On the Importance of Representing Snow Over Sea-Ice for Simulating the Arctic Boundary Layer

Gabriele Arduini1, Sarah Keeley1, Jonathan J. Day1, Irina Sandu1, Lorenzo Zampieri2,3, and Gianpaolo Balsamo1

1European Centre for Medium-Range Weather Forecasts, Reading, UK, 2Now at National Center for Atmospheric Research, Boulder, CO, USA, 3Alfred Wegener Institute, Bremerhaven, Germany

Abstract Correctly representing the snow on sea-ice has great potential to improve cryosphere-atmosphere coupling in forecasting and monitoring (e.g., reanalysis) applications, via improved modeling of surface temperature, albedo and emissivity. This can also enhance the all-weather all-surface coupled data assimilation for atmospheric satellite radiances. Using wintertime observations from two Arctic field campaigns, SHEBA and N-ICE2015, and satellite data, we can explore the merits of different approaches to represent the snow over sea-ice in a set of 5-day coupled forecasts. Results show that representing the snow insulation effects is essential for capturing the wintertime surface temperature variability over sea-ice and its response to changes in the atmospheric forcing. Modeling the snow over sea-ice improves the representation of strong cooling events, reduces surface temperature biases in clear-sky conditions and improves the simulation of surface-based temperature inversions. In clear-sky conditions, when using a multi-layer snow scheme the root-mean-squared error in the surface temperature is reduced by about 60% for both N-ICE2015 and SHEBA. This study also highlights the role of compensating errors in different components of the surface energy budget in the Arctic boundary layer. During warm air intrusions, errors in the surface temperature increase when cloud phase and cloud radiative processes are misrepresented in the model, inducing large errors in the net radiative energy at the surface. This work indicates that numerical weather prediction systems can fully benefit from a better representation of snow over sea-ice, for example, with multi-layer snow schemes, combined with improvements to other boundary layer processes including mixed phase clouds.

Plain Language Summary Correctly representing the snow on sea-ice in coupled numerical weather prediction models has great potential to improve weather forecast and climate monitoring applications, such as climate reanalyses, which are usually produced using such systems. In this study two different methodologies to account for the effect of the snowpack accumulating over the sea-ice in 5-day global forecasts are compared. All experiments are performed with the ECMWF Integrated Forecast System. The presence of snow over sea-ice enables an improved representation of extreme low temperatures occurring in the Arctic, and so reducing surface temperature errors in clear-sky conditions. This study also investigates feedbacks between errors in processes controlling the surface temperature evolution in the wintertime Arctic boundary layer, and how more realistic representations of the snowpack accumulating on sea-ice are affected by these. This work indicates that numerical weather prediction systems can fully benefit from a better representation of snow over sea-ice, for example, with a multi-layer snow scheme, combined with further improvements to other boundary layer processes.

1. Introduction

Snow is a key element of the climate system due to its large impacts on the surface-atmosphere exchanges of heat and moisture at high latitudes (Thackeray et al., 2019). Over polar regions, snow accumulates both over land and sea-ice surfaces, therefore leading to complex feedbacks between the atmosphere, snow, ice and ocean system. Snow and ice act as thermal insulators between the warm ocean and the atmosphere above, with the snow having a larger impact on conductive heat fluxes per unit thickness due to its lower thermal conductivity compared to ice (Maykut, 1978; Maykut & Untersteiner, 1971). During wintertime, the presence of snow over sea-ice leads to an effective thermal decoupling between the atmosphere and ice below, by reducing the heat transfer between the two media and therefore limiting the sea ice basal growth (Maykut & Untersteiner, 1971; Merkouriadi et al., 2017; Sturm & Massom, 2012). The reduction of the amount of heat transferred from the ocean through the ice and snow to the atmosphere also impacts the atmospheric boundary layer, and in particular...
its thermal stratification (Sterk et al., 2013; Vihma et al., 2014). As the solar radiation increases during spring and summertime, the high albedo of the snowpack reduces the absorption of shortwave radiation at the surface, delaying the onset of sea-ice melt (Granskog et al., 2017). Conversely, the formation of ponds due to the melting of the snowpack can increase the pace at which sea ice is melting via a reduction of the surface albedo, which increases the solar radiation absorption (Fiocco et al., 2010). Melt pond formation can be affected by the spatial variability of snow, which is largely controlled by wind drift redistribution and pressure ridges (Liston et al., 2018; Petrich et al., 2012). Snow-to-ice conversion can significantly affect the ice mass balance for Antarctic sea-ice, and to some extent also in the Arctic (Granskog et al., 2017). Therefore, correctly representing the snow layer over sea-ice has great potential to improve cryosphere-atmosphere coupling in weather forecasting and climate monitoring applications.

Alongside the challenges associated with the representation of the snow physical processes, modeling the snow accumulating over sea-ice introduces additional complexities and computational costs of the snow scheme over sea-ice compared to the land; for example, the prognostic variables need to be advected together with the ice (Liston et al., 2020). Snow observations over sea-ice are limited and mainly based on climatologies (Warren et al., 1999), which may not be representative of the current conditions (Stroeve & Notz, 2018). As a result, the snow models implemented in climate models over sea ice are relatively simple compared to those used to model the snow over land (Hunke et al., 2010). However, more physically complex models might be required to represent the atmospheric boundary layer over wintertime sea-ice. Persson, Bao, and Michelson (2002) investigated how the complexity of snow and ice models impact the near-surface and boundary layer structure in a clear-sky case-study using the Penn State/NCAR Mesoscale Model (MM5) model. Their numerical results showed that multiple layers for representing the snow and sea-ice were required to thermally insulate the atmosphere from the ocean underneath and correctly simulate the near-surface temperature cooling in response to changes in the atmospheric forcing. Batrak et al. (2018) evaluated the impact of a simple thermodynamic sea-ice scheme in the HARMONIE-AROME configuration of the ALADIN-HIRLAM NWP system (Aire Limitée Adaptation dynamique Développement InterNational–High Resolution Limited Area Model, Bengtsson et al., 2017), considering also the effect of the overlying snowpack, compared to using persistent ice conditions throughout the forecast. Their results indicate that a 1-dimensional ice thermodynamic scheme, without considering the snowpack on the ice, improves near-surface temperature forecasts in coastal regions compared to persistence, whilst considering also the insulation effect of the snowpack had a negative impact on scores. Interestingly, the comparison with satellite surface temperature observations in the Central Arctic indicated that the experiment accounting for snow on top of sea-ice reduced the warm bias present in the persistence and in the experiment with the thermodynamic ice model only. Comparing model data with in situ observations from the N-ICE2015 drift north of the Svalbard Archipelago, Batrak and Muller (2019) also found that accounting for the snow on top of sea-ice and a variable ice thickness improves the surface temperature in the HARMONIE-AROME.

In many global weather forecasting applications, the evolution of the marine cryosphere, and in particular the thermodynamic coupling between the ice-snow surface and the atmosphere, has often been simplified due to the computational cost and difficulties in accurately initializing the processes that need to be represented. For instance, the current surface scheme of the ECMWF Integrated Forecasting System (IFS) does not represent snow over sea-ice and its impact on the atmospheric variables. The IFS coupling with the Louvain-la-Neuve sea-ice model (LIM2, Fichefet & Maqueda, 1997) only updates the sea ice fraction. The thermodynamic calculations remain part of the atmospheric surface scheme, implying that the accumulated snow in the sea-ice model does not affect the thermodynamic state of the atmospheric forecast (Keeley & Mogensen, 2018; Mogensen et al., 2012). Therefore, when snow is present in reality, the surface temperature in the atmospheric forecast shows a persistent positive bias, due to the overestimation of the heat flux conducted from the underlying ocean toward the surface (Batrak & Muller, 2019; Persson et al., 2017). This issue is common to other global NWP systems used for medium-range forecasts and climate reanalyses, which are produced with such systems (Batrak & Muller, 2019).

Errors in representing the surface energy fluxes do not only result in errors in surface temperatures, but can also induce errors in the boundary layer characteristics. The wintertime Arctic boundary layer is characterized by two well-defined states (Persson et al., 1999; Stramler et al., 2011). The first one corresponds to clear-sky conditions, during which net longwave radiative heat loss at the surface results in strong temperature inversions; the second one is associated with cloudy conditions during which longwave fluxes at the surface are in balance and the atmospheric boundary layer is near-isothermal (Persson et al., 2017; Persson, Fairall, et al., 2002; Shupe...
et al. (2014) showed that most climate models fail in representing the cloudy state as observed during the SHEBA campaign (Perovich et al., 1999; Stramler et al., 2011). This was explained by the current challenges of representing liquid water in supersaturated clouds formed during moist-air intrusions into the Arctic. The authors also explored in an idealized framework the impact of a lower heat conduction between the snowpack and the ice underneath, showing that this can potentially increase the strength of the simulated temperature inversion in models.

Graham et al. (2017) compared the wintertime Arctic states from the ECMWF reanalysis ERA-Interim (ERA-I, Dee et al., 2011), with observations from the SHEBA campaign, which took place in the Beaufort sea sector of the Arctic basin between 1998 and 1999, and the N-ICE2015 campaign. Their results indicate that the wintertime Arctic states observed during SHEBA and N-ICE2015 have similar characteristics, and that ERA-I simulates a preferred winter state associated with clear skies, however underestimating the mode of the net longwave during both campaigns. Also, they showed that the reanalysis poorly represents the winter state produced by cloudy conditions and the associated surface longwave flux balance. More recently, Graham et al. (2019) analyzed the performances of six reanalyses products compared to N-ICE2015 observations. Their results indicate little improvements between ERA-I and the more recent ERA5 reanalysis (Hersbach et al., 2020), in terms of near-surface weather parameters. Biases in 2-m temperature were found to be even larger in ERA5 compared to ERA-I (Graham et al., 2019). When evaluating the frequency distribution of surface net longwave radiation in the two reanalyses with the observations used by Graham et al. (2019), ERA5 shows similar issues to ERA-I, underestimating the mode of net longwave radiation in the clear-sky state (see Figure A1 in Appendix A).

The coupling between different earth system components (ocean, ice, and atmosphere) in a NWP system can be done in different ways (West et al., 2016). One possibility is to compute the surface energy balance in the sea-ice module and pass the fluxes back to the atmospheric component. This method is used and described by Smith et al. (2018) for the Global Environmental Multiscale atmospheric model (Côté, Desmarais, et al., 1998; Côté, Gravel, et al., 1998), but without assessing in detail the impact on the sea-ice surface temperature. The coupling interface can also be considered within the ice layer via a conductive flux, as used by the UK MetOffice and described by West et al. (2016). When considering these possibilities for the IFS, additional technical elements in the coupling should be considered for the data assimilation systems, which run separately for the atmosphere and ocean. To consider the effect of the snow over sea-ice surfaces on the atmosphere, we have implemented a hybrid method, which consists of a thermodynamic coupling between the surface temperature predicted by the ocean-ice model and the skin temperature of the atmospheric model. This means that the snow parametrizations of the sea-ice model are used to represent the skin temperature variations in the atmospheric model, therefore also constraining the surface energy flux computations. Another approach that we have implemented is the explicit representation in the atmospheric model of the snow over sea-ice surfaces, in a similar way to what it is done for seasonal snow over land. Those approaches allow an improved consistency in the surface fluxes exchanged between the sea-ice/ocean and atmospheric models. The latter approach can be beneficial in a NWP framework in which initial conditions for short and medium-range forecasts are generated using incremental variational assimilation techniques (Rabier et al., 1998). Within this framework, running an ocean model at incremental spatial resolutions introduces technical and scientific challenges. Moreover, the thermodynamic coupling with a sea-ice model may also introduce additional scientific challenges related to coupled data assimilation methodologies (Zhang et al., 2020).

In this paper, we investigate the impact of the representation of snow over sea-ice on the Arctic surface energy fluxes and boundary layer characteristics in the ECMWF IFS. This is an important question for NWP applications, but also for climate reanalyses, such as ERA5 or ERA-I, that are produced with NWP systems. The latter are also used as boundary conditions for uncoupled ocean and sea ice simulations (see for instance Zampieri et al. [2021]). The challenges associated with the way snow heat transfer processes and the ice-snow-atmosphere coupling are represented are evaluated. Single-layer and multi-layer snow schemes of different physical complexities and using different coupling strategies between the sea-ice and atmospheric models are tested. In situ observations from SHEBA and N-ICE2015 field campaigns are used to evaluate at the process-level the impact of different representations of the snow over sea-ice on surface parameters. This includes the surface skin temperature, and typical wintertime relationships between the surface energy fluxes and the vertical structure of the Arctic boundary layer temperature. We also discuss the role of compensating errors between mixed-phase clouds and surface processes in the Arctic boundary layer. Then, we evaluate the impact of representing snow
over sea-ice over the wider Arctic region using satellite surface temperature observations and ECMWF high-resolution analysis data.

In Section 2 the setup of the IFS and the ways in which the effect of the snow over sea-ice is considered in the atmospheric forecasts is presented; Section 3 reports a description of the observations used in the evaluation. Section 4 details the process-level analysis of the simulations using the SHEBA and N-ICE2015 observations. In Section 5 a pan-Arctic assessment of the snow over sea-ice is performed. Conclusions from this work, perspectives and future applications are outlined in Section 6.

2. Model Setup

Global coupled 5-day forecasts with the IFS cycle 45r1 (operational from June 2018 to June 2019) were performed from 1 January to 31 January 1998 and from 21 January 2015 to 21 February 2015, using a TCo1279L137 grid setup (horizontal resolution of 9 km and 137 vertical levels), with the lowest model level being located at about 10 m above the surface. The same grid configuration is currently used in operational (HRES) 10-day forecasts at ECMWF. Forecasts were initialized every day at 00UTC from the ERA5 reanalysis (Hersbach et al., 2020). As the IFS cycle used in this work differs from the one used in ERA5 (i.e., cycle 41r2), forecasts at day 2 and day 5 are analyzed in the following to account for a spin-up of the forecast fields. In addition to that, coupled 5-day forecasts covering the January and February 2015 are performed for a subset of the configurations investigated using a horizontal resolution of about 25 km (TCo399).

The NEMO ocean model and the sea-ice model LIM2, coupled to the IFS, run at a resolution of 1/4°, using a sequential coupling procedure as described in Mogensen et al. (2012). LIM2 is a dynamic-thermodynamic ice model with a single category ice thickness and a 3-layer thermodynamic model. The ice thickness is discretized using two vertical layers and a single layer is used to simulate the snow on top (Fichefet & Maqueda, 1997). The single-layer snow model in LIM2 uses a fixed snow density (300 kg m$^{-3}$) and thermal conductivity (0.22 W m$^{-1}$ K$^{-1}$). The latter is about the snow thermal conductivity of a snowpack with density of 300 kg m$^{-3}$, when determined using usual formulas relating the snow thermal conductivity to snow density (see for instance Calonne et al. [2011]). Also, this quantity, as many others in coupled Earth System Models used seamlessly from medium-range to climate time-scales, requires some tuning to avoid unrealistic model drifts, as for instance reported by Döscher et al. (2022). The snow scheme simulates the snow to ice conversion due to the snow load on the ice.

In the IFS, the sub-grid scale surface heterogeneities are taken into account in the HTESSEL surface model (Balsamo et al., 2011) using a tiling approach. This means that a grid-box is subdivided in fractions (“tiles”), each describing a different surface type, and the surface energy budget is solved for each tile using the coupling approach proposed by Best et al. (2004). Water surfaces are subdivided into land lakes, ocean and sea ice fractions. In the ice module of HTESSEL, the thickness of sea-ice is set constant to a value of 1.5 m and discretized using four layers.

In the current dynamical coupling implemented between LIM2 and IFS, as used in the operational model, only the sea-ice cover fraction evolution determined by LIM2 is coupled to the IFS. The surface energy balance is then solved over the sea-ice tile of HTESSEL, without accounting for the snow accumulation over the ice. This configuration is used as a control simulation and referred to as CTL (see Table 1). Simulations with no snow over sea-ice, as in the operational IFS, and with different representation of the snowpack over sea-ice are performed and compared (see Table 1). Details on the different simulations with the snow over sea-ice are given in the following subsections.

| Exp name | Snow | Ice thickness | Description |
|----------|------|---------------|-------------|
| CTL      | None | 1.5 m         | Control with no snow over sea-ice |
| Tight    | SL   | Dynamic       | Thermodynamic coupling with LIM2 |
| SnowSL   | SL   | 1.5 m         | Snow single-layer over HTESSEL’s sea-ice tile |
| SnowML   | ML   | 1.5 m         | Snow multi-layer over HTESSEL’s sea-ice tile |

- **Table 1** List of Experiments Performed and Analyzed, See Section 2.1 and 2.2 for More Details
2.1. IFS Snow Schemes

The operational snow scheme of the ECMWF IFS, included in HTESSEL, is a single-layer snow scheme (snowSL) computing the evolution in time of snow mass, temperature, density and albedo (Dutra et al., 2010). A detailed description and evaluation over land of snowSL is reported in Dutra et al. (2010).

A multi-layer snow scheme (snowML) was recently added in HTESSEL and coupled to the atmospheric model (Arduini et al., 2019). In snowML, the snowpack is discretized in a variable number of vertical layers, with a maximum number of active snow layers set to five. Additionally, snowML also computes the prognostic evolution of the liquid water within the snowpack. A detailed description and evaluation of snowML and comparison with snowSL over land is reported in Arduini et al. (2019).

In the sensitivity simulations, snow, if present, can also cover the fraction of a grid box that is occupied by the sea-ice in HTESSEL (see Figure 1a). However, the representation of snow over sea-ice surfaces is handled differently to that of snow on land. The main differences concern the temporal evolution of the snow mass, albedo and
roughness throughout the forecast. A correct representation of the radiative (albedo) and mechanical (roughness) properties of snow over sea-ice would require the parametrization of physical processes which are not considered in the current formulation of the HTESSEL snow schemes, for instance the formation of melt ponds. In the absence of such processes, monthly climatological values of sea ice albedo are used (in the same way as in the current operational system), which are derived from climatological monthly values (Ebert & Curry, 1993). For the roughness length, the same empirical formulation currently implemented for the sea-ice is used, which is based on the ice concentration (Bidlot et al., 2014). Snow variables are not advected with the sea-ice within HTESSEL, constraining the calculation of the snow mass balance. The snow mass balance is not computed over the sea-ice, and at each coupling time step the sea-ice cover fraction and snow depth from the LIM2 model are exchanged. The snow depth is then converted to snow mass assuming a constant snow density.

To initialize snow temperatures, different initialization procedures are used for snowSL and snowML. For snowSL, the ERA5 skin temperature is used as a proxy for the snow temperature. For snowML, an initialization procedure adapted from Arduini et al. (2019) is used over sea-ice surfaces. This consists of an algorithm deriving the 5-layer snow fields from the skin and ice top temperature (Arduini et al., 2019). A limitation of this initialization strategy is the propagation of possible ERA5 biases to the forecast initial state.

The approach described enables us to consider the thermodynamic effect of snow over sea-ice at time-scales relevant for NWP applications, without the additional complexity of 3-dimensional snow fields. The treatment of snow over lake-ice was not considered here, and will be investigated in future work.

2.2. Thermodynamic Coupling With the Sea-Ice Model

We evaluate the impact of including the thermodynamic component of the sea-ice model coupling in the IFS forecasts, on top of the existing ice cover coupling (hereafter “Tight”). Tight coupling enforces the skin temperature of the sea-ice model LIM2 within the atmospheric model, enabling to take into account both the effect of a spatially varying ice thickness and of the snowpack in insulating the atmospheric boundary layer from the warm ice and ocean underneath. The skin temperature at the surface of the atmospheric model is held fixed between coupling timesteps and only updated from LIM2 (see Figure 1b). In practice this requires the thermodynamic calculations on the sea ice tile in the IFS surface module to be disabled, and the top ice layer temperature is fixed to the surface temperature computed by LIM2 at every coupling time step, when the ocean and atmosphere models are coupled. The skin temperature of the IFS is then constrained by solving the surface energy balance while setting the thermal conductivity between the skin layer and the top ice layer to a very large number. This procedure fixes the skin temperature in the atmospheric model to that of LIM2. This “tight” thermodynamic coupling between earth system components ensures that the effects of spatially and temporally variable sea-ice thicknesses and the presence of snow on the evolution of the sea-ice/atmosphere interface are accounted for when performing coupled forecasts with IFS.

3. Description of Observations

3.1. SHEBA Campaign

The SHEBA ice station was deployed in the Beaufort Sea, north of Alaska, on 2 October 1997 and ended in the northern Chukchi Sea on 11 October 1998 (Perovich et al., 1999). In this study, we focus on observations collected between 1 January 1998 and 31 January 1998, which was analyzed by Persson et al. (2017) and contains a similar number of days as used for the N-ICE2015 campaign described later. The ice thickness slowly increased during this period, from 2.24 m to 2.40, with a mean value of 2.3 m. The average snow depth during this period was about 0.23 m near the on-ice 18 m meteorological tower, about 300 m from the ship (Perovich et al., 2003). The ice concentration was near 100%, though some leads were present (Overland et al., 2000). Extensive meteorological measurements were made, including all terms of the surface energy budget (Persson, Fairall, et al., 2002). This study uses hourly observations of skin temperature, 2- and 10-m air temperatures from the meteorological tower, and downwelling and upwelling longwave radiation from the tower site. In addition, water vapor, cloud liquid water and ice water paths, derived from a microwave radiometer and a Ka-band radar located on the ship (Intrieri et al., 2002; Shupe et al., 2005; Uttal et al., 2002), are used.
3.2. N-ICE2015 Campaign

The N-ICE2015 campaign was held from 15 January 2015 to June 2015, by different (non-continuous) ice drifts of the Norwegian Research Vessel “R/V Lance.” In this work, observations obtained during the first drift are used, from 15 January to 21 February 2015. This corresponds to a period in which the vessel drifted with an ice floe from an area north of Svalbard southwards, until the sea-ice broke up (Cohen et al., 2017). The average ice thickness during the drift period was 1.2 ± 0.33 m (Rösel et al., 2018) and the average snow depth was about 0.41 ± 0.19 m (Granskog et al., 2017). In this work, hourly observations of surface temperature are used together with hourly averaged observations of 2- and 10-m temperature, collected from a meteorological mast co-located with the radiation and turbulent flux measurement station (Cohen et al., 2017; Walden et al., 2017). Surface temperature observations are derived from the measured upward longwave flux following the methodology reported in Batrak and Muller (2019). Total column water vapor observations are derived from radiosoundings launched twice daily (Cohen et al., 2017).

During N-ICE2015 there were periods for which the sea-ice fraction was not equal to one over the IFS grid box (see Figure A3 for ERA5 and ERA-I). This can lead to mismatch in the comparison between a grid-box average skin temperature and the in situ observations, the latter being representative of the sea-ice fraction only. For this reason, the model skin temperature (representative of the entire grid-box) was rescaled by the sea-ice cover fraction and assuming the ocean temperature at the freezing point of saline water.

3.3. Satellite Surface Temperature Observations

Detailed in situ observations from the two observation campaigns are complemented with satellite observations of surface temperature over ice, which allow to assess the quality of the forecasts for wider spatiotemporal scales. The surface temperature over sea-ice is evaluated using satellite-derived ice surface temperature observations (Dybkjær et al., 2012; Høyer et al., 2014), available from the Copernicus Marine Environmental Monitoring Service (CMEMS). The observations required a pre-processing to remove the cloud-covered pixels. In the original data set, for cloud-covered pixels a spatial interpolation using the neighboring pixels was applied to fill spatial gaps. Therefore, the processed data set consists of observations in clear-sky conditions only. This is an important aspect for the evaluation reported in Section 5. Also, it should be noted that the satellite observations provide information on the daily mean surface temperature, because the data set is formed by a composite of many passages of the satellite within a day over the same area. Therefore the satellite observations are compared to the model’s daily mean skin temperature. The comparison of the model’s data with this data set covers the time period between 1 January 2015 and 28 February 2015, during the N-ICE2015 time period.

4. Impact of Snow Over Sea-Ice During SHEBA and N-ICE2015

4.1. Analysis of the Temporal Variability

During the period of the SHEBA campaign considered in this study, that is, January 1998, relatively clear sky periods are interrupted by warm air intrusions from mid-latitudes, for instance between 8 and 11 January, as described by Persson et al. (2017). The warm air intrusion is characterized by the presence of mixed phase clouds that lead to an increase in the downwelling longwave radiation at the surface (LW↓), rapidly changing the value of the skin temperature (Tsk) by 10–15°C (see Figures 2a and 2b). The CTL forecasts do not capture the large variations in the skin temperature observed during SHEBA (see Figure 2a). During periods with low LW↓, associated with relatively clear-sky periods, CTL generally overestimates the temperature by several degrees. In clear-skies, when LW↓ is well represented by the model, errors in the skin temperature can be driven by errors in the heat conducted from the ice and ocean underneath, as well as in the sensible heat fluxes between the surface and the atmosphere. In particular for very stable conditions, the turbulent transfer coefficient is largely reduced over the sea-ice surface (Grachev et al., 2007). If this is not correctly represented in the model, turbulent heat fluxes can be overestimated and so the skin temperature (Vignon et al., 2017). However, during these periods turbulent sensible heat fluxes are small and well represented by the model, with a mean bias of about −2.5 W m⁻² (not shown). Therefore, the overestimation of Tsk in clear-skies is primarily due to the overestimation of heat conducted through the ice. This results from the absence of the thermal insulation effect of the snow, and to a lesser extent from the misrepresentation of the sea ice thickness (discussed later), in agreement with previous work (Batrak & Muller, 2019; Persson, Bao, & Michelson, 2002). This leads to a mean bias of 4.2°C when only
clear-sky periods are considered (see Table 2). On the other hand, during warm air intrusions $T_{sk}$ is underestimated in the model, linked to an underestimation of $LW_{↓}$. Overall, the model shows a negative bias in $LW_{↓}$ of $-12 \, \text{W m}^{-2}$ (see Table 2). The compensation between positive biases in $T_{sk}$, occurring in clear-skies, and negative biases, occurring in cloudy periods, results in a positive bias of $1.5^{\circ} \text{C}$ in all cloud conditions (see Table 2).

Considering the N-ICE2015 period, CTL largely overestimates the skin temperatures during periods characterized by a weak cloud radiative forcing (see Figure 3), even after the rescaling of $T_{sk}$ by the sea-ice fraction (see Section 3.2). For those periods, the bias of $T_{sk}$, corrected by the sea-ice fraction, is of $8^{\circ} \text{C}$ (see Table 2). As opposed to SHEBA, the $LW_{↓}$ is generally overestimated in clear-skies, whereas it is well simulated during warm air intrusion events, for instance between 2 and 6 February (see Figure 3b), with an overall positive bias in this quantity of $18 \, \text{W m}^{-2}$ (see Table 2). The relatively good simulation of $LW_{↓}$ in warm-air intrusions results in a better simulation of the $T_{sk}$ during such periods in CTL during N-ICE2015, compared to SHEBA. Only a limited number of sensible heat flux observations are available for the time period considered. After 5 February, the sea-ice fraction is often smaller than 1, making a quantitative comparison of turbulent fluxes challenging.

The differences in the simulation of $LW_{↓}$ between N-ICE2015 and SHEBA lead to very different values of the $T_{sk}$ overall bias between the two campaigns, being around $5.9^{\circ} \text{C}$ in the former case and about $1.5^{\circ} \text{C}$ for the latter (see Table 2). The reasons for the different sign of the $LW_{↓}$ between the two campaigns, and the link between clouds and surface temperature errors, will be further discussed in Section 4.3 and 4.4. Similar errors in both $T_{sk}$ and $LW_{↓}$ are found in the ERA5 data, which are used to initialize the forecasts (see Figure A2, Appendix A).

We now consider the role of the snow on ice with the three experiments: Tight; snowSL and snowML (see Table 1). The $T_{sk}$ variability is closer to that of the observations compared to CTL, with a better representation of strong cooling events, for instance during SHEBA on 5 January and during N-ICE2015 between 22 and 23
January 2015 (see Figures 2a and 3a), and more generally during clear-sky periods. Considering only clear-skies, RMSEs are largely reduced in all snow-on-ice experiments, for example, by 60% in snowML (see Table 2). The increase of the thermal insulation due to the snowpack enables a more realistic thermal decoupling between the atmospheric boundary layer and the warm ocean and ice underneath, therefore increasing the sensitivity of the skin temperature to variations in the radiative forcing (see Figure 4). In snowML, a thin top snow layer with a reduced thermal inertia further increases the sensitivity of skin temperature to the radiative forcing, compared to the other snow-on-ice experiments (see Figures 4d and 4h).

This sensitivity can be further quantified by the slope parameter $\alpha$, which is the coefficient of the linear regression between $T_{sk}$ and the atmospheric radiative forcing (measured by $LW_{\downarrow}$), and thus is measuring changes of $T_{sk}$ in response to changes in the atmospheric radiative forcing (for more details see Day et al. [2020]). The observed value of $\alpha$ is 0.164 and 0.186, for SHEBA and N-ICE2015 respectively, indicating differences between the two campaigns. CTL largely underestimates the sensitivity of $T_{sk}$, with values of $\alpha = 0.118$ and $\alpha = 0.141$ for SHEBA and N-ICE2015, respectively (see Figure 4). For both periods, the sensitivity to the forcing increases when snow over sea ice is taken into account, being closer to the observed value for all the snow-on-ice experiments. For SHEBA, values of $\alpha$ are 0.142, 0.16, and 0.176 for Tight, snowSL and snowML, respectively; for N-ICE2015 values of $\alpha$ are 0.169, 0.183, 0.207 for Tight, snowSL and snowML, respectively.

Differences in the response of the skin temperature to variations in the radiative forcing may also be affected by differences in ice thickness and snow depth between the model and observations. During SHEBA, for CTL, snowSL and snowML, the average ice thickness is underestimated by 0.70 m, whereas during N-ICE2015 it is overestimated by 0.3 m; regarding the snow depth, this is underestimated by 0.09 and 0.05 m during SHEBA and N-ICE2015, respectively (not shown). For N-ICE2015, an overestimation of the ice thickness may lead to an overestimation of the thermal decoupling between the surface temperature and the ocean underneath, therefore increasing the value of $\alpha$. However, given the relatively deep snowpack during N-ICE2015, errors in snow depth and ice thickness are likely to play a minor role. For SHEBA the negative biases in ice thickness and snow depth may contribute in reducing the value of $\alpha$, although for the time-scales considered in this work, for thick ice

### Table 2

| Variable | Statistics (°C) | SHEBA | N-ICE2015 |
|----------|----------------|-------|-----------|
| $T_{sk}$ | Bias           | 1.5 (4.2) | 5.9 (8.0) |
|          | RMSE           | 3.8 (4.5) | 7.4 (8.5) |
| $LW_{\downarrow}$ | Bias | −12.0 | 18.2 |
|          | RMSE           | 22.3 | 32.9 |

Note. For N-ICE2015, the corrected skin temperature accounting for the sea-ice fraction of the grid-box is used. Values with a * indicate that differences with respect to the value of CTL are not statistically significant to a 95% confidence level. The values in brackets are statistics computed for clear-sky periods, the latter defined as $LW_{\downarrow} < 160$ W m$^{-2}$ and $\left| LW_{\downarrow}^{\text{obs}} - LW_{\downarrow}^{\text{mod}} \right| < 10$ W m$^{-2}$.
covered with snow the effect of the heat flux from the ocean underneath to the surface might be less than when compared to thinner ice conditions (Maykut, 1978).

For the snow-on-ice experiments, the closer agreement with observations in the sensitivity of $T_{sk}$ indicates a more realistic representation of surface processes. This means that given a correct radiative forcing, $T_{sk}$ is better simulated in the experiments with snow over sea-ice, as also demonstrated by smaller values of the bias and RMSE in clear-skies in Table 2. However, it also means that this variable is more sensitive to errors in the $LW_{\downarrow}$. This is particularly evident during SHEBA, for which the errors in $T_{sk}$ increase for the snow-on-ice experiments during misrepresented warm-air intrusion events, characterized by large errors in the downwelling longwave radiation (see for instance between 8 January and 10 January 1998, Figure 2). Considering all cloud conditions, for SHEBA the RMSE of skin temperature at day 2 is significantly reduced only for Tight. For SnowSL and SnowML, RMSEs are not significantly different with respect to CTL, even though it is increased in SnowML (see Table 2). For N-ICE2015 both the bias and RMSE of skin temperature are largely improved for the snow-on-ice experiments, with the largest improvement for snowML, where the mean error of $T_{sk}$ is reduced by 76% compared to CTL (see Table 2).

Considering the effect of the snow over sea-ice appears thus to improve the skin temperature simulation in clear-skies. It also improves the response of the skin temperature to the radiative forcing, as diagnosed in Figure 4 and by the value of $\alpha$. These results are consistent between SHEBA and N-ICE2015. Errors in the representation of

![Figure 3](https://example.com/figure3.png)
clouds in driving skin temperature errors during warm air intrusions and differences between the two case-studies are analyzed in Section 4.3 and 4.4.

4.2. Impact of Snow on the Representation of Arctic Winter Boundary-Layer States

The previous section highlights how process-based analysis can help the evaluation of different model components irrespective of compensating errors at a specific time or location. To further expand this concept, the ability of the forecasting system to represent the Arctic winter states is evaluated in this section. In the following, the net longwave radiative flux, $LW_n = LW_i + LW_i^\uparrow$ (with $LW_i^\uparrow$ being the upwelling longwave radiative heat flux), $LW_i$ and the temperature inversion strength ($\Delta T$), are used to characterize the Arctic winter states. In this work, $\Delta T$ is defined as the difference between the skin temperature and the temperature measured at 10 m (Persson, Fairall, et al., 2002; Walden et al., 2017).

Both SHEBA and N-ICE2015 datasets display the typical bimodal distribution of $LW_n$ characterizing the Arctic boundary layer (see Figure 5), as already shown in previous studies (Graham et al., 2017; Persson et al., 1999). One state corresponds to a nearly zero $LW_n$ and weak atmospheric stability ($\Delta T \approx 0$), which is usually associated with cloudy conditions; the second state corresponds to negative values of $LW_n$ (net cooling at the surface) and strong near-surface inversions ($\Delta T > 0$), which is usually associated with conditions with a weak cloud radiative forcing, for example, during clear-sky conditions (Persson et al., 1999). For simplicity, the two states will be called “cloudy” and “clear-sky” in the following.

For both campaigns, CTL roughly simulates the two Arctic states, however underestimating the mode of the clear-sky state, implying that the model simulates excessive longwave radiative heat loss at the surface (see Figure 5). The underestimation of the mode of $LW_n$ in the clear-sky state can be directly associated to the overestimation of the $LW_i$ (i.e., the skin temperature), when the $LW_i$ is minimum due to the low cloud radiative forcing, as described in Section 4.

Considering the distributions of the $LW_i$ and $\Delta T$ together enables the analysis of the temperature inversion strength in the two winter states irrespectively to biases in the net longwave radiation (Figure 6). As mentioned

![Figure 4. Scatter plot of skin temperature (y-axis) and downwelling longwave radiation (x-axis), for SHEBA (top row) and N-ICE2015 (bottom row) from observations (black symbols) compared to the coupled forecast experiments with (a and e) bare sea-ice (CTL), (b and f) the tight coupling with LIM2 (Tight), (c and g) the single-layer snow scheme (SnowSL) and (d and h) the multi-layer snow scheme (SnowML).]
Figure 5. Histogram of net longwave heat flux at the surface for (a) SHEBA and (b) N-ICE2015 and from the coupled forecast experiments with bare sea-ice (CTL, red line), the tight coupling with LIM2 (Tight, magenta), the single-layer snow scheme (snowSL, cyan) and the multi-layer snow scheme (snowML, blue) and observations (black line) for SHEBA (top row) and N-ICE2015 (bottom row). The period 1 January 1998 to 31 January 1998 and 21 January 2015 to 20 February 2015 are considered for SHEBA and N-ICE2015, respectively.

Figure 6. Histogram of downwelling longwave heat flux at the surface and inversion strength, defined as the difference between $T_{10m}$ and $T_{sw}$, for SHEBA (top row) and N-ICE2015 (bottom row) and from the coupled forecast experiments with bare sea-ice (second column), the tight coupling with LIM2 (third column), the single-layer snow scheme (fourth column) and the multi-layer snow scheme (fifth column).
previously, the clear-sky state is associated with low values of the $LW_\text{f}$, whereas the cloudy state is associated with large values of the $LW_\text{f}$.

Figure 6 indicates that CTL largely overestimates the clustering of the inversion strength occurrences in the clear sky state around a small range of $\Delta T$ values. The fact that the model's states are clustered around a small range of $\Delta T$ values indicates that the $T_\text{sk}$ hardly responds to variations in the radiative forcing. During SHEBA, for values of $125 < LW_\text{f} < 150$ W m$^{-2}$, the observed $\Delta T$ varies between 1 K to almost 4 K, whereas for CTL $0 < \Delta T < 1$ K; during N-ICE2015, the lack of sensitivity of CTL is even more pronounced, as simulated $\Delta T$ of approximately 1 K can occur across the whole range of $LW_\text{f}$ values.

The experiments with snow over sea-ice improve the representation of the relationship between $LW_\text{f}$ and $\Delta T$, in particular for the clear-sky state (see Figure 6). For this state, the snow on ice experiments have a more realistic “spread” of the distributions, showing a continuous increase of occurrences of strong surface-based temperature inversions as the complexity of the snow scheme increases, from Tight to SnowML, both for SHEBA and N-ICE2015 (see Figure 6). Also, the mode of $LW_\text{f}$ for the clear-sky state decreases from the less to the more complex snow scheme, with snowML having the best agreement with the observations (see Figure 5). As expected, the clear-sky state is largely affected by changes in the thermal insulation properties of the underlying surface and the skin temperature response to the radiative forcing. In this case, the change in the snow scheme largely impacts the representation of the distribution of the $LW_\text{f}$ and $\Delta T$, and of their relationship. On the other hand, for the cloudy state, the snow on ice experiments show minor improvements compared to CTL. This state is both controlled by the cloud radiative forcing, which drives to a large extent the $LW_\text{f}$, and the skin temperature, driving the $LW_\text{f}$. Therefore, a change in the snow scheme can only have a partial effect on this state.

4.3. Cloud and Water Vapor Simulation During SHEBA and N-ICE2015

The differences of model's performance, as measured by the RMSE and bias in Table 2 for the SHEBA and N-ICE2015 experiment are mainly related to the difference in the representation of the $LW_\text{f}$. The main drivers of the $LW_\text{f}$ variability in the Arctic are mixed-phase clouds, which are analyzed in more detail in this section. During SHEBA, all experiments underestimate the frequency of occurrences of the peak at 220 W m$^{-2}$, whilst overestimating the occurrences of the peak corresponding to a clear-sky radiative forcing, at about 130 W m$^{-2}$ (see Figure 7a). During N-ICE2015, all experiments underestimate the occurrences of $LW_\text{f} < 140$ W m$^{-2}$, corresponding to clear-skins, while overestimating the frequency of cases with $LW_\text{f} > 140$ W m$^{-2}$ (see Figure 7b). This can be due to a general overestimation of water vapor in the model in clear-sky periods at the N-ICE2015 location, or an overestimation of cloudiness and their radiative properties.

The link between $LW_\text{f}$ and water phase simulation by the model can be investigated considering the frequency distributions of liquid, ice, and vapor water paths. During SHEBA the occurrence of liquid water clouds $<10$ g m$^{-2}$ is overestimated by the model, whereas it is underestimated for values $>10$ g m$^{-2}$, compared to observations (see Figure 8). Liquid water clouds are more radiatively opaque than ice clouds, so having a greater impact on the downwelling longwave radiative flux compared to ice clouds (Persson et al., 2017; Shupe & Intrieri, 2004). The total water vapor path is also skewed toward lower values, indicating a general dryness of the atmosphere compared to observations (see Figure 8). The underestimation of liquid and ice water clouds, as well as of water vapor, explains the underestimation of $LW_\text{f}$ during SHEBA (see Figure 2), in agreement with previous studies comparing ERA-I with SHEBA data (Graham et al., 2017).

As opposed to SHEBA, during N-ICE2015 the water vapor path distribution in the model agrees well with the observed one (see Figure 8f), although there are differences in the relative frequency of each bin in particular for values smaller than 3 kg m$^{-2}$. Due to the lack of direct observations of cloud water and ice phases during N-ICE2015, the model's distribution during this campaign is compared with the one during SHEBA. The comparison of the distributions indicates that during N-ICE2015 the occurrences of both ice and liquid cloud water in the model is larger than during SHEBA (see Figure 8). The overestimation of the frequency of cases with $LW_\text{f} > 140$ W m$^{-2}$ during N-ICE2015 suggests that also the occurrence of liquid and ice water is overestimated. This hypothesis can be partially evaluated by stratifying the histogram of $LW_\text{f}$ for cases with low values of cloud liquid and ice water path. Computing the distribution of $LW_\text{f}$ considering only periods with values of cloud liquid water $<1$ g m$^{-2}$, brings the mode of $LW_\text{f}$ to a value of about 130 W m$^{-2}$, in closer agreement with the observations.
However, the lack of observations of cloud phase to compare with does not allow us to verify if the model overestimates either the amount of cloud liquid and ice water or the radiative properties of such clouds.

4.4. Discussion

When all cloud conditions are considered, snowML shows the best and worst agreement with the observations, respectively during N-ICE2015 and SHEBA. This results from a compensation of errors occurring in different processes at those two different locations. As discussed earlier, the mean LW↓ bias is negative and positive during SHEBA and N-ICE2015, respectively (see Table 2). Assuming that the errors in Tsk are only due to LW↓, a negative bias in this variable should be associated with a negative bias in Tsk, and vice versa. Table 2 shows that for the experiments with snow over sea ice, Tsk and LW↓ always have the same sign of the bias, whereas CTL shows a positive bias during SHEBA associated with a negative bias in LW↓. This indicates that during SHEBA the large errors in CTL due to the absence of the snow (which leads to an overestimation of Tsk), outweigh the errors in the cloud phase (which leads to an underestimation of Tsk), leading to a small positive bias in Tsk. In this case, when looking at overall statistics snowML is the most penalized scheme, being the most sensitive to the atmospheric radiative forcing and its errors (see Figure 4). This compensating mechanism is also demonstrated by considering the statistics for clear-sky periods only, showing an increase of both bias and RMSE for CTL, whereas for SnowML errors are reduced and lower than in CTL. During N-ICE2015, the positive LW↓ bias in the CTL simulation acts in the same direction as the errors that are due to the absence of snow, therefore increasing the positive bias in Tsk. The snow on ice experiments, accounting for the thermal insulation effect of the snow, reduce the skin temperature biases. It is interesting to note that snowML overestimates the sensitivity of Tsk to the radiative forcing compared to the observed one (see Figure 4). This may be due to an overestimation of the thermal decoupling between the snow surface and the ice and ocean underneath (as already shown over land by Day et al. [2020]), and, for N-ICE2015, to the overestimation of ice thickness. For N-ICE2015, those errors can partly compensate for the positive bias in LW↓, therefore leading to an overall best performance of the snowML experiment, in terms of bias and RMSE.
We can only speculate on the causes of the differences of simulated liquid water clouds between SHEBA and N-ICE2015. A possibility is a systematic difference in the model’s climate between the SHEBA and N-ICE2015 ships’ locations, that is between the North Atlantic Arctic sector and the Beaufort Sea. Spatial maps of the mean value of the simulated liquid water path for CTL at day 2 indicate that the model simulates larger liquid water paths in the Atlantic sector compared to the central and Beaufort sea regions, during both observational campaign periods (see Figure 9). This can be due to challenges for the model in maintaining supercooled liquid water in clouds traveling far from the initial moisture source, as those observed at the SHEBA location in January 1998 (Persson et al., 2017). However, the evaluation of cloud microphysics processes in the IFS across the Arctic region will require additional in situ observations and should be evaluated in future work.

5. Impact on the Wider Arctic Region

In this section, the simulations performed at a lower horizontal resolution and covering the winter period of 2015 are analyzed (see Section 2) to assess over the wider Arctic region the impact of considering the snow over sea-ice both at the surface and in the lower troposphere.

In the CTL experiment the daily mean skin temperature has a positive bias of several degrees over the Arctic region, compared to the satellite observations (see Figure 10). Such errors are consistent with what was reported in previous studies (Batrak & Muller, 2019) and with the analysis in the previous sections. The magnitude of the bias is larger in the Central Arctic and near Greenland, whereas it is lower in the Siberian and Alaskan side. The bias is largely reduced in the snow on ice experiments, with areas of near-zero bias for snowML in the East Siberian and Beaufort sea regions both at day 2 and day 5 (see Figure 10 and Table 3 for a quantitative comparison). It should be pointed out here that compensating errors among different processes (as discussed in the previous sections), can still play a role, affecting the $T_{sk}$ biases of the different simulations.

The spatial pattern of the biases is consistent, to some extent, with a larger snow accumulation in the central Arctic and near Greenland, compared to the Siberian side, as simulated in LIM2 (not shown). However, the spatial variability of snow depth alone does not fully explain the spatial variability of the biases of CTL, in particular on the
Eastern side of the Arctic basin. It should be noted that the snow depth accumulating in LIM2 can have biases due to errors in the precipitation amount in the atmospheric model, as well as in the LIM2 snow parametrizations. Another factor which can influence the surface temperature variability is the ice thickness, which is fixed to 1.5 m in the current HTESSEL sea-ice module. Such thickness implies a negligible heat flux from the warmer ocean underneath

Figure 9. Spatial map of the mean total cloud liquid water path (TCLW, g m \(^{-2}\)) for the SHEBA (left) and N-ICE2015 (right) time periods for the experiment with bare sea-ice (CTL) at day 2 of the forecast. The red and magenta dots correspond to the SHEBA and N-ICE2015 drift positions, respectively, during the considered time period of each field campaign.

Figure 10. Bias of simulated daily average skin temperature with respect to the satellite surface temperature observations of the coupled forecast with bare sea-ice (CTL), the tight coupling with LIM2 (Tight), the single-layer (snowSL) and multi-layer (snowML) snow schemes at a lead time of 2 days (top row) and 5 days (bottom row) for the period 1 January 2015 to 28 February 2015.
With LIM2 (Tight), the Single-Layer (SnowSL) and the Multi-Layer Coupled Forecast Experiments With Bare Sea-Ice (CTL), the Tight Coupling Averaged Over the Sea-Ice Covered Region Northward 60°N, for the Satellite Surface Temperature Observations at Different Lead Times, Bias and RMSE of Simulated Daily Skin Temperature With Respect to the ECMWF operational analysis. The fact that the value of the bias around the location of the SHEBA observations differs from that reported in the analysis of Section 4 can be due to differences in the model climates between the two time periods. Also, the satellite observations are representative of clear-sky only periods, whereas the model daily mean skin temperature is relevant for all conditions, possibly amplifying the positive biases. However, another possibility is related to the uncertainties of the satellite observations. A crude estimate of the uncertainties in the satellite observations can be assessed by comparing the satellite skin temperature retrieved at the N-ICE2015 pixel with the in situ observation, as shown by the black stars in Figure 3. For this comparison, the satellite data were rescaled by the sea-ice cover fraction at the N-ICE2015 pixel, similarly to what was done for the model data (see Section 3.2). At the same time, it is worth noticing that in situ observations are representative of a single and very localized ice type (likely level ice) and do not account for the heterogeneity of the sea ice surface, which is captured by remote sensing products. The satellite observations hardly capture the variability of the in situ observations. The RMSE of the satellite observations compared to the in situ, the latter considered as the truth, is about 4.7°C, the standard deviation of the error is 6.3°C and the bias is 0.9°C. This suggests that differences of the order of a few Kelvin between the model skin temperature and the satellite observations should be interpreted with caution.

The previous analysis indicates that the insulation effect of snow over sea-ice largely affects the surface temperature representation over the Arctic region. Therefore, one might expect an impact also in the upper-air geopotential field. As an illustration of the sensitivity of the upper-air fields to changes at the surface, Figure 11 shows biases of geopotential height at 850 hPa for CTL, snowML (in which the ML snow scheme is used both over land and over sea-ice) and snowML-Land (in which the ML snow scheme is used only over land, with no account for snow over the sea-ice), with respect to the ECMWF operational analysis, at a lead time of 48 and 120 hr. Representing snow over sea-ice leads to a general reduction of the geopotential height in the central Arctic region, compared to CTL and snowML-Land, due to the increased cooling in the lower part of the atmospheric boundary layer. Both CTL and snowML-Land have a positive bias in this region, therefore indicating a positive impact of snowML on the geopotential height. However, the evaluation against the ECMWF operational analysis should be taken with caution over the central Arctic region, because of the caveats with assimilating observations in polar regions (Lawrence, Bormann, et al., 2019).

### Table 3

| Lead time | Statistics (°C) | CTL | Tight | SnowSL | SnowML |
|-----------|----------------|-----|-------|--------|--------|
| Day 2     | Bias           | 7.8 | 6.1   | 4.6    | 3.2    |
|           | RMSE           | 9.0 | 8.1   | 7.4    | 7.4    |
| Day 5     | Bias           | 7.3 | 5.6   | 4.5    | 3.1    |
|           | RMSE           | 8.7 | 7.9   | 7.5    | 7.5    |

6. Conclusions

In this paper, we have investigated the impact of the representation of snow over sea-ice on the Arctic surface energy fluxes and boundary layer characteristics in global coupled forecasts performed with the ECMWF IFS model. Two coupling methodologies were compared: a thermodynamic coupling between the sea-ice model LIM2 and the IFS, bypassing the surface temperature computation in the land-surface model, and an explicit representation of snow over the sea-ice tile of the IFS. In the latter configuration, a simpler single-layer snow scheme (comparable to the one used in LIM2) and a more sophisticated multi-layer snow scheme were tested. The key conclusions from the analysis are outlined below:

1. Accounting for the snow over sea-ice enables strong cooling events and extreme low temperatures occurring in the Arctic to be better represented. The presence of snow increases the thermal insulation from the warm ocean and ice underneath, therefore reducing the heat conducted through the sea-ice and snow to the atmosphere. A multi-layer snow scheme enables a faster response of the snow temperature to variations in the atmospheric forcing, due to the reduced thermal inertia of a thin top snow layer. This improves the skin
2. The current IFS does not represent strong surface-based inversions and their relationship to the downwelling longwave radiative flux at the surface over the Arctic sea-ice. This is because of the issues in simulating strong cooling events due to the overestimation of the heat conducted through the ice. Considering the effect of snow over sea-ice improves the representation of surface-based inversions and their relationship to the downwelling surface longwave radiation, improving the representation of the two main Arctic winter boundary-layer states. The improvements are most evident when using a more complex multi-layer snow scheme.

3. Compensating errors between cloud and surface processes in a coupled model are a key factor affecting the near-surface forecast biases in the Arctic boundary layer. The comparison with the N-ICE2015 data indicate that the model overestimates the downwelling longwave radiation in clear-skies, which may be due to an overestimation of both liquid and ice clouds in this region. Systematic errors in radiative fluxes may increase errors in the simulated skin temperature when a more realistic representation of the surface is used. For the SHEBA period considered, this was shown to be due to the misrepresentation of mixed-phase clouds.

4. The comparison over a 2-month winter period with satellite data over the Arctic region indicates that considering snow over sea-ice reduces the wintertime bias of skin temperature by several degrees, compared to the experiment with no snow over sea-ice. Explicitly representing the snowpack in the atmospheric model shows a smaller bias compared with the observations, even when using a simpler single-layer snow scheme, at the time-scales investigated. The comparison of satellite data with in situ observations indicates however that errors in the satellite data can be of several degrees in situations with high synoptic variability. Therefore the interpretation of small model biases (of the order of 1 K) compared to these observations should be done with caution.
The results presented in this work highlight the need for future forecasting systems to account for snow over sea-ice. This is not only important for short and medium-range forecasts but also for seasonal and climate time scales, as the change in upwelling longwave radiation can modify the heat conducted through the ice layer and therefore modify the sea-ice growth. These aspects will be investigated in future work in combination with the complex sequence of processes affecting wintertime sea-ice growth (Persson et al., 2017).

The results indicate that even simpler single-layer snow schemes can enable the first order effect of the presence of the snow over the ice layer to be captured. Indeed, another important aspect highlighted in this work regards the importance of cloud phase errors in driving surface temperature responses. A multi-layer snow scheme is by construction more sensitive to such errors and therefore penalized when evaluating atmospheric variables in deterministic forecasts, compared to simpler snow schemes. However, the larger sensitivity in the near-surface temperature can be an added value in an ensemble framework, as already shown over land surfaces by Day et al. (2020), if the ensemble distribution correctly samples the uncertainties in the model parametrizations and initial conditions. An increase in the forecast spread can be particularly beneficial for the Arctic and Antarctic regions, which show a large lack of spread in the ECMWF system (Lawrence, Farnan, et al., 2019).

The different conclusions that one can draw from the evaluation based on the SHEBA and N-ICE2015 campaigns respectively, also highlight the need of using multiple observational datasets, covering different temporal, spatial, and meteorological conditions, to evaluate model developments in complex regions like the Arctic. As the physical complexity of the components of coupled models increase, it is fundamental the availability of co-located observations, covering long time periods, enabling the evaluation of the physical processes controlling the evolution of the Arctic surface and boundary layer in a holistic way. The year-long comprehensive MOSAiC data set (Shupe et al., 2022) will likely prove useful in future process evaluations of coupled models in the Arctic environment.

The current work mainly focuses on the impact of the thermal insulation effect of the snow on near-surface variables over sea-ice. Future work is needed for a more thorough evaluation of turbulent sensible heat fluxes and its coupling to the surface. This should also consider the impact of turbulent mixing processes in the Arctic boundary layer temperature structure.

An important aspect for NWP systems is the initialization of the forecasts. The simulations analyzed in this work were initialized using ERA5 fields, but an operational implementation will require a consistent initialization of the snow fields over sea-ice in a data assimilation framework, like the 4D-VAR assimilation system used in the ECMWF IFS. Future work should explore the merits of both the thermodynamic coupling (Tight) and the explicit representation of snow over sea-ice in the atmospheric model (snowSL and snowML) as part of a data assimilation system. This would enable the assessment of the added value of a consistent initialization of the snow fields for forecast skills in the Arctic and beyond.

Snow observations widely covering the Arctic region both spatially and temporally are required to constrain the forecast initial conditions. The results presented in this study indicate the large sensitivity of the surface temperature to the presence or absence of the snowpack. Model biases in the snowpack cover can therefore lead to large errors in the surface-atmosphere coupling, in particular during transition seasons. Future satellite missions like the Copernicus Imaging Microwave Radiometer (CIMR, Braakmann-Folgmann & Donlon, 2019), measuring the snow cover and depth over the Arctic sea-ice, can potentially help in bridge this gap and constrain the snowpack over sea-ice for weather predictions.

Appendix A: Analysis of ERA-Interim and ERA5 During SHEBA and N-ICE2015

This section contains an evaluation of ERA5 and ERA-I data using the observations from SHEBA and N-ICE2015. The histogram of the LW and the time-series for ERA-I and ERA5 show similar characteristics to the CTL simulation discussed in Section 4. Both reanalyses underestimate the mode of the clear-sky state by more than 15 W m$^{-2}$ during SHEBA and N-ICE2015 and there are little improvements between ERA-I and ERA5 in the simulation of the Arctic winter states (Figure A1). Both reanalyses do not capture the large variability of $T_s$ during the two field campaigns (see Figures A2 and A3), as in the CTL simulation, and hardly simulate the strong cooling events and the extreme low temperatures occurring during SHEBA between 13 January 1998 and 25 January 1998. In Figure A3b, the OSI-SAF sea-ice concentration is displayed as reference.
Figure A1. Histogram of net longwave heat flux at the surface (W m\(^{-2}\)), for ERA-Interim (red), ERA5 (blue) and observations (black) for SHEBA (top panel) and N-ICE2015 (bottom panel). The period 1 January 1998 to 31 January 1998 and 21 January 2015 to 20 February 2015 are considered for SHEBA and N-ICE2015, respectively.

Figure A2. Concatenated day 1 (t + 1 to t + 24) forecasts for the period 2 January 1998 to 30 January 1998 of (a) skin temperature (\(T_{sk}\) °C), (b) downwelling longwave radiation (\(LW_{\downarrow}\) W m\(^{-2}\)) from SHEBA observations (black), ERA-Interim (red) and ERA5 (blue).
Figure A3. Concatenated day 1 ($t + 1$ to $t + 24$) forecasts for the period 21 February 2015 to 20 February 2015 of (a) skin temperature ($T_{sk}$, °C), (b) sea-ice cover (c) downwelling longwave radiation ($LW_\downarrow$, W m$^{-2}$) from N-ICE2015 in situ observations (black), satellite DMI observations (black symbol), ERA-Interim (red) and ERA5 (blue). For the coupled experiments, a modified skin temperature based on the sea-ice cover fraction is displayed with dashed lines (see Section 3.2 for more details). In (b), OSI-SAF sea-ice concentration (red dot) is displayed as reference.

Data Availability Statement

The SHEBA observations are available at https://data.eol.ucar.edu/dataset/list?project=73. The N-ICE2015 observations are available at https://www.npolar.no/prosjekter/n-ice2015/. The satellite-derived ice surface temperature observations are available from CMEMS at https://resources.marine.copernicus.eu/product-detail/SEAICE_ARC_PHY_CLIMATE_L4_MY_011_016/INFORMATION. The IFS code used to produce the simulations analysed in this work is not freely available. The model data used in this work are available on zenodo.org, https://doi.org/10.5281/zenodo.6576248.

Acknowledgments

The work described in this paper has received funding from the European Union's Horizon 2020 Research Innovation programme through grant agreement No. 727862 APPLICATE. The content of the paper is the sole responsibility of the author(s) and it does not represent the opinion of the European Commission, and the Commission is not responsible for any use that might be made of information contained. We thank the two anonymous reviewers for their comments and suggestions which greatly improved the manuscript.

References

Arduini, G., Balsamo, G., Dutra, E., Day, J. J., Sandu, I., Boussetta, S., & Haiden, T. (2019). Impact of a multi-layer snow scheme on near-surface weather forecasts. Journal of Advances in Modeling Earth Systems, 11(12), 4687–4710. https://doi.org/10.1029/2019ms001725

Balsamo, G., Pappenberger, F., Dutra, E., Viterbo, P., & Van den Hurk, B. (2011). A revised land hydrology in the ECMWF model: A step towards daily water flux prediction in a fully-closed water cycle. Hydrological Processes, 25(7), 1046–1054. https://doi.org/10.1002/hyp.7808

Batrak, Y., Kourzeneva, E., & Homleid, M. (2018). Implementation of a simple thermodynamic sea ice scheme, SICE version 1.0-38h1, within the ALADIN–HIRLAM numerical weather prediction system version 38h1. Geoscientific Model Development, 11(8), 3347–3368. https://doi.org/10.5194/gmd-11-3347-2018

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

Bengtsson, L., Andrae, U., Aspelien, T., Batrak, Y., Calvo, J., de Rooy, W., et al. (2017). The HARMONIE–AROME model configuration in the ALADIN–HIRLAM NWP system. Monthly Weather Review, 145(5), 1919–1935. https://doi.org/10.1175/mwr-d-16-0147.1

Best, M., Beljaars, A., Polcher, J., & Viterbo, P. (2004). A proposed structure for coupling tiled surfaces with the planetary boundary layer. Journal of Hydrometeorology, 5(6), 1271–1278. https://doi.org/10.1175/jhm-382.1

Bidlot, J., Keeley, S., & Mogensen, K. (2014). Towards the inclusion of sea ice attenuation in an operational wave model. In Proceedings of the 22nd EAHIR International Symposium on Ice. Braakmann-Folgmann, A., & Donlon, C. (2019). Estimating snow depth on Arctic sea ice using satellite microwave radiometry and a neural network. The Cryosphere, 13(9), 2421–2438. https://doi.org/10.5194/tc-13-2421-2019

Calonne, N., Flin, F., Morin, S., Lesaffre, B., du Roscoat, S. R., & Geindreau, C. (2011). Numerical and experimental investigations of the effective thermal conductivity of snow. Geophysical Research Letters, 38(23), L23501. https://doi.org/10.1029/2011gl049234
Persson, P. O. G., Fairall, C., Andreas, E., Guest, P., & Perovich, D. (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 21–1–SHE 21–35. https://doi.org/10.1029/2000jc000705

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

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, 5th Conference on Polar Meteorology and Oceanography* (pp. 10–15).

Petrich, C., Eicken, H., Polashenski, C. M., Sturm, M., Harbeck, J. P., Perovich, D. K., & Finnegan, D. C. (2012). Snow dunes: A controlling factor of melt pond distribution on Arctic sea ice. *Journal of Geophysical Research, 117*(C9), C09029. https://doi.org/10.1029/2012jc008192

Pithan, F., Meidereos, 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

Rabier, F., Thépaut, J.-N., & Courtier, P. (1998). Extended assimilation and forecast experiments with a four-dimensional variational assimilation system. *Quarterly Journal of the Royal Meteorological Society, 124*(550), 1861–1887. https://doi.org/10.1002/qj.4971245005

Ricker, R., Hendricks, S., Helm, V., Skourup, H., & Davidson, M. (2014). Sensitivity of Cryosat-2 Arctic sea-ice freeboard and thickness on radar-waveform interpretation. *The Cryosphere, 8*(4), 1607–1622. https://doi.org/10.5194/tc-8-1607-2014

Rösel, A., Ikin, P., King, J., Divine, D., Wang, C., Granskog, M. A., et al. (2018). Thin sea ice, thick snow, and widespread negative freeboard observed during N-ICE2015 north of Svalbard. *Journal of Geophysical Research: Oceans, 123*(2), 1156–1176. https://doi.org/10.1002/2017jc012965

Shupe, M. D., & Intrieri, J. M. (2004). Cloud radiative forcing of the Arctic surface: The influence of cloud properties, surface albedo, and solar zenith angle. *Journal of Climate, 17*(3), 616–628. https://doi.org/10.1175/1520-0442(2004)017<0616:crfota>2.0.co;2

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

Shupe, M. D., Uttal, T., & Matrosov, S. Y. (2005). Arctic cloud microphysics retrievals from surface-based remote sensors at SHEBA. *Journal of Applied Meteorology, 44*(10), 1544–1562. https://doi.org/10.1175/jam2297.1

Smith, G. C., Bélanger, J.-M., Roy, F., Pellerin, P., Ritchie, H., Onu, K., et al. (2018). Impact of coupling with an ice–ocean model on global medium-range NWP forecast skill. *Monthly Weather Review, 146*(4), 1157–1180. https://doi.org/10.1175/mwr-d-17-0157.1

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

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

Stroeve, J., & Notz, D. (2018). Changing state of Arctic sea ice across all seasons. *Environmental Research Letters, 13*(10), 103001. https://doi.org/10.1088/1748-9326/aade56

Sturm, M., & Massom, R. (2012). Snow on sea-ice. In D. Thomas & G. Dieckmann (Eds.), *Sea ice* (2nd ed.). Wiley-Blackwell Publishing.

Thackeray, C. W., Derksen, C., Fletcher, C. G., & Hall, A. (2019). Snow and climate: Feedbacks, drivers, and indices of change. *Current Climate Change Reports, 5*(4), 322–333. https://doi.org/10.1007/s40641-019-00143-w

Uttal, T., Curry, J. A., McPhee, M. G., Perovich, D. K., Moritz, R. E., Maslanka, J. A., et al. (2002). Surface heat budget of the Arctic Ocean. *Bulletin of the American Meteorological Society, 83*(2), 255–276. https://doi.org/10.1175/1520-0477(2002)083<0255:sbsbota>2.3.co;2

Vignon, E., Hourdin, F., Genthon, C., Gallée, H., Bazile, E., Lefebvre, M.-P., et al. (2017). Antarctic boundary layer parameterization in a general circulation model: 1-D simulations facing summer observations at Dome C. *Journal of Geophysical Research: Atmospheres, 122*(13), 6818–6843. https://doi.org/10.1002/2017jd026802

Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjemström, M., et al. (2014). Advances in understanding and parameterization of small-scale physical processes in the marine Arctic climate system: A review. *Atmospheric Chemistry and Physics, 14*(17), 9403–9450. https://doi.org/10.5194/acp-14-9403-2014

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

Warren, S. G., Rigor, I. G., Untersteiner, N., Radionov, V. F., Brazzygin, N. N., Aleksandrov, Y. I., & Colony, R. (1999). Snow depth on Arctic sea ice. *Journal of Climate, 12*(6), 1814–1829. https://doi.org/10.1175/1520-0442(1999)012<1814:sdoa>2.0.co;2

West, A. E., McLaren, A. J., Hewitt, H. T., & Best, M. J. (2016). The location of the thermodynamic atmosphere–ice interface in fully coupled models—a case study using JULES and CICE. *Geoscientific Model Development, 9*(3), 1125–1141. https://doi.org/10.5194/gmd-9-1125-2016

Zampieri, L., Kauker, F., Fröhle, J., Sumata, H., Hunke, E. C., & Goessling, H. F. (2021). Impact of sea-ice model complexity on the performance of an unstructured-mesh sea-ice/ocean model under different atmospheric forcings. *Journal of Advances in Modeling Earth Systems, 13*(5), e2020MS002438. https://doi.org/10.1029/2020ms002438

Zhang, S., Liu, Z., Zhang, X., Wu, X., Han, G., Zhao, Y., et al. (2020). Coupled data assimilation and parameter estimation in coupled ocean–atmosphere models: A review. *Climate Dynamics, 54*(11–12), 5127–5144. https://doi.org/10.1007/s00382-020-05275-6