwave radiative transfer are represented as a soft and interactive boundary condition.

3. Results

With the initialization, boundary conditions, and large-scale forcing thus defined, an experimental configuration is obtained with the following key characteristics;

- The simulated domain represents a statistical, homogeneously forced downscaling of the mean atmospheric column as observed daily at 11:00 UTC at the *Polarstern* during *MOSAiC*;
- The recent history of the low-level air mass is represented through Lagrangian forcing;
- Time-composite forcing means the simulations can quasi-equilibrate, depending on the proximity to a balance in the prognostic budget equations;
- Resolved, small-scale boundary layer processes are free to evolve below the thermal inversion;
- Mixed phase clouds are allowed to spin up and interact with radiation, resolved dynamics, and prognostic aerosol.

This model configuration was arrived at after extensive testing on both smaller and larger grids. A three-dimensional volume rendering of a mixed-phase cloud that results from this setup is shown in Figure 1 to illustrate that the turbulent dynamics in which these clouds are embedded are resolved to a high degree with this setup.

The presentation of the results with this model setup is subdivided into three parts. Section 3.1 provides an overview of the successfully simulated cases during the year-long *MOSAiC* drift. Section 3.2 presents the statistical evaluation of the model output against a year of *MOSAiC* data, and includes a brief initial discussion on data comparability. Section 3.3 focuses in more detail on five selected days, to gain insight into the typical behavior of single simulations and to explore the potential use of the generated library of simulations for further scientific research.

3.1. Simulation Overview

Figure 2 gives an overview of the days during the *MOSAiC* drift that were successfully simulated in the PRODUCTION setup featuring the highest resolution and the largest domain, as described in Table 2. Here, success implies two things:

- The data needed for the standardized setup is complete and available at 11:00 UTC;
- No numerical issues occurred, and the simulation was completed successfully.

A subset of non-simulated days can be distinguished that is more or less randomly distributed in time. On these days, typically, an observational data set is missing that is part of the experimental setup. Two longer, continuous periods also exist that are not simulated; one in May-June 2020 and another in July-August. These coincide with the *Polarstern* not being at the ice floe or *MetCity* not being operational. If not stated otherwise, all results discussed below represent the PRODUCTION setup.

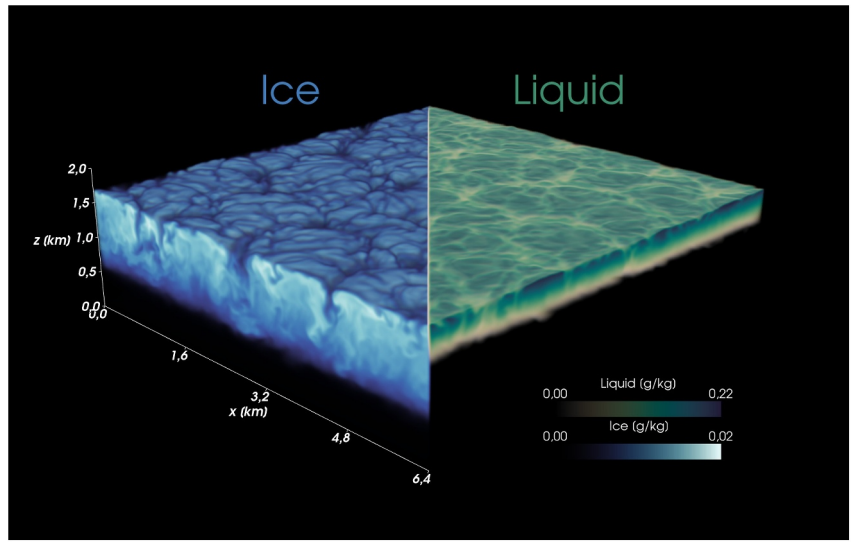

Figure 1. Volume rendering of the *20191101* case (see Table 3) after 1.5 h of the full horizontal domain. The two quantities shown are the specific mixing ratios for ice (blue) and liquid (green) water. For visual reasons, the ice is only shown on the left, and the liquid only on the right.

3.2. Statistical Evaluation

The objective of this section is to evaluate the library of daily LES experiments against the year-long record of *MOSAiC* data. The general approach is to assess the difference between simulated values and their measured counterparts for selected variables to identify both strengths and weaknesses of the library of simulations concerning resolved aspects of the Arctic boundary layer and the SEB. When interpreting the simulations and comparing them to measurements, three important considerations should be made, as briefly discussed here.

The first consideration is the state of equilibrium of the simulated and observed boundary layer. When the sum of all sources and sinks in the bulk ABL budget is non-zero, a net tendency exists that can gradually warm, cool, moisten, or dry the layer. Due to our lack of knowledge of these tendencies in nature, it remains unknown if the observed ABL is close to equilibrium or not. The simulated ABL can similarly be out of equilibrium, for example, through uncertainties in the applied forcings. However, what helps in model-observation comparability is the well-known slow adjustment of Arctic boundary layers, reflecting that net bulk tendencies are small compared to individual budget terms (Neggers et al., 2019). Another convenient factor is the observed persistence of ABL-related phenomena such as clouds (Morrison et al., 2012; Shupe, 2011; Shupe et al., 2006; Stramler et al., 2011), which implies that the model budget should be pretty similar to the observed one when initializing with observed cloud profiles.

For these reasons, it is assumed a priori that the simulated ABL is, in principle, comparable to its observed equivalent. When doing so, one still needs to determine the optimal time point after initialization at which a comparison can be made. On the one hand, turbulent-cloudy processes need to spin up properly; on the other,

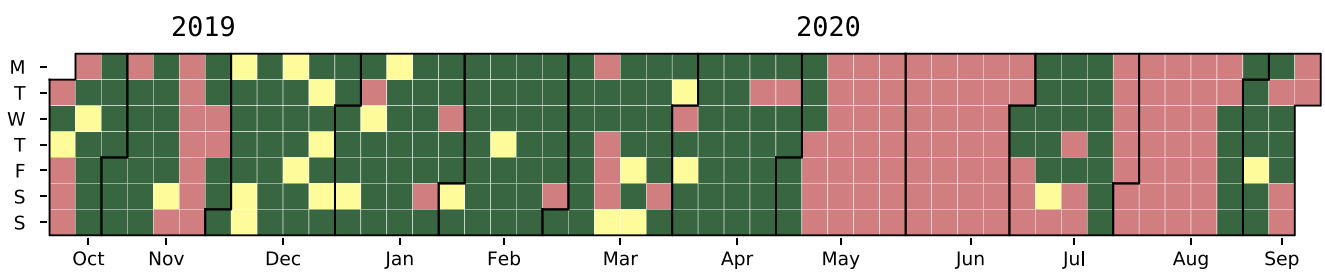

Figure 2. Overview of simulated case studies by date. Green indicates a successful simulation of the state of the boundary around 11:00 UTC at the *Polarstern*. Days on which simulations were unsuccessful because of a lack of observational data are marked in red. Yellow indicates unsuccessful days due to model failure.

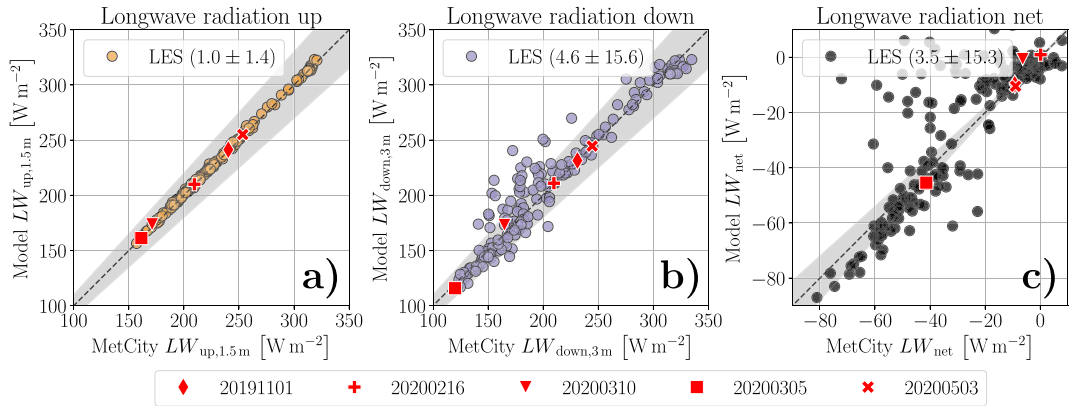

Figure 3. Scatter plots of simulated vs. observed (a) upward, (b) downward, and (c) net longwave radiative fluxes. LES results (ordinate) are plotted against *MetCity* tower measurements (abscissa). Each dot represents one of the 190 simulated cases. The simulation is sampled at $t = 1.5$ hr after initialization, while the measurement represents the 15 min average after radiosonde launch time. The dashed line indicates the one-to-one diagonal. The gray shading indicates the ≤10% difference area, for reference. Values in the legend show $\mu \pm \sigma$, with μ and σ being the mean and standard deviation of simulated minus measured values, respectively. The five cases highlighted are further discussed in Section 3.3.

significant drift from the desired state can happen and needs to be minimized. Careful equilibration analysis of all simulations yielded an optimal time point of 1.5 h after initialization. More details about this analysis can be found in Appendix C.

The second important consideration in the evaluation effort is the *independence of the observational data* from the numerical experiments. What is relevant is that the latter are already closely based on *MOSAiC* data sets. Ideally, one should evaluate resolved processes in the LES against completely independent data sets. This is indeed the case for a large subset of evaluation data sets, such as the surface radiative and turbulent energy fluxes.

However, in principle, cloud properties are far less independent from the simulations, as these are initialized with observed cloud profiles. Nonetheless, it should be noted that clouds in the model can significantly evolve during the simulated period before the sampling time point due to strong interactions between thermodynamics, turbulence, and radiation. Near surface and coupled clouds can also be shaped by surface fluxes. Accordingly, it is far from trivial that clouds remain unchanged from their initial state. This makes comparison to the observed cloud properties meaningful; a close agreement reflects model skill in maintaining the turbulent cloud layer.

The third consideration is about *uncertainties* in the observational data sets and how to deal with them in the model evaluation. A complicating factor is that these uncertainties also flow into the model setup through the initialization and boundary conditions. How these errors percolate into the final simulation, and to what degree, is hard to disentangle and even harder to isolate from other error sources such as numerics and the prescribed forcings. For these reasons, it was decided to exclude error bars in the evaluation plots. A thorough error analysis is for now considered future work.

3.2.1. Radiative Energy Fluxes

3.2.1.1. Longwave Radiation

Figure 3 shows scatter plots of the simulated versus observed near-surface longwave radiative fluxes at the *MetCity* site located on the ice floe near the *Polarstern* (C. Cox et al., 2023; C. J. Cox et al., 2023; Riihimaki, 2021). Included are the (a) upward, (b) downward, and (c) net fluxes at 1 – 3 m height.

The simulated upward flux reproduces the observed values to a high degree, in part because the surface skin temperature used in the simulations is derived from these radiation measurements and kept as a fixed boundary condition. The good agreement confirms that the radiation scheme accurately represents the radiative energy emitted by the surface. A slight warm bias exists, which is possibly introduced by the use of the double skin temperature as described in Section 2.3.3. The radiation scheme also uses this combined skin temperature

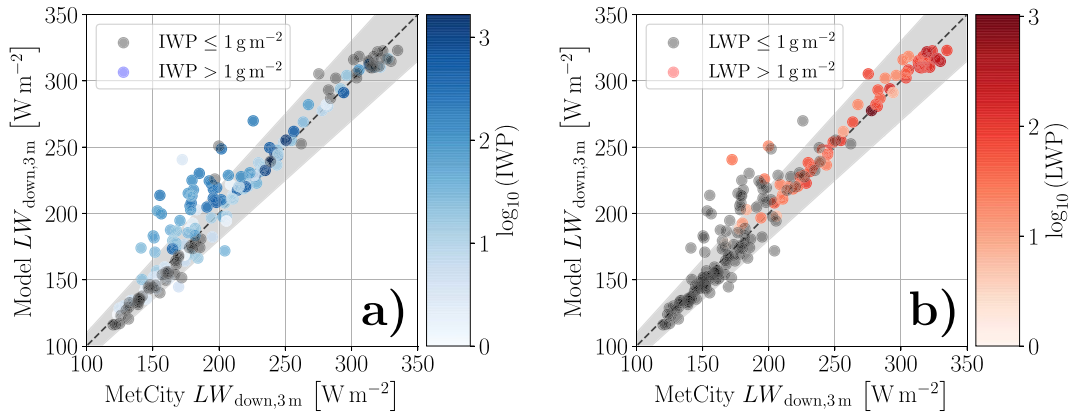

Figure 4. Similar to Figure 3b, but now shown with data points color-coded based on the simulated (a) ice water path (IWP) and (b) liquid water path (LWP). Gray dots indicate water path values below a minimal threshold of 1 g m^{-2}, situations with minimal to no clouds.

(weighted by the sea-ice fraction) as the surface emission temperature for the wider area, while the measured fluxes purely reflect the local temperature of the (colder) sea ice on which they were made.

For the longwave downward flux (that indirectly reflects cloud presence) there is also a good agreement, especially for the upper and lower tails of the distribution (Figure 3b). This agreement indicates that the model does a reasonable job of simulating both the atmospheric temperature and the vertical location and presence of clouds. In the lower intermediate range, some positive outliers exist. To gain insight, Figure 4 shows the same data but now shaded according to the (a) ice water path (IWP) and the (b) liquid water path (LWP). This additional information reveals that good agreement exists for (a) cold and cloud-free conditions ($LW_{down} \leq 120$ W m^{-2}) as well as (b) warm and optically thick liquid clouds ($LW_{down} \geq 200$ W m^{-2}). The intermediate range suffers from a well-defined bias for relatively high IWP values, suggesting too high emissivity by ice clouds in the model. A deeper investigation of high-bias days revealed that the error is introduced by parametric assumptions in the *DALES* radiation scheme concerning the effective diameter of ice crystals. An effective diameter range of 20 – 120 μm is applied, which was originally developed for cirrus clouds. Ice crystals in the Arctic can be observed to exceed that value (Shupe et al., 2006). Capping ice diameters at a too-small value leads to the overestimation of the optical thickness of ice clouds. Additionally, the radiative effective radius of ice hydrometeors in the radiation scheme is set equal to the mean particle radius calculated by the bulk microphysics scheme. This estimation might introduce further uncertainties. These sensitivities are discussed further in Section 4.

Combining the upward and downward fluxes yields the net flux (Figure 3c). The data shows the bimodal distribution typical of the Arctic, which is reproduced to a reasonable degree by the simulations and is explored in more detail in Section 3.2.4. While a general agreement exists with the observations, as expressed by the relatively low bias compared to the mean signal, the error introduced by the biases in the longwave downward radiation for ice clouds now materializes much more pronouncedly. Analysis indicates that the model particularly struggles in situations with optically thin ice clouds, claiming unrealistically low magnitude net radiative fluxes close to 0 W m^{-2}.

3.2.1.2. Shortwave Radiation

Figure 5 compares the simulated and observed near-surface shortwave radiative fluxes. Non-zero values indicate polar spring and summer and reflect model performance during the melt season. When interpreting these results, it is important to realize that *DALES* only considers purely vertical radiative transfer.

This assumption is commonly made in most atmospheric models, including most present-day LES codes (e.g. (B. Stevens et al., 2005; Skamarock et al., 2021), with only a few notable exemptions (e.g., studies using *DALES* (Veerman et al., 2020), *MicroHH* (Tijhuis et al., 2023) or *PALM* (Krč et al., 2021)). However, it is also a significant simplification of reality that excludes three-dimensional radiative effects, which might introduce biases for high solar zenith angles, as are typical for the Arctic during polar day. With this in mind, the model biases in

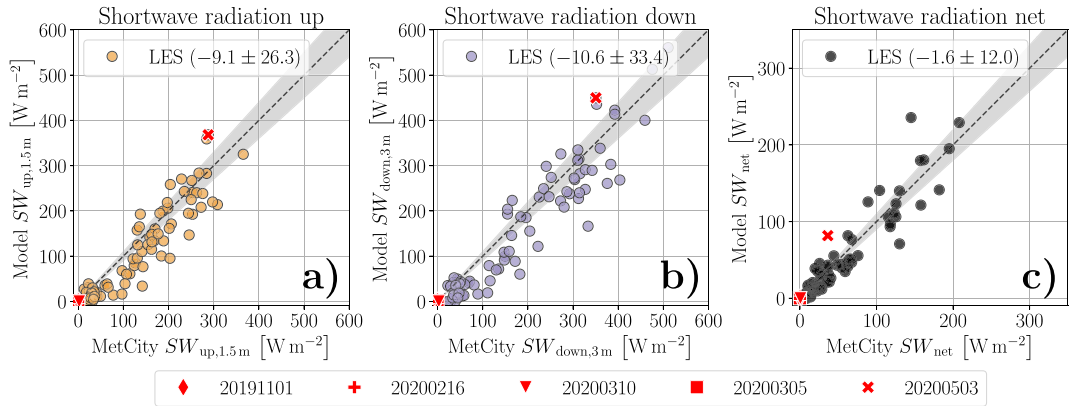

Figure 5. Similar to Figure 3 but now showing the (a) upward, (b) downward, and (c) net shortwave radiative flux at the surface.

shortwave fluxes are largest at small values and decrease significantly as the shortwave fluxes increase. In terms of energy flux units, the bias and spread are somewhat larger compared to the longwave fluxes. For all three fluxes, the bias is negative, suggesting that too little shortwave energy is entering the system. This might be due to the exclusion of three-dimensional effects, a hypothesis that needs testing in future research. One notices that the distribution of data points in the upward and downward flux figures is structurally similar. In combination with the accuracy of the upward flux and in association with the resulting net flux, one concludes that the prescribed albedo boundary condition must be close to the true value. This skill was only arrived at after including the satellite-based albedo measurements (not shown).

3.2.2. Water Paths

Two well-known and critical success metrics for reproducing clouds in the virtual domain are the liquid water path (LWP) and the ice water path (IWP). Figure 6 shows the comparison of simulated and measurement-based values for (a) IWP and (b) LWP. Focusing first on the ice, the general magnitude of measured IWP is well reproduced by the model. Measurements and model data agree for most cases. Nonetheless, several simulations under- or overestimate the amount of ice in the column. When interpreting these outliers it should be taken into account that the simulations are initialized with the observed profiles of liquid and ice cloud water. Accordingly, in these cases, the offsets arise during the simulated period, which points at the model experiencing difficulties

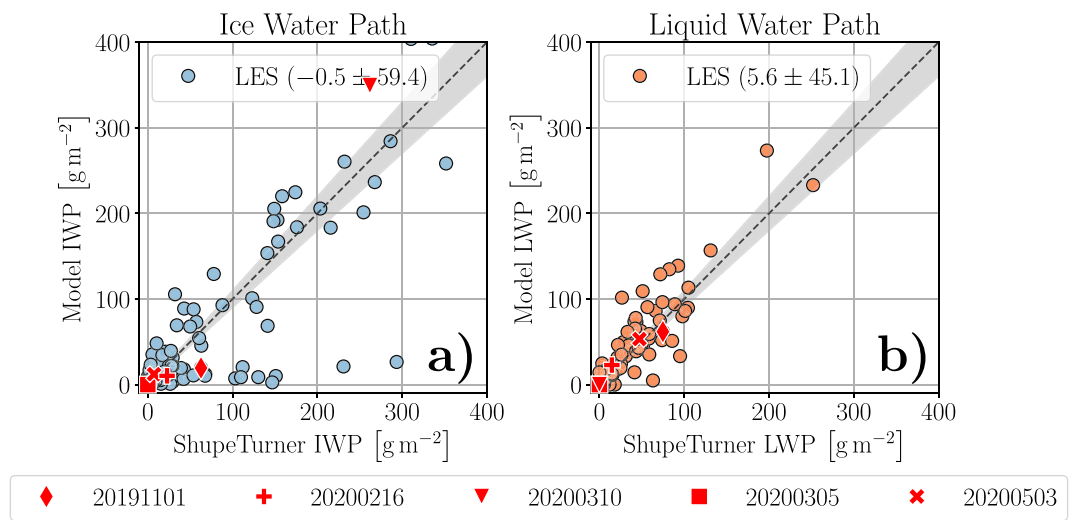

Figure 6. Scatter plot of simulated versus observed (a) Liquid Water Path (LWP) and (b) Ice Water Path (IWP). The LES is plotted against the *ShupeTurner* data set. Sampling, notation, and plotting style are analogous to the previous figures.

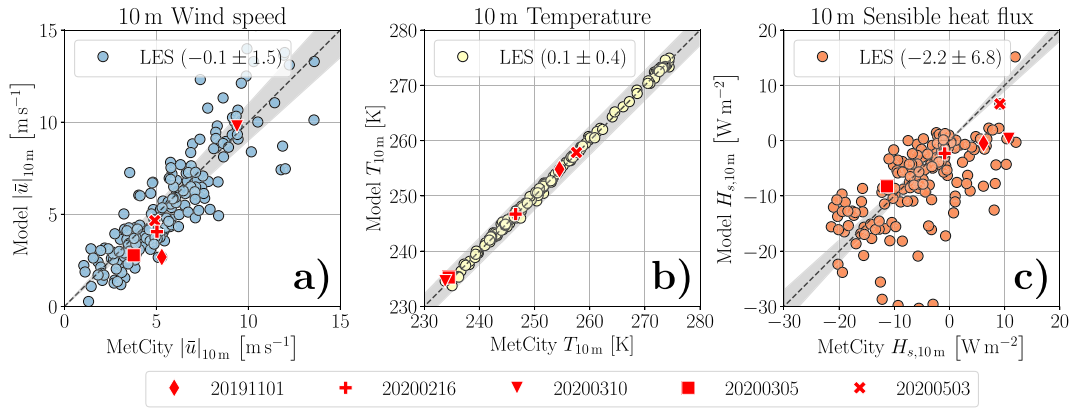

Figure 7. Scatter plots of simulated versus observed meteorological properties at 10 m height, including (a) wind speed, (b) temperature, and (c) bulk sensible heat flux. Sampling, notation, and plotting style are analogous to the previous figures. In (c) a positive sign indicates an upward sensible heat flux.

maintaining the right ice amount. Possible causes for this trend include (a) lack of resolution to fully resolve cloud-forming dynamical processes in the upper atmosphere, (b) uncertainties in the applied large-scale forcings, and (c) lack of skill in the microphysics scheme. At this point, it is not clear which of these causes apply. However, encouraging aspects are that the model always maintains at least some ice and that the outliers seem randomly scattered around the correct mean as calculated over the full drift.

The measured LWP is mostly well captured by the simulations. A few cases underestimate the amount of liquid. This underestimate can happen when freezing rates are overestimated by the model and liquid clouds glaciate too rapidly. The most significant relative errors for LWP and IWP occur in the region $\leq$50 g m^{-2}. Closer inspection of these cases suggested that in situations with really thin ice and liquid clouds, the model is increasingly sensitive to the initial conditions. On the positive side, it can be considered a success that the model can capture and reproduce these thin clouds in the first place, to some degree.

3.2.3. Near-Surface Meteorology

Figure 7 shows the simulated versus observed (a) wind speed, (b) temperature, and (c) sensible heat flux at 10 m height. The wind speed in the simulation compares well to the measurements, which is especially gratifying since the wind profile given to each simulation as initial condition stems from ERA5 reanalysis data. The MOSAiC radiosonde data were assimilated into ERA5, but simulation forcings and near-surface turbulence can cause differences from the initial profile after the spin-up phase. For these reasons, a good agreement is not trivial. Nonetheless, the relative error can be high for low wind speeds, and the impact on near-surface transport processes is significant. The simulation is biased toward lower values for wind speeds $\geq$10 m s^{-1}. Low-level jets are a common phenomenon in the high Arctic (López-García et al., 2022), and their underestimation in the LES might explain this small bias.

Figure 7b shows that the simulated 10 m temperatures closely resemble the observations. This is arguably not surprising, given (a) the prescribed observed skin temperature as measured at *MetCity* and (b) the radiosonde temperature profile being part of the model initialization. The close agreement at least confirms that the simulation has not drifted away from the measured thermodynamic state after the spin-up period.

Figure 7c evaluates the simulated 10 m sensible heat flux (H_s). By convention, a positive sign indicates an upward flux. In general, the observed distribution and orientation of the data points in this space are reproduced, but the mean bias and spread are relatively large compared to the mean signal. The mean bias is negative, expressing a general underestimation of the flux of sensible heat between surface and atmosphere. Significant negative biases occur on individual days, which are responsible for most of the mean bias and spread. On some days, the underestimation reaches more than 20 W m^{-2}. While these may seem large excursions, on average the bias and spread are still comparable to other components of the SEB, such as the net radiative fluxes (see Figures 3c and 5c).

The surface H_s is a key player in the SEB, an important aspect of the Arctic climate system, which is crucial to correctly represent in models. Accordingly, it is imperative to gain an understanding of these large differences between the model and measurements. Before claiming model shortcomings, it is relevant to first consider data comparability. The flux measurements are point samples, and can reflect very local conditions close to the instrument site. In contrast, the LES data represent means over a larger area. It can be speculated that small-scale, local spatial variability in H_s could be a prime cause for the scatter in Figure 7c, a problem well known from previous studies at meteorological sites (Doran et al., 1998; Tang et al., 2019) including the polar regions (Bintanja, 2000).

What the model and observational product do share is their reliance on a bulk method. The observed fluxes were calculated using the method developed by Fairall et al. (1996), making use of turbulent transfer coefficients and roughness length estimated from the Surface Heat Budget of the Arctic Ocean (SHEBA) campaign (Andreas et al., 2003). For a more detailed description refer to C. Cox et al. (2023) and C. J. Cox et al. (2023). As described in Section 2.2.4, the simulations also make use of stability functions based on SHEBA data (Grachev et al., 2007). But the use of roughness lengths based on daily *MOSAiC* measurements in the simulations is a key difference. Adopting the transfer coefficients and roughness lengths improved the simulations from the first attempts in terms of the H_s (not shown). However, potential differences with the SHEBA roughness lengths could explain some of the remaining outlying data points. The impact of ERA5-derived sea-ice concentration and of the updated stability functions are further discussed in Section 4.

More causes can be thought of to explain the differences in the H_s. This includes first and foremost the possible inapplicability of Monin-Obukhov similarity theory over sea ice in stable wintertime conditions (Heisel & Chamecki, 2023; Liu et al., 2024). An interesting alternative data source for comparison with the LES could be the eddy-covariance flux product that is part of the *MOSAiC* data archive (C. Cox et al., 2023; C. J. Cox et al., 2023). Further investigating these research questions is for now considered future work. The library of LES experiments presented in this study can inform this effort by identifying and highlighting situations for which large differences occur.

3.2.4. The Bimodal Arctic

The SEB of the Arctic is dominated by two primary modes. The first reflects cold, clear, and stable conditions, while the second represents warmer, cloudy states with often neutral or weakly unstable near-surface conditions. This bimodal nature of the Arctic climate system is a long-known phenomenon (Sverdrup, 1933) and has been measured extensively during previous drift campaigns like SHEBA (Persson et al., 1999, 2002; Shupe & Intrieri, 2004; Stramler et al., 2011). The bimodal state has been intensely researched in recent years, given its importance in Arctic Amplification and sea ice melt. In particular, transitions between the dominant modes have been a topic of interest. Previous modeling efforts have contributed to our insight, for example, by linking the bimodality to mixed-phase cloud persistence (Morrison et al., 2012) and large-scale dynamics (Neggers et al., 2019; X. Zhang et al., 2022). Also, well-defined bimodality has been used as a metric to test the skill of climate and single-column models (Pithan et al., 2014, 2016; Solomon et al., 2023). Because small-scale dynamics are resolved in LES instead of being parameterized, one expects a priori that the library of *MOSAiC* simulations, as discussed here, should have some skill in reproducing this feature. The availability of 190 realizations under a broad range of atmospheric conditions should provide a sufficient sample size to reproduce the bimodal distribution. These questions are addressed in this section.

Figure 8 shows the two-dimensional probability density function (PDF) of near-surface longwave net radiation (LW_{net}) and the near-surface sensible heat flux (H_s), as (a) sampled from the year-long library of LES experiments and (b) as observed during *MOSAiC*. These two variables were previously discussed individually in Sections 3.2.1 and 3.2.3, respectively, but are combined here. The bimodality is visible in the measurements, with the modes best defined along the LW_{net} axis. A variety of states is observed between the cloudy mode (near the origin) and the clear mode (bottom left). For convenience, the PDFs are also shown separately as side panels, to better determine and compare the mode locations.

For comparing the model to the measurements the two modes are now considered individually. The cloud-free mode ($LW_{net} \leq -30$ W m^{-2}) is nearly always accompanied by a negative H_s, which can be as low as -30 W m^{-2} for extremely cold surface conditions. The model correctly reproduces this mode, although the diagonally-shaped maximum is shifted somewhat towards more negative LW_{net} values. The particularly

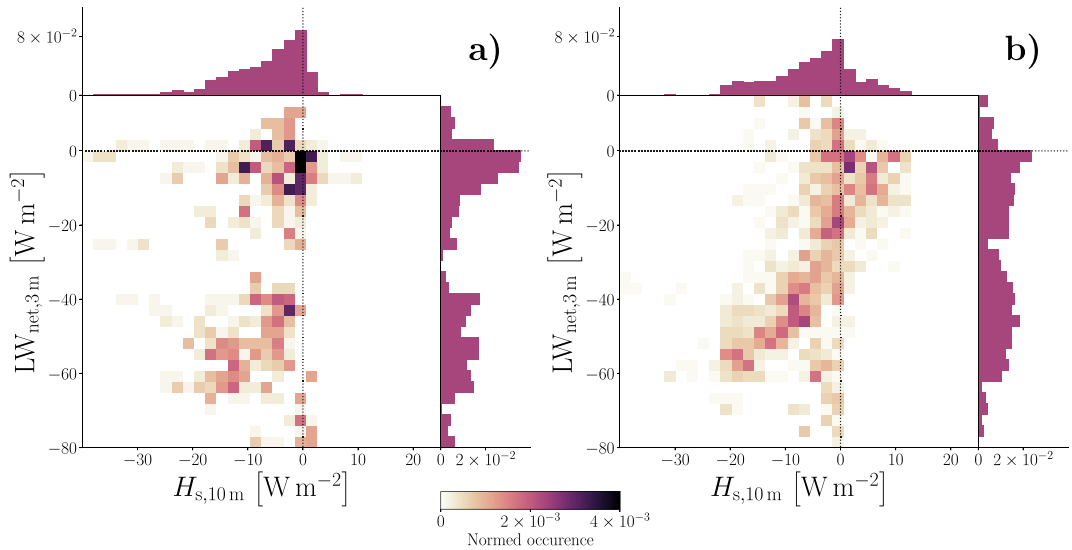

Figure 8. Two-dimensional histograms of normalized occurrence of atmospheric states characterized by the near-surface net longwave radiation (ordinate) and sensible heat flux at 10 m height (abscissa). Panel (a) shows the LES results in which each data point of the given quantities is averaged over 15 min and the whole domain from every simulated case. Panel (b) shows the measured values from *MetCity* (C. Cox et al. (2023); C. J. Cox et al. (2023); Fairall et al. (1996), "*bulk_Hs*_10m") between 10:45 UTC and 11:15 UTC daily.

interesting cluster of cases with extremely low $\mathrm{LW_{net}}$ but near-zero H_s is also well captured by the model. The cloudy state ($\mathrm{LW_{net}} > -30\ \mathrm{W\ m^{-2}}$) in the LES is situated near the observed one, with a maximum occurrence of $\mathrm{LW_{net}}$ around 0 $\mathrm{W\ m^{-2}}$. However, the LES fails to capture the observed weakly positive H_s occurrences and in general, is biased towards lower heat flux values (as discussed previously in Section 3.2.3). Part of a possible explanation for this underestimation could be the absence of leads in the simulations, which are known to locally boost the sensible and latent heat flux (Li et al., 2020). Additional sensitivity experiments at this location of the phase space could be informative, especially because C. Cox et al. (2023) and C. J. Cox et al. (2023) note that they are likely to also miss lead effects in their bulk heat flux calculations.

A closer comparison of the simulated and observed $\mathrm{LW_{net}}$ side panels suggests that the LES slightly underestimates the number of intermediate states between the two dominant modes. This is also visible in the two-dimensional phase space. What causes the absence of these intermediate states in the LES is not yet fully understood, and is currently under investigation. Part of the issue could be related to the overestimation of longwave radiation emitted by ice clouds discussed in Section 3.2.1. In cases with solely optically thin ice clouds, the optical thickness of the atmosphere is overestimated and therefore the results appear closer to the cloudy mode than they were measured. More possible causes include the absence of mesoscale variability in the numerical realizations due to spatially homogeneous forcing and small domain size, or an underestimation of cloud-transitional states due to characteristics of the experimental setup.

3.3. Case Studies

The statistical evaluation against the full year of data in Section 3.2 provides confidence in the basic skill of the LES in reproducing the cloud-radiative climate and SEB of the central Arctic. This section explores a subset of five simulations in more detail. Each simulation serves as a case study of a different Arctic boundary layer, selected to accentuate results of particular interest. Table 3 gives an overview of these cases. The goal is to document resolved boundary-layer processes especially turbulence associated features such as vertical structure, inversions, and mixed-phase clouds. This is motivated by the potential use of these virtually resolved data sets in future research efforts. For a more detailed discussion on vertical structure in the Arctic boundary layer refer to Jozef et al. (2024). Additionally, the five cases are highlighted in Figures 3 and 5–7.

Table 3
Overview of Exemplary Case Studies Shown in Section 3.3

Name	Date	Description
20191101	01 November 2019	Single layer mixed-phase cloud
20200216	16 February 2020	Mixed-phase cloud with thin ice cloud above
20200310	10 March 2020	Extremely cold icy atmosphere
20200305	05 March 2020	Extremely cold and stable conditions
20200503	03 May 2020	Mixed-phase cloud during polar day

3.3.1. Vertical Structure

Figure 9 shows the potential temperature Θ, the ice q_i, liquid q_l and vapor q_v specific humidity profiles for the five cases. Here, q_i is the sum of ice crystal, graupel and snow specific humidities. Analogous to previous sections, the simulations are evaluated after 1.5 hr run time. The observational data sets served as initial conditions for the simulations. The value of comparing the simulation to its initial conditions is to judge how well the numerical experiment can keep the measured atmosphere in the virtual domain after the spin-up phase.

The first case simulates the atmosphere measured around the *Polarstern* on 2020-05-03 showcased in Figures 9a–9c. In contrast to the following simulations this case is during polar day, that is, solar radiation has a significant effect on the atmospheric energy budget (see also Figure 5). The cloudy mixed-layer is decoupled from the surface and extends down to 300 m. It is topped by a temperature and humidity inversion (see panel (a) and (c) respectively) around 750 m height which are maintained throughout the simulation time. The models ability to keep the inversion layers in a dynamic equilibrium is a clear success. The right balance between the entrainment of warmer air above into the cloud layer and cloud-top cooling (leading to a colder cloud, more frozen precipitation) is crucial in simulating the longevity of mixed-phase clouds (Morrison et al., 2012; Solomon et al., 2011). With a decoupled cloud layer these terms must be close to equilibrium to explain the long lifespan of these clouds. There are minimal differences in the vertical distribution of liquid and ice water content between simulation and measurement: The liquid cloud layer extends further down and cloud top height is lower by 30 m. The total amount of liquid water is reproduced, even though its distribution is bottom-heavy. While the model underestimates the total amount of ice, simulation and measurement agree on the order of magnitude. Overall the representation of ice and liquid water is a positive result, especially considering the low amount of mass present in the atmosphere. The next simulation captures the atmospheric conditions on 2019-11-01 at 11:00 UTC. Figures 9d–9f give an overview of the results. The boundary layer contains a well-mixed layer decoupled from the surface, as seen from the potential temperature in panel (d). The temperature inversion at 1750 m caps a mixed-phase cloud seen in panel (e). It contains significant mass of liquid at the height 1200 – 1750 m and a minimal amount of ice at 600 – 1750 m. Here, the mixed-phase cloud is maintained, and the measured liquid and ice water content is reproduced. Only the ice layer around 250 m measured by the *ShupeTurner* data sets is not found in the simulation. While turbulent processes intensify, the temperature and humidity profile remain nearly unchanged except the creation of a minimal humidity inversion layer at cloud top by the simulation. Only a slight elevation and cooling of the temperature inversion is observed, as expected for a typical Arctic mixed-phase cloud (Morrison et al., 2012).

The simulation for 2020-02-16 at 11:00 UTC shares several characteristics with the *20191101* and *20200503* cases. As seen in Figures 9g and 9i, a decoupled mixed-phase cloud is measured at the *Polarstern* and reproduced by the LES. Additionally, in the spin-up time between the measurement and the evaluation of the simulation, the temperature inversion has lifted slightly while the cloud layer has cooled and water vapor has been depleted. In contrast to the second case, the cloud sits closer to the surface and is topped with a significantly stronger humidity inversion layer around 700 m. Further, there is a thin ice cloud ($q_i < 0.02$ g kg^{-1}) around 4.5 km height (not shown) that weakly interacts with the mixed-phase cloud via longwave radiation. Analog to Case *20200503*, a clear success is the ability to maintain this humidity inversion right above the cloud top. In contrast to the measurement, the simulated liquid water profile has a maximum slightly below the cloud top. This is likely due to the simple adiabatic profile shape assumed in the observation-based estimate, compared to the simulation which takes time-dependent microphysical processes and turbulent mixing into account.

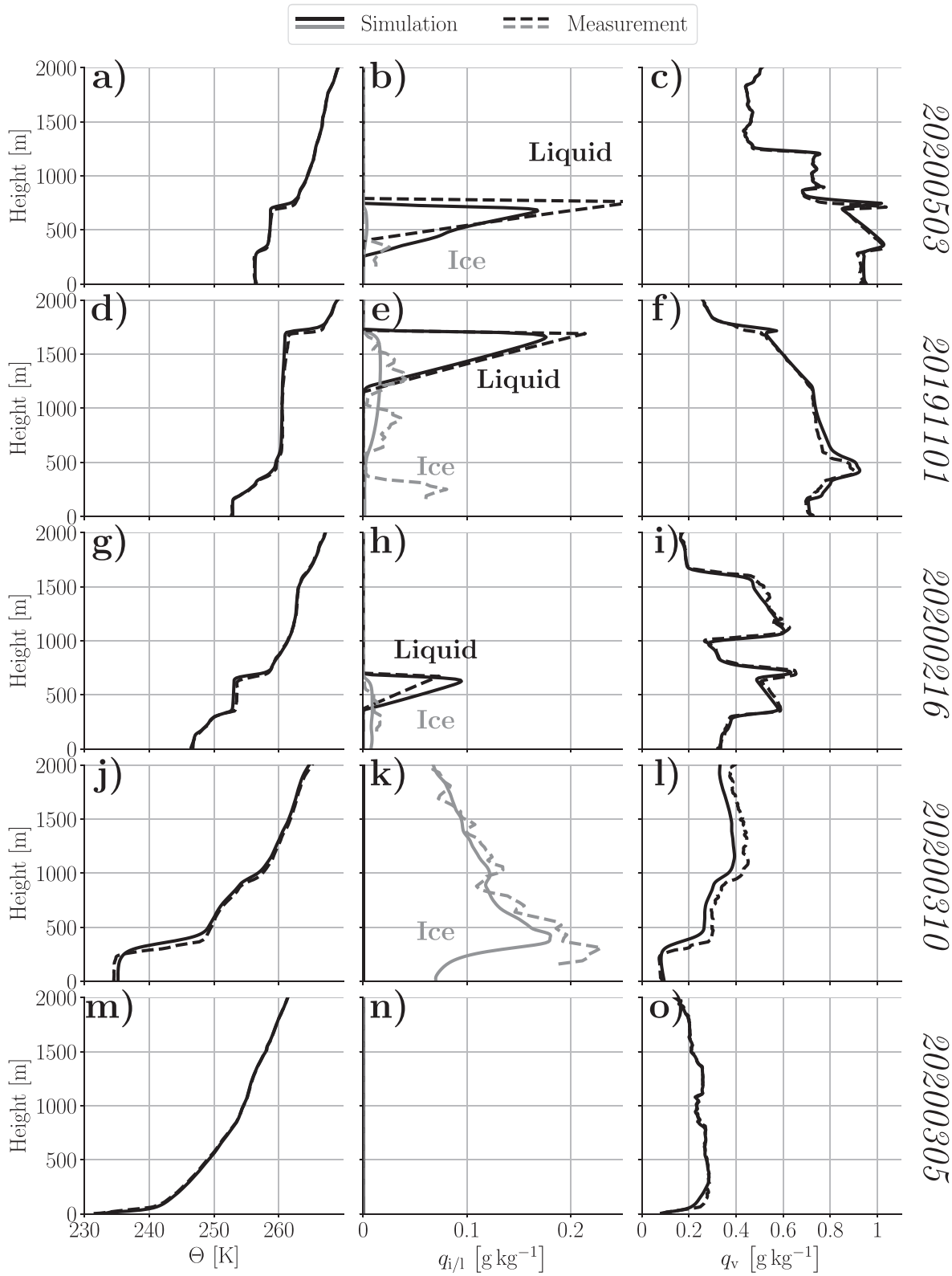

Figure 9.

The next case discussed simulates 2020-03-10 at 11:00 UTC. It is characterized by an extremely low boundary layer temperature, as can be seen in Figure 9j. Further, panel (k) shows a significant amount of ice, especially considering the low amount of total moisture contained in the lower atmosphere (panel (l)). Compared to the other cloudy cases the cloud ice is substantial over the full depth of the layer up to 2 km. This suggests large-scale synoptic drivers of cloud mass, which are prescribed in our experimental design, and not so much local generation by resolved small scale turbulent processes. Despite the assumptions made, this simulation maintains an adequate amount of ice in the atmosphere while also depleting some of the water vapor during the 1.5 h run time. Additional research is needed to evaluate if the model can reproduce realistic freezing rates and precipitation, as these are the main local source and sink processes for ice under extremely cold conditions.

The final case shown simulates 2020-03-05 at 11:00 UTC. In contrast to the previous cases it is characterized by extremely stable conditions, especially close to the surface, as shown in Figure 9m. In addition to that, it contains no cloud water (panel (n)) and little vapor humidity (panel (o)) of 0.2–0.25 g kg^{-1} between 50 m and 2000 m. Values closer to the surface are even smaller. Traditionally, LES struggle to resolve the small-scale dynamics in stable conditions (Couvreux et al., 2020) as the only source of turbulent energy is generated by shear and surface roughness without any significant contribution of buoyant dynamics. With the chosen resolutions for the simulations this is expected to be no different in this work. Nonetheless, the initial thermodynamic profiles are maintained and these extremely stable cases can serve as control simulations to improve LES performance in future research.

3.3.2. Mixed-Phase Cloud

In Figure 10 the *20191101* case is shown in more detail. Panel a–c show the same vertical cross-section for the vertical velocity w, and the liquid and ice water specific humidities q_l and q_i. This is only one of dozens of single-layer mixed-phase clouds simulated in the collection of cases. In comparison to the previous discussion on the averaged vertical profiles, this exemplary case is shown to give inside into the resolved turbulence and clouds for this case. The vertical velocity shows the up- and downdrafts in the cloud. Note, that here the downdrafts are stronger in magnitude than the updrafts, since cloud turbulence is mainly driven by cloud-top cooling. This skewness of the vertical velocity distribution is documented in past measurements (Shupe et al., 2013). As seen before, the typical vertical distribution of ice and liquid in mixed-phase clouds is shown: A relatively shallow liquid cloud is the source region for the formation of ice crystals that fall below the liquid cloud base. In this plot, the heterogeneity in the horizontal directions can be observed. There is a visible positive correlation between the vertical velocity and the liquid and ice water. Updrafts supply plentiful moisture to form and grow both cloud liquid and ice, while downdrafts entrain unsaturated air from above into the cloud. A detailed investigation of the interaction between turbulent, thermodynamic, radiative, and microphysical processes in Arctic mixed-phase clouds will be a future topic of research.

4. Discussion

In the creation of the year-long library of daily LES experiments, some decisions had to be made about the numerical setup, the inclusion of data sets, and the usage of various subgrid-scale parameterizations. All of these decisions affect the eventual results presented and evaluated in the previous section. Gaining insight into these impacts is part of an overarching objective of this study, which is to assess the feasibility of turbulence-resolving LES daily for long continuous periods in the Arctic theater. To our knowledge, this has not been attempted or achieved before, and can still be considered a "terra incognita" (Wyngaard, 2004). In this section, the impact of various aspects is discussed in more detail to convey this experience to the scientific community and inform possible future efforts of a similar nature.

Figure 9. Analysis of the vertical profiles of potential temperature Θ, ice and water specific mixing-ratio $q_{i/l}$, and water vapor specific mixing-ratio q_v for the cases *20200503* (a–c), *20191101* (d–f), *20200216* (g–i), *20200310* (j–l), and *20200305* (m–o). q_i is the sum of ice crystal, graupel and snow specific humidities. Shown are the simulation results averaged over 15 min in time and over the whole domain in space after 1.5 h run time (solid) and observations at radiosonde launch time around 11:00 UTC (dashed). Observed Θ and q_v is derived from radiosonde measurements; $q_{i/l}$ is provided by the *ShupeTurner* data set. *20200216* contains a thin ice cloud ($q_i < 0.02$ g kg^{-1}) in 4.5 km height (not shown).

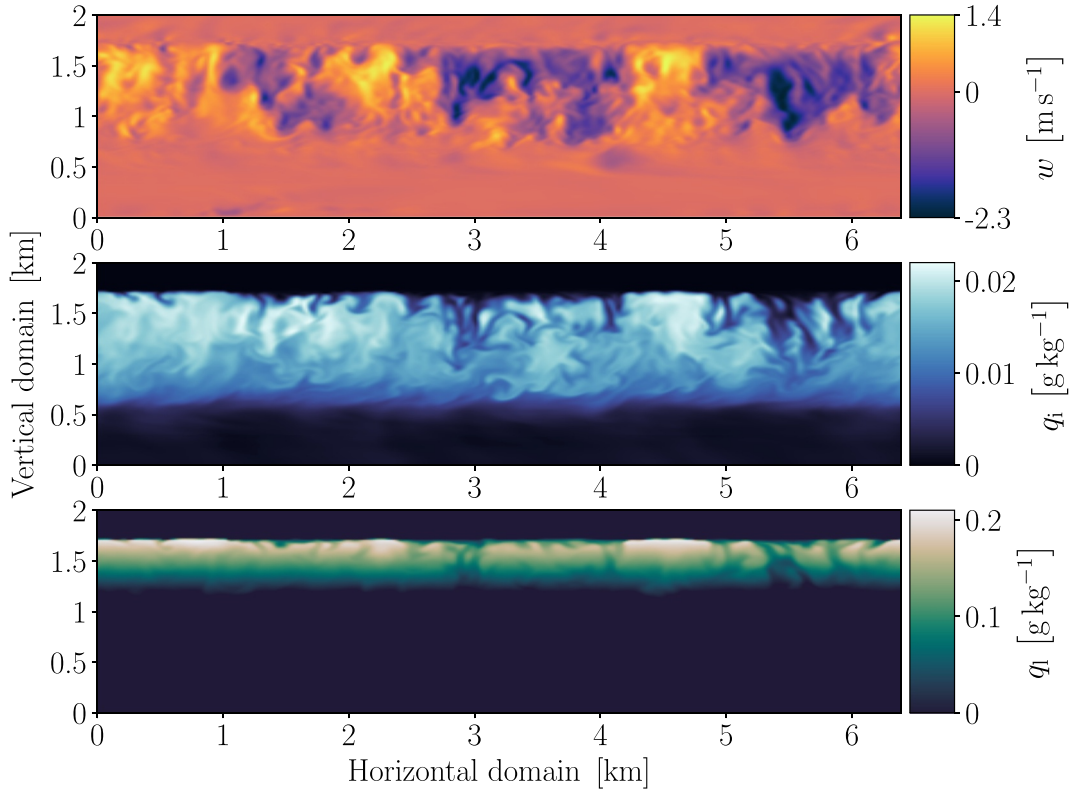

Figure 10. *20191101*: Cross-section plots of vertical velocity w, ice specific humidity q_i and liquid specific humidity q_l. The model output after 1.5 h is shown.

4.1. Grid and Domain

Existing limitations in computational resources put limits on the feasible domain size, grid size, and spatial resolution of LES. This becomes even more pressing when a high number $(\sim 10^2)$ of experiments is to be performed.

The first step was to adopt a vertically telescoping grid, as described in Section 2.3.1. For the PRODUCTION runs the horizontal domain size of 6.4 km × 6.4 km was chosen, considered large enough to contain smaller mesoscale features but small enough to allow resolutions sufficient for resolving Arctic turbulence. A key step that proved crucial for arriving at an optimized setup was to use small 32 × 32 horizontal grid simulations for testing early versions of the model configuration. This was inspired by previous research that adopted a similar strategy for Arctic LES (Neggers et al., 2019). In Figure 11a this TEST setup is compared to the PRODUCTION setup in terms of downward longwave radiative flux with a good agreement between the two. The same result applies to the other radiative fluxes (not shown). It is evident that the small grid TEST setup, in all its limitations, is still capable of reproducing year-average radiative climate to a high degree. This suggests good enough skill to spin up turbulence and maintain liquid clouds, which are reflected in the longwave flux. The conclusion from this encouraging result is that using small-grid test setups to develop, calibrate, and optimize full LES runs in the Arctic can be effective, not just scientifically but also economically, in terms of reducing computational cost and turnover-time of numerical experiments.

4.2. Representation of the Surface

The representation of the surface significantly affects the simulations. To achieve a workable experimental setup, a few simplifications were made. First, the surface is assumed to be homogeneous, a strong simplification in a region with a notoriously heterogeneous topography (Castellani et al., 2014; Mchedlishvili et al., 2023). The sea ice surface is often covered by snow, acting as an insulator and as a source of blowing ice crystals (Wagner et al., 2022) and a variety of aerosols (Held et al., 2011; Park et al., 2019).

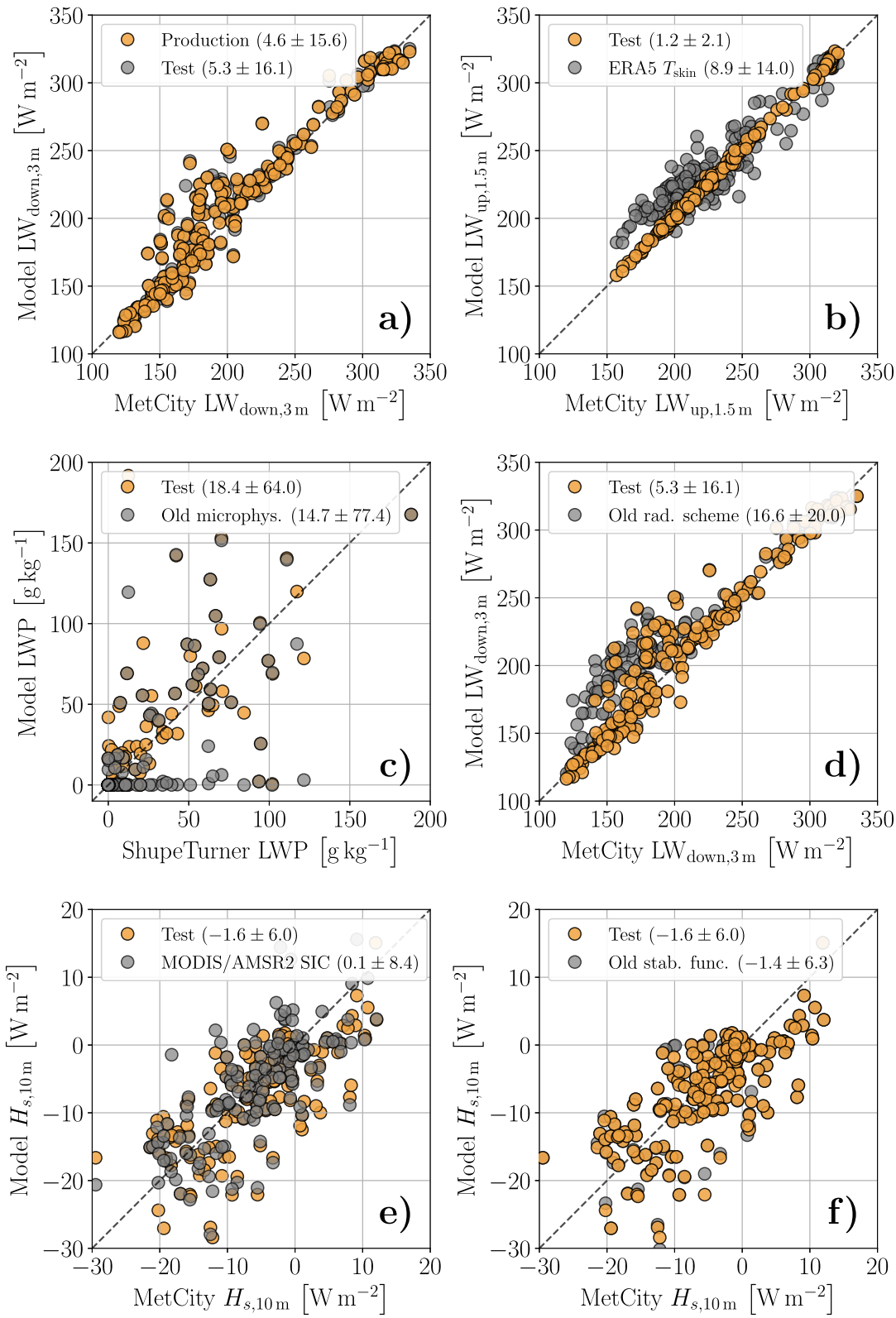

Figure 11.

While acknowledging these impacts, limitations in data availability on these aspects motivated our decision not to represent them in the control setup. Instead, as a guiding principle, observational data was integrated on only three key surface variables; skin temperature, roughness, and albedo. Open-water effects are included by adapting the skin temperature and humidity to the sea-ice fraction. Figure 11b demonstrates the beneficial impact of integrating the locally observed surface skin temperature as measured at the *MetCity* site instead of using ERA5 reanalysis data. The in situ observations are more precise in time and space for the acute situation around the *Polarstern*, and remove most of the spread introduced by the ERA5 values. Further, the reanalysis values introduce a known well-defined warm bias for colder surface skin temperatures (Herrmannsdörfer et al., 2023). This is evident from the positive shift in emitted longwave radiation for values $\leq$250 W m^{-2}.

4.3. Microphysics

In early test simulations, the default *DALES* microphysics scheme struggled to maintain liquid and frozen cloud mass for conditions below −15 ° C. A few microphysical processes were found to cause these issues. As described in Section 2.2.2, the parameterization for heterogeneous freezing is not valid for those temperature ranges and had to be limited to the freezing rate for conditions at −15 ° C. In addition, the maximum number concentration for primary ice crystals produced by deposition-nucleation is artificially limited to 200 L^{-1}. Figure 11c shows that under cold conditions, the microphysics scheme in its default setting leads to a full loss of liquid for a significant subset of days, especially for cases with thin ice clouds measured. This loss of liquid stems from an overestimation of the ice production, quickly depleting any liquid contained in the column. Introducing the mentioned fixes in the microphysics scheme enabled simulations that create and maintain mixed-phase clouds throughout the year, including the Arctic winter.

4.4. Ice-Radiation Interaction

Another key step toward a realistic simulation of ice clouds was to separate the treatment of liquid and frozen cloud mass in the radiation scheme. The default *DALES* radiation scheme treated all cloud mass in the same way, a crude simplification that was abandoned for the final set of runs. Figure 11d compares the near-surface longwave downward radiation between the original and adjusted radiation scheme, using the Test setup.

Adding separate ice-radiation interaction reduces the mean bias toward higher emissions for colder atmospheres with the default scheme, in particular in the presence of ice clouds, discussed in Section 3.2.1. The adjustment does not completely remove the bias. Note, that the cause of this bias seems to stem from exceeding the parameterization limits, but the cause for the large radiative effective radius of the ice crystals might be twofold. First, the ice crystals might be larger than the limits of the radiations scheme (Ryan, 2000), but second, the effective radiative diameter of non-spherical ice particles depends on the distribution of particle shapes (Eichler et al., 2009; Fu & Liou, 1993; Mitchell et al., 2011; Yang & Fu, 2009).

In Section 2.2.3, it is assumed that the mean particle radius calculated by the bulk microphysics scheme can be used with proportional ratio 1 as the radiative effective radius in the radiation scheme. The resulting bias in the longwave radiation statistics indicates that this assumption is not fully correct. Improving the approximations of ice-cloud radiation emission and their dependence on the distribution of crystal size and shapes is a goal for future research (Cairo et al., 2023; Ham et al., 2017). These results suggest that the unique impact of ice clouds on radiative transfer can not be ignored, and should be accounted for in model studies of the Arctic climate system. This is in agreement with studies of climate models (Fan et al., 2023; Waliser et al., 2011) that also conclude that further research is required to fully understand and address these shortcomings (Wang et al., 2020).

Figure 11. Scatter plots of simulated versus observed variables with different simulation setups. LES results (ordinate) are plotted against *MetCity* (radiation/fluxes) or *ShupeTurner* (LWP) measurements (abscissa). Each dot represents one of the 190 simulated cases. The simulations are sampled at $t = 1.5$ h after initialization, the measurements are averaged for 15 min after the radiosonde launch. The dashed line indicates the one-to-one diagonal. Each panel shows the results for two different simulation setups. Panel (a) shows the results in downwards longwave radiation for the Production and the Test setup. Analogously, panel (b) shows the upwards longwave radiation for the Test setup and a variation of it with the surface skin temperature derived from ERA5 data. Panel (c) shows the liquid water path for the Test setup and a variation with the unchanged microphysics scheme. Panel (d) shows the downwards longwave radiation for the Test setup and a variation with the original radiation scheme. Panel (e) and (f) show the near-surface sensible heat flux for the Test setup and variations with (e) the sea-ice fraction derived from the MODIS/AMSR2 satellite product and (f) with the unchanged stability functions for stable conditions. A positive sign in H_s indicates an upwards flux. The MODIS/AMSR2 run contains only 155 cases.

4.5. Sea-Ice Fraction

Small variations in the sea-ice fraction can have a strong impact on the calculated surface heat fluxes. This is especially true in Arctic winter caused by strong differences between sea-ice and ocean temperatures. In a compromise between the goal to use as many observational data sets as possible and data availability it was decided to use ERA5-derived sea-ice fraction as a boundary condition instead of an individual satellite product (see Section 2.1). To quantify the impact of this consideration Figure 11e shows the near-surface sensible heat flux for simulations with (a) ERA5 sea-ice fraction and (b) sea-ice fraction from the MODIS-AMSR2 product (Ludwig et al., 2020). Note that the number of simulated cases reduces by 35 when using the satellite product. In ii) the bias is reduced from -1.6 to 0.1 W m^{-2} but two issues persist: First, a significant number of cases still show the wrong sign when compared to *MetCity* measurement and second, the error spread increased from ± 6.0 to ± 8.4 W m^{-2}. Concluding, it was prioritized to increase the number of simulable cases over the usage of MODIS-AMSR2 data without a pronounced difference in average performance even though individual cases might be strongly influenced.

4.6. Stability Functions

In an attempt to improve model performance for the calculated surface fluxes, the stability functions derived by Grachev et al. (2007) for stable conditions during the SHEBA campaign were adopted. To quantify the impact of this choice, Figure 11f compares the results of the near-surface sensible heat flux in the Test setup for (a) the initially used stability functions (Heus et al., 2010) and (b) the newly introduced functions. Since only the stability functions under stable conditions were changed, a number of simulations are identical in both setups. The overall impact is minimal. This result excludes the stability function closure as a possible reason for the substantial model-obs differences, and would promote the other causes mentioned earlier; local impacts hindering a one-to-one comparison of model to observations, the limited applicability of standard Monin-Obukhov similarity theory in this regime (Heisel & Chamecki, 2023; Liu et al., 2024) or uncertainties in the sea-ice fraction data. Further investigation of this topic are considered future research.

4.7. Other Impacts

Other aspects of the experimental setup might have a significant impact but are not discussed here for the sake of brevity. These include the treatment of other relevant microphysical processes, aerosol, and the profiles of cloud liquid and ice water content used for model initialization. The influence of large-scale forcings such as advection and subsidence was also not assessed in this study, even though it can be significant (Savre et al., 2015). Some studies have suggested large-scale forcing could play a role in establishing the intermediate state of the bimodal Arctic (Morrison et al., 2012; Neggers et al., 2019; X. Zhang et al., 2022). Investigation of these impacts is ongoing, making use of targeted sensitivity experiments from the library of control runs as described in this study.

5. Summary, Conclusions, and Outlook

This study presents a year-long library of daily high-resolution large-eddy simulations (LES) of the atmospheric boundary layer as observed during the recent *MOSAiC* drift campaign in the central Arctic. A specific target is to deeply integrate a multitude of measurements into the experimental configuration. A dedicated standardized model setup for Arctic conditions is developed and employed for this purpose. An overarching science objective is to provide a virtual domain with resolved small-scale turbulence and clouds, which can help in interpreting local measurements made at the *Polarstern* and *MetCity*. To achieve the desired model skill, various sub-grid parameterizations in the LES code for microphysical processes and cloud-radiative interaction had to be adapted to yield realistic results under central Arctic conditions. The production runs are statistically evaluated against a year of *MOSAiC* data. Three individual case studies are investigated more closely, to assess the model's ability to represent key boundary-layer features including vertical structure and inversions, mixed-phase clouds, and turbulence. The impact of various key aspects of the model and experimental setup is assessed.

The main scientific conclusions coming out of this research can be itemized as follows:

- The statistical evaluation of the library of LES runs demonstrates that the surface radiative fluxes, the surface sensible heat flux, near-surface meteorology, and mixed-phase clouds are reproduced under a broad range of meteorological conditions.
- The bimodality in the net longwave surface flux typical of the Arctic is well reproduced, with an underestimation of intermediate states.
- An overestimation of downward longwave radiation under ice clouds is identified, which is traced to artificial parametric limitations on the maximum effective diameter of ice crystals.
- Downward shortwave radiation during polar day is slightly underestimated and speculated to be related to the use of one-dimensional radiative transfer in the model.
- On occasion the surface sensible heat flux, which is an interactive lower boundary condition in the simulations, is underestimated or falsely calculated as downward flux.
- Integrating a multitude of *MOSAiC* data into the initial- and boundary conditions proved crucial for achieving a good statistical agreement. This in particular applies to surface data, sonde data, and value-added cloud products.
- The detailed investigation of three case studies suggests that initialized liquid and ice cloud layers are maintained long enough to survive in the simulations, indicating that cloud-radiation-turbulence interactions are well captured in this setup.
- Small grid test simulations are shown to be a viable tool for configuring and optimizing full grid LES experiments of Arctic boundary layers.

The obtained results, and the library of high-resolution turbulence-resolving numerical experiments itself, create new research opportunities but also raise some new science questions. An obvious new scientific opportunity is to use the simulations to gain insight into Arctic Amplification at a process level. The three-dimensional model output at high frequencies allows for the sampling of aspects of small-scale physics and dynamics that are still impossible to measure. A prime example relevant for Arctic climate change is the sampling of tendencies of all terms in the energy and water budgets of the Arctic boundary layer and the surface. Targeted perturbation experiments can be conducted to test hypotheses, such as the role of small-scale physical processes in climate feedback mechanisms. First results of a boundary-layer budget analysis based on the LES data set for *MOSAiC* as presented in this study was recently published as part of Linke et al. (2023) (Figure 10).

Concerning the LES experiments, the library could still be expanded significantly; only $\sim$25 % of radiosonde launches were used. Applying the standardized setup to generate more simulations at other sonde launch times is straightforward. While this would represent a significant computational effort, the benefit would be an improved sample size of meteorological conditions throughout the year. The range of observational data sets integrated into the LES experiments could also still be expanded, which might further improve model skills. In addition, for various already integrated data sets alternative products could be used. Assessing the impact of these actions is a future research topic, and could further inform subsequent high-resolution model efforts of a similar nature in the Arctic.

In this study, the model evaluation was limited to a few key *MOSAiC* data sets, selected based on relevance but also on availability for the full drift. More such observational data products are available; in addition, a considerable number of special measurements were made during the campaign that are only available for short Intensive Observation Periods (IOP). These include tethered balloon data (Akansu et al., 2023; Lonardi et al., 2022; Pilz et al., 2023), unmanned aerial vehicle (UAV) measurements (de Boer et al., 2022; Egerer et al., 2023), among others. On the one hand, the LES output for these periods can play a role in providing context for these observational data sets in terms of small-scale variability surrounding the sites. On the other, some of these special observational data are also useful for evaluating the LES on resolved processes such as small-scale variability in turbulence, clouds and aerosol. Such evaluation of resolved properties against other special data sets including IOP data is a logical next step, providing further insight into LES model skill. This work is currently in progress.

A significant part of the work behind this study went into deriving adequate forcing data sets. The content and format of these daily forcings conform to the input requirements of most single-column models. A good example

is the horizontal homogeneity of forcing profiles. With resolved LES results now available to accompany these forcings, and in combination with the extensive *MOSAiC* data set, this creates a rich testing ground for improving larger-scale weather and climate models in the central Arctic. The option exists to test such models for both the whole drift or for a subset of days, to gain further process-level understanding and inform parameterization improvement. The library of cases based on MOSAiC data also invites model intercomparison studies for LES codes themselves or using the generated data to train machine learning models.

Appendix A: Vertical Grid

In this section, the choice of the vertical grid is detailed. The heights of vertical levels of the grid are calculated as follows:

$$\Delta z^k = \begin{cases} \Delta z_{\text{Start}}, & \text{if } k \leq k_T \\ \Delta z_{\text{Start}} \cdot (1+s)^{k-(k_T+1)}, & \text{if } k_M > k > k_T \\ \Delta z_{\text{End}}, & \text{if } k \geq k_M \end{cases}$$

$$z^k = \begin{cases} z^{k-1} + \Delta z^k, & \text{if } k > 1 \\ \frac{\Delta z_{\text{Start}}}{2}, & \text{if } k = 0 \end{cases}$$

$$k_T = 120, \quad s = 0.0125, \quad \Delta z_{\text{Start}} = 10\,\text{m},$$

$$k_M = 260, \quad \Delta z_{\text{End}} = 185\,\text{m}$$

This produces a grid with a maximum height of 11857.23 m. The lower 1.2 km are resolved with a fixed grid spacing of 10 m. From there, the grid spacing expands according to the given formula until it reaches $\Delta z_{\text{max}} = 185$ m at around 7 km height. From there the grid spacing is kept fixed again until the top of the domain. Figure A1 shows the height and grid spacing depending on the grid point.

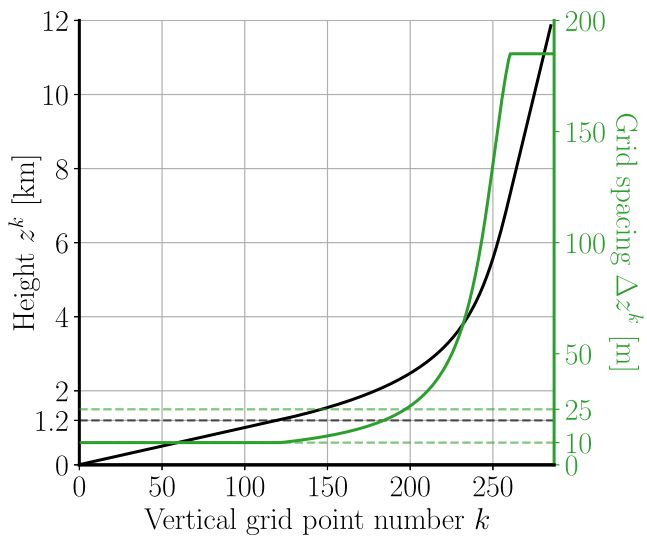

Figure A1. Vertical grid height and spacing of the standardized simulation setup.

Appendix B: Radiosonde Data

The radiosonde variables (Dahlke et al., 2023) used here are the height z, pressure p, temperature T, and the relative humidity RH. From these Θ_v and q_v are calculated using Tetens formula (Murray, 1967; Tetens, 1930):

$$\Theta = T \cdot \left(p_{\mathrm{ref}}/p\right)^{R_d/c_p} \tag{B1}$$

$$e_{\mathrm{sat,l}} = e_{\mathrm{sat,0}} \cdot \exp\left[a_t \cdot (T - T_m)/(T - b_t)\right] \tag{B2}$$

$$q_{\mathrm{sat,l}} = \frac{R_d/R_v \cdot e_{\mathrm{sat,l}}}{p - (1 - R_d/R_v) \cdot e_{\mathrm{sat,l}}} \tag{B3}$$

$$q_v = \mathrm{RH} \cdot q_{\mathrm{sat,l}} \tag{B4}$$

$$\Theta_v = \Theta \cdot \left(1 + 0.61 \cdot q_v\right) \tag{B5}$$

here, $a_t = 17.27$ (dimensionless) and $b_t = 35.86\,\mathrm{K}$ are dimensionless parameterization constants, $e_{\mathrm{sat,0}} = 610.78\,\mathrm{Pa}$ the reference vapor pressure constant, $R_d = 287.04\,\mathrm{J\,kg^{-1}\,K^{-1}}$ and $R_v = 461.5\,\mathrm{J\,kg^{-1}\,K^{-1}}$ the gas constants for dry air and water vapor, $c_p = 1004\,\mathrm{J\,kg^{-1}\,K^{-1}}$ the specific heat capacity, $T_m = 273.16\,\mathrm{K}$ the melting temperature of ice and $p_{\mathrm{ref}} = 1000\,\mathrm{hPa}$ the reference pressure. Θ is the potential temperature, e_{sat} the saturation water vapor pressure with respect to liquid and q_{sat} the corresponding saturation specific humidity.

A further correction to the derived sonde water vapor specific humidity $\left(q_v\right)$ profile is made: Thin, high-level ice clouds frequently occur in the Arctic atmosphere - in areas where the air is saturated with water vapor concerning ice. Small measurement uncertainties can lead to situations where $q_v > q_{\mathrm{sat,i}}$ and therefore place unrealistic amounts of cloud ice in the simulation domain. This is prevented by setting,

$$q_v(z) = \min\left[q_{\mathrm{sat,i}}(z),\ q_v(z)\right], \quad \text{for} \quad z \geq z_{\mathrm{cut}} \tag{B6}$$

here, $q_{\mathrm{sat,i}}$ is the saturation vapor pressure concerning ice calculated analogously to Equations B2 and B3 with $a_{t,i} = 21.8746$ and $b_{t,i} = 7.66\,\mathrm{K}$. The cutoff length z_{cut} is chosen to divide the vertical domain into areas of turbulent ice production in the boundary layer (e.g., in mixed-phase clouds) and areas of deposition ice production. It is calculated as the layer of the highest occurring liquid water content q_l above a threshold:

$$z_{\mathrm{cut}} = \max_z \left\{z \,|\, q_l(z) \geq 0.01\,\mathrm{g\,kg^{-1}}\right\} \tag{B7}$$

here, q_l is derived as described in Section 2.3.2.

Appendix C: Evaluation Point

To ensure comparability between all cases, a common point in time to evaluate the simulations has to be found. Different issues have to be considered: First, the grand goal of this work is to be able to compare the simulations to *MOSAiC* measurements. The initial conditions are derived once at the start of the simulation and, besides weak nudging, do not influence the further evolution of the atmosphere in the virtual domain. Naturally, after some time, the conditions will have changed so much due to internal turbulent, radiative, and microphysical processes that there is no value in comparing the LES to the measurements anymore. This motivates an early time point for evaluation. But, it has to be considered that the simulation needs time to "spin up." Initially, all variables in the LES are homogeneously distributed in the horizontal directions. Their vertical distribution and wind lead to turbulent mixing in the domain, ultimately creating a realistic spatial distribution of the prognostic variables.

Quantifying the right point in time, considering both of these issues, is difficult to achieve since the atmospheric state varies strongly from case to case. Figure C1 shows the distribution of the absolute change in total turbulent kinetic energy (TKE) in the collection of cases dependent on time. For better comparison, the values are normalized by the average value in the shown interval. Generally speaking, the spin-up phase is associated with a quick change in TKE, since turbulent processes are still developing. A small change in TKE signifies a period of

Acknowledgments
We thank the reviewers for their constructive and extensive feedback on this study. We gratefully acknowledge the funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—Projektnummer 268020496—TRR 172, within the Transregional Collaborative Research Center "ArctiC Amplification: Climate Relevant Atmospheric and SurfaCe Processes, and Feedback Mechanisms $(\mathcal{AC})^3$." Matthew D. Shupe was supported by the NOAA Global Ocean Monitoring and Observing Program (FundRef https://doi.org/10.13039/100018302), the Physical Sciences Laboratory (NA22OAR4320151), and a Mercator Fellowship with $(\mathcal{AC})^3$. Hersbach et al. (2018a, 2018b) was downloaded from the Copernicus Climate Change Service (C3S) (2023). The results contain modified Copernicus Climate Change Service information 2020. Neither the European Commission nor ECMWF is responsible for any use that may be made of the Copernicus information or data it contains. We gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for providing computing time on the GCS Supercomputer JUWELS at the Jülich Supercomputing Centre (JSC) under projects VIRTUALLAB and RCONGM. We furthermore thank the Regional Computing Center of the University of Cologne (RRZK) for providing computing time on the DFG-funded (Funding number: INST 216/512/1FUGG) High-Performance Computing (HPC) system CHEOPS as well as support. This work used resources of the Deutsches Klimarechenzentrum (DKRZ) granted by its Scientific Steering Committee (WLA) under project ID 1339. Data used in this manuscript were produced as part of the international Multidisciplinary drifting Observatory for the Study of Arctic Climate (*MOSAiC*) with the tag MOSAiC20192020. We thank all persons involved in the expedition of the Research Vessel Polarstern during MOSAiC (AWI_PS122_00). Radiation data was obtained from the Atmospheric Radiation Measurement (ARM) User Facility, a U.S. Department of Energy Office of Science User Facility managed by the Biological and Environmental Research Program. We thank the Institute of Environmental Physics, University of Bremen for the provision of the merged MODIS and AMSR2 sea-ice concentration data set at https://seaice.uni-bremen.de/data/modis_amsr2 (last access [22 May 2024]). Surface roughness data for the MOSAiC drift as used in the LES case configuration were kindly provided by Michael Gallagher (NOAA/CIRES), and are included in the case configuration files in Schnierstein et al. (2024). Special thanks go out to

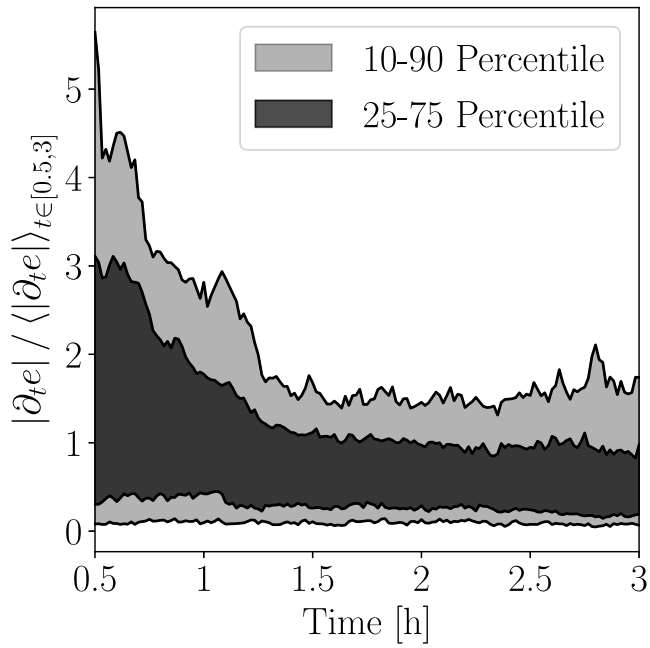

Figure C1. Distribution of normalized absolute change in total turbulent kinetic energy of all simulated cases depending on simulation time.

quasi-equilibrium, where the state of the atmosphere is not shifting significantly. Therefore, the point in time for the evaluation of the simulations is chosen to be 1.5 h.

Data Availability Statement

One of the key objectives of this work is to make all the created data available. The model code, input, and configuration, the generated forcing files, and the selected model output can be found in Schnierstein et al. (2024). The forcing files are compatible with typical single-column model (SCM) forcings because *DALES* is run with horizontally homogenous initial- and boundary conditions. The standard model output contains extensive data averaged in space and time. Some output is also available as three-dimensional fields. For a more detailed description, please refer to the accompanying documentation. Additional output can always be generated. Please get in touch if needed for your research. Finally, the *DALES* version used for the final runs is 4.3 with an extension for mixed-phase microphysics. The specific version used for this work is uploaded in Schnierstein et al. (2024). For the official *DALES* repository please refer to Arabas et al. (2021). The combination of forcing files and simulation framework makes the results fully reproducible.

References

Akansu, E. F., Dahlke, S., Siebert, H., & Wendisch, M. (2023). Evaluation of methods to determine the surface mixing layer height of the atmospheric boundary layer in the central arctic during polar night and transition to polar day in cloudless and cloudy conditions. *Atmospheric Chemistry and Physics*, *23*(24), 15473–15489. https://doi.org/10.5194/acp-23-15473-2023

Andreas, E. L., Fairall, C. W., Grachev, A. A., Guest, P. S., Horst, T. W., Jordan, R. E., & Persson, P. O. G. (2003). Turbulent transfer coefficients and roughness lengths over sea ice: The Sheba results. Retrieved from https://ams.confex.com/ams/7POLAR/techprogram/paper_60666.htm

Ansmann, A., Ohneiser, K., Engelmann, R., Radenz, M., Griesche, H., Hofer, J., et al. (2023). Annual cycle of aerosol properties over the central arctic during mosaic 2019–2020 – Light-extinction, CCN, and INP levels from the boundary layer to the tropopause. *Atmospheric Chemistry and Physics*, *23*(19), 12821–12849. https://doi.org/10.5194/acp-23-12821-2023

Arabas, S., Axelsen, S., Attema, J., Beets, C., Boeing, S. J., de Bruine, M., et al. (2021). dalesteam/dales: Dales 4.3 [Software]. *Zenodo*. https://doi.org/10.5281/zenodo.4604726

Baran, A. J. (2005). The dependence of cirrus infrared radiative properties on ice crystal geometry and shape of the size-distribution function. *Quarterly Journal of the Royal Meteorological Society*, *131*(607), 1129–1142. https://doi.org/10.1256/qj.04.91

Bintanja, R. (2000). Surface heat budget of Antarctic snow and blue ice: Interpretation of spatial and temporal variability. *Journal of Geophysical Research*, *105*(D19), 24387–24407. https://doi.org/10.1029/2000JD900356

Cairo, F., Krämer, M., Afchine, A., Di Donfrancesco, G., Di Liberto, L., Khaykin, S., et al. (2023). A comparative analysis of in situ measurements of high-altitude cirrus in the tropics. *Atmospheric Measurement Techniques*, *16*(20), 4899–4925. https://doi.org/10.5194/amt-16-4899-2023

Sandro Dahlke for providing early access to enhanced radiosonde data (Dahlke et al., 2023), and to Amy Solomon and Ola Persson for fruitful and insightful discussions, that significantly improved our LES experiments. Finally, we thank Hannes Griesche and Patric Seifert for the conversations about cloud measurements. Open Access funding enabled and organized by Projekt DEAL.

Castellani, G., Lüpkes, C., Hendricks, S., & Gerdes, R. (2014). Variability of arctic sea-ice topography and its impact on the atmospheric surface drag. *Journal of Geophysical Research: Oceans*, *119*(10), 6743–6762. https://doi.org/10.1002/2013JC009712

Chylik, J., Chechin, D., Dupuy, R., Kulla, B. S., Lüpkes, C., Mertes, S., et al. (2023). Aerosol impacts on the entrainment efficiency of arctic mixed-phase convection in a simulated air mass over open water. *Atmospheric Chemistry and Physics Discussions*, *2023*(8), 1–42. https://doi.org/10.5194/acp-23-4903-2023

Couvreux, F., Bazile, E., Rodier, Q., Maronga, B., Matheou, G., Chinita, M. J., et al. (2020). Intercomparison of large-eddy simulations of the Antarctic boundary layer for very stable stratification. *Boundary-Layer Meteorology*, *176*(3), 369–400. https://doi.org/10.1007/s10546-020-00539-4

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]. *NSF Arctic Data Center*. https://doi.org/10.18739/A2PV6B83F

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

Creamean, J. (2022). Arctic ice nucleation sampling during mosaic. *Atmospheric Radiation Measurement (ARM) user facility*. https://doi.org/10.5439/1798162

Creamean, J. M., Kirpes, R. M., Pratt, K. A., Spada, N. J., Maahn, M., de Boer, G., et al. (2018). Marine and terrestrial influences on ice nucleating particles during continuous springtime measurements in an arctic oilfield location. *Atmospheric Chemistry and Physics*, *18*(24), 18023–18042. https://doi.org/10.5194/acp-18-18023-2018

Dahlke, S., Shupe, M. D., Cox, C. J., Brooks, I. M., Blomquist, B., & Persson, P. O. G. (2023). Extended radiosonde profiles 2019/09-2020/10 during MOSAiC Legs PS122/1 - PS122/5 [Dataset]. *PANGAEA*. Retrieved from https://doi.pangaea.de/10.1594/PANGAEA.961881

Dai, A., Luo, D., Song, M., & Liu, J. (2019). Arctic amplification is caused by sea-ice loss under increasing CO2. *Nature Communications*, *10*(1), 121. https://doi.org/10.1038/s41467-018-07954-9

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

Deardorff, J. W. (1980). Stratocumulus-capped mixed layers derived from a three-dimensional model. *Boundary-Layer Meteorology*, *18*(4), 495–527. https://doi.org/10.1007/bf00119502

de Boer, G., Calmer, R., Jozef, G., Cassano, J. J., Hamilton, J., Lawrence, D., et al. (2022). Observing the central arctic atmosphere and surface with university of Colorado uncrewed aircraft systems. *Scientific Data*, *9*(1), 439. https://doi.org/10.1038/s41597-022-01526-9

Derbyshire, S., Beau, I., Bechtold, P., Grandpeix, J.-Y., Piriou, J.-M., Redelsperger, J.-L., & Soares, P. (2004). Sensitivity of moist convection to environmental humidity. *Quarterly Journal of the Royal Meteorological Society: A Journal of the Atmospheric Sciences, Applied Meteorology and Physical Oceanography*, *130*(604), 3055–3079. https://doi.org/10.1256/qj.03.130

de Roode, S. R., Frederikse, T., Siebesma, A. P., Ackerman, A. S., Chylik, J., Field, P. R., et al. (2019). Turbulent transport in the gray zone: A large eddy model intercomparison study of the constrain cold air outbreak case. *Journal of Advances in Modeling Earth Systems*, *11*(3), 597–623. https://doi.org/10.1029/2018MS001443

de Roode, S. R., Sandu, I., van der Dussen, J. J., Ackerman, A. S., Blossey, P., Jarecka, D., et al. (2016). Large-eddy simulations of EUCLIPSE–GASS Lagrangian stratocumulus-to-cumulus transitions: Mean state, turbulence, and decoupling. *Journal of the Atmospheric Sciences*, *73*(6), 2485–2508. https://doi.org/10.1175/JAS-D-15-0215.1

Dimitrelos, A., Ekman, A. M., Caballero, R., & Savre, J. (2020). A sensitivity study of arctic air-mass transformation using large eddy simulation. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD031738. https://doi.org/10.1029/2019JD031738

Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., & Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, *116*, 140–158. Advanced Along Track Scanning Radiometer (AATSR) Special Issue. https://doi.org/10.1016/j.rse.2010.10.017

Doran, J. C., Hubbe, J. M., Liljegren, J. C., Shaw, W. J., Collatz, G. J., Cook, D. R., & Hart, R. L. (1998). A technique for determining the spatial and temporal distributions of surface fluxes of heat and moisture over the southern Great Plains cloud and radiation testbed. *Journal of Geophysical Research*, *103*(D6), 6109–6121. https://doi.org/10.1029/97JD03427

ECMWF. (2016). *Ifs documentation cy41r2 - Part II: Data assimilation*. Author. https://doi.org/10.21957/vhe0zlxr8

ECMWF. (2021). *Ifs documentation cy47r3 - Part IV physical processes*. Author. https://doi.org/10.21957/eyrpir4vj

Egerer, U., Cassano, J. J., Shupe, M. D., de Boer, G., Lawrence, D., Doddi, A., et al. (2023). Estimating turbulent energy flux vertical profiles from uncrewed aircraft system measurements: Exemplary results for the mosaic campaign. *Atmospheric Measurement Techniques*, *16*(8), 2297–2317. https://doi.org/10.5194/amt-16-2297-2023

Egerer, U., Ehrlich, A., Gottschalk, M., Griesche, H., Neggers, R. A. J., Siebert, H., & Wendisch, M. (2021). Case study of a humidity layer above arctic stratocumulus and potential turbulent coupling with the cloud top. *Atmospheric Chemistry and Physics*, *21*(8), 6347–6364. https://doi.org/10.5194/acp-21-6347-2021

Eichler, H., Ehrlich, A., Wendisch, M., Mioche, G., Gayet, J.-F., Wirth, M., et al. (2009). Influence of ice crystal shape on retrieval of cirrus optical thickness and effective radius: A case study. *Journal of Geophysical Research*, *114*(D19), D19203. https://doi.org/10.1029/2009JD012215

Fairall, C. W., Bradley, E. F., Rogers, D. P., Edson, J. B., & Young, G. S. (1996). Bulk parameterization of air-sea fluxes for tropical ocean-global atmosphere coupled-ocean atmosphere response experiment. *Journal of Geophysical Research*, *101*(C2), 3747–3764. https://doi.org/10.1029/95JC03205

Fan, C., Chen, Y.-H., Chen, X., Lin, W., Yang, P., & Huang, X. (2023). A refined understanding of the ice cloud longwave scattering effects in climate model. *Journal of Advances in Modeling Earth Systems*, *15*(10), e2023MS003810. https://doi.org/10.1029/2023MS003810

Fridlind, A. M., & Ackerman, A. S. (2018). Chapter 7 - Simulations of Arctic mixed-phase boundary layer clouds: Advances in understanding and outstanding questions. In C. Andronache (Ed.), *Mixed-phase clouds* (pp. 153–183). Elsevier. https://doi.org/10.1016/B978-0-12-810549-8.00007-6

Fridlind, A. M., van Diedenhoven, B., Ackerman, A. S., Avramov, A., Mrowiec, A., Morrison, H., et al. (2012). A FIRE-ACE/SHEBA case study of mixed-phase arctic boundary layer clouds: Entrainment rate limitations on rapid primary ice nucleation processes. *Journal of the Atmospheric Sciences*, *69*(1), 365–389. https://doi.org/10.1175/JAS-D-11-052.1

Fu, Q., & Liou, K. N. (1993). Parameterization of the radiative properties of cirrus clouds. *Journal of the Atmospheric Sciences*, *50*(13), 2008–2025. https://doi.org/10.1175/1520-0469(1993)050⟨2008:POTRPO⟩2.0.CO;2

Grachev, A. A., Andreas, E. L., Fairall, C. W., Guest, P. S., & Persson, P. O. G. (2007). Sheba flux–profile relationships in the stable atmospheric boundary layer. *Bound.-Lay. Meteorol*, *124*(3), 315–333. https://doi.org/10.1007/s10546-007-9177-6

Gustafson Jr, W. I., Vogelmann, A. M., Li, Z., Cheng, X., Dumas, K. K., Endo, S., et al. (2020). The large-eddy simulation (LES) atmospheric radiation measurement (ARM) symbiotic simulation and observation (lasso) activity for continental shallow convection. *Bulletin of the American Meteorological Society*, *101*(4), E462–E479. https://doi.org/10.1175/bams-d-19-0065.1

Ham, S.-H., Kato, S., & Rose, F. G. (2017). Examining impacts of mass-diameter (M-D) and area-diameter (A-D) relationships of ice particles on retrievals of effective radius and ice water content from radar and Lidar measurements. *Journal of Geophysical Research: Atmospheres*, *122*(6), 3396–3420. https://doi.org/10.1002/2016JD025672

Heisel, M., & Chamecki, M. (2023). Evidence of mixed scaling for mean profile similarity in the stable atmospheric surface layer. *Journal of the Atmospheric Sciences*, *80*(8), 2057–2073. https://doi.org/10.1175/JAS-D-22-0260.1

Held, A., Brooks, I. M., Leck, C., & Tjernström, M. (2011). On the potential contribution of open lead particle emissions to the central arctic aerosol concentration. *Atmospheric Chemistry and Physics*, *11*(7), 3093–3105. https://doi.org/10.5194/acp-11-3093-2011

Herrmannsdörfer, L., Müller, M., Shupe, M. D., & Rostosky, P. (2023). Surface temperature comparison of the Arctic winter MOSAiC observations, ERA5 reanalysis, and MODIS satellite retrieval. *Elementa: Science of the Anthropocene*, *11*(1), 00085. https://doi.org/10.1525/elementa.2022.00085

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018a). Era5 hourly data on pressure levels from 1940 to present. ECMWF.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2018b). *Era5 hourly data on single levels from 1940 to present*. Copernicus climate change service (C3S) climate data store (CDS).

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., et al. (2023). *Era5 hourly data on pressure levels from 1940 to present*. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). https://doi.org/10.24381/cds.bd0915c6

Heus, T., van Heerwaarden, C. C., Jonker, H. J. J., Pier Siebesma, A., Axelsen, S., van den Dries, K., et al. (2010). Formulation of the Dutch atmospheric large-eddy simulation (dales) and overview of its applications. *Geoscientific Model Development*, *3*(2), 415–444. https://doi.org/10.5194/gmd-3-415-2010

Hundsdorfer, W., Koren, B., van Loon, M., & Verwer, J. (1995). A positive finite-difference advection scheme. *Journal of Computational Physics*, *117*(1), 35–46. https://doi.org/10.1006/jcph.1995.1042

Istomina, L., Marks, H., Huntemann, M., Heygster, G., & Spreen, G. (2020). Improved cloud detection over sea ice and snow during arctic summer using meris data. *Atmospheric Measurement Techniques*, *13*(12), 6459–6472. https://doi.org/10.5194/amt-13-6459-2020

Jeffery, C. A., & Austin, P. H. (1997). Homogeneous nucleation of supercooled water: Results from a new equation of state. *Journal of Geophysical Research*, *102*(D21), 25269–25279. https://doi.org/10.1029/97JD02243

Jenkins, M., & Dai, A. (2021). The impact of sea-ice loss on arctic climate feedbacks and their role for arctic amplification. *Geophysical Research Letters*, *48*(15), e2021GL094599. https://doi.org/10.1029/2021GL094599

Jozef, G. C., Cassano, J. J., Dahlke, S., Dice, M., Cox, C. J., & de Boer, G. (2024). An overview of the vertical structure of the atmospheric boundary layer in the central arctic during mosaic. *Atmospheric Chemistry and Physics*, *24*(2), 1429–1450. https://doi.org/10.5194/acp-24-1429-2024

Kiszler, T., Ebell, K., & Schemann, V. (2023). A performance baseline for the representation of clouds and humidity in cloud-resolving ICON-LEM simulations in the arctic. *Journal of Advances in Modeling Earth Systems*, *15*(5), e2022MS003299. https://doi.org/10.1029/2022MS003299

Klein, S. A., McCoy, R. B., Morrison, H., Ackerman, A. S., Avramov, A., de Boer, G., et al. (2009). 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*(641), 979–1002. https://doi.org/10.1002/qj.416

Konoshonkin, A., Borovoi, A., Kustova, N., Okamoto, H., Ishimoto, H., Grynko, Y., & Förstner, J. (2017). Light scattering by ice crystals of cirrus clouds: From exact numerical methods to physical-optics approximation. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *195*, 132–140. https://doi.org/10.1016/j.jqsrt.2016.12.024

Koontz, A., Uin, J., Andrews, E., Enekwizu, O., Hayes, C., & Salwen, C. (2020). Cloud condensation nuclei particle counter (aosccn2colaavg). *Atmospheric Radiation Measurement (ARM) user facility*. https://doi.org/10.5439/1323894

Krč, P., Resler, J., Sühring, M., Schubert, S., Salim, M. H., & Fuka, V. (2021). Radiative transfer model 3.0 integrated into the palm model system 6.0. *Geoscientific Model Development*, *14*(5), 3095–3120. https://doi.org/10.5194/gmd-14-3095-2021

Li, X., Krueger, S. K., Strong, C., Mace, G. G., & Benson, S. (2020). Midwinter arctic leads form and dissipate low clouds. *Nature Communications*, *11*(1), 206. https://doi.org/10.1038/s41467-019-14074-5

Linke, O., Quaas, J., Baumer, F., Becker, S., Chylik, J., Dahlke, S., et al. (2023). Constraints on simulated past arctic amplification and lapse-rate feedback from observations. *Atmospheric Chemistry and Physics Discussions*, *2023*, 1–37. https://doi.org/10.5194/acp-2022-836

Liou, K.-N., Fu, Q., & Ackerman, T. P. (1988). A simple formulation of the delta-four-stream approximation for radiative transfer parameterizations. *Journal of the Atmospheric Sciences*, *45*(13), 1940–1948. https://doi.org/10.1175/1520-0469(1988)045⟨1940:ASFOTD⟩2.0.CO;2

Liu, C., Yang, Q., Gao, Z., Shupe, M. D., Han, B., Zhang, H., et al. (2024). The role of non-local effects on surface sensible heat flux under different types of thermal structures over the arctic sea-ice surface. *Geophysical Research Letters*, *51*(4), e2023GL106753. https://doi.org/10.1029/2023GL106753

Lonardi, M., Pilz, C., Akansu, E. F., Dahlke, S., Egerer, U., Ehrlich, A., et al. (2022). Tethered balloon-borne profile measurements of atmospheric properties in the cloudy atmospheric boundary layer over the arctic sea ice during mosaic: Overview and first results. *Elementa: Science of the Anthropocene*, *10*(1), 000120. https://doi.org/10.1525/elementa.2021.000120

López-García, V., Neely, I., Ryan, R., Dahlke, S., & Brooks, I. M. (2022). Low-level jets over the Arctic Ocean during MOSAiC. *Elementa: Science of the Anthropocene*, *10*(1), 00063. https://doi.org/10.1525/elementa.2022.00063

Louis, J.-F. (1979). A parametric model of vertical eddy fluxes in the atmosphere. *Boundary-Layer Meteorology*, *17*(2), 187–202. https://doi.org/10.1007/bf00117978

Ludwig, V., Spreen, G., & Pedersen, L. T. (2020). Evaluation of a new merged sea-ice concentration dataset at 1 km resolution from thermal infrared and passive microwave satellite data in the arctic. *Remote Sensing*, *12*(19), 3183. https://doi.org/10.3390/rs12193183

Maturilli, M., Sommer, M., Holdridge, D. J., Dahlke, S., Graeser, J., Sommerfeld, A., et al. (2022). MOSAiC radiosonde data (level 3) [Dataset]. *PANGAEA*. https://doi.org/10.1594/PANGAEA.943870

McFarquhar, G. M., & Heymsfield, A. J. (1998). The definition and significance of an effective radius for ice clouds. *Journal of the Atmospheric Sciences*, *55*(11), 2039–2052. https://doi.org/10.1175/1520-0469(1998)055⟨2039:TDASOA⟩2.0.CO;2

Mchedlishvili, A., Lüpkes, C., Petty, A., Tsamados, M., & Spreen, G. (2023). New estimates of pan-arctic sea ice–atmosphere neutral drag coefficients from ICESat-2 elevation data. *The Cryosphere*, *17*(9), 4103–4131. https://doi.org/10.5194/tc-17-4103-2023

Michaelis, J., Lüpkes, C., Zhou, X., Gryschka, M., & Gryanik, V. M. (2020). Influence of lead width on the turbulent flow over sea ice leads: Modeling and parametrization. *Journal of Geophysical Research: Atmospheres*, *125*(15), e2019JD031996. https://doi.org/10.1029/2019JD031996

Middlemas, E. A., Kay, J. E., Medeiros, B. M., & Maroon, E. A. (2020). Quantifying the influence of cloud radiative feedbacks on arctic surface warming using cloud locking in an earth system model. *Geophysical Research Letters*, *47*(15), e2020GL089207. https://doi.org/10.1029/2020GL089207

Mitchell, D. L. (2002). Effective diameter in radiation transfer: General definition, applications, and limitations. *Journal of the Atmospheric Sciences*, *59*(15), 2330–2346. https://doi.org/10.1175/1520-0469(2002)059⟨2330:EDIRTG⟩2.0.CO;2

Mitchell, D. L., Lawson, R. P., & Baker, B. (2011). Understanding effective diameter and its application to terrestrial radiation in ice clouds. *Atmospheric Chemistry and Physics*, *11*(7), 3417–3429. https://doi.org/10.5194/acp-11-3417-2011

Morrison, H., de Boer, G., Feingold, G., Harrington, J., Shupe, M. D., & Sulia, K. (2012). Resilience of persistent arctic mixed-phase clouds. *Nature Geoscience*, *5*(1), 11–17. https://doi.org/10.1038/NGEO1332

Morrison, H., Zuidema, P., Ackerman, A. S., Avramov, A., De Boer, G., Fan, J., et al. (2011). Intercomparison of cloud model simulations of arctic mixed-phase boundary layer clouds observed during sheba/fire-ace. *Journal of Advances in Modeling Earth Systems*, *3*(2). https://doi.org/10.1029/2011ms000066

Murray, F. W. (1967). On the computation of saturation vapor pressure. *Journal of Applied Meteorology and Climatology*, *6*(1), 203–204. https://doi.org/10.1175/1520-0450(1967)006⟨0203:OTCOSV⟩2.0.CO;2

Neggers, R. A. J., Ackerman, A. S., Angevine, W. M., Bazile, E., Beau, I., Blossey, P. N., et al. (2017). Single-column model simulations of subtropical marine boundary-layer cloud transitions under weakening inversions. *Journal of Advances in Modeling Earth Systems*, *9*(6), 2385–2412. https://doi.org/10.1002/2017MS001064

Neggers, R. A. J., Chylik, J., Egerer, U., Griesche, H., Schemann, V., Seifert, P., et al. (2019). Local and remote controls on arctic mixed-layer evolution. *Journal of Advances in Modeling Earth Systems*, *11*(7), 2214–2237. https://doi.org/10.1029/2019MS001671

Neggers, R. A. J., Siebesma, A. P., & Heus, T. (2012). Continuous single-column model evaluation at a permanent meteorological supersite. *Bulletin of the American Meteorological Society*, *93*(9), 1389–1400. https://doi.org/10.1175/BAMS-D-11-00162.1

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

Niehaus, H., Spreen, G., Birnbaum, G., Istomina, L., Jäkel, E., Linhardt, F., et al. (2023). Sea ice melt pond fraction derived from sentinel-2 data: Along the mosaic drift and arctic-wide. *Geophysical Research Letters*, *50*(5), e2022GL102102. https://doi.org/10.1029/2022GL102102

Ong, C. R., Koike, M., Hashino, T., & Miura, H. (2022). Modeling performance of scale-amps: Simulations of arctic mixed-phase clouds observed during sheba. *Journal of Advances in Modeling Earth Systems*, *14*(6), e2021MS002887. https://doi.org/10.1029/2021MS002887

Ovchinnikov, M., Ackerman, A. S., Avramov, A., Cheng, A., Fan, J., Fridlind, A. M., et al. (2014). Intercomparison of large-eddy simulations of arctic mixed-phase clouds: Importance of ice size distribution assumptions. *Journal of Advances in Modeling Earth Systems*, *6*(1), 223–248. https://doi.org/10.1002/2013MS000282

Park, K., Kim, I., Choi, J.-O., Lee, Y., Jung, J., Ha, S.-Y., et al. (2019). Unexpectedly high dimethyl sulfide concentration in high-latitude arctic sea ice melt ponds. *Environmental Sciences: Processes & Impacts*, *21*(10), 1642–1649. https://doi.org/10.1039/C9EM00195F

Peng, S., Yang, Q., Shupe, M. D., Xi, X., Han, B., Chen, D., et al. (2023). The characteristics of atmospheric boundary layer height over the Arctic Ocean during mosaic. *Atmospheric Chemistry and Physics*, *23*(15), 8683–8703. https://doi.org/10.5194/acp-23-8683-2023

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., Uttal, T., Intrieri, J., Fairall, C., Andreas, E., & Guest, P. (1999). Observations of large thermal transitions during the arctic night from a suite of sensors at sheba. In *Preprints, fifth conference on polar meteorology and oceanography* (pp. 10–15).

Philipp, D., Stengel, M., & Ahrens, B. (2020). Analyzing the arctic feedback mechanism between sea ice and low-level clouds using 34 years of satellite observations. *Journal of Climate*, *33*(17), 7479–7501. https://doi.org/10.1175/JCLI-D-19-0895.1

Pilz, C., Lonardi, M., Egerer, U., Siebert, H., Ehrlich, A., Heymsfield, A. J., et al. (2023). Profile observations of the arctic atmospheric boundary layer with the beluga tethered balloon during mosaic. *Scientific Data*, *10*(1), 534. https://doi.org/10.1038/s41597-023-02423-5

Pincus, R., & Stevens, B. (2009). Monte Carlo spectral integration: A consistent approximation for radiative transfer in large eddy simulations. *Journal of Advances in Modeling Earth Systems*, *1*(2). https://doi.org/10.3894/JAMES.2009.1.1

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., & Mauritsen, T. (2014). Arctic amplification dominated by temperature feedbacks in contemporary climate models. *Nature Geoscience*, *7*(3), 181–184. https://doi.org/10.1038/ngeo2071

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

Pruppacher, H., & Klett, J. (1996). *Microphysics of clouds and precipitation*. Springer. Retrieved from https://books.google.de/books?id=1mXN_qZ5sNUC

Rabe, B., Heuzé, C., Regnery, J., Aksenov, Y., Allerholt, J., Athanase, M., et al. (2022). Overview of the MOSAiC expedition: Physical oceanography. *Elementa: Science of the Anthropocene*, *10*(1), 00062. https://doi.org/10.1525/elementa.2021.00062

Randall, D. A., & Cripe, D. G. (1999). Alternative methods for specification of observed forcing in single-column models and cloud system models. *Journal of Geophysical Research*, *104*(D20), 24527–24545. https://doi.org/10.1029/1999jd900765

Rantanen, M., Karpechko, A. Y., Lipponen, A., Nordling, K., Hyvärinen, O., Ruosteenoja, K., et al. (2022). The arctic has warmed nearly four times faster than the globe since 1979. *Communications Earth & Environment*, *3*(1), 168. https://doi.org/10.1038/s43247-022-00498-3

Rauterkus, R., & Ansorge, C. (2020). Cloud-top entrainment in mixed-phase stratocumulus and its process-level representation in large-eddy simulation. *Journal of the Atmospheric Sciences*, *77*(12), 4109–4127. https://doi.org/10.1175/JAS-D-19-0221.1

Reisner, J., Rasmussen, R. M., & Bruintjes, R. T. (1998). Explicit forecasting of supercooled liquid water in winter storms using the mm5 mesoscale model. *Quarterly Journal of the Royal Meteorological Society*, *124*(548), 1071–1107. https://doi.org/10.1002/qj.49712454804

Riihimaki, L. (2021). Radiation instruments on ice (iceradriihimaki). https://doi.org/10.5439/1608608

Ryan, B. F. (2000). A bulk parameterization of the ice particle size distribution and the optical properties in ice clouds. *Journal of the Atmospheric Sciences*, *57*(9), 1436–1451. https://doi.org/10.1175/1520-0469(2000)057⟨1436:ABPOTI⟩2.0.CO;2

Sandu, I., & Stevens, B. (2011). On the factors modulating the stratocumulus to cumulus transitions. *Journal of the Atmospheric Sciences*, *68*(9), 1865–1881. https://doi.org/10.1175/2011JAS3614.1

Savre, J., & Ekman, A. M. L. (2015). Large-eddy simulation of three mixed-phase cloud events during ISDAC: Conditions for persistent heterogeneous ice formation. *Journal of Geophysical Research: Atmospheres*, *120*(15), 7699–7725. https://doi.org/10.1002/2014JD023006

Savre, J., Ekman, A. M. L., Svensson, G., & Tjernström, M. (2015). Large-eddy simulations of an arctic mixed-phase stratiform cloud observed during ISDAC: Sensitivity to moisture aloft, surface fluxes and large-scale forcing. *Quarterly Journal of the Royal Meteorological Society*, *141*(689), 1177–1190. https://doi.org/10.1002/qj.2425

Schalkwijk, J., Jonker, H. J. J., Siebesma, A. P., & Bosveld, F. C. (2015). A year-long large-eddy simulation of the weather over CABAUW: An overview. *Monthly Weather Review*, *143*(3), 828–844. https://doi.org/10.1175/MWR-D-14-00293.1

Schnierstein, N., Chylik, J., Neggers, R., & Shupe, M. D. (2024). A year in LES: Standardized daily high-resolution Large Eddy Simulations of the Arctic boundary layer and clouds during the MOSAiC drift [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.10491362

Screen, J. A., & Simmonds, I. (2010). The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, *464*(7293), 1334–1337. https://doi.org/10.1038/nature09051

Seifert, A., & Beheng, K. D. (2006a). A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteorology and Atmospheric Physics*, *92*(1–2), 45–66. https://doi.org/10.1007/s00703-005-0112-4

Seifert, A., & Beheng, K. D. (2006b). A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 2: Maritime vs. continental deep convective storms. Meteor. *Atmospheric physics*, *92*(1–2), 67–82. https://doi.org/10.1007/s00703-005-0113-3

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

Shupe, M. D. (2011). Clouds at arctic atmospheric observatories. part ii: Thermodynamic phase characteristics. *Journal of Applied Meteorology and Climatology*, *50*(3), 645–661. https://doi.org/10.1175/2010JAMC2468.1

Shupe, M. D., & 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., Matrosov, S. Y., & Uttal, T. (2006). Arctic mixed-phase cloud properties derived from surface-based sensors at sheba. *Journal of the Atmospheric Sciences*, *63*(2), 697–711. https://doi.org/10.1175/JAS3659.1

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

Shupe, M. D., 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). https://doi.org/10.1525/elementa.2021.00060

Shupe, M. D., Rex, M., Dethloff, K., Damm, E., Fong, A. A., Gradinger, R., et al. (2020). Arctic report card 2020: The MOSAiC expedition: A year drifting with the arctic sea ice. *Arctic Report Card*. https://doi.org/10.25923/9G3V-XH92

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

Siebesma, A. P., Bretherton, C. S., Brown, A., Chlond, A., Cuxart, J., Duynkerke, P. G., et al. (2003). A large eddy simulation intercomparison study of shallow cumulus convection. *Journal of the Atmospheric Sciences*, *60*(10), 1201–1219. https://doi.org/10.1175/1520-0469(2003)60⟨1201:ALESIS⟩2.0.CO;2

Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Liu, Z., Berner, J., et al. (2021). A description of the advanced research WRF model version 4.3; no. NCAR/TN556+ STR.

Sobel, A. H., & Bretherton, C. S. (2000). Modeling tropical precipitation in a single column. *Journal of Climate*, *13*(24), 4378–4392. https://doi.org/10.1175/1520-0442(2000)013<4378:mtpias>2.0.co;2

Solomon, A., Shupe, M. D., Persson, P. O. G., & Morrison, H. (2011). Moisture and dynamical interactions maintaining decoupled arctic mixed-phase stratocumulus in the presence of a humidity inversion. *Atmospheric Chemistry and Physics*, *11*(19), 10127–10148. https://doi.org/10.5194/acp-11-10127-2011

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

Spreen, G., Kaleschke, L., & Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research*, *113*(C2), C02S03. https://doi.org/10.1029/2005JC003384

Stevens, B., Moeng, C.-H., Ackerman, A. S., Bretherton, C. S., Chlond, A., de Roode, S., et al. (2005). Evaluation of large-eddy simulations via observations of nocturnal marine stratocumulus. *Monthly Weather Review*, *133*(6), 1443–1462. https://doi.org/10.1175/MWR2930.1

Stevens, R. G., Loewe, K., Dearden, C., Dimitrelos, A., Possner, A., Eirund, G. K., et al. (2018). A model intercomparison of CCN-limited tenuous clouds in the high arctic. *Atmospheric Chemistry and Physics*, *18*(15), 11041–11071. https://doi.org/10.5194/acp-18-11041-2018

Stramler, K., Del Genio, A., & Rossow, W. B. (2011). Synoptically driven arctic winter states. *Journal of Climate*, *47*(6), 1747–1762. https://doi.org/10.1175/2010JCLI3817.1

Stuecker, M. F., Bitz, C. M., Armour, K. C., Proistosescu, C., Kang, S. M., Xie, S.-P., et al. (2018). Polar amplification dominated by local forcing and feedbacks. *Nature Climate Change*, *8*(12), 1076–1081. https://doi.org/10.1038/s41558-018-0339-y

Sverdrup, H. (1933). The Norwegian north polar expedition with the "MAUD", 1918-1925: Scientific results. Geofysisk Institutt. In *co-operation with other institutions and A. S. John Griegs Boktrykkeri*.

Tang, S., Xie, S., Zhang, M., Tang, Q., Zhang, Y., Klein, S. A., et al. (2019). Differences in eddy-correlation and energy-balance surface turbulent heat flux measurements and their impacts on the large-scale forcing fields at the arm SGP site. *Journal of Geophysical Research: Atmospheres*, *124*(6), 3301–3318. https://doi.org/10.1029/2018JD029689

Taylor, P. C., Hegyi, B. M., Boeke, R. C., & Boisvert, L. N. (2018). On the increasing importance of air-sea exchanges in a thawing arctic: A review. *Atmosphere*, *9*(2), 41. https://doi.org/10.3390/atmos9020041

Tetens, O. (1930). über einige meteorologische begriffe. *Zeitschrift für Geophysik*, *6*, 297–309.

Thackeray, C. W., & Hall, A. (2019). An emergent constraint on future Arctic sea-ice albedo feedback. *Nature Climate Change*, *9*(12), 972–978. https://doi.org/10.1038/s41558-019-0619-1

Tijhuis, M., van Stratum, B. J. H., Veerman, M. A., & van Heerwaarden, C. C. (2023). An efficient parameterization for surface shortwave 3d radiative effects in large-eddy simulations of shallow cumulus clouds. *Journal of Advances in Modeling Earth Systems*, *15*(1), e2022MS003262. https://doi.org/10.1029/2022MS003262

Van der Dussen, J., De Roode, S., Ackerman, A., Blossey, P., Bretherton, C., Kurowski, M., et al. (2013). The GASS/EUCLIPSE model intercomparison of the stratocumulus transition as observed during ASTEX: LES results. *Journal of Advances in Modeling Earth Systems*, *5*(3), 483–499. https://doi.org/10.1002/jame.20033

van der Linden, S. J. A., van de Wiel, B. J. H., Petenko, I., van Heerwaarden, C. C., Baas, P., & Jonker, H. J. J. (2020). A businger mechanism for intermittent bursting in the stable boundary layer. *Journal of the Atmospheric Sciences*, *77*(10), 3343–3360. https://doi.org/10.1175/JAS-D-19-0309.1

van Laar, T. W., Schemann, V., & Neggers, R. A. J. (2019). Investigating the diurnal evolution of the cloud size distribution of continental cumulus convection using multi-day les. *Journal of the Atmospheric Sciences*, *76*(3), 729–747. https://doi.org/10.1175/JAS-D-18-0084.1

Veerman, M. A., Pedruzo-Bagazgoitia, X., Jakub, F., Vilà-Guerau de Arellano, J., & van Heerwaarden, C. C. (2020). Three-dimensional radiative effects by shallow cumulus clouds on dynamic heterogeneities over a vegetated surface. *Journal of Advances in Modeling Earth Systems*, *12*(7), e2019MS001990. https://doi.org/10.1029/2019MS001990

Wagner, D. N., Shupe, M. D., Cox, C., Persson, O. G., Uttal, T., Frey, M. M., et al. (2022). Snowfall and snow accumulation during the mosaic winter and spring seasons. *The Cryosphere*, *16*(6), 2373–2402. https://doi.org/10.5194/tc-16-2373-2022

Waliser, D. E., Li, J.-L. F., L'Ecuyer, T. S., & Chen, W.-T. (2011). The impact of precipitating ice and snow on the radiation balance in global climate models. *Geophysical Research Letters*, *38*(6). https://doi.org/10.1029/2010GL046478

Wang, Y., Su, H., Jiang, J. H., Xu, F., & Yung, Y. L. (2020). Impact of cloud ice particle size uncertainty in a climate model and implications for future satellite missions. *Journal of Geophysical Research: Atmospheres*, *125*(6), e2019JD032119. https://doi.org/10.1029/2019JD032119

Wendisch, M., Brückner, M., Crewell, S., Ehrlich, A., Notholt, J., Lüpkes, C., et al. (2023). Atmospheric and surface processes, and feedback mechanisms determining arctic amplification: A review of first results and prospects of the (ac)3 project. *Bulletin of the American Meteorological Society*, *104*(1), E208–E242. https://doi.org/10.1175/BAMS-D-21-0218.1

Wyngaard, J. C. (2004). Toward numerical modeling in the "Terra Incognita". *Journal of the Atmospheric Sciences*, *61*(14), 1816–1826. https://doi.org/10.1175/1520-0469(2004)061<1816:tnmitt>2.0.co;2

Yang, P., & Fu, Q. (2009). Dependence of ice crystal optical properties on particle aspect ratio. *Journal of Quantitative Spectroscopy and Radiative Transfer*, *110*(14), 1604–1614. XI Conference on Electromagnetic and Light Scattering by Non-Spherical Particles: 2008. https://doi.org/10.1016/j.jqsrt.2009.03.004

Zhang, M., Bretherton, C. S., Blossey, P. N., Austin, P. H., Bacmeister, J. T., Bony, S., et al. (2013). CGILS: Results from the first phase of an international project to understand the physical mechanisms of low cloud feedbacks in single column models. *Journal of Advances in Modeling Earth Systems*, *5*(4), 826–842. https://doi.org/10.1002/2013MS000246

Zhang, X., Schneider, T., Shen, Z., Pressel, K. G., & Eisenman, I. (2022). Seasonal cycle of idealized polar clouds: Large eddy simulations driven by a GCM. *Journal of Advances in Modeling Earth Systems*, *14*(1), e2021MS002671. https://doi.org/10.1029/2021MS002671