



RESEARCH ARTICLE

10.1029/2025MS004945

Rapid Summertime Sea Ice Melt in a Coupled Numerical Weather Prediction System

Key Points:

- Sea ice melts too fast in 5-day coupled forecasts from August 2022, verified using in situ aircraft observations and satellite-based products
- The assimilated sea-ice fraction product is biased low, leading to accelerated sea ice melt through the ice-albedo feedback mechanism
- The passage of an Arctic cyclone provides a short-term dynamical forcing of the sea ice that is reasonably well represented in the forecasts

Correspondence to:

C. Barrell,
c.barrell@uea.ac.uk

Citation:

Barrell, C., Renfrew, I. A., Bennett, M., Elvidge, A. D., Weiss, A., Methven, J., et al. (2025). Rapid summertime sea ice melt in a coupled numerical weather prediction system. *Journal of Advances in Modeling Earth Systems*, 17, e2025MS004945. <https://doi.org/10.1029/2025MS004945>

Received 7 JAN 2025
Accepted 28 AUG 2025

Author Contributions:

Conceptualization: Christopher Barrell, Ian A. Renfrew, Andrew D. Elvidge, Alexandra Weiss, John Methven
Data curation: Christopher Barrell, Miriam Bennett, Alexandra Weiss, John Methven, Ed Blockley, Dan Copsey, Tim Graham, Thomas Krumpen
Formal analysis: Christopher Barrell, Miriam Bennett, Thomas Krumpen
Funding acquisition: Ian A. Renfrew, John Methven, Thomas Krumpen
Investigation: Christopher Barrell, Ian A. Renfrew, Miriam Bennett, Andrew D. Elvidge, Alexandra Weiss, John Methven, Thomas Krumpen
Methodology: Christopher Barrell, Ian A. Renfrew, Andrew D. Elvidge,

Christopher Barrell¹ , Ian A. Renfrew¹ , Miriam Bennett¹ , Andrew D. Elvidge¹ , Alexandra Weiss², John Methven³ , Ed Blockley⁴ , Dan Copsey⁴ , Tim Graham⁴ , and Thomas Krumpen⁵ 

¹School of Environmental Sciences, University of East Anglia, Norwich, UK, ²British Antarctic Survey, Cambridge, UK, ³Department of Meteorology, University of Reading, Reading, UK, ⁴Met Office, Exeter, UK, ⁵Helmholtz Centre for Polar and Marine Research, Alfred Wegener Institute, Bremerhaven, Germany

Abstract Coupled Numerical Weather Prediction (NWP) models have only recently been implemented for short-term environmental prediction and both challenges and benefits are evident in polar regions. Their simulation of surface exchange over sea ice depends on the model's sea-ice characteristics, however these are hard to constrain due to a lack of in situ and accurate remotely sensed observations. We focus on the Fram Strait region during peak melt conditions and during the passage of an Arctic cyclone: very challenging conditions for coupled NWP. We use in situ aircraft observations from the *Arctic Summertime Cyclones* field campaign in July–August 2022, plus satellite products, to evaluate a set of 5-day forecasts from the Met Office Unified Model. Our model set ups are based on operational GC4 (Global Coupled 4) and developmental GC5 (Global Coupled 5) configurations, which use the CICE5.1 and SI³ sea-ice models respectively. We find a combination of deficiencies in the simulated sea-ice field, due to initialization and modeling problems. An initially low concentration of sea ice results in excessive absorption of shortwave radiation by the ocean, leading to excessive basal melting of the sea ice, and further sea-ice loss; leading to relatively poorly simulated sea-ice fields in general. In contrast, the passage of an Arctic cyclone and its impact on sea-ice velocities are captured well. Although we demonstrate several deficiencies in the short-term forecasts of two state-of-the-art coupled NWP models, we also find promising aspects of model performance and some clear benefits from a fully coupled atmosphere-ice-ocean system.

Plain Language Summary Weather prediction in the Arctic requires an accurate representation of sea ice as it plays a key role in the exchange of momentum, heat and moisture between the surface and the atmosphere. We investigate a challenging set of conditions for weather forecasting: the passage of an Arctic cyclone over Fram Strait in the European Arctic during peak summertime sea ice melting. We use observations made during the *Arctic Summertime Cyclones* field campaign in July–August 2022 to evaluate forecasts from the Met Office Unified Model that feature ocean and sea ice model components that interact (are “coupled”) with the atmosphere. We find discrepancies in the simulated sea ice field that result from issues in the satellite observations fed into the models and model biases. A lack of sea ice results in increased heat absorption by the ocean, after which the warmer water melts the sea ice faster, forming a feedback loop known as the ice-albedo feedback. The passing cyclone also drives changes in the sea ice cover, but we find that these effects are generally simulated well in the forecasts. Overall, despite the issues discussed, we find that using such coupled models are advancing weather prediction in the Arctic.

1. Introduction

The Arctic marine environment is inherently a coupled system, with the atmosphere, ocean and sea ice all interacting and significantly affecting development over a range of timescales—from hours to decades. To predict these interactions, in particular the feedbacks between weather systems and the ocean/sea-ice surface, requires a coupled environmental prediction system. Coupled Numerical Weather Prediction (NWP) systems have started to be used as operational forecast systems for short-to-medium term predictions in the last few years. Here we present one of the first studies evaluating performance in the Arctic, making use of new aircraft-based observations over sea ice in Fram Strait and northeast of Greenland from our *Arctic Summertime Cyclones* field campaign in July–August 2022.

© 2025 The Author(s). Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

Ed Blockley, Dan Copsey, Tim Graham,
Thomas Krumpen

Project administration: Ian A. Renfrew,
Andrew D. Elvidge, Alexandra Weiss,
John Methven, Thomas Krumpen

Resources: Ian A. Renfrew,
Alexandra Weiss, John Methven,
Ed Blockley, Dan Copsey, Tim Graham,
Thomas Krumpen

Software: Christopher Barrell,
Miriam Bennett, Andrew D. Elvidge,
Ed Blockley, Dan Copsey, Tim Graham,
Thomas Krumpen

Supervision: Ian A. Renfrew, Andrew
D. Elvidge

Validation: Ian A. Renfrew, Ed Blockley,
Dan Copsey, Tim Graham,
Thomas Krumpen

Visualization: Christopher Barrell

Writing – original draft:
Christopher Barrell

Writing – review & editing:
Christopher Barrell, Ian A. Renfrew,
Andrew D. Elvidge, Alexandra Weiss,
John Methven, Ed Blockley, Dan Copsey,
Tim Graham, Thomas Krumpen

The distribution and surface properties of sea ice play a critical role in determining the characteristics and evolution of a variety of weather systems, primarily by affecting small-scale atmospheric processes such as surface exchange and atmospheric boundary layer (ABL) mixing (Vihma et al., 2014). Idealized numerical modeling has shown that the sea-ice distribution significantly affects downstream surface fluxes and ABL circulations (Gryschka et al., 2008; Liu et al., 2006; Müller et al., 2023; Spensberger & Spengler, 2021), the development of polar lows (Sergeev et al., 2018) and the development of ABL clouds (Brooks et al., 2017; Spensberger & Spengler, 2021). Sea-ice aerodynamic roughness is key in determining the surface exchange of momentum, heat and moisture (Elvidge et al., 2016, 2021), with its parameterization critically affecting atmospheric drag, ABL properties over the MIZ and downstream ABL evolution (Elvidge et al., 2023; Renfrew, Elvidge et al., 2019). Sea-ice characteristics also affect synoptic-scale weather systems such as Arctic cyclones (Parker et al., 2022); for example, Valkonen et al. (2021) and Finocchio et al. (2022) found Arctic cyclones had higher energy and were more impactful when sea-ice concentrations were lower.

In turn, weather systems affect the sea ice—particularly in the spring, summer and autumn. They can transiently accelerate or decelerate the seasonal cycle of melt and growth that solar radiation dictates via warm- or cold-air advection, cloud development, the advection of sea ice and via additional oceanic mixing (Clancy et al., 2022; Finocchio et al., 2022; Peng et al., 2021). Strong Arctic cyclones can cause rapid ice loss events (e.g., Cavallo et al., 2025; Kriegsmann & Brümmer, 2014; Lukovich et al., 2021; Zhang et al., 2013); indeed, a study of hindcasts with dynamic sea ice has demonstrated that the sea ice is less predictable on weather timescales during such rapid ice loss events (McGraw et al., 2022). The impact of Arctic cyclones appears seasonally dependent:

- Triggering the onset of melt in spring—due to warm-air advection and increased cloud cover, leading to an increase in downward longwave radiation flux that raises the surface temperature above the freezing point (Persson, 2012);
- Slowing ice loss in early summer—when their associated cloud cover prevents solar radiation reaching the surface (Finocchio et al., 2020);
- Destroying sea ice in late summer—due to warm-air advection (Fearon et al., 2021) and turbulent mixing of relatively warm water from beneath the ocean mixed layer (Stern et al., 2020), when the sea-ice pack is at its most diffuse and prone to drift and break up in response to winds (Lei et al., 2020).

In short, the case for a coupled environmental prediction system on weather timescales has become clear over recent years and studies examining their performance are now vital for their development in an operational setting.

Until recently, much of the focus on Arctic environmental prediction has been on seasonal to interannual forecasting of sea ice. Addressing questions such as “how much Arctic sea ice will there be next September?”; or “when will the Arctic become ice free?” (Blanchard-Wrigglesworth & Bushuk, 2019; Blanchard-Wrigglesworth et al., 2011; Bushuk et al., 2024; Jahn et al., 2016, 2024). Coupled seasonal forecasting systems have been used for more than a decade at some forecasting centers, with mixed results for their predictions of Arctic sea-ice extent (Zampeiri et al., 2018). Arctic sea ice thickness has been shown to be a crucial predictor of sea-ice extent in a coupled seasonal forecast system (Day et al., 2014). Some skill has been demonstrated in forecasting systems up to about 1.5 months. However, some systems do not have any more skill than climatology; indeed, the range in forecast skill over sub-seasonal to seasonal timescales is enormous (Zampeiri et al., 2018). Some skill in predicting pan-Arctic sea-ice area has been demonstrated at longer timescales, however summer minimum sea-ice extent predictions are limited by the “Arctic spring predictability barrier”—predictions initialized before the onset of melting in springtime have lower skill due to enhanced sea-ice feedback processes and atmospheric variability, whereas predictions initialized after this can skilfully predict summer sea ice (Bushuk et al., 2020, 2022). Generally, sea-ice prediction appears more challenging during summer than during winter (Goessling et al., 2016; Zampeiri et al., 2018); partly due to the poorer retrievals of sea-ice fraction from satellite sensors during melt conditions. Satellite products primarily use passive microwave radiometers which have difficulty distinguishing between melt ponds and open water, leading to larger biases in sea-ice fraction during the melt season (Kern et al., 2016, 2020; Zhao et al., 2021).

The increasing use of the Arctic for shipping, tourism, resource extraction, and military purposes is now increasing demand for short-range environmental forecasts (Jung et al., 2016; Meier et al., 2014). In particular, for forecasts of weather, sea ice and ocean waves for the coming few days at a spatial and temporal resolution that is useful for end users. This demand has motivated the development of coupled NWP systems—incorporating

atmosphere, ocean and sea-ice models—for weather timescales (less than ~ 10 days). Coupled NWP systems have become operational at some forecasting centers in the last few years: including across all resolutions at the European Center for Medium-Range Weather Forecast (ECMWF) in 2018 (Day et al., 2022; Keeley & Mogensen, 2018); and at the Met Office in 2022. The ECMWF coupled prediction system is generally more skilful than the equivalent uncoupled system for the Arctic region over weather timescales, but not all the time and not in all regions (Day et al., 2022; Keeley & Mogensen, 2018). Coupled NWP is still in its infancy and the one or two studies evaluating their performance in the Arctic thus far demonstrate many challenges as well as benefits.

Here, we perform a case study evaluation of a 10-day summertime period using two coupled NWP configurations of the Met Office NWP system: a version of the Global Coupled model version 4 (GC4) with the same scientific configuration as the operational forecast system, which uses the Community Ice Code 5.1 (CICE5.1) sea ice model; and a developmental version of the Global Coupled model version 5 (GC5) pre-operational forecast system, which utilizes the recently developed Sea Ice modeling Integrated Initiative (SI³) sea ice model—part of the NEMO modeling framework. Note a GC5 configuration is also planned to be the climate model configuration used for CMIP7 (Blockley et al., 2024). This is the first study to evaluate short-range NWP from a model system that uses SI³, and one of the first to evaluate coupled NWP for the Arctic region using the Met Office's forecasting system (see also Barrell, 2023). We investigate the model predictions of sea ice, the surface energy budget, and the near-surface atmosphere, evaluating against our novel low-level aircraft observations and satellite-based analyses. We focus on the Fram Strait region of the European Arctic during peak melt conditions and during the passage of an Arctic Cyclone. These represent very challenging conditions for coupled NWP. We demonstrate the coupled NWP system has clear benefits: being able to accurately represent sea-ice movement and thickness, and the surface energy budget when the sea ice fraction is correct. However, we also demonstrate some significant model deficiencies: the coupled initialization set up results in biased sea-ice concentration, with the ensuing melt rate enhanced by ice-albedo feedback, which compromises regional sea-ice predictions. Our study highlights some recent progress in coupled NWP and explores some of the key remaining challenges.

2. Data and Methods

2.1. Met Office Unified Model

The Met Office's Unified Model is the atmospheric model used operationally for weather forecasting over timescales from hours to seasons, globally and over limited areas, and is also the atmospheric component of the UK climate models, for example, the Hadley Centre Global Environment Model 3 (HadGEM3) and the UK Earth System Model 1 (UKESM1). It can be coupled to ocean and sea-ice models, with the GC4 configuration (Xavier et al., 2024a)—consisting of GA8/GL9 (for atmosphere/land), GO6 (for ocean; Storkey et al., 2018) and GSI8 (for sea ice; Ridley et al., 2018)—becoming operational in May 2022. For the atmospheric component we use a horizontal resolution of N320 (40 km) and 70 vertical levels. Tests with a resolution of N1280 (10 km), the current operational resolution, showed no substantial difference to the N320 simulations for our case study and are not analyzed further here. The GA8/GL9 scientific configuration is very similar to previous versions of the atmospheric model (see Walters et al., 2019). The ocean model component is based on Nucleus for European Modeling of the Ocean 3.6 (NEMO3.6) (Madec & the NEMO team, 2017) and the sea ice model is CICE5.1 (Hunke et al., 2015). The developmental GC5 configuration (Xavier et al., 2024b), planned as the next operational configuration for weather forecasting and climate modeling (e.g., CMIP7), consists of GA8/GL9 and GOSI9 (Guiavarc'h et al., 2024; Blockley et al., 2024). The ocean model component is based on NEMO v4.0.4 (Madec & the NEMO team, 2019) and the sea ice model component is the recently developed SI³ (Vancoppenolle et al., 2023), with coupling to the atmosphere model described by Blockley et al. (2024). Here we use the same model resolutions as the GC4 simulations. GC4 and GC5 use the Ocean Atmosphere Sea Ice Soil Model Coupling Toolkit (OASIS3-MCT) coupler (Craig et al., 2017) to exchange information between the ocean and atmosphere model grids.

For both configurations, the simulations are initialized using restart files adapted from the same operational forecast system, which used ocean and sea ice initial conditions from the Forecast Ocean Assimilation Model (FOAM) (Blockley et al., 2014) and atmosphere initial conditions from the operational analysis. In terms of sea ice, this includes assimilated sea-ice concentration observations from the Ocean and Sea Ice Satellite Application Facility 401-b data set (OSI-SAF-401-b) (see Section 2.4). Note, sea-ice thickness is not routinely assimilated due to a lack of available observations; instead, the ice thickness distribution across 5 subgrid thickness categories

evolves following Thorndike et al. (1975). Likewise, melt-pond fraction is not assimilated, with melt ponds simulated using the topographic scheme developed by Flocco and Feltham (2007). These schemes are essentially the same in the CICE5.1 and SI³ sea ice models, albeit coded differently. Assimilated ocean data include satellite sea level anomaly, satellite and in situ sea surface temperature. In situ ocean observations include data from drifting buoys, moored buoys, ships, Argo floats, marine mammals, CTD/XBT profiles and gliders (Barbosa Aguiar et al., 2024). Although there are some profiles available under sea ice, they generally have little influence in data assimilation; here sea surface temperature essentially holds at freezing temperature with excess heat used to melt sea ice (or grow ice if colder than freezing temperature).

Sea-ice albedo in GC4 and GC5 is determined by the Joint UK Land Environment Simulator (JULES) model component (Best et al., 2011), through a process that incorporates the relative fractions of bare ice, snow and melt ponds, as well as ice thickness, surface temperature, pond depth and refrozen pond lids, for direct and diffuse bands of visible and near-infrared shortwave radiation. However, the models differ in how the penetration of incident shortwave radiation is treated. In GC4 no shortwave penetrates the sea-ice surface; instead, the albedo is adjusted using an adaptation of the Semtner (1976) scheme to approximate the effect of internal shortwave scattering (Rae et al., 2015). In GC5 this scheme has been further modified to allow a portion of the shortwave radiation to penetrate the surface, which is passed to the sea ice model via the coupler (Blockley et al., 2024). Another improvement in SI³ is a time-evolving salinity that is used in freshwater exchanges and thermodynamic processes, in contrast to a single bulk salinity in the CICE5.1 code. Additionally, SI³ uses a parameterized representation of the floe size distribution in calculating lateral melting, whilst CICE5.1 assumes a fixed floe size of 200 m.

2.2. Arctic Summertime Cyclones Field Campaign

The Arctic Summertime Cyclones (ASC) field campaign took place in July–August 2022 out of Longyearbyen, Svalbard, in concert with the RALI-THINICE field campaign (see Rivière et al., 2024 for an overview). The core objectives were to observe Arctic cyclones, their boundary-layer and cloud structures, and their interactions with tropopause polar vortices, and the sea-ice surface during summertime conditions. A British Antarctic Survey Twin Otter aircraft, equipped with the Meteorological Airborne Science INstrumentation (MASIN), was deployed. The instruments pertinent to this study include a turbulence probe and high-frequency humidity and temperature sensors from which turbulent fluxes of momentum, heat and moisture are derived (after Petersen & Renfrew, 2009; Elvidge et al., 2016). Pyranometers and pyrometers measured the downwelling and upwelling shortwave and longwave radiation. An infrared thermometer observed surface brightness temperature and a video camera recorded surface and cloud conditions. On two of the ASC flights (Flights 375 and 379) the high frequency humidity instrument failed, so latent heat fluxes were calculated using the COARE 3.0 bulk flux algorithm (Fairall et al., 2003). A more detailed description of the instrumentation is provided in Rivière et al. (2024), as well as in King et al. (2008), Fiedler et al. (2010), and Renfrew, Pickart et al. (2019). To make use of field observations to interrogate model simulations we define a study region that encompasses Fram Strait and the area around the northeast of Greenland (specifically -40 – 40°E and 75 – 88°N ; Figure 1).

Sea-ice fraction is estimated via shortwave radiation measurements (i.e., the albedo) following Elvidge et al. (2016). This method involves setting tie points for the *no ice transition*, between open water and sea ice, and the *all ice transition*, between continuous ice and the appearance of some water. Previously this method has been used for wintertime campaigns, when the albedo of sea ice is relatively constant (implying the tie points can be set as constant values) and the surface temperature can be used as a separate measurement of surface type (Elvidge et al., 2016, 2021). In the summertime, distinguishing sea ice, melt ponds and open water via shortwave radiation and surface temperature data is much more challenging. The albedo of melt-ponds and water can be similar (when averaging over the pyranometer footprint) and at this time the surface temperatures are not dramatically different.

To accurately determine the surface type, we subjectively analyzed video footage of all relevant flights, recording the fraction of sea ice and melt ponds every minute. Fractions were recorded to the nearest 5% and regularly benchmarked by objective image color analysis. The video analysis was used to define appropriate tie points, so that a sea-ice fraction could then be derived from the 1 Hz shortwave radiation measurements. We estimate an uncertainty in the sea-ice fraction timeseries from ± 1 standard deviation of the observed albedo at 1 Hz, when the sea-ice fraction is >0.8 . For reference, mean sea-ice fraction estimates [with uncertainty] for the four flights used

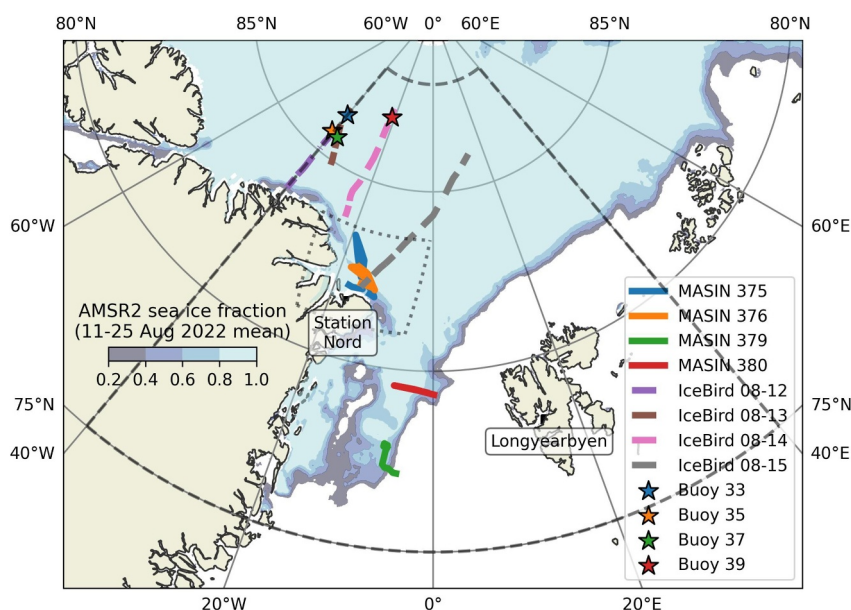


Figure 1. Overview map of the European Arctic showing NE Greenland and Svalbard, along with the mean sea-ice fraction during the study period (shaded contours). Overlaid are low-level aircraft science legs from MASIN (thick solid lines) and IceBird (thick dashed lines), and the deployment locations of the CALIB buoys (stars). The thin dashed line shows the ASC study region boundary and the thin dotted line shows the region displayed in Figure 5.

here are: Flight 375 = 0.88 [0.69 0.97], Flight 376 = 0.76 [0.61 0.90], Flight 379 = 0.75 [0.57 0.96], Flight 380 = 0.81 [0.68 0.93].

2.3. IceBird Summer 2022 Campaign

Conveniently, the Alfred Wegener Institute sea-ice survey programme, named the IceBird Summer 2022 campaign, overlapped in time with the ASC campaign (Krumpen et al., 2022). The primary objective of this regular survey programme is to make sea-ice thickness and morphology measurements using an electromagnetic sounding instrument (EM-Bird) and a scanning laser altimeter flown onboard a research aircraft (see Krumpen et al., 2025 for an overview of the programme). Four flights were carried out between 12 and 15 August 2022, based out of Station Nord, Greenland, targeting the sea ice north of Greenland (Figure 1). During these flights, four Compact Air-Launched Ice Beacon (CALIB) buoys were deployed enabling the measurement of ice velocities every hour.

2.4. Satellite Sea-Ice Concentration Products

The European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Application Facility (OSI SAF) produce the OSI-401-b data set, hereafter referred to as OSISAF-401 (EUMETSAT, 2005; Tonboe & Lavelle, 2016). It makes use of observations from the Special Sensor Microwave Imager/Sounder (SSMIS) satellite instrument in the 19 and 37 GHz channels using the OSI SAF algorithm. The product is available on a 10 km grid, however the footprint at these frequencies is in the order of 30–70 km, and each observation node is scaled using a Gaussian weighting with a radius of influence of 75 km. This produces smooth gradients in sea-ice concentration that are unable to capture leads or features $< \sim 100$ km in width, which has been shown to influence atmosphere-ice-ocean interactions when used in a reanalysis product (Renfrew et al., 2021).

The Advanced Microwave Scanning Radiometer 2 (AMSR2) sea-ice concentration analysis used here is produced by the University of Bremen using the ARTIST Sea Ice algorithm (Spreen et al., 2008). It uses the 89 GHz channel, which provides a higher spatial resolution than the lower frequencies used for the OSISAF-401 product but is more sensitive to cloud liquid water and water vapor. The swath data is resampled onto a 6.25 km polar stereographic grid to provide a daily gridded product.

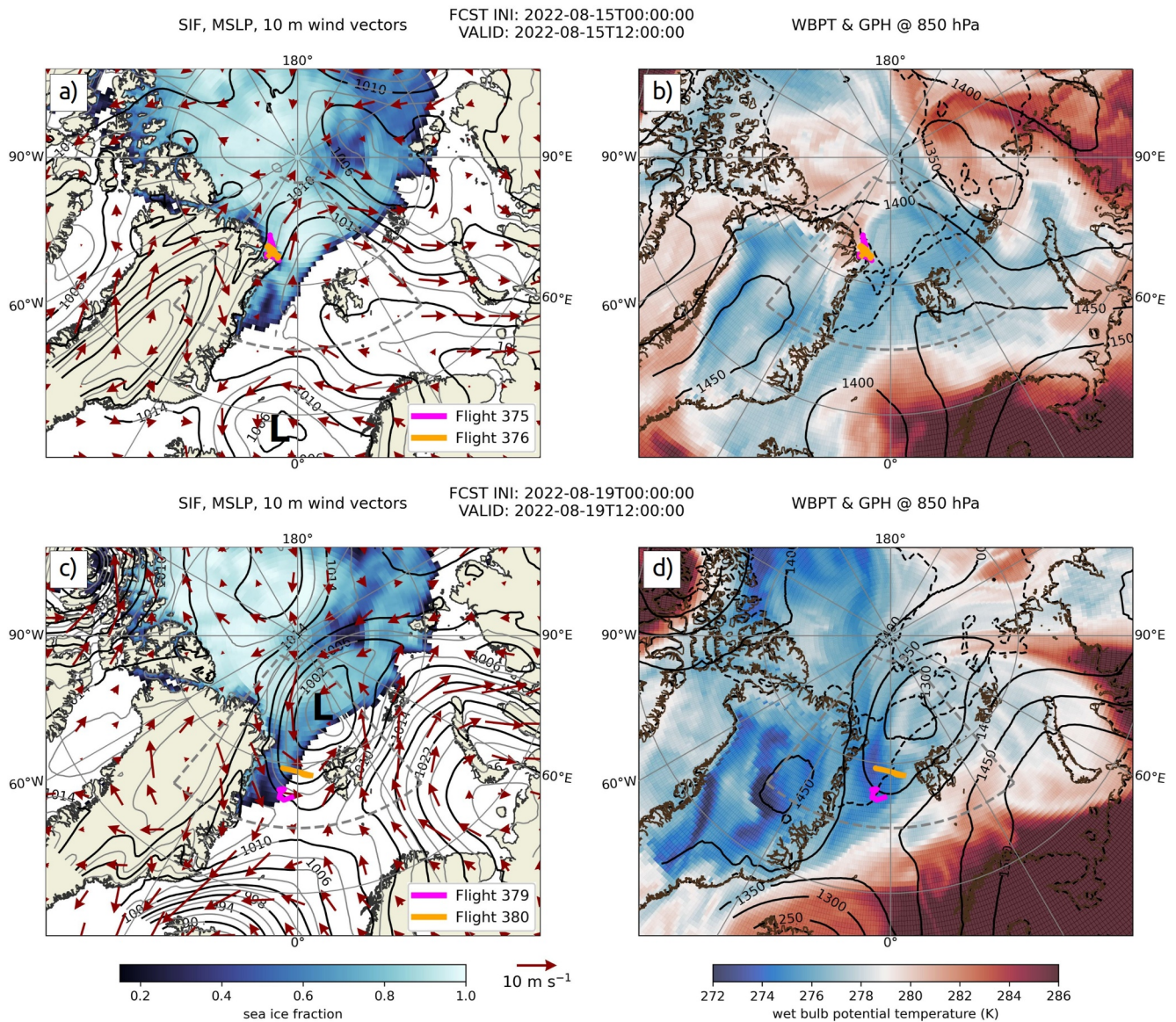


Figure 2. Synoptic overview during the passage of Cyclone 3 (labeled L) from GC5 forecasts on 15/08/2022 12Z (T + 12 hr) and 19/08/2022 12Z (T + 12 hr) with tracks of Flights 375 and 376 (top) and Flights 379 and 380 (bottom) overlaid; panels (a, c) show sea-ice fraction (shading), mean sea level pressure (contours) and 10-m wind (vectors); (b, d) wet bulb potential temperature at 850 hPa (shading) and geopotential height at 850 hPa (contours). All panels highlight the ASC study region marked by the gray dashed line.

Note, the terms sea-ice concentration (0%–100%) and sea-ice fraction (0–1) are often used interchangeably so we use sea-ice fraction hereon for consistency.

2.5. Case Study Period

During the field campaign four cyclones were observed, of which Cyclone 3 was the longest lived and most intensively observed and thus forms the focus of our 15-day case study period (11–25 August 2022). Cyclogenesis occurred on 15 August to the north-east of Iceland (Figures 2a and 2b), after which Cyclone 3 developed as a typical warm-cored mid-latitude cyclone and moved northwards over the Greenland Sea and then into Fram Strait (Figures 2c and 2d). Cyclone 3's mean sea-level-pressure minimum was 990 hPa on the 18 August 2022. After which it transitioned to a cold core axisymmetric vortex, as is typical of Arctic cyclones (Croard et al., 2023). Following which, it continued northwards over the Arctic sea-ice pack, before looping back toward the study region (see Figure 4 in Rivière et al., 2024 for the cyclone track).

Table 1
Key Information About the Four ASC Flights Used in This Study

Flight	Date and time (UTC)	Minutes at low level	Flight comments
375	15/8/22 15:33–18:41	85	N and NW of Cap Nord over the MIZ in southerly tip jet flow; Cyclone 3 over the Iceland and Greenland Seas
376	16/8/22 09:39–12:35	85	N and NW of Cap Nord over the MIZ in southerly tip jet flow; Cyclone 3 over the Iceland and Greenland Seas
379	19/8/22 07:05–12:05	40	W of Longyearbyen over the MIZ in northerly flow; Cyclone 3 north of Fram Strait
380	20/8/22 10:00–14:20	40	WNW of Longyearbyen over the MIZ in northerly flow; Cyclone 3 north of Fram Strait

Interactions with the sea-ice surface were observed by two pairs of research flights (Table 1 and Figure 1). On the 15–16 August, Flights 375 and 376 to the north-east of Greenland were in southerly tip jet flows—associated with a northward barrier wind along the orography of eastern Greenland—ahead of Cyclone 3 which was located in the Iceland and Greenland Seas at this time (Figures 2a and 2b). Then on the 19–20 August, Flights 379 and 380 were over the MIZ off the east coast of Greenland in northerly flows and the cold sector of Cyclone 3 which was located north of Fram Strait at this time (Figures 2c and 2d). The science legs of all four flights took place in cloud free conditions.

3. Sea Ice Verification

We first evaluate sea ice predictions for the entire Arctic and the ASC study region, then proceed with in depth evaluation of other variables using the field observations. Sea-ice extent (SIE) is calculated as the total area of grid cells where the sea-ice fraction ≥ 0.15 . Sea-ice area (SIA) is calculated as the sum of the products of sea-ice fraction and grid cell area in each grid cell. All calculations are performed by transforming the data onto the model grid and using a common land mask that ignores any grid cells which do not contain data from all the products.

Given the relatively large difference between AMSR2 and OSISAF-401 SIE, at first glance Figure 3a shows relatively good agreement between simulated Arctic SIE and the two satellite products, particularly for the first few days. Despite the simulated sea-ice loss being faster than observed, the predictions are much more realistic than if the initial sea ice field was fixed throughout each forecast (typically referred to as sea-ice persistence) during this period of Arctic wide sea-ice melt. However, limiting the analysis to just the ASC study region (Figure 3b) highlights a substantial regional bias; the models simulate a significant loss of sea ice in Fram Strait, which is not observed. We suggest that the observed SIE change is largely due to the passing Arctic cyclone (cf. Figure 2). An increase in SIE from around the 12th to 16th August is due to southerly winds that buffer the rate of ice loss through the southern part of Fram Strait, while ice accumulates to the north driven by the Transpolar Drift. Followed by a decrease in SIE from 16 to 19th August, due to northerly winds as the cyclone moves to the northeast and ice is exported through Fram Strait. An increase in SIE starts again from around 19–20th August, when a southerly wind regime is re-established. The model simulations fail to capture the regional evolution in sea ice.

Evaluating Arctic SIA in Figure 3c suggests poorer model performance than for Arctic SIE. The simulated SIA decreases rapidly and is out of the range of observations in all the forecasts. Note while the OSISAF-401 SIE is larger than AMSR2, its SIA is substantially smaller than AMSR2, implying the OSISAF-401 has much more low-concentration sea ice. This is consistent with Renfrew et al. (2021) and Barrell (2023), who showed that, compared to AMSR2, a smoother sea ice concentration gradient in OSISAF-401 results in polynyas and the marginal ice zone containing excess low concentration sea ice. As noted in Section 2.4, the OSISAF-401 product relies on a sensor with a larger footprint and a Gaussian smoothing step as part of its retrieval algorithm.

The simulated rapid decrease in SIA across the Arctic suggests a problem in the models' representation of melt during this period, with this problem emphasized when focusing on the ASC study region (Figure 3d). The pattern of SIA change is similar to that for SIE, as would be expected, with a similarly large bias in ice loss in the models.

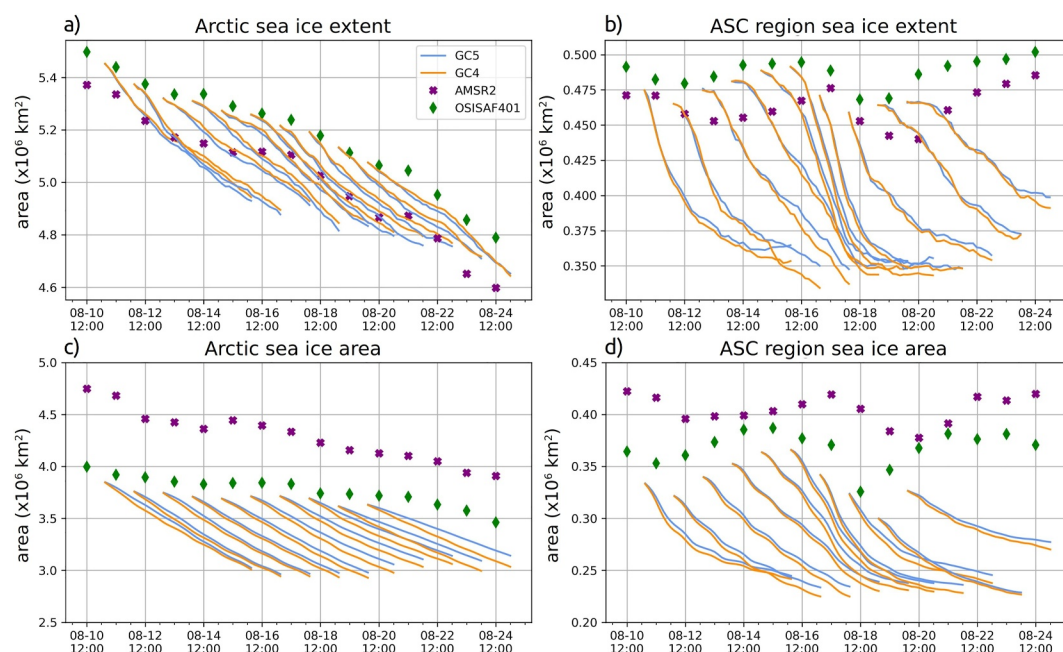


Figure 3. Timeseries of sea-ice extent (top, a, b) and sea-ice area (bottom, c, d) showing 10 GC5 and GC4 model simulations and satellite-based observations from AMSR2 and OSISAF-401 (as used to initialize the model). The left column is for the whole Arctic; the right column is for the ASC region.

SIE is more commonly used to measure sea-ice coverage as it circumvents the uncertainty in estimating sea-ice fraction during summertime conditions, when melt ponds pose a particular challenge for satellite based passive microwave sensors (Kern et al., 2016). Examining extent only, rather than both extent and area, disguises problems in areas of lower ice concentration that are illustrated here. Comparing the two model simulations, it is clear they are very similar for all the forecasts. For SIA, the GC5 simulations are only slightly better than the GC4 simulations.

To quantify the skill in the sea-ice simulations we calculate the Integrated Ice Edge Error (IIEE) metric following Goessling et al. (2016). The IIEE sums the symmetric difference between the areas enclosed by the forecast and observed ice edge at each time step, that is, the combined area of overestimate and underestimate. We compute the IIEE for the coupled simulations using an ice-edge threshold of 0.15 and compare against a persistent sea ice field (generated from the earliest output timestep of $T + 3$ hr). Figure 4 shows IIEE against lead time for both the Arctic and ASC study regions using both OSISAF-401 and AMSR2 as truth. Across the Arctic, the IIEE grows at a similar rate in the coupled simulations and persistence (Figures 4a and 4c); although the simulations are slightly worse than persistence when using OSISAF-401 as verification. For the ASC study region, the coupled simulations are markedly worse than persistence, which is relatively consistent regardless of lead time (Figures 4b and 4d). During this period, an initial SIE overestimation in the East Siberian, Laptev and Kara Seas tends to remain relatively constant over the forecasts, while sea ice retreats rapidly north of Svalbard, in Fram Strait, around the Canadian Archipelago and in the Beaufort Sea (not shown). This pattern of sea-ice retreat is similar regardless of the wind pattern. These results show the European Arctic region to be particularly challenging for short-term sea-ice prediction in summer.

We now investigate the spatial characteristics of the sea ice fields using our aircraft observations. Figure 5a shows good agreement between the daily analysis field from AMSR2 and sea-ice fraction observed during Flight 375. A similarly good agreement is found when using Flight 376 observations (not shown). Figure 5b illustrates a lack of detail in the sea-ice fraction in the OSISAF-401 product and consequently poorer agreement with the aircraft observations. Figure 5c, from a GC5 simulation at $T + 12$ hr, shows the sea-ice fraction is already substantially reduced compared to the satellite products and compares poorly to the aircraft observations. The difference between the GC5 simulations and AMSR2 at $T + 12$ hr (Figure 5d) shows a generally reduced sea-ice fraction, except over the large coastal polynyas north of Greenland which are not well represented by the relatively smooth

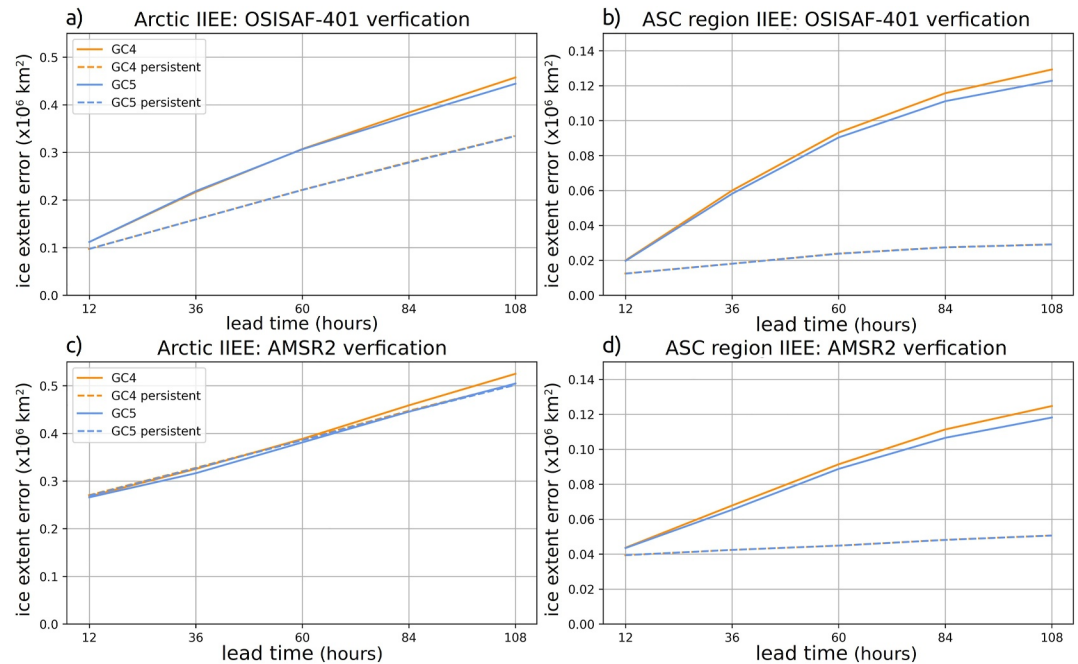


Figure 4. Sea-ice area predictive skill against lead time over the 10 GC5 (blue solid line) and GC4 (orange solid line) model simulations using the Integrated Ice Edge Error metric, with OSISAF 401 (top, a, b) and AMSR2 (bottom, c, d) as the verification products. The left column is for the whole Arctic; the right column is for the ASC region. Note, in each panel the results for persistent sea ice fields are very similar, thus the orange dashed line (GC4) sits underneath the blue dashed line (GC5).

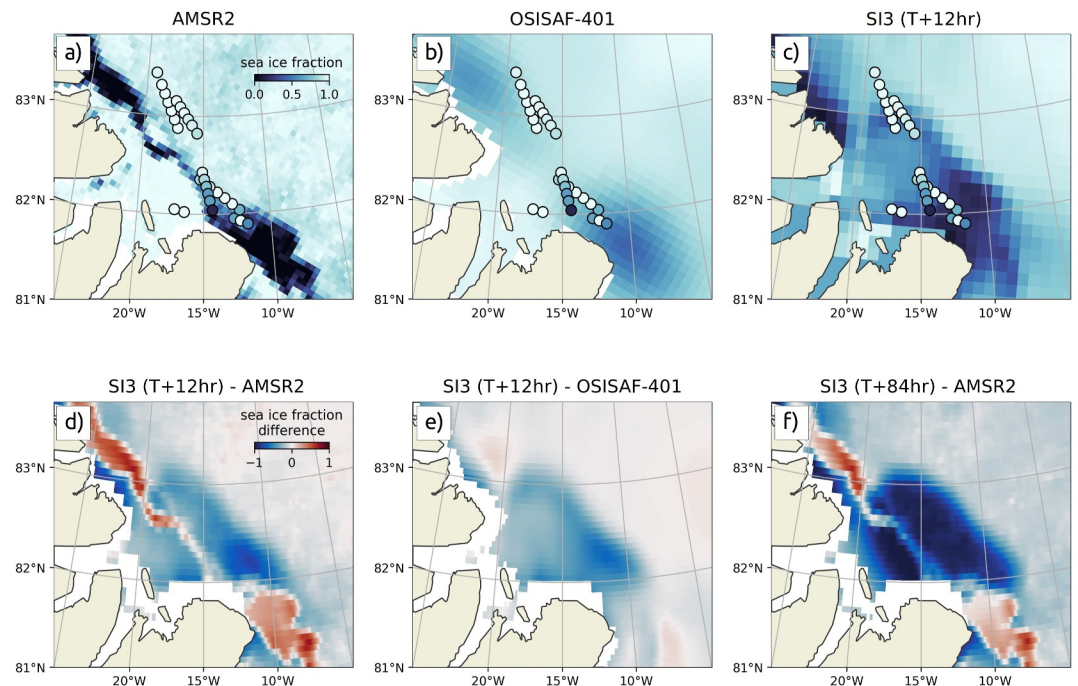


Figure 5. The top row shows sea-ice fraction on native grids at 12 UTC on 15 August 2022 in: (a) AMSR2, (b) OSISAF-401 and (c) SI³ at T + 12 hr with estimated sea-ice fraction from Flight 375 overlaid. The bottom row shows differences between the simulated sea-ice fraction in SI³ and satellite products: (d) SI³ at T + 12 hr—AMSR2, (e) SI³ at T + 12 hr—OSISAF-401, and (f) SI³ at T + 84 hr—AMSR2.

model fields. A similar pattern is seen for the difference between GC5 and OSISAF-401, with a surprising amount of ice loss seen over just 12 hr (Figure 5e); although the pattern is less extreme, as would be expected given the model uses this product during initialization. Figure 5f shows the difference between GC5 and AMSR2 at the same verification time but for an earlier initialization time (so at $T + 84$ hr). It illustrates a substantial loss of sea ice in the simulation (consistent with Figures 3 and 4), yet still with poor representation of the coastal polynya. Generally, the forecasts for this period have too little sea ice and the rate of ice loss is too rapid.

4. Sea Ice Model Diagnostics

We now investigate the causes of the discrepancies in the simulated sea ice by evaluating diagnostics from the SI^3 component of the GC5 model. In Figure 6a, near the northwest coast of Greenland where MASIN Flights 375 and 376 took place, there is strong dynamic transport of ice out of this area acting to open the coastal polynya at $T + 12$ hr. This corresponds with the strong southerly low-level jet noted off Cap Nord at this time (cf. Figure 2). Over the entire study region, the strong positive and negative dynamic ice mass tendencies roughly balance out. At $T + 108$ hr, key parts of the spatial pattern of the dynamic forcing are reversed, in accordance with the changing flow regime associated with the passage of the cyclone (Figure 6b). Now, in the specific region of Flights 375 and 376 under the northerly wind regime, the dynamic forcing acts to transport the sea ice toward the Greenland coast closing the coastal polynya.

In contrast, Figure 6c displays a more consistent thermodynamic ice loss, which is enhanced in the marginal ice zone. Later in the forecast, at $T + 108$ hr, the magnitude of the thermodynamic ice mass change is reduced but remains negative (Figure 6d). The primary components of the thermodynamic forcing are the net downward heat flux at the ocean-ice surface and the sensible heat flux at the ice base from relatively warm water (all heat fluxes are positive downward); these are shown in the bottom 4 panels of Figure 6. At $T + 12$ hr the downwards atmospheric heat flux is entirely positive and is intensified at the ice edge, where the low ice fraction reflects less shortwave radiation (Figure 6e). However, at $T + 108$ hr a swath of upwards heat flux occurs over Fram Strait due to cold-air outbreak conditions and a small amount of snow falling in the simulation, which raises the albedo and stops the surface melt (Figure 6f). The sensible heat flux at the ice base remains similar at the early and late forecast time steps with an ice heat gain of around 50 W m^{-2} or greater in the ASC study region (Figures 6g and 6h).

Figure 7 shows the dynamic and thermodynamic contributions to ice mass change in SI^3 along the track of the MASIN flights (cf. Figure 1) with simulations initialized on the day of each flight. Thermodynamic ice loss generally dominates, apart from times of increased dynamic ice loss around the location of the low-level jet in the middle of Flight 375 and over a region that was passed twice during the roughly reciprocal legs in Flight 380. During Flights 375 and 376, off the northeast Greenland coast, the dynamic forcing compounds the thermodynamic ice loss. Deeper investigation suggests that the nearby low-level jet in the model may be more meridional than observed, indicating an error in wind direction rather than magnitude (not shown), that is also seen in the higher resolution (10 km) tests. Flights 379 and 380 find positive dynamic and negative thermodynamic ice-mass contributions, illustrating ice export through Fram Strait (under these flights) and continuous melting.

Thermodynamic ice loss is stronger when the sea-ice fraction is between 0.2 and 0.6 and reduces when the sea-ice fraction is above 0.8 or approaches 0. The three largest contributors to thermodynamic ice mass change are also shown: bottom melt dominates, while the contributions from top melt and lateral melt are very small. In other words, thermodynamic ice loss is being driven by bottom melt from a relatively warm ocean. We have already established that the model's sea-ice fraction is generally biased low, we postulate that this bias is allowing excessive solar heating of the ocean, which then melts the base of the sea ice, that is, the ice-albedo effect is accelerating sea-ice loss. We shall investigate this further in Section 6, but first we complete our evaluation of the model's sea ice in Section 5.

5. Sea-Ice Thickness and Drift Speed

In Figure 8 the sea-ice thickness measured along the four IceBird flights to the north of Greenland is compared with that simulated by GC5, using forecasts initialized on the day of each flight. In general, the model sea ice compares relatively well with the observations, although it is consistently biased too thin. The model biases are -1.4 , -0.5 , -1.0 , and -0.8 m for the flights on 12/13/14/15 August 2022 respectively. The model does not successfully simulate the range of observed sea-ice thickness and appears limited to a maximum thickness of

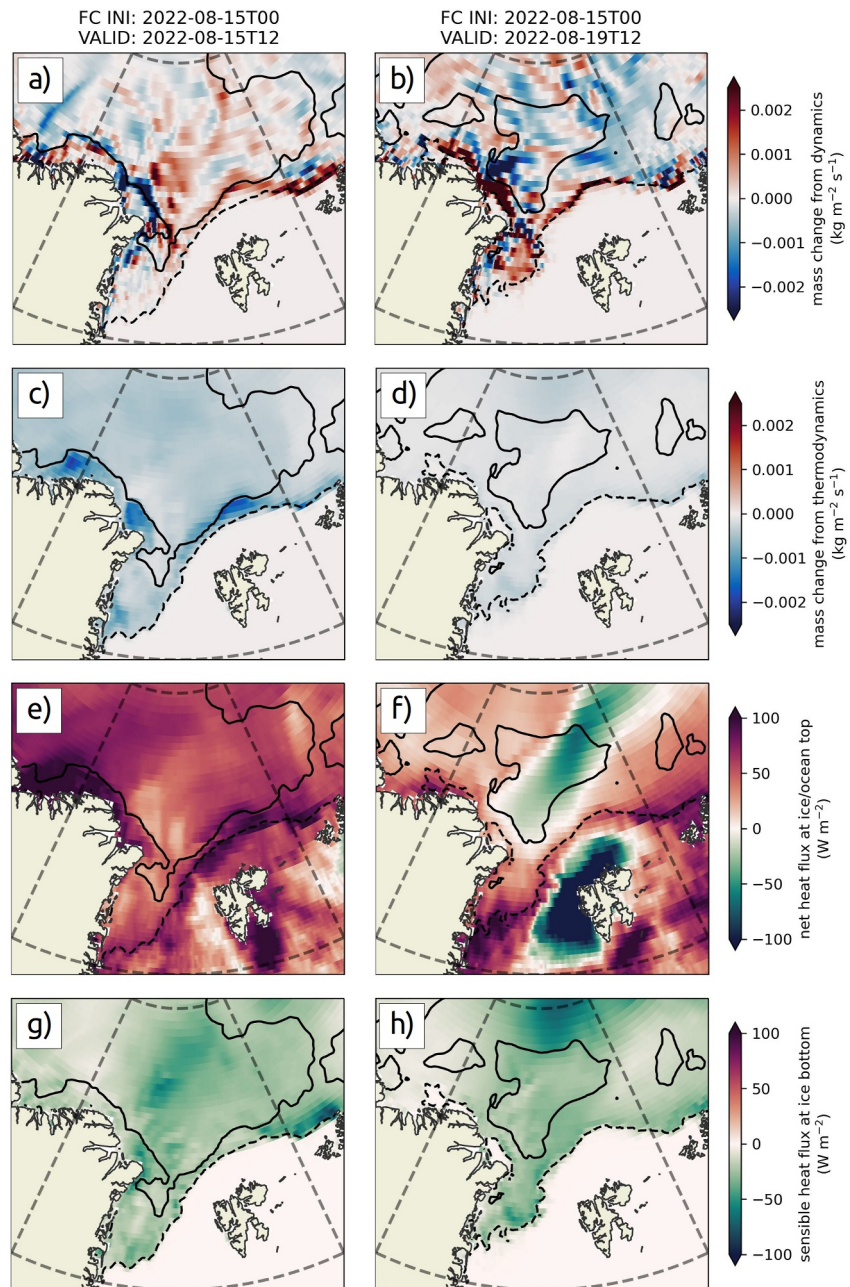


Figure 6. Sea ice mass change tendencies from GC5-SI³ due to dynamics in (a, b) and to thermodynamics in (c, d). The thermodynamic tendencies due to net downward heat flux at the top surface of the ice and ocean are shown in (e, f); and the sensible heat flux at the base of the ice in (g, h). The model is initialized at 00 UTC on 15 August 2022, and the left column is for $T + 12$ hr, while the right column is for $T + 108$ hr. Dashed and solid black lines show 0.15 and 0.8 sea-ice fraction contours, respectively, while gray dashed lines show the boundary of the ASC study region.

around 2.5 m, while observed maximums were 4–5 m. Spatial investigation finds that the largest differences are in the early parts of the IceBird flights on 12 and 15 August (not shown), when the aircraft was over the marginal ice zone (cf. Figure 2). On these days the MIZ has thicker sea ice, possibly due to convergence and rafting, which appears to be less well represented in the simulations. Given that the simulations were initialized 10–12 hr before each flight, and no ice thickness data is assimilated, it is unlikely that substantial convergence would have occurred in the model by these verification times. Note the results are essentially the same for GC4 (not shown).

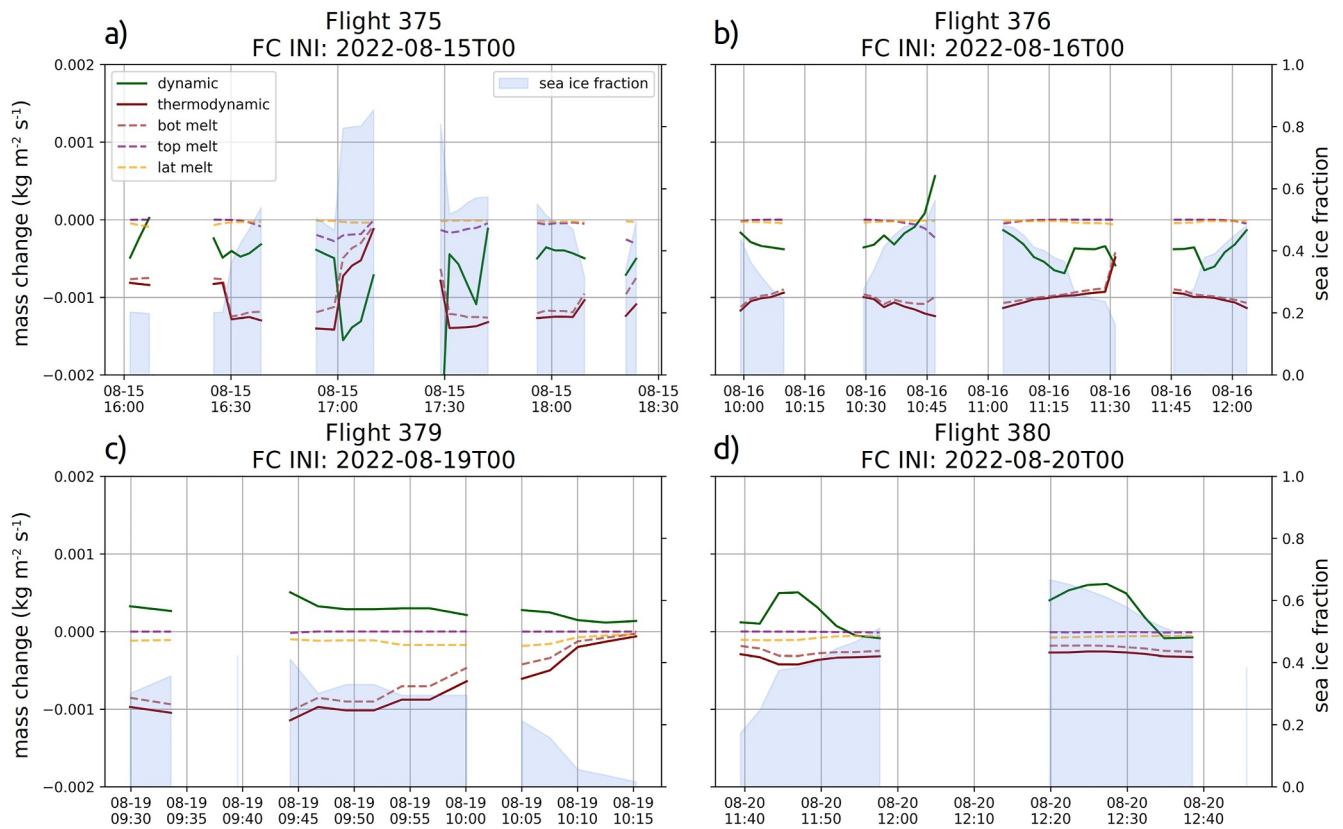


Figure 7. Sea ice mass tendencies in SI3 interpolated spatially and temporally to the tracks of MASIN Flights 375, 376, 379, and 380 in (a–d), respectively, using simulations initialized on the day of each flight as indicated. Lines show the model tendencies and shading shows model sea-ice fraction—see Figure 1 for spatial reference.

Sea-ice velocities have been calculated from the hourly positional data from the CALIB buoys deployed during the IceBird flights. A conservative uncertainty estimate has been calculated by applying ± 10 m to the positions, though the true positional accuracy is likely < 3 m (Yastrebova et al., 2021). These observations enable an

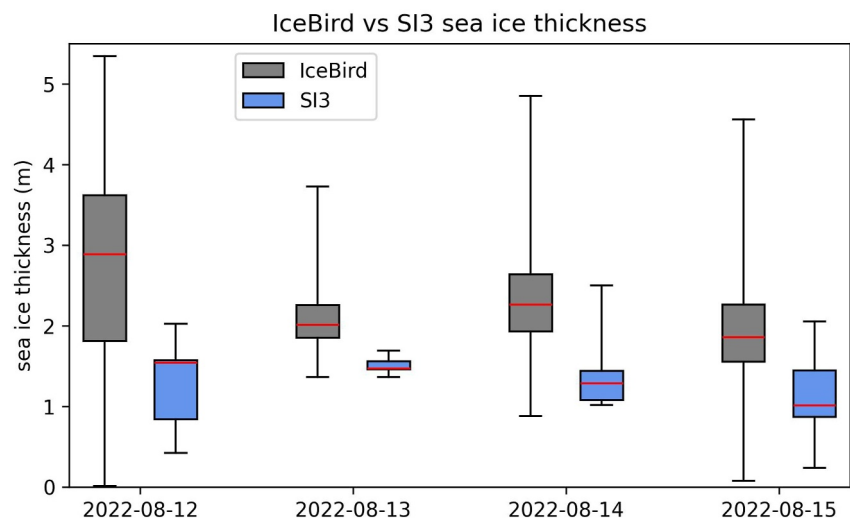


Figure 8. Box and whisker plot illustrating the sea-ice thickness distribution measured during the IceBird summer 2022 campaign flights (using 1-min means; gray boxes) and simulated by the SI³ component of GC5 forecasts initialized on the day of each flight interpolated to each flight track (blue boxes); red lines show median values.

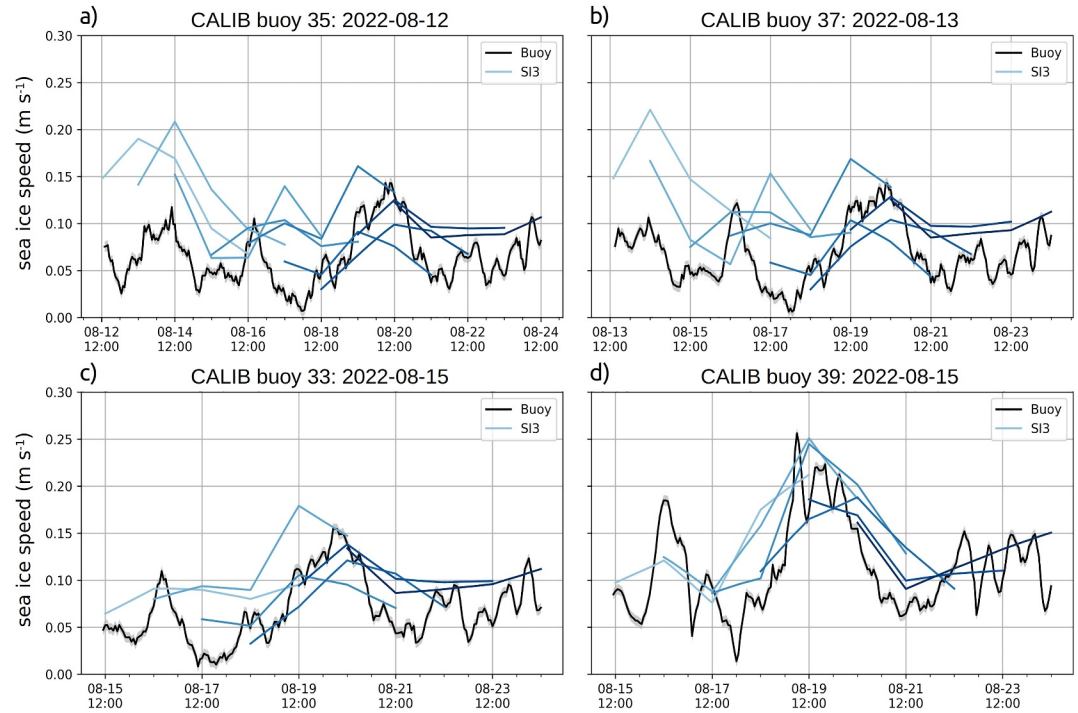


Figure 9. Sea-ice drift speeds calculated from the CALIB buoys (black lines), with a conservative uncertainty estimate (gray shading). Overlaid are simulated sea-ice drift speeds along the track of each buoy from the SI³ component of GC5 (blue lines; increasing darkness shows later initialization). Panels show speeds for buoy numbers: (a) 35, (b) 37, (c) 33 and (d) 39 (see Figure 1 for locations).

evaluation of the sea-ice speed along the track of the buoys. Figure 9 shows sea-ice speeds from observations and from the SI³ component of the GC5 simulations. GC5 appears to overestimate the drift speeds earlier in the period, when the wind regime at the first two buoys was generally southerly, becoming more aligned with observations later on (Figures 9a and 9b). Later in the period, at the second two buoys, GC5 agrees well with the observations during a southeasterly wind (Figures 9c and 9d). In general, the sea-ice speed evaluation suggests the dynamical processes are well simulated during this period.

6. Surface Energy Budget

So far, we have shown that the simulated sea-ice fraction is biased low in the ASC region, and the largest overall contribution to sea-ice loss is melting at the ice base. Observations from the IceBird summer 2022 campaign suggest that, despite some discrepancies, the sea ice is sufficiently well represented for an evaluation of the surface energy budget (SEB). We examine GC5 simulations initialized on the day of each flight using the MASIN observations with the aim of further investigating the veracity of the underlying physical processes simulated by the model.

The net surface heat flux budget, F_0 , at the top surface of the ocean and ice can be written as:

$$F_0 = -F_{CT} = F_{\downarrow SW} - F_{\uparrow SW} + F_{\downarrow LW} - F_{\uparrow LW} + F_{SH} + F_{LH},$$

$$F_0 = -F_{CT} = (1 - \alpha) F_{\downarrow SW} + F_{\downarrow LW} - F_{\uparrow LW} + F_{SH} + F_{LH},$$

where F_{CT} is the conductive flux at the surface between the ocean/ice and the atmosphere $F_{\downarrow SW}$ and $F_{\uparrow SW}$ are the downward and upward shortwave fluxes, $F_{\downarrow LW}$ and $F_{\uparrow LW}$ are the downward and upward longwave fluxes, F_{SH} is the downward atmospheric sensible heat flux, F_{LH} is the downward latent heat flux and $\alpha = F_{\uparrow SW}/F_{\downarrow SW}$ is the surface albedo. Note we do not have observations of the conductive heat flux, F_{CT} . Here we use the atmospheric fluxes observed during the ASC campaign to estimate the total heat flux absorbed at the ocean-ice surface. The surface albedo is calculated from the SW fluxes for both the model and the observations (allowing a direct

comparison). Note that the aircraft pyranometer footprint is on the order of 100 m (for the altitudes flown) and the observations are time averaged into 150-s runs (~ 9 km).

Figure 10 provides an evaluation of the key SEB variables in GC5 during flights 375 and 376. Generally, the model simulates the SEB relatively well, however there are some biases that relate to a frequent lack of sea ice in the model. Even at short lead times, sea-ice fraction is considerably lower than observed by the aircraft and in AMSR2 (consistent with Figures 3 and 5). On all but one of the low-level aircraft legs the model sea-ice thickness is <0.25 m, which is significantly lower than measured in the IceBird flights (Figure 8); unfortunately, there are no thickness measurements from the MASIN flights. Melt-pond fraction compares reasonably well on the two most northerly legs of Flight 375, where model sea-ice fraction is highest. Otherwise, it is lower than observed, particularly during Flight 376, which was during the morning and, despite it being polar summer, many model melt ponds develop a frozen lid overnight in the cooler temperatures which did not occur in reality. Despite the discrepancies in melt-pond fraction, the mean ocean-ice surface temperatures are simulated well, although slightly warmer in the model than observed during these flights. This is partly due to the larger fraction of ocean in the model and is further evidence that too much heat is being absorbed by the ocean. Modeled air temperature is often higher than observed in Flight 375, while the reverse is true in Flight 376. This is also likely to be due to the timing of these flights—the model's diurnal variability may be exaggerated. Time series observations from the sea ice would be required to investigate this further, so this is reserved for a future study.

During both flights, observed surface albedo is compared with both the grid box mean albedo (including open water) and the albedo of the sea ice surface only (not including open water). The grid box mean albedo is substantially lower than observed, as expected given the much lower sea-ice fraction. However, one would expect the model sea-ice albedo to roughly agree with the observed albedo when observed sea-ice fraction is ~ 1 , but it does not, it is also substantially lower. Further surface albedo evaluation is beyond this study but is planned. Simulated and measured $F_{\downarrow SW}$ agrees well, but $F_{\uparrow SW}$ does not, as a result of the model consistently underestimating sea-ice fraction, leading to an overestimate in F_{SWnet} . In contrast, the simulated and measured longwave radiative fluxes agree very well. As noted earlier, these flights took place in cloudless conditions, so there is no additional complexity due to clouds to consider.

In Flight 375 there is good agreement in the sensible and latent heat fluxes though notably during the third leg, when the model sea-ice fraction and melt-pond fraction were highest, there is a positive bias in downwards sensible heat flux, while observed sensible heat flux was $\sim 0 \text{ W m}^{-2}$. In Flight 376 the model turbulent heat fluxes tend to be negative (upwards), associated with the upwards temperature gradient, while they were observed to be around $\sim 0 \text{ W m}^{-2}$. In both flights the turbulent heat fluxes are relatively small corresponding to low wind speeds. Overall, the total heat flux toward the surface was strongly positive ($\sim 100 \text{ W m}^{-2}$) in the model, while it was around 0 W m^{-2} in the observations. This is a substantial difference. The primary reason for this bias is the lack of sea ice leading to a $F_{\uparrow SW}$ flux that is too low in magnitude.

Figure 10 also displays the SEB evaluation for Flights 379 and 380 over the MIZ to the east of Greenland over Fram Strait (cf. Figure 1 for location). Generally, the results for these two flights corroborate those to the northeast of Greenland. In brief, the model simulates too little sea ice, a lack of melt ponds, a good agreement in surface temperature, a low air temperature, a low surface albedo associated with the F_{SWnet} bias, close agreement in longwave radiation, and high F_{SH} in Flight 379 but otherwise close agreement in turbulent fluxes. The result is again a consistent high bias in total downward heat flux, albeit generally smaller in magnitude than the earlier flights as the F_{SWnet} bias is smaller. Note, the results for GC4 (not shown) are very similar to those presented for GC5.

The model surface energy budget results can be summarized by examining key variables for each flight as a function of increasing forecast lead time. Figure 11a shows a consistent low bias in sea-ice fraction across all 10 forecast periods, which grows with lead time. Reflecting the low sea-ice fraction, the surface albedo is also biased low on initialization, and this grows with lead time (Figure 11b). As a result, the total heat flux to the surface is positively biased by around 100 W m^{-2} at short lead times and the bias varies between $50\text{--}150 \text{ W m}^{-2}$ at longer lead times (Figure 11c). For reference, observational uncertainty is on the order of 10 W m^{-2} for the turbulent flux measurements and 3% for the radiant fluxes (Fiedler et al., 2010). The bias in surface temperature was generally near zero, apart from a small positive bias for Flight 376 (Figure 11d). Conversely, the model near-surface air temperature was generally too cold with biases varying substantially (Figure 11e). Considering the small biases in surface temperature and turbulent heat fluxes, the errors in near-surface air temperature are likely not related to

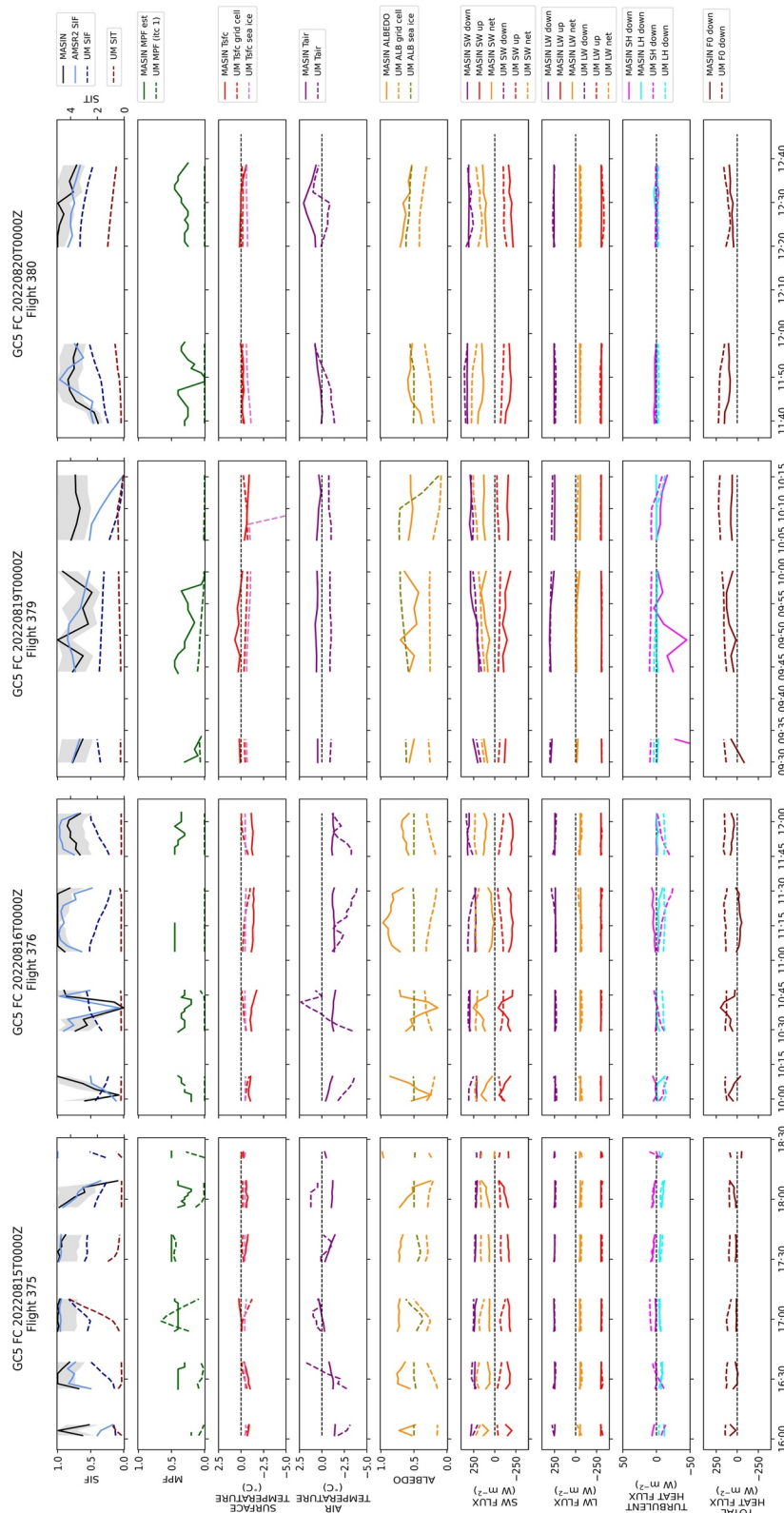


Figure 10. Comparison of observations (solid lines) and model output (dashed lines) for variables related to the Surface Energy Balance. Data are for Flights 375 (first column), 376 (second column), 379 (third column) and 380 (fourth column), and model outputs are from forecasts initialized at 00 UTC 15/08/2022 ($T + 16-18.5$ hr), 00 UTC 16/08/2022 ($T + 10-12$ hr), 00 UTC 19/08/2022 ($T + 9.5-10.25$ hr) and 00UTC 20/08/2022 ($T + 11.5-12.75$ hr) respectively. Panels show (as labeled): sea-ice fraction estimated from the aircraft (black with shading showing observational uncertainty), from AMSR2 (light blue) and from GC5 (dark blue), with sea-ice thickness (dark red, no measurements); melt pond fraction (green); surface temperature (red) and sea-ice surface temperature from GC5 (pink); near-surface air temperature (purple); average surface albedo derived from SW radiative fluxes (yellow) and sea-ice albedo (olive green); SW radiative fluxes down (purple), up (red) and net (yellow); LW radiative fluxes down (purple), up (red) and net (yellow); sensible (magenta) and latent (cyan) downward turbulent heat fluxes; and total downward heat flux (brown) calculated from the above components as described in the text.

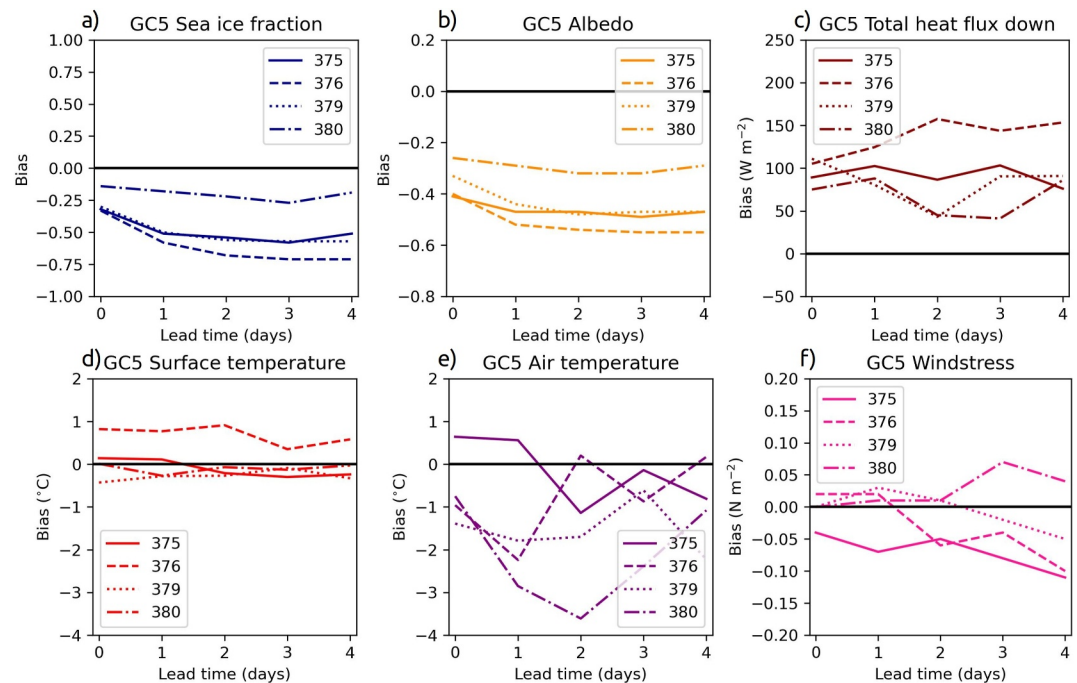


Figure 11. Summary of biases for selected SEB variables as a function of lead time evaluated using the four research flights and for GC5 output (10 simulations). Panels show (a) sea-ice fraction; (b) melt-pond fraction; (c) surface albedo (ratio of incoming and outgoing shortwave); and (d) total heat flux down.

local surface interactions but to advection. Surface wind stress bias is generally small, albeit growing with lead time (Figure 11f).

Further details of the average model bias and RMSE across the four ASC flights is provided in Table 2. With increasing lead time, there is generally a growth in bias and RMSE for sea-ice fraction, melt-pond fraction, albedo and F_{LH} . Total heat flux error is toward 100 W m^{-2} and changes little with lead time. F_{SH} bias hovers around zero. Surface temperature was generally accurate, though bias improves with lead time as sea ice declines. This could be expected as measured surface temperature was typically around 0°C over both ice and open water, which makes error diagnosis difficult. Likewise, the fact that F_{SH} and F_{LH} were generally simulated well is likely linked to their small magnitude due to low winds and temperature gradient. In contrast, near-surface air temperature was less accurate; the average error is smallest at the shortest lead times and more variable at longer lead times, probably linked to the fidelity of the forecast cyclone. The simulated surface wind stress agrees well with flight observations. Further understanding of the sources of error requires evaluation of the simulated atmospheric boundary layer and of the passing cyclone. Evaluation of simulated mesoscale features, such as low-level jets, is a subject of ongoing investigation but outside the scope of this paper.

Table 2

Summary of Bias and Root-Mean-Square-Error (RMSE, in Brackets) With Lead Time for a Selection of Key SEB Variables, Evaluated for the Four ASC Research Flights and Averaged

Bias (RMSE)	SIF	MPF	Albedo	Total HFX (W m^{-2})	SHFX (W m^{-2})	LHFX (W m^{-2})	Ta ($^\circ\text{C}$)	Ts ($^\circ\text{C}$)	Wind-stress (N m^{-2})
T+0 d	−0.27 (0.33)	−0.24 (0.28)	−0.35 (0.37)	95.15 (105.15)	4.81 (17.47)	−2.69 (5.25)	−0.62 (1.53)	0.14 (0.58)	−0.01 (0.07)
T+1 d	−0.44 (0.48)	−0.18 (0.30)	−0.43 (0.45)	98.93 (110.60)	2.34 (17.69)	−2.89 (6.23)	−1.58 (2.28)	0.08 (0.56)	0.00 (0.08)
T+2 d	−0.50 (0.54)	−0.28 (0.31)	−0.45 (0.47)	83.11 (100.21)	−1.82 (17.97)	−4.93 (6.73)	−1.56 (2.33)	0.09 (0.59)	−0.02 (0.08)
T+3 d	−0.53 (0.57)	−0.30 (0.32)	−0.46 (0.48)	94.66 (104.17)	1.88 (15.56)	−5.47 (7.42)	−1.00 (1.66)	−0.04 (0.54)	−0.02 (0.10)
T+4 d	−0.50 (0.54)	−0.30 (0.33)	−0.45 (0.47)	101.60 (110.43)	−1.30 (16.13)	−4.44 (7.06)	−0.99 (1.87)	−0.01 (0.62)	−0.06 (0.11)

Note. The number of data points (150-s means or $\sim 9 \text{ km}$) for each flight are: 30, 33, 15, and 17 for flights 375, 376, 379, and 380 respectively (95 in total).

It is worth remembering that the SEB evaluation in this section is based on four flights, so a relatively small sample over a small area, and one where the model sea-ice errors are largest. That said, our aircraft observations allow a detailed assessment of the model's performance in this challenging region and pinpoint the biases in SEB one may expect when sea-ice representation is lacking fidelity.

7. Discussion

The most commonly used measure of sea-ice coverage is the sea-ice extent (SIE), as its insensitivity to concentration enables consistency across measurement platforms and over the seasonal cycle. However, this study illustrates that evaluating sea ice by SIE alone provides an incomplete picture. For short-term forecasting, and particularly a regional focus, using sea-ice area (SIA) is advantageous due to the dependence of physical processes on sea-ice fraction. Examining both provides a broad picture and retains the key spatial information on sea-ice fraction that is critical to ocean-atmosphere exchange as illustrated by Notz (2014) for climate model evaluation. In our case study period, the bias in simulated SIA is much larger than that for SIE, and the distribution of that bias points to where model deficiencies lie and what processes may be responsible. We find the primary issue is the significant low bias in sea-ice fraction on model initialization for both GC5 and GC4. The low bias leads to a reduced surface albedo and too much heat flux into the ocean, which results in further basal sea-ice melt; that is, the ice-albedo feedback mechanism amplifies the initial bias substantially, even over these 5-day forecasts. This highlights the importance of accurate initial conditions and the representation of physical processes that contribute to feedbacks in models.

The sea-ice initialization bias is partly due to overly smooth gradients in the OSISAF-401 sea-ice fraction product and the lack of a representative ice edge, as shown by Renfrew et al. (2021); and a low concentration bias during summertime melt conditions when there are melt ponds. The position of the ice edge on initialization was also found to be a large source of error in coupled wintertime NWP simulations by Day et al. (2022). Melt ponds increase the difficulty of measuring sea-ice fraction, as it is difficult for passive microwave sensors to differentiate between them and open water. This can result in satellites actually measuring the “sea-ice surface fraction”—the fraction of ice after subtracting both leads and melt ponds—instead of the sea-ice fraction (which includes melt ponds). This summertime satellite retrieval issue has been investigated in several recent studies. Kern et al. (2016) examined the sensitivity of a range of sea-ice fraction retrieval algorithms to melt-pond fraction (using data where the sea-ice fraction from MODIS was greater than 90%) and found that the version of the Bootstrap and the Bristol algorithms used by OSI SAF were the most sensitive to melt ponds. For example, the Bootstrap algorithm underestimated sea-ice fraction by 14% (for 100% sea ice with 40% melt ponds). Zhao et al. (2021) found that SSMI-based sea-ice fraction products using the OSI SAF retrieval algorithms during high melt-pond fractions underestimated ice concentration by 20%, with error particularly large in the marginal ice zone. Kern et al. (2020) and Song and Minnett (2024) showed that while the OSI SAF algorithm underestimated sea-ice fraction during summer melt conditions (by 6%), the algorithm used to generate the AMSR2 data employed within this study tended to overestimate sea-ice fraction (by 1%). Kern et al. (2020) commented, “*we do not know what the algorithms actually measure in summer (actual sea-ice concentration or net sea-ice surface fraction), and whichever they measure the accuracy is poor compared to winter conditions.*” They go on to suggest that measuring “net surface ice fraction” (not including ponded ice) would suit current passive microwave instruments better, which could be combined with separate melt pond fraction data sets using visible-to-near-infrared satellite instruments to generate accurate summer sea ice fraction data. Reducing summertime bias in retrieval algorithms could provide a substantial benefit to operational forecasting in the Arctic and should be a priority for the community.

Clearly, these summertime sea-ice retrieval challenges make the evaluation of numerical models very difficult. It is a reason that many studies prefer to use sea-ice extent rather than area; however, as we have demonstrated, only examining extent could give a misleading impression of model skill. Our results suggest that substantial improvements in summertime sea-ice simulation at NWP timescales could be achieved from a better utilisation of available sea-ice products at the data assimilation stage.

Observations from the IceBird campaign show that the simulated sea-ice thickness in GC5 is reasonably accurate, albeit biased by between -0.5 and -1.4 m. Simulations evaluated over a 50-year period suggest that GC5 produces thicker sea ice than GC3.1, closer to the PIOMAS reanalysis, due to changes in sea-ice dynamics that result from different atmosphere-ice drag schemes (Blockley et al., 2024). The 5-day forecasts we evaluate here

have the same atmosphere-ice drag scheme (Renfrew, Elvidge et al., 2019). Note we find no practical difference in ice thickness change between GC5 and GC4, remembering that the initial sea-ice fields are from the same source. We found the largest model sea-ice thickness errors to be in the MIZ near the Greenland coast, where the IceBird measured thicker ice, likely due to convergence and rafting. However, the negative bias in the models is likely due to the excessive ice melt in this region as discussed above.

The ice thickness distribution plays an important role in determining ice melt. In the topographic melt pond scheme, melt ponds form preferentially on thin, level ice, reducing albedo and increasing melt rate. Thinner ice also has less thermal inertia and allows more solar energy to pass through, warming the ocean quicker and leading to more rapid melt. Given the smaller volume of the ice where ice is thinner, these processes quickly lead to open water, which further increases energy absorption at the surface in a positive feedback process. In GC4 solar radiation does not penetrate the sea ice, an adjustment to surface albedo is applied to represent internal scattering based on Semtner (1976); this lack of penetrating radiation has been shown to result in high levels of melting at the top ice surface (Keen et al., 2021; West & Blockley, 2024). Thus, for GC5, the Semtner scheme has been modified to calculate the amount of penetrating solar radiation and pass this to the sea ice model (Blockley et al., 2024). Below a thickness of 0.3 m, this formulation adjusts the surface albedo based on the ice thickness, to take account of the internal scattering of shortwave radiation in the visible part of the spectrum; the remaining shortwave radiation is then passed to the sea-ice model via the coupler. For example, while bare ice >0.3 m thick has a prescribed albedo of 0.78 (in the visible spectrum), if the ice thickness is reduced to 0.2 m or 0.1 m the visible albedo is reduced to 0.71 or 0.52 respectively (calculated offline). This approximation was shown by Semtner (1976) to produce a positive bias in ice thickness in winter and a negative bias in summer; West and Blockley (2024) show that too much ice growth in winter and too much ice melt in summer is common in CMIP6 models. However, sensitivity experiments that effectively switch this Semtner scheme on and off in the most recent SI³ configuration suggest that the impact is minimal owing to the relative scarcity of thin, bare ice in the model, which is generally either covered by snow or pond. The amount of solar radiation penetration and heat conduction through the ice to the ocean was small in our simulations in the location of our flights, however, they are not always insignificant so should be a consideration in model development.

While we comment extensively on the substantial bias in model surface albedo, it must be remembered that we are comparing the measured/modeled broadband shortwave radiation. We have not been able to go into detail about the component radiation bands, or influence of surface characteristics (e.g., ice thickness, snow depth, melt-pond depth) as we don't have sufficiently detailed in situ observations. Nonetheless, the error we find in the surface albedo must be due to some combination of the fractions and albedo parameters of bare ice, melt ponds, and snow. Melt-pond fraction was simulated reasonably well when the sea-ice fraction was accurate; however, this was only in a handful of data points; we have limited observations to analyze this issue in more detail. In any case, the lack of shortwave radiation being reflected upwards due to low sea-ice fraction is more critical in these simulations. Further investigation, with more data over continuous sea ice, will be required to evaluate the parameterization of sea-ice albedo more robustly.

Our two MASIN flights north of Greenland are situated at a hot spot for very rapid model sea ice loss giving us the opportunity to interrogate model processes related to the sea ice mass budget. Strong melting from basal heat fluxes is driven by the low sea-ice fraction on initialization, leading to strong ocean heat uptake from insolation. This thermodynamic change is compounded by the dynamic forcing due to strong winds as a cyclone moves past, which drives ice northwards away from the flight areas. The simulated wind stress agrees well with observations, and at two separate sites further north the sea-ice speed also agrees well with buoy observations, implying that dynamical processes are generally well simulated. This adds weight to our suggestion that thermodynamic processes are the primary driver of model discrepancies. The two flights further south in Fram Strait corroborate that thermodynamic processes are the primary source of error, as the dynamic ice mass tendencies are either around neutral or positive at these locations, while ice is still lost too quickly.

We have demonstrated that the heat flux at the base of the sea ice is a critical term, and this is dependent on the upper ocean temperature. Unfortunately, where sea ice is present there are usually no ocean temperature data available for assimilation. So, it is plausible the mixed-layer ocean is simply initialized too warm—we don't have ocean observations to check for these cases. One issue that was identified in an earlier version of the Met Office Coupled Model, GC2, is “initialization shock” and a sudden ice melt at the first few time steps caused by relatively warm under-ice ocean temperatures that result from differences in physics between the analysis system and

the coupled model (Barrell, 2023; Johns et al., 2021). Improvements have been made to minimize initialization shock in more recent GC versions; however, the intrinsic problem remains and we think the strong thermodynamic ice melt in early forecast timesteps seen here suggests this warrants further attention.

It is well documented that Arctic cyclones can play a role in rapid ice loss events, such as the “Great Arctic Cyclone of August 2012” (Simmonds & Rudeva, 2012; Zhang et al., 2013) and in ice loss more generally (Cavallo et al., 2025; Finocchio et al., 2020). Our findings suggest that compounding thermodynamic and dynamic model biases affected ice loss during the simulated cyclone case, particularly in the ice export region of Fram Strait. Despite the rapid sea-ice melt, many aspects of the Arctic cyclone are qualitatively simulated well, for example, the rate of simulated sea-ice loss reached near zero during the rebound in satellite-measured sea-ice area in the lee of the cyclone (Figure 3). This retreat in sea ice poleward followed by an advance southward, driven by the cyclone's winds, is in line with the cyclone composite findings of Clancy et al. (2022). In addition, there is good agreement between simulated and observed sea-ice drift speeds (Figure 9) and surface windstress (Figure 11f). That said, the wind was generally weak during the broader case study period, apart from a low-level mesoscale jet northeast of Greenland. The limited measurements we have of this jet suggest it was not strong enough in the simulations and at an incorrect angle, which may compound the errors in the sea ice distribution due to melting, again this is a subject for future investigations.

More generally, our case study investigation provides a comprehensive analysis of a typical late summer Arctic cyclone as described by Finocchio et al. (2020, 2022). In common with some of their findings, we find an accelerated ice loss from the initial poleward warm-air intrusion, reduced surface albedo due to previous melting, sea ice divergence, and clear skies in the lee of the cyclone with strong shortwave insolation. We illustrate that this combination of forcings and their associated feedbacks remains a challenge for models. In keeping with our findings, Stern et al. (2020) found that their coupled NWP system (Navy-ESPC) also demonstrated strong basal sea-ice melt in a simulation of the strong Arctic cyclone in 2012 due to turbulent entrainment of relatively warm water at 15–35 m depth. It is clear that further investigations, likely using several modeling systems, should consider the role and accuracy of ocean temperature, stratification, mixed layer depth, sub-grid-scale mixing, salinity and ice base roughness in coupled NWP in the summertime Arctic.

Thinking more broadly, the examination of summertime Arctic cyclone events via detailed case studies using coupled models could help explain summertime biases in sea ice simulation within climate models. Systematic model biases causing sea ice melt that is too fast in the summertime reduces confidence in the predicted timings of an ice-free Arctic that are frequently discussed by the climate community. However, understanding systematic model biases also provides an avenue for reducing the large uncertainty in predictions of an ice-free Arctic as discussed by Jahn et al. (2024) and others. Further work is needed to understand these biases in more detail and the root causes of overestimated ice loss in climate simulations.

8. Conclusions

We have evaluated coupled NWP simulations made using the Met Office's GC4 and GC5 model configurations by comparing with case study observations from northeast of Greenland and Fram Strait during summertime sea-ice melt conditions and under the influence of an Arctic cyclone. Our focus has been on determining the causes of excessive sea-ice retreat during these forecasts, compared to satellite products and our in situ aircraft observations.

We highlight the limitations of using sea-ice extent for model evaluation. Although this metric is commonly used, during these summertime conditions it disguises biases in the sea-ice fraction field that are critical for the representation of atmosphere-ice-ocean exchange. We demonstrate that using sea-ice area (or fraction) can be more informative for short-range and regional sea ice verification, particularly in summertime.

Following a comprehensive evaluation of the simulated sea ice and surface energy budget, we find that the main source of forecast error is a substantial low bias in the model's sea-ice fraction on initialization that becomes exacerbated over time. We find that the sea-ice drift speed is simulated well, and the sea-ice thickness agrees reasonably well with observations away from the ice edge, but poorly near the ice edge due to the rapid ice melt. The OSISAF-401 sea-ice product i.e. assimilated suffers from reduced accuracy during melting conditions (e.g., Kern et al., 2020; Zhao et al., 2021) and near the ice edge (Renfrew et al., 2021). Here, a low sea-ice fraction triggers the ice-albedo feedback mechanism, leading to substantial heat uptake in the upper ocean and rapid ice

loss even over 5-day forecasts. This has direct consequences for the prediction of surface-layer meteorology and fluxes. Such conditions present a significant challenge for operational forecasting in the Arctic, particularly around the marginal ice zone where shipping and economic activities are increasing in summer as the sea ice retreats. Our results suggest that improving sea-ice data assimilation is a must for improving summertime Arctic environmental prediction.

A key aspect of this study is that we evaluate sea-ice melt during peak melt season, when feedback processes are very active. A critical factor in the ice-albedo feedback process is the rate of change of surface albedo (e.g., Winton, 2008), which is mostly dependent on sea-ice fraction in summer. Our results demonstrate nonlinear amplification of sea-ice melt, due to the combined effects of the initial sea-ice fraction bias and over-active model processes, which compromise predictability over just a few days. It is possible that the issues we describe here in 5-day forecasts are linked to the common issue of accelerated summertime sea-ice melt in climate models due to the dependence on the same physical equations. West and Blockley (2024) showed that CMIP6 models tend to overestimate summertime sea-ice melt. We show corroborating results using versions of the CICE and SI³ sea ice models that are used in similar form in many CMIP6 climate models. Addressing this possible model weakness is underway.

Despite the problems noted, the implementation of the SI³ model in GC5 is successful in that it provides some improvement in the simulation of sea ice in the Arctic basin and Fram Strait; including providing a skilful representation of atmosphere-ice-ocean exchange where sea-ice fraction is accurate. Continued model development is in progress, including improvements in ocean and sea-ice data assimilation (Mignac et al., 2024). However, further work should seek to integrate more accurate sea-ice fraction observations into the data assimilation process, assess the simulation of the upper ocean and ocean-ice heat exchange and carry out a thorough review of the veracity of the ice surface albedo and melt-pond schemes. To note we have ongoing work assessing the sensitivity of sea-ice melt to specific thermodynamic and dynamic processes, with the aim of highlighting potential model improvements that should hone the representation of key processes in the coupled atmosphere-ice-ocean system.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Data Availability Statement

The Arctic Summertime Cyclones MASIN flight data is archived at the Centre for Environmental Data Analysis (CEDA) at <https://data.ceda.ac.uk/badc/arcticcyclones>. Met Office simulations are archived on the Met Office MASS archive and can be shared on request. IceBird summer 2022 campaign data is archived on PANGAEA (Krumpfen et al., 2024). All processed data and code used to produce the figures in this article are available online from Barrell (2025).

Acknowledgments

We thank the Arctic Summertime Cyclones project team funded by the Natural Environment Research Council (NERC) (NE/T006773/1, NE/T006811/1, NE/T00682X/1), the operations staff of the British Antarctic Survey Twin Otter aircraft, particularly Dan Beeden for facilitating the visit to Station Nord, and the RALI-THINICE project team for their generous research cooperation. We also thank the Alfred Wegener Institute IceBird survey team for sharing their observations. EWB acknowledges funding from the Met Office Advancing Arctic meteorological and oceanographic Capabilities & services project, which is supported by the Department for Science, Innovation & Technology (DSIT). We thank the reviewers for their constructive feedback.

References

- Barbosa Aguiar, A., Bell, M. J., Blockley, E., Calvert, D., Crocker, R., Inverarity, G., et al. (2024). The Met Office Forecast Ocean Assimilation Model (FOAM) using a 1/12-degree grid for global forecasts. *Quarterly Journal of the Royal Meteorological Society*, 150(763), 3827–3852. <https://doi.org/10.1002/QJ.4798>
- Barrell, C. (2023). *Air-sea interaction and the influence of sea ice during cold-air outbreaks over the subpolar North Atlantic Ocean*. Doctoral thesis. University of East Anglia. Retrieved from <https://ueaeprints.uea.ac.uk/id/eprint/97235/>
- Barrell, C. (2025). Rapid summertime sea ice melt in a coupled numerical weather prediction system [Dataset]. *Journal of Advances in Modeling Earth Systems*. Zenodo. <https://doi.org/10.5281/zenodo.15635774>
- Best, M. J., Pryor, M., Clark, D. B., Rooney, G. G., Essery, R. L. H., Ménard, C. B., et al. (2011). The Joint UK Land Environment Simulator (JULES), model description—Part 1: Energy and water fluxes. *Geoscientific Model Development*, 4(3), 677–699. <https://doi.org/10.5194/GMD-4-677-2011>
- Blanchard-Wrigglesworth, E., Armour, K. C., Bitz, C. M., & Deweaver, E. (2011). Persistence and inherent predictability of Arctic Sea Ice in a GCM ensemble and observations. *Journal of Climate*, 24(1), 231–250. <https://doi.org/10.1175/2010JCLI3775.1>
- Blanchard-Wrigglesworth, E., & Bushuk, M. (2019). Robustness of Arctic sea-ice predictability in GCMs. *Climate Dynamics*, 52(9–10), 5555–5566. <https://doi.org/10.1007/s00382-018-4461-3>
- Blockley, E., Fiedler, E., Ridley, J., Roberts, L., West, A., Copsey, D., et al. (2024). The sea ice component of GC5: Coupling SI³ to HadGEM3 using conductive fluxes. *Geoscientific Model Development*, 17(17), 6799–6817. <https://doi.org/10.5194/GMD-17-6799-2024>
- Blockley, E. W., Martin, M. J., McLaren, A. J., Ryan, A. G., Waters, J., Lea, D. J., et al. (2014). Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new Global FOAM forecasts. *Geoscientific Model Development*, 7(6), 2613–2638. <https://doi.org/10.5194/GMD-7-2613-2014>

- Brooks, I. M., Tjernström, M., Persson, P. O. G., Shupe, M. D., Atkinson, R. A., Canut, G., et al. (2017). The turbulent structure of the Arctic summer boundary layer during the arctic summer cloud-ocean study. *Journal of Geophysical Research: Atmospheres*, 122(18), 9685–9704. <https://doi.org/10.1002/2017JD027234>
- Bushuk, M., Ali, S., Bailey, D. A., Bao, Q., Batté, L., Bhatt, U. S., et al. (2024). Predicting September Arctic Sea Ice: A multimodel seasonal skill comparison. *Bulletin of the American Meteorological Society*, 105(7), E1170–E1203. <https://doi.org/10.1175/BAMS-D-23-0163.1>
- Bushuk, M., Winton, M., Bonan, D. B., Blanchard-Wrigglesworth, E., & Delworth, T. L. (2020). A mechanism for the Arctic Sea Ice spring predictability barrier. *Geophysical Research Letters*, 47(13), e2020GL088335. <https://doi.org/10.1029/2020GL088335>
- Bushuk, M., Zhang, Y., Winton, M., Hurlin, B., Delworth, T., Lu, F., et al. (2022). Mechanisms of regional Arctic Sea Ice predictability in two dynamical seasonal forecast systems. *Journal of Climate*, 35(13), 4207–4231. <https://doi.org/10.1175/JCLI-D-21-0544.1>
- Cavallo, S. M., Frank, M. C., & Bitz, C. M. (2025). Sea ice loss in association with Arctic cyclones. *Communications Earth & Environment*, 6(1), 1–9. <https://doi.org/10.1038/s43247-025-02022-9>
- Clancy, R., Bitz, C. M., Edward, B. W., McGraw, M. C., & Cavallo, S. M. (2022). A cyclone-centered perspective on the drivers of asymmetric patterns in the atmosphere and sea ice during Arctic Cyclones. *Journal of Climate*, 35(1), 73–89. <https://doi.org/10.1175/JCLI-D-21-0093.1>
- Craig, A., Valcke, S., & Coquart, L. (2017). Development and performance of a new version of the OASIS coupler, OASIS-MCT-3.0. *Geoscientific Model Development*, 10(9), 3297–3308. <https://doi.org/10.5194/GMD-10-3297-2017>
- Croard, H. L., Methven, J., Harvey, B., Keeley, S. P. E., Volonté, A., & Hodges, K. I. (2023). A climatology of summer-time arctic cyclones using a modified phase space. *Geophysical Research Letters*, 50(22), e2023GL105993. <https://doi.org/10.1029/2023GL105993>
- Day, J. J., Hawkins, E., & Tietsche, S. (2014). Will Arctic sea ice thickness initialization improve seasonal forecast skill? *Geophysical Research Letters*, 41(21), 7566–7575. <https://doi.org/10.1002/2014GL061694>
- Day, J. J., Keeley, S., Arduini, G., Magnusson, L., Mogensen, K., Rodwell, M., et al. (2022). Benefits and challenges of dynamic sea ice for weather forecasts. *Weather and Climate Dynamics*, 3(3), 713–731. <https://doi.org/10.5194/WCD-3-713-2022>
- Elvidge, A. D., Renfrew, I. A., Brooks, I. M., Srivastava, P., Yelland, M. J., & Prytherch, J. (2021). Surface heat and moisture exchange in the marginal ice zone: Observations and a new parameterization scheme for weather and climate models. *Journal of Geophysical Research: Atmospheres*, 126(17), e2021JD034827. <https://doi.org/10.1029/2021JD034827>
- Elvidge, A. D., Renfrew, I. A., Edwards, J. M., Brooks, I. M., Srivastava, P., & Weiss, A. I. (2023). Improved simulation of the polar atmospheric boundary layer by accounting for aerodynamic roughness in the parameterization of surface scalar exchange over sea ice. *Journal of Advances in Modeling Earth Systems*, 15(3), e2022MS003305. <https://doi.org/10.1029/2022MS003305>
- Elvidge, A. D., Renfrew, I. A., Weiss, A. I., Brooks, I. M., Lachlan-Cope, T. A., & King, J. C. (2016). Observations of surface momentum exchange over the marginal ice zone and recommendations for its parametrisation. *Atmospheric Chemistry and Physics*, 16(3), 1545–1563. <https://doi.org/10.5194/acp-16-1545-2016>
- EUMETSAT Ocean and Sea Ice Satellite Application Facility (OSI SAF). (2005). Global Sea Ice Concentration (SSMIS), OSI-401-d. https://doi.org/10.15770/EUM_SAF_OSI_NRT_2004
- Fairall, C. W., Bradley, E. F., Hare, J. E., Grachev, A. A., & Edson, J. B. (2003). Bulk parameterization of air-sea fluxes: Updates and verification for the COARE algorithm. [https://doi.org/10.1175/1520-0442\(2003\)016<0571:BPOASF>2.0.CO;2](https://doi.org/10.1175/1520-0442(2003)016<0571:BPOASF>2.0.CO;2)
- Fearon, M. G., Doyle, J. D., Ryglicki, D. R., Finocchio, P. M., & Sprenger, M. (2021). The role of cyclones in moisture transport into the Arctic. *Geophysical Research Letters*, 48(4), e2020GL090353. <https://doi.org/10.1029/2020GL090353>
- Fiedler, E. K., Lachlan-Cope, T. A., Renfrew, I. A., & King, J. C. (2010). Convective heat transfer over thin ice covered coastal polynyas. *Journal of Geophysical Research*, 115(C10), 10051. <https://doi.org/10.1029/2009JC005797>
- Finocchio, P. M., Doyle, J. D., & Stern, D. P. (2022). Accelerated sea ice loss from late summer cyclones in the new Arctic. *Journal of Climate*, 35(23), 7751–7769. <https://doi.org/10.1175/JCLI-D-22-0315.1>
- Finocchio, P. M., Doyle, J. D., Stern, D. P., & Fearon, M. G. (2020). Short-term impacts of Arctic summer cyclones on sea ice extent in the marginal ice zone. *Geophysical Research Letters*, 47(13), e2020GL088338. <https://doi.org/10.1029/2020GL088338>
- Flocco, D., & Feltham, D. L. (2007). A continuum model of melt pond evolution on Arctic sea ice. *Journal of Geophysical Research*, 112(C8). <https://doi.org/10.1029/2006JC003836>
- Goessling, H. F., Tietsche, S., Day, J. J., Hawkins, E., & Jung, T. (2016). Predictability of the Arctic sea ice edge. *Geophysical Research Letters*, 43(4), 1642–1650. <https://doi.org/10.1002/2015GL067232>
- Gryschka, M., Drüe, C., Etling, D., & Raasch, S. (2008). On the influence of sea-ice inhomogeneities onto roll convection in cold-air outbreaks. *Geophysical Research Letters*, 35(23), 23804. <https://doi.org/10.1029/2008GL035845>
- Guiavarch, C., Storkey, D., Blaker, A. T., Blockley, E., Megann, A., Hewitt, H. T., et al. (2024). GOSI9: UK global ocean and sea ice configurations. <https://doi.org/10.5194/EGUSPHERE-2024-805>
- Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., & Elliott, S. (2015). *CICE: The Los Alamos Sea ice model documentation and software user's manual version 5.1* LA-CC-06-012. T-3 Fluid Dynamics Group (Vol. 675, p. 15). Los Alamos National Laboratory. Retrieved from <http://oceans11.lanl.gov/trac/CICE/wiki>
- Jahn, A., Holland, M. M., & Kay, J. E. (2024). Projections of an ice-free Arctic Ocean. *Nature Reviews Earth & Environment*, 5(3), 164–176. <https://doi.org/10.1038/s43017-023-00515-9>
- Jahn, A., Kay, J. E., Holland, M. M., & Hall, D. M. (2016). How predictable is the timing of a summer ice-free Arctic? *Geophysical Research Letters*, 43(17), 9113–9120. <https://doi.org/10.1002/2016GL070067>
- Johns, T. C., Blockley, E. W., & Ridley, J. K. (2021). Causes and consequences of sea ice initialization shock in coupled NWP Hindcasts with the GC2 climate model. *Monthly Weather Review*, 149(7), 2239–2254. <https://doi.org/10.1175/MWR-D-20-0341.1>
- Jung, T., Gordon, N. D., Bauer, P., Bromwich, D. H., Chevallier, M., Day, J. J., et al. (2016). Advancing polar prediction capabilities on daily to seasonal time scales. *Bulletin of the American Meteorological Society*, 97(9), 1631–1647. <https://doi.org/10.1175/BAMS-D-14-00246.1>
- Keeley, S., & Mogensen, K. (2018). Dynamic sea ice in the IFS. *ECMWF Newsletter*, 156, 23–29. <https://doi.org/10.21957/aska25furb>
- Keen, A., Blockley, E., Bailey, D. A., Boldingh Debernard, J., Bushuk, M., Delhay, S., et al. (2021). An inter-comparison of the mass budget of the Arctic sea ice in CMIP6 models. *The Cryosphere*, 15(2), 951–982. <https://doi.org/10.5194/TC-15-951-2021>
- Kern, S., Laverne, T., Notz, D., Toudal Pedersen, L., & Tonboe, R. (2020). Satellite passive microwave sea-ice concentration data set inter-comparison for Arctic summer conditions. *The Cryosphere*, 14(7), 2469–2493. <https://doi.org/10.5194/TC-14-2469-2020>
- Kern, S., Rösel, A., Toudal Pedersen, L., Ivanova, N., Saldo, R., & Tage Tonboe, R. (2016). The impact of melt ponds on summertime microwave brightness temperatures and sea-ice concentrations. *The Cryosphere*, 10(5), 2217–2239. <https://doi.org/10.5194/TC-10-2217-2016>
- King, J. C., Lachlan-Cope, T. A., Ladkin, R. S., & Weiss, A. (2008). Airborne measurements in the stable boundary layer over the Larsen Ice Shelf, Antarctica. *Boundary-Layer Meteorology*, 127(3), 413–428. <https://doi.org/10.1007/S10546-008-9271-4/METRICS>
- Kriegsmann, A., & Brümmer, B. (2014). Cyclone impact on sea ice in the central Arctic Ocean: A statistical study. *The Cryosphere*, 8(1), 303–317. <https://doi.org/10.5194/TC-8-303-2014>

- Krumpen, T., Birnbaum, G., Ludwig, V., & Petersen, C. (2022). IceBird summer 2022 campaign—Sea ice surveys with polar 6 from station Nord. *Campaign Report*. Retrieved from <https://hdl.handle.net/10013/epic.563637c4-3648-41e6-a033-e2ac7560582e>
- Krumpen, T., Haas, C., & Hendricks, S. (2024). Laser altimeter data obtained over sea-ice during the IceBird Summer campaign in Aug, 2022 [Dataset]. *PANGAEA*. <https://doi.org/10.1594/PANGAEA.970711>
- Krumpen, T., von Albedyll, L., Bünger, H. J., Castellani, G., Hartmann, J., Helm, V., et al. (2025). Smoother sea ice with fewer pressure ridges in a more dynamic Arctic. *Nature Climate Change*, 15(1), 66–72. <https://doi.org/10.1038/s41558-024-02199-5>
- Lei, R., Gui, D., Hutchings, J. K., Heil, P., & Li, N. (2020). Annual cycles of sea ice motion and deformation derived from buoy measurements in the western Arctic Ocean over two ice seasons. *Journal of Geophysical Research: Oceans*, 125(6), e2019JC015310. <https://doi.org/10.1029/2019JC015310>
- Liu, A. Q., Moore, G. W. K., Tsuboki, K., & Renfrew, I. A. (2006). The effect of the sea-ice zone on the development of boundary-layer roll clouds during cold air outbreaks. *Boundary-Layer Meteorology*, 118(3), 557–581. <https://doi.org/10.1007/S10546-005-6434-4>
- Lukovich, J. v., Stroeve, J. C., Crawford, A., Hamilton, L., Tsamados, M., Heorton, H., & Massonnet, F. (2021). Summer extreme cyclone impacts on Arctic sea ice. *Journal of Climate*, 34(12), 4817–4834. <https://doi.org/10.1175/JCLI-D-19-0925.1>
- Madec, G., & the NEMO team. (2017). NEMO ocean engine—Version 3.6. Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL) No 27, ISSN No 1288-1619. <https://doi.org/10.5281/zenodo.3248739>
- Madec, G., & the NEMO team. (2019). NEMO ocean engine—Version 4.0.1. Notes du Pôle de modélisation de l'Institut Pierre-Simon Laplace (IPSL) No 27, ISSN No 1288-1619. <https://doi.org/10.5281/zenodo.3878122>
- McGraw, M. C., Blanchard-Wrigglesworth, E., Clancy, R. P., & Bitza, C. M. (2022). Understanding the forecast skill of rapid Arctic sea ice loss on subseasonal time scales. *Journal of Climate*, 35(4), 1179–1196. <https://doi.org/10.1175/JCLI-D-21-0301.1>
- Meier, W. N., Hovelsrud, G. K., van Oort, B. E. H., Key, J. R., Kovacs, K. M., Michel, C., et al. (2014). Arctic sea ice in transformation: A review of recent observed changes and impacts on biology and human activity. *Reviews of Geophysics*, 52(3), 185–217. <https://doi.org/10.1002/2013RG000431>
- Mignac, D., Waters, J., Lea, D. J., Martin, M. J., While, J., Weaver, A. T., et al. (2024). Updates to the Met Office's global ocean-sea ice forecasting system including model and data assimilation changes. <https://doi.org/10.5194/EGUSPHERE-2024-3143>
- Müller, M., Batrak, Y., Dinessen, F., Grote, R., & Wang, K. (2023). Challenges in the description of sea ice for a kilometer-scale weather forecasting system. *Weather and Forecasting*, 38(7), 1157–1171. <https://doi.org/10.1175/WAF-D-22-0134.1>
- Notz, D. (2014). Sea-ice extent and its trend provide limited metrics of model performance. *The Cryosphere*, 8(1), 229–243. <https://doi.org/10.5194/TC-8-229-2014>
- Parker, C. L., Mooney, P. A., Webster, M. A., & Boisvert, L. N. (2022). The influence of recent and future climate change on spring Arctic cyclones. *Nature Communications*, 13(1), 1–14. <https://doi.org/10.1038/s41467-022-34126-7>
- Peng, L., Zhang, X., Kim, J. H., Cho, K. H., Kim, B. M., Wang, Z., & Tang, H. (2021). Role of intense Arctic storm in accelerating summer sea ice melt: An in situ observational study. *Geophysical Research Letters*, 48(8), e2021GL092714. <https://doi.org/10.1029/2021GL092714>
- Persson, P. O. G. (2012). Onset and end of the summer melt season over sea ice: Thermal structure and surface energy perspective from SHEBA. *Climate Dynamics*, 39(6), 1349–1371. <https://doi.org/10.1007/s00382-011-1196-9>
- Petersen, G. N., & Renfrew, I. A. (2009). Aircraft-based observations of air-sea fluxes over Denmark Strait and the Irminger Sea during high wind speed conditions. *Quarterly Journal of the Royal Meteorological Society*, 135(645), 2030–2045. <https://doi.org/10.1002/QJ.355>
- Rae, J. G. L., Hewitt, H. T., Keen, A. B., Ridley, J. K., West, A. E., Harris, C. M., et al. (2015). Development of the global sea ice 6.0 CICE configuration for the Met Office Global Coupled model. *Geoscientific Model Development*, 8(7), 2221–2230. <https://doi.org/10.5194/GMD-8-2221-2015>
- Renfrew, I. A., Pickart, R. S., Våge, K., Moore, G. W. K., Bracegirdle, T. J., Elvidge, A. D., et al. (2019). The Iceland Greenland seas project. In *Bulletin of the American Meteorological Society* (Vol. 100, pp. 1795–1817). American Meteorological Society. <https://doi.org/10.1175/BAMS-D-18-0217.1>
- Renfrew, I. A., Barrell, C., Elvidge, A. D., Brooke, J. K., Duscha, C., King, J. C., et al. (2021). An evaluation of surface meteorology and fluxes over the Iceland and Greenland Seas in ERA5 reanalysis: The impact of sea ice distribution. *Quarterly Journal of the Royal Meteorological Society*, 147(734), 691–712. <https://doi.org/10.1002/QJ.3941>
- Renfrew, I. A., Elvidge, A. D., & Edwards, J. M. (2019). Atmospheric sensitivity to marginal-ice-zone drag: Local and global responses. *Quarterly Journal of the Royal Meteorological Society*, 145(720), 1165–1179. <https://doi.org/10.1002/QJ.3486>
- Ridley, J. K., Blockley, E. W., Keen, A. B., Rae, J. G. L., West, A. E., & Schroeder, D. (2018). The sea ice model component of HadGEM3-GC3.1. *Geoscientific Model Development*, 11(2), 713–723. <https://doi.org/10.5194/GMD-11-713-2018>
- Rivière, G., Delanoë, J., Doyle, J. D., Methven, J., Barrell, C., Fearon, M., et al. (2024). The THINICE field campaign: Interactions between arctic cyclones, tropopause polar vortices, clouds, and sea ice in summer. *Bulletin of the American Meteorological Society*, 105(12), E2330–E2354. <https://doi.org/10.1175/BAMS-D-23-0143.1>
- Semtner, A. J., Jr. (1976). A model for the thermodynamic growth of sea ice in numerical investigations of climate. *Journal of Physical Oceanography*, 6(3), 379–389. [https://doi.org/10.1175/1520-0485\(1976\)006<0379:AMFTTG>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0379:AMFTTG>2.0.CO;2)
- Sergeev, D. E., Renfrew, I. A., & Spengler, T. (2018). Modification of polar low development by orography and sea ice. *Monthly Weather Review*, 146(10), 3325–3341. <https://doi.org/10.1175/MWR-D-18-0086.1>
- Simmonds, I., & Rudeva, I. (2012). The great Arctic cyclone of August 2012. *Geophysical Research Letters*, 39(23). <https://doi.org/10.1029/2012GL054259>
- Song, K., & Minnett, P. J. (2024). Evaluation of summertime passive microwave and reanalysis sea-ice concentration in the central Arctic. *Earth and Space Science*, 11(1), e2023EA003214. <https://doi.org/10.1029/2023EA003214>
- Spensberger, C., & Spengler, T. (2021). Sensitivity of air-sea heat exchange in cold-air outbreaks to model resolution and sea-ice distribution. *Journal of Geophysical Research: Atmospheres*, 126(5), e2020JD033610. <https://doi.org/10.1029/2020JD033610>
- Spren, G., Kaleschke, L., & Heygster, G. (2008). Sea ice remote sensing using AMSR-E 89-GHz channels. *Journal of Geophysical Research*, 113(C2), 2–03. <https://doi.org/10.1029/2005JC003384>
- Stern, D. P., Doyle, J. D., Barton, N. P., Finocchio, P. M., Komaromi, W. A., & Metzger, E. J. (2020). The impact of an intense cyclone on short-term sea ice loss in a fully coupled atmosphere-ocean-ice model. *Geophysical Research Letters*, 47(4), e2019GL085580. <https://doi.org/10.1029/2019GL085580>
- Storkey, D., Blaker, A. T., Mathiot, P., Megann, A., Aksenov, Y., Blockley, E. W., et al. (2018). UK Global Ocean GO6 and GO7: A traceable hierarchy of model resolutions. *Geoscientific Model Development*, 11(8), 3187–3213. <https://doi.org/10.5194/GMD-11-3187-2018>
- Thorndike, A. S., Thorrock, D. A., Maykut, G. A., & Colony, R. (1975). The thickness distribution of sea ice. *Journal of Geophysical Research*, 80(33), 4501–4513. <https://doi.org/10.1029/JC080I033P04501>

- Tonboe, R., & Lavelle, J. (2016). The EUMETSAT OSI SAF sea ice concentration algorithm: Algorithm theoretical basis document documentation change record.
- Valkonen, E., Cassano, J., & Cassano, E. (2021). Arctic cyclones and their interactions with the declining sea ice: A recent climatology. *Journal of Geophysical Research: Atmospheres*, 126(12), e2020JD034366. <https://doi.org/10.1029/2020JD034366>
- Vancoppenolle, M., Rousset, C., & Blockley, E., & Group. (2023). SI3—Sea ice modelling integrated initiative—The NEMO sea ice engine, technical report. Retrieved from <https://zenodo.org/record/7534900>
- Vihma, T., Pirazzini, R., Fer, I., Renfrew, I. A., Sedlar, J., Tjernström, 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>
- Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., et al. (2019). The Met Office Unified model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. *Geoscientific Model Development*, 12(5), 1909–1963. <https://doi.org/10.5194/GMD-12-1909-2019>
- West, A. E., & Blockley, E. W. (2024). CMIP6 models overestimate sea ice melt, growth & conduction relative to ice mass balance buoy estimates. <https://doi.org/10.5194/GMD-2024-121>
- Winton, M. (2008). Sea ice–albedo feedback and nonlinear Arctic climate change. In E. T. DeWeaver, C. M. Bitz, & L.-B. Tremblay (Eds.), *Arctic Sea Ice Decline: Observations, Projections, Mechanisms, and Implications*. American Geophysical Union. <https://doi.org/10.1029/180GM09>
- Xavier, P., Willett, M., Graham, T., Earnshaw, P., Copsey, D., Narayan, N., et al. (2024a). Assessment of the Met Office Global Coupled model version 4 (GC4) configurations—The Forecasting Research Technical Report No: 66. <https://doi.org/10.62998/uzui3766>
- Xavier, P., Willett, M., Graham, T., Earnshaw, P., Copsey, D., Narayan, N., et al. (2024b). Assessment of the Met Office Global Coupled model version 5 (GC5) configurations. Retrieved from <https://centaur.reading.ac.uk/112173/>
- Yastrebova, A., Hoyhtya, M., Boumard, S., Lohan, E. S., & Ometov, A. (2021). Positioning in the Arctic Region: State-of-the-Art and Future Perspectives. *IEEE Access*, 9, 53964–53978. <https://doi.org/10.1109/ACCESS.2021.3069315>
- Zampieri, L., Goessling, H. F., & Jung, T. (2018). Bright prospects for arctic sea ice prediction on subseasonal time scales. *Geophysical Research Letters*, 45(18), 9731–9738. <https://doi.org/10.1029/2018GL079394>
- Zhang, J., Lindsay, R., Schweiger, A., & Steele, M. (2013). The impact of an intense summer cyclone on 2012 Arctic sea ice retreat. *Geophysical Research Letters*, 40(4), 720–726. <https://doi.org/10.1002/grl.50190>
- Zhao, J., Yu, Y., Cheng, J., Guo, H., Li, C., & Shu, Q. (2021). Influence of melt ponds on the SSMIS-based summer sea ice concentrations in the arctic. *Remote Sensing*, 13(19), 3882. <https://doi.org/10.3390/rs13193882>