The influence of non-static sea ice on Antarctic and Southern Ocean numerical weather prediction

Zhaohui Wang[1], Alexander D. Fraser[1,2], Phil Reid[3,2], Richard Coleman[1,2], and Siobhan O'Farrell[4]

[1] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia.
[2] Australian Antarctic Program Partnership, Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania, Australia.
[3] Australian Bureau of Meteorology, Hobart, Tasmania, Australia.
[4] CSIRO Oceans and Atmosphere, Aspendale, Victoria, Australia.

Corresponding author: Zhaohui Wang (zhaohui.wang@utas.edu.au)

ABSTRACT

Although operational weather forecasting centres are increasingly using global coupled atmosphere-ocean-ice models to replace atmosphere-only models for short-term (10-day) weather forecasting, the influence of sea ice on such forecasting has yet to be fully quantified, especially in the Southern Ocean. To address this gap, a polar-specific version of the Weather Research and Forecasting model is implemented within a circumpolar Antarctic domain to investigate the impact of daily updates of sea-ice concentration on short-term weather forecasting. A statistically-significant improvement in near-surface atmospheric temperature and humidity is shown from +48 hours to +192 hours when assimilating daily sea-ice concentration into the model. Improvement in model performance is enhanced from July to September, which is the period of late sea-ice advance. Regionally, model improvement is shown to occur in most sea-ice regions, although the improvement is strongest in the Ross Sea and Weddell Sea sectors. The surface heat balance also shows remarkable improvement in outgoing radiative heat fluxes and both sensible and latent heat fluxes after 48 hours. This research demonstrates the non-negligible effect of including daily updates of sea-ice concentration in numerical weather forecasting and indicates the necessity of implementing a fully coupled atmosphere-ocean-ice model in operational high-latitude southern hemisphere weather forecasting.

SIGNIFICANCE STATEMENT

Antarctic sea ice plays important key roles in modulating Southern Ocean weather and climate processes. Accurate representation of sea-ice properties is one of the keys for improving predictive skill in polar atmospheric forecasts. However, sea-ice representation is relatively basic in widely-used global and regional forecast models, with static sea ice properties throughout the forecast. Hence, we used a polar-optimised weather forecasting model to evaluate the impacts of including daily updated sea-ice concentrations on short-term weather forecasting. We found the forecast skill of near-surface temperature and humidity show the most significant improvements. This indicates that it is important to include non-static sea ice in Antarctic short-term operational weather forecasting.

1. Introduction

Atmospheric numerical weather prediction (NWP) is the primary tool used for real-time forecasting of short-term weather conditions (generally out to +10 days). NWP is achieved using an atmospheric model which employs a number of dynamic/thermodynamic governing

equations, numerical computing methods and appropriate parameterisations of some physical processes (Phillips 1971). NWP provides forecasts of the fundamental atmospheric variables, such as temperature, winds, surface pressure and precipitation for the next several days (Mass and Kuo 1998). Many countries' operational weather services run global NWP models, however, results from these models are often regionally downscaled to higher-resolution (spatial and temporal) forecast models (or limited-area models, LAMs) that are better able to represent localised fine-scale meteorological conditions. LAMs typically nest within a global model which provides the boundary conditions for larger spatial scale guidance.

Weather forecasting in the polar regions presents extra complexities in comparison to mid- and lower-latitudes (Jung and Matsueda 2016). Forecast guidance obtained from NWP models is generally better at lower latitudes rather than at the poles, with strong interannual variability in performance over high latitudes (Jung et al. 2016). This is particularly true in the Antarctic region (Powers et al. 2012). Model initialisation over Antarctica and the Southern Ocean is problematic, given the sparsity of in-situ observations. Polar-orbiting satellites provide the potential for improving our observational base, but with these observational systems there are difficulties in distinguishing surface features, such as snow and ice-covered surfaces. Given that NWP is an "initial value problem" (Al-Yahyai et al. 2010), a lack of observational data is considered one of the prime reasons for poor model performance in these regions. However, distinct and fundamental polar processes, which are not necessarily included in global models, also add to the complexity and possible poor model performance over the Antarctic (Wilby and Wigley 1997). For example, global NWP models are tuned to simulate mid-latitude planetary boundary layers. However, Antarctica has a very shallow and stable boundary layer which is difficult to simulate in global NWP models. Such high-latitude processes can best be represented in a regional model (Tastula et al. 2012).

An example of a regional model that is specifically designed to simulate atmospheric processes over Antarctica is the Antarctic Mesoscale Prediction System (AMPS), which is used in support of the U.S. Antarctic Program. AMPS is based on the polar-optimised version of the Weather Research and Forecasting model (Polar WRF) and provides a real-time atmosphere forecast product in six Antarctic domains, ranging from the entire Antarctic region with a 24 km grid size, to the McMurdo regional area with ~1 km grid resolution (Powers et al. 2012). The AMPS project shows improved performance, compared to global

NWP models, in providing forecast guidance for air, land and sea transportation and navigation purposes in the Antarctic region. Aside from providing operational forecasts, Polar WRF is often used in research to better understand polar-specific atmospheric processes, for example the atmospheric boundary layer dynamics (Vignon et al. 2018), Antarctic cyclonic activities (Pezza et al. 2016; Uotila et al. 2011) and katabatic winds (Wille et al. 2017).

Another process important in the polar regions is cryosphere interactions – in particular, interactions involving sea ice. Sea ice plays a key role in the Antarctic weather and climate system as a modulator of atmospheric processes and ocean-ice-atmosphere interaction (Massom & Lubin, 2006; Simpkins et al., 2012). Antarctic sea ice is characterised by strong seasonality, changing dramatically from ~ 19 million km^2 at maximum in late September (austral late winter/early spring peak) to ~ 3 million km^2 at minimum in late February (austral late summer trough) each year (Parkinson 2019; Eayrs et al. 2019). Over a 10-day timescale, sea ice is highly dynamic, and in 2018, which is the year that this study focuses on, the largest magnitude of 10-day sea-ice extent retreat can change by up to around 3 million km^2 (Figure 1). Sea ice has an insulative effect between the relatively warm ocean and cold atmosphere, and its presence or absence highly modifies heat and momentum exchange and water vapour transport (Cassano et al., 2016; Massom and Stammerjohn 2010), especially when snow-covered. The presence of sea ice and snow cover also significantly modulates both the longwave and shortwave radiation balance due to its high albedo surface and its much lower surface temperature than the open water (Thorndike et al. 1975).

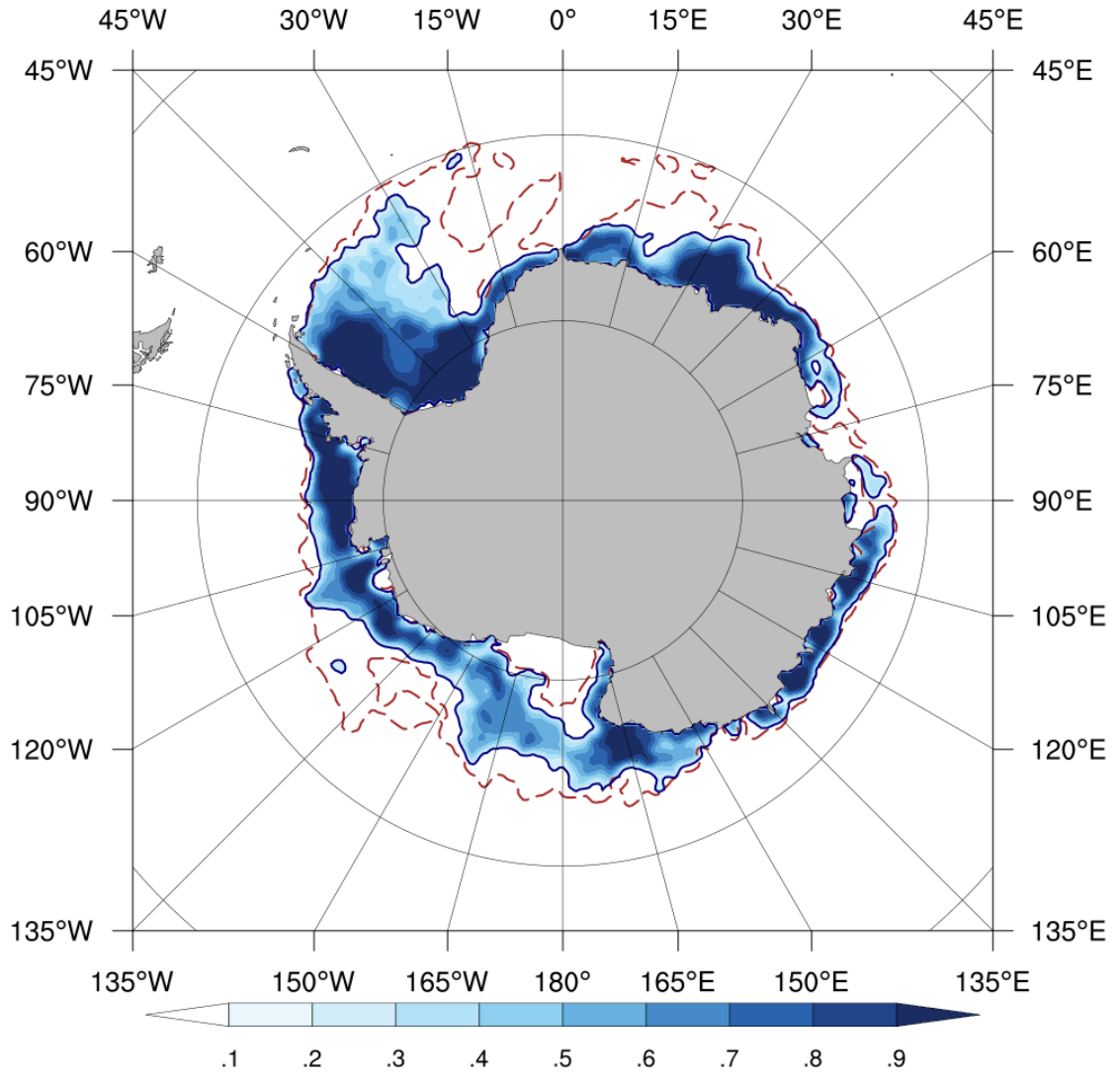


Figure 1: Sea-ice concentration on the 26th of December 2018 (blue shading), and the sea-ice edge on the 17th (red dashed line) and 26th (solid blue line) of December, illustrating a period of rapid retreat.

Because of its key role in modulating weather and climate processes, accurate representation of sea ice is an important variable to improve predictive skill in polar atmospheric forecasts. However, many global NWP models still use a static sea-ice cover – one where the representation of sea ice remains fixed as a slab of constant thickness and fixed concentration -- for the duration of the forecast. Some operational weather services are developing fully coupled NWP systems, where the atmospheric model is coupled to an ocean/sea-ice model (Smith et al. 2018; Day et al. 2022). Coupled regional modelling over the Arctic is being developed and is showing some skill (Smith et al. 2013, 2021). Because global NWP models do not have atmospheric polar modifications, it is difficult to quantify the impacts on an

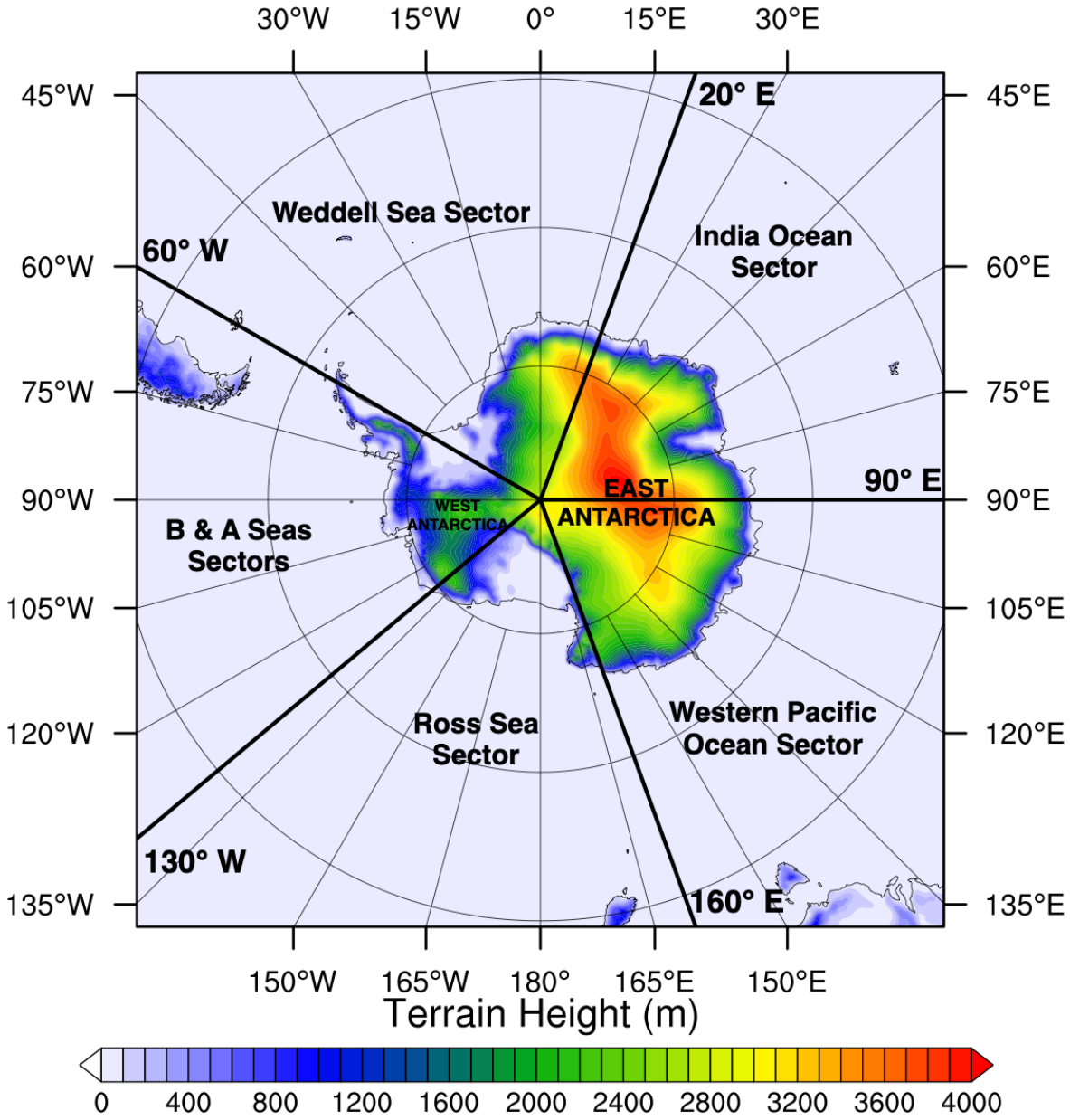
improved atmospheric forecast performance, or otherwise, on the inclusion of Antarctic-specific sea-ice coverage within a coupled global NWP model. At this stage there is no operational, fully coupled regional model optimised for the Antarctic region.

The impact of non-static (in this case, daily-updated) Antarctic sea-ice concentration distribution on the synoptic-scale performance of a regional model over a circum-Antarctic domain is investigated. We objectively quantify how the inclusion of an updating sea-ice field within a polar-optimised model might improve short-term weather forecasts over Antarctica and the Southern Ocean. In particular, the focus of the study is to: a) explore whether or not forecast accuracy increases when non-static sea ice is implemented in the Polar WRF model; b) characterise how errors propagate in space and time when using the unrealistic (static) sea-ice representation; c) determine which time period (e.g., sea-ice advance, retreat, etc.) shows the greatest improvement when non-static sea-ice representation is implemented in this atmospheric regional model; d) spatially characterise the forecast improvement when a daily-updated sea-ice concentration field is prescribed; and e) determine why these improvements happen, i.e., are the improvements ascribed to radiation, heat flux, or other parameters? By answering these questions, a detailed analysis of the effects of non-static sea ice in an atmospheric model will be given. The results provide the impetus to move towards operational regional coupled modelling and provide a baseline against which to compare future model performance.

2. Methods

The main objective is to investigate how the addition of a daily-updated sea-ice field in an Antarctic-specific atmospheric regional circulation model may improve, or otherwise, model forecast performance at a synoptic scale. The research domain covers the entire Antarctic region, including the region of maximum sea-ice extent (Figure 2), using a polar stereographic projection. Domain corners reach 30° S while the latitude at the midpoints of

149 the outer boundaries reaches 45° S. The experiment is based on the year 2018, which was the
150 most complete calendar year of sea-ice coverage at the start of this experiment.
151



152
153 Figure 2: The Polar WRF domain used for this study. Sectors are based on the widely used
154 regions for sea ice analysis (Zwally et al. 2002).
155

156 *a. Data*

157 The fifth-generation atmospheric reanalysis of the European Centre for Medium-Range
158 Weather Forecasts (ERA5; Hersbach et al., 2020) is used for forecast evaluation. ERA5
159 provides hourly atmospheric reanalysis data at a high spatial resolution (0.25° x 0.25°). The
160 regional NWP model used here (described below) requires initial and lateral boundary

conditions, such as wind, temperature, geopotential height and relative humidity. Boundary conditions are obtained from ERA5 and updated every six hours. Surface boundary conditions of sea-ice concentration and sea surface temperature are optional fields for use in the regional model used here. For consistency between the initial and boundary conditions, six-hourly sea-ice concentration distribution is also taken from ERA5, itself based on the EUMETSAT Data Center (EDC) Ocean and Sea Ice Satellite Application Facility (OSI SAF) operational dataset, although this field only updates once per day (in line with the underlying dataset). In both the model experiments described below, sea surface temperature (SST) is updated from ERA5 every six hours. For the lower boundary conditions, the elevation data are obtained from the Radarsat Antarctic Mapping Project Digital Elevation Model (RAMP-DEM; Liu et al., 2001) with a horizontal spatial resolution of 1 km.

The use of reanalysis data as the reference to evaluate model performance is not an ideal situation, particularly over Antarctica, due to the inherent biases in reanalysis model (Zhu et al. 2021; Tetzner et al. 2019). However due to the difficulties associated with validation of modelled products in this region characterised by a lack of observational data, this evaluation method is commonly used in operational verification (e.g., Eerola 2013; Schroeter et al., 2019). While comparison with observational data at Antarctic and mid-latitude stations can provide a more independent validation, observational station data in Antarctica are relatively sparse and have limitations for verification of spatial variability (Ebert et al. 2013). Also, due to the site-specific nature of station data, stations may not be representative of grid cells, especially at coastal sites. Thus, in this study, ERA5 is treated as the “real world” reference when comparing to Polar WRF model output, with the acknowledgement that this situation is not ideal.

ERA5 is basically an assimilation scheme with short-term forecasts for some selected variables. It is important to understand the limitations and advantages of using ERA5 as a forecast comparison. Previous research has widely validated ERA5 over Antarctica. Zhu et al. (2021) suggests that ERA5 is appropriate for climatological studies of Antarctic near-surface temperatures. By validating with automatic weather station (AWS) observations, Tetzner et al. (2019) suggests that ERA5 represents the magnitude and variability of near surface temperature and wind regimes with high accuracy over the Antarctic Peninsula. Dong et al. (2020) found that ERA5 has the best performance for magnitude and internal variability of near-surface wind speed simulation among the six recent global reanalysis products that

they compared. However, there also exist conflicting studies: Roussel et al. (2020) found that ERA5 has a high bias in precipitation simulations, and the seasonal cycles are not well-reproduced in Antarctica. With these studies as background, we use ERA5 as a reference noting the various possible issues.

b. Model

Polar WRF is a modification of the Weather Research and Forecasting (WRF) model aimed to better represent polar processes (Skamarock et al. 2008; Hines and Bromwich 2008). The WRF model is developed and maintained by the USA National Center for Atmospheric Research (NCAR) and other collaborative organisations, which is commonly used for atmospheric research ranging from large eddy-resolving to global scale and has wide applications (Skamarock et al. 2019). Polar WRF is developed and maintained by The Ohio State University's Polar Meteorology Group (PMG) as a code supplement to the standard WRF model (Hines et al. 2015). The Polar WRF model has been widely-used in both the Arctic and Antarctic to better understand polar-specific atmospheric processes, for example, atmospheric boundary layer dynamics (Vignon et al. 2018), the Ross Ice Shelf air stream (Seefeldt and Cassano 2012) and katabatic winds (Nigro and Cassano 2014).

Polar modifications to the standard WRF model include polar-specific improvements to the longwave flux, emissivity and freezing point of polar sea water, and thermal conductivity of the permanent snow and ice exceeding 20 cm in depth (Bromwich et al. 2013). In addition, fractional sea-ice concentration, which can be input from external datasets, is now coupled into the standard WRF model since WRF Version 3.1 (Skamarock et al. 2008). Furthermore, the sea-ice thickness, sea-ice/snow albedo and the snow depth can also now be specified as an input (e.g., from the US National Snow and Ice Data Center; Hines et al., 2015). The main modification of Polar WRF is optimising the Noah LSM for better representation of snow and sea-ice processes and heat transfer in polar regions (Bromwich et al. 2009).

The choice of physics parameterisation used within this study are mainly based on the latest practice used by the PMG, which has been tested as a mature and appropriate physical scheme combination for operational Antarctic NWP. The Morrison double-moment scheme (Morrison et al. 2009) is selected as the microphysics option. The Kain-Fritsch (KF) scheme (Kain 2004) is implemented for cumulus cloud parameterisation and updated at every model time step. For radiation schemes, the Rapid Radiative Transfer Model for GCMs (RRTMG;

Clough et al., 2005) is chosen as the parameterisation for both shortwave and longwave radiation. This parameterisation shows an improved radiation performance in polar regions compared to the prior version (Hines et al. 2015). The shortwave and longwave radiation fields are updated every 30 minutes. We use the Mellor-Yamada-Nakanishi-Niino Level 2.5 PBL scheme (MYNN; Nakanishi and Niino 2006) for the planetary boundary layer and update at every time step. The Nakanishi and Niino PBL's surface layer scheme (Nakanishi and Niino 2006) is used as the corresponding atmospheric surface layer. The land surface scheme is the Unified Noah Land Surface Model (Chen and Dudhia 2001) with polar optimisation modified by the PMG (Hines and Bromwich 2008). Previous research has shown that a higher model top pressure layer (i.e., representing a higher altitude) gives a better representation of gravity wave propagation (Bromwich et al. 2005), so an upper model top pressure value of 3 hPa was used for this study. A staggered vertical grid on 71 full-η levels is set up on WRF hybrid vertical coordinates from the sea surface to 3 hPa with vertical velocity damping within the top 8 km of the model to enhance vertical stability. Sea-ice albedo and thickness are set to uniform, circumpolar values of 0.8 and 1 m respectively, recommended by Xue et al. (2022). Snow depth on sea ice was initialised to 5 cm following Valkonen et al. (2014). It can be increased or decreased by precipitation or melting during the simulation but always maintains a value of at least 5 cm. Table 1 describes the main schemes and parameters that were implemented in the Polar WRF model configuration.

Model version	Polar WRF 4.1.1
Vertical coordinate	WRF hybrid vertical coordinate
Vertical resolution	71 levels up to 3 hPa. Vertical velocity damping is applied in the top 8 km
Horizontal grid	330 x 349 grid for the whole Antarctic region
Horizontal resolution	30 km grid cell size
Sea ice	ERA5 0.25° sea-ice fraction
Initial and boundary conditions	ERA5 0.25° reanalysis with 6-hourly intervals
Terrain field	1 km Radarsat Antarctic Mapping Project Digital Elevation Model (RAMP-DEM)
Longwave/shortwave Radiation	Rapid Radiative Transfer Model for GCMs (RRTMG)

Boundary layer	Mellor-Yamada-Nakanishi-Niino Level 2.5 PBL scheme (MYNN)
Surface layer	Nakanishi and Niino PBL surface layer scheme (MYNN)
Land surface option	Unified Noah Land Surface Model (LSM) with Polar optimisation
Microphysics	Morrison double-moment scheme
Cumulus parameterisation	Kain-Fritsch (KF)
Spin up	First 24 h used as model spin-up time
Time step	60 seconds; not adaptive

Table 1: Overview of the main physical schemes and parameters of Polar WRF used in the experiments.

Figure 2 shows the Polar WRF domain used for this study. As briefly described above, it is a single domain with a polar stereographic projection, centred at the South Pole. The model domain has a 30 km horizontal resolution with 330x349 grid points, covering the entire Antarctic sea-ice zone at maximum extent (approximately matching the spatial coverage of the National Snow and Ice Data Center (NSIDC) southern hemisphere polar stereographic projection). Ideally the domain would encompass the entire southern hemisphere. However the polar-specific boundary layer physics used within Polar WRF are inconsistent with the relatively turbulent boundary layers found over mid-latitude land masses (Edwards et al. 2020).

We compared two model experiments over a 10-day forecast period: (a) static sea ice (denoted *PWstatic*); and (b) daily-updated sea ice (denoted *PWupdate*). Each experiment is initiated approximately every five days in 2018 for the whole year, i.e., runs initiated on the 1st, 6th, 11th, 16th, 21st, 26th of each month, covering all phases of sea-ice advance and retreat within one calendar year. Prior to each 10-day forecast, a spin-up for 24 hours was performed to allow thermal-dynamic balance to be achieved, following Bromwich et al. (2013) and Wilson et al. (2011). Previous research has shown that the planetary boundary layer in Antarctic regions requires at least 12 hours spin-up time to reach quasi-steady state (Parish and Cassano 2003), and Hines & Bromwich (2008) found minimal difference between a 12-hour and 24-hour spin-up for Polar WRF. Considering that a relatively long forecast (10 days) is performed in this study, a 24-hour spin-up time is chosen, i.e., a total of 11 days model running period in each case. To ensure that all the meteorological conditions are identical between the two experiments after spin-up, all the parameterisations and

initial/boundary conditions stay the same, including the sea-ice concentration and SST updates, throughout the first 24 hours (spin-up). All 24-hour spin-ups are discarded in the following analysis.

Following a case study comparing the results of nudging and non-nudging experiments, the nudging technique was not implemented in this study, because its influence on error propagation is not in line with our experimental aims. A detailed discussion about nudging versus non-nudging is given in Appendix A.

c. Evaluation

The two model experiments are evaluated by comparing hourly forecast output from Polar WRF against hourly ERA5 data for several surface and near-surface variables, as well as on pressure levels throughout the atmosphere. Key parameters at the near-surface level, such as the 10 m winds (U10m/V10m) and 2 m air temperature (T2m) and dewpoint (TD2m), are compared hourly throughout the 10-day forecast period. On pressure levels, geopotential height, U (eastward) and V (northward) winds, temperature, and relative humidity at 37 levels are selected to investigate the upper-level model performance. The ERA5 data are interpolated to the same grid as the Polar WRF model output using spline interpolation.

Four commonly used validation metrics are chosen for evaluation: mean error (ME; commonly referred to as "bias"), mean absolute error (MAE), root-mean-squared error (RMSE) and the Pearson's correlation coefficient (CORR). The statistical relationship between WRF output and ERA5 is measured by CORR. Trends are not removed in our analyses because the time periods are too short to include climate trends. The diurnal cycle is also not removed because the cycle is an important assessment aspect in this study.

3. Results and discussion

a. Model performance and quality control

The annual average model forecast performance using updated sea ice (*PWupdate*) for the five near-surface variables and three assessment metrics is shown in Figure 3. The RMSE, MAE and bias of the variables are calculated in the first 48 hours (after spin-up) of each forecast period against ERA5, then these variables are averaged for the 72 forecast periods to

provide an annual-average performance. For T2m, the performance is found to be very good over the Southern Ocean (as expected due to regular updating of SST), with <1 K MAE and RMSE. Relatively larger errors are observed over the Antarctic continent. The model tends to overestimate (positive bias) T2m around the Transantarctic Mountains (1 ~ 4 K) and underestimate along the Antarctic coast (1 ~ 5 K). TD2m has a similar tendency but shows lower errors on the Antarctic continent and slightly larger RMSE over the Southern Ocean. These validation metrics are broadly similar to previous studies (Valkonen et al. 2014; Hines et al. 2019), where the RMSE in T2m is 1.7 ~ 2.8 K (Valkonen et al. 2014) and 2.0 ~ 2.7 K (Hines et al. 2019), respectively. The errors obtained within the first 48 hours of the forecast period shows the model has the ability to accurately represent the near-surface temperature and humidity. The Polar WRF-simulated surface pressure (PSFC) corresponds well to that of ERA5 on the Antarctic continent, even though noise is present, with a relatively large difference between Polar WRF-simulated and ERA5 surface pressure around the Antarctic continent coastlines and other regions with steep orography. This noise appears to be partly caused by the spline interpolation method used to interpolate the ERA5 pressure to the WRF grid. The generally accurate simulation of PSFC indicates that Polar WRF can simulate synoptic-scale pressure systems. U10m and V10m also show good agreement with ERA5. U10m is underestimated (negative bias) along coastal regions (by 1 ~ 2 m/s), while V10m is slightly overestimated by a similar magnitude. The magnitude of errors (MAE and RMSE) of near-surface winds are smaller than 3 m/s throughout almost the entire domain.

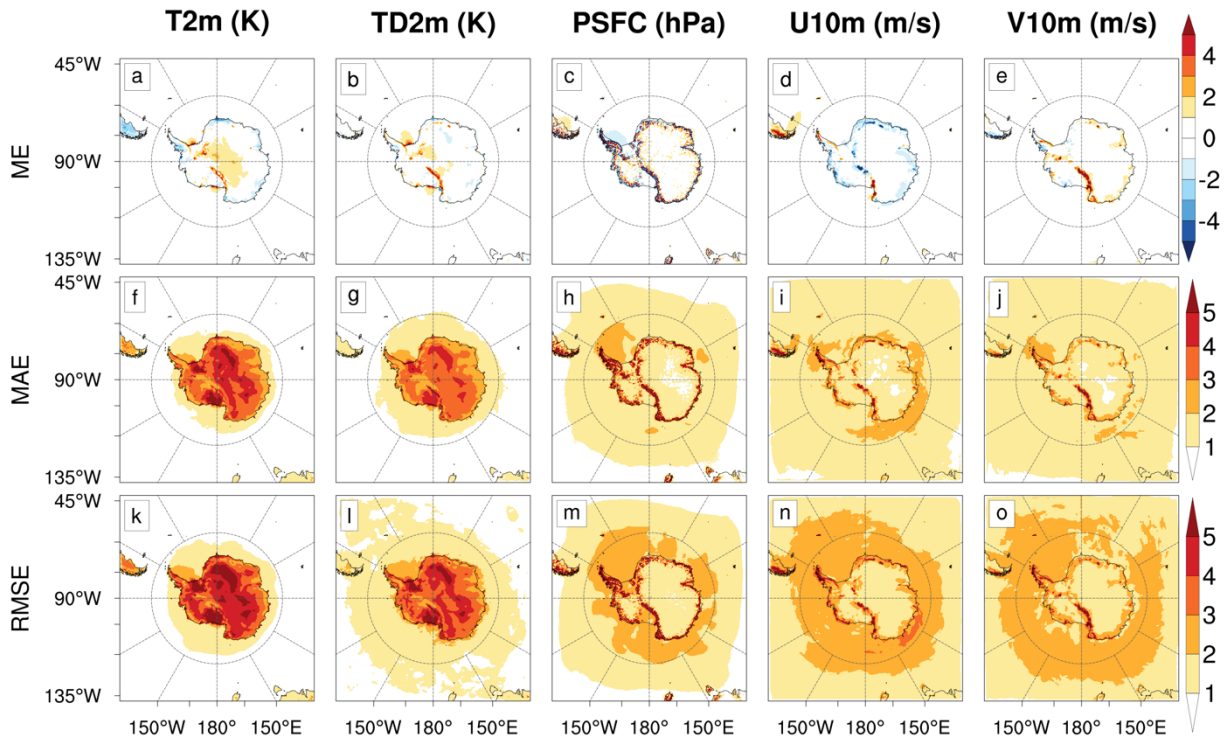


Figure 3: Model validation results showing the annual average bias, MAE and RMSE in T2m, TD2m, PSFC, U10m and V10m. These metrics are calculated using the *PWupdate* experiment output against the ERA5 reanalysis. The metrics were calculated using 72 model runs (throughout 2018) for the first 48 hours of each forecast period (after spin-up).

b. Influence of updating sea ice on the near-surface variables

Following the above validation, which indicated that this configuration of Polar WRF is appropriate for addressing the aims of this study, the *PWupdate* and *PWstatic* experiments are compared with ERA5 reanalysis for the near-surface variables (Table 2). Of all the surface variables, T2m and TD2m show the most significant improvement when daily-updated fractional sea ice is implemented. Figure 4 shows the difference of RMSE for T2m and TD2m between the simulations with updated sea ice and static sea ice, which were averaged across the 72 model runs. In almost all sea-ice regions during the first six days, the *updating* of sea ice has a positive impact on the T2m and TD2m forecasts. For many regions, the difference becomes statistically significant (99% confidence level) immediately, with marked improvement in the Ross Sea, Weddell Sea, and a part of the Indian Ocean sector. The improvement in the seasonal average of T2m (Figure 5 a) is mainly contributed between June and September, corresponding to the period of late sea-ice advance (Figure 5 d). Updated sea-ice concentration values appear to make a useful contribution to the model during this period of intense heat transfer from the ocean to the atmosphere. TD2m results also indicate

significant improvements when using updated sea ice (Figure 4 f - j). The regional pattern of TD2m is very similar to that of T2m.

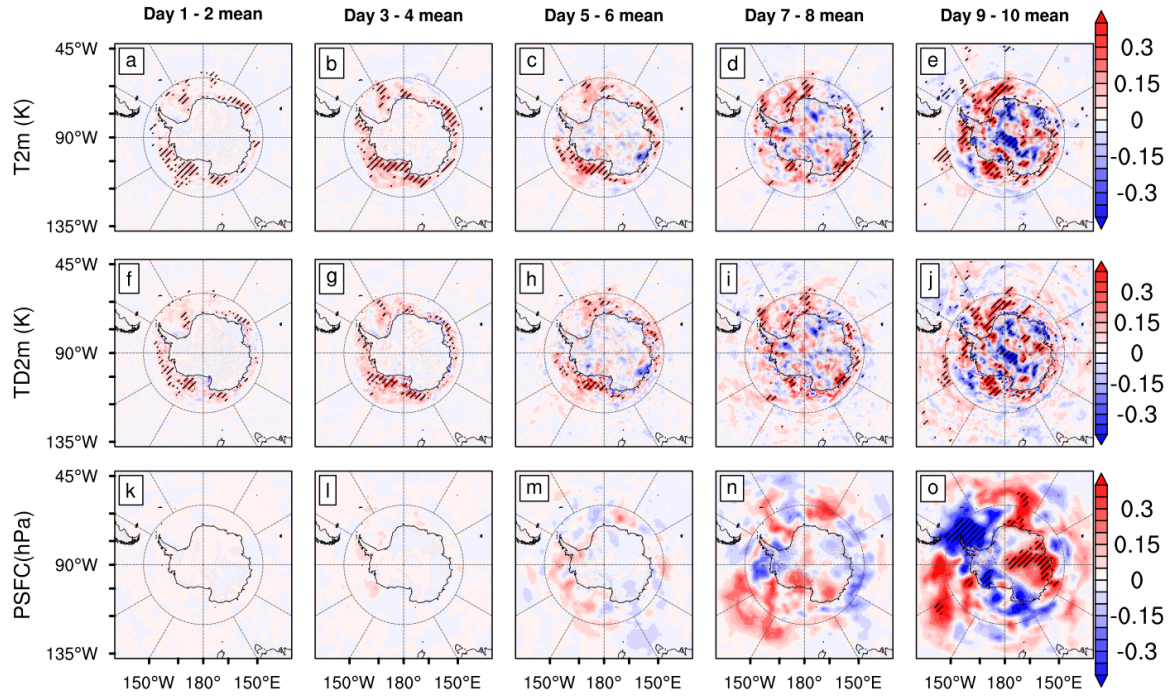


Figure 4: The difference (*PWstatic* minus *PWupdate*) of annual average RMSE for T2m, TD2m and PSFC between Polar WRF with updated sea ice (*PWupdate*) and static sea ice (*PWstatic*) when compared with ERA5 reanalysis data. The metrics were calculated from the average of 72 model runs during 2018 (without spin-up) for every 48-hour forecasts. Here red shading indicates that updated sea ice outperforms static sea ice. Stippled areas indicate statistical significance at the 99% level.

Figure 4 also shows the influence of updated sea ice on PSFC (bottom row). Unlike the temperature and humidity variables, the improvement for PSFC appears to be spatially mixed, but with overall improvement of results (Table 2). The impacts are not statistically significant (99% confidence level) until day 9 to 10 of the forecast period (Figure 4 o), while distinct improvements can be seen in T2m and TD2m within one-to-two days (Figure 4 a and f). From the day 5 to 6 averaged RMSE, some small improvements are found in the Bellingshausen Sea, Ross Sea, and Weddell Sea, although the impacts are not statistically significant. Two regions of reduction in performance over the sea-ice region of the Weddell Sea sector and the Antarctic continent in the Western Pacific Ocean sector are visible from day 5 to 6 (Figure 4 m). After this time, regions of positive and negative modifications gradually expand towards the open ocean and the Antarctic continent during day 7 to 8

(Figure 4 n). Table 2 shows that *PWupdate* generally has a slightly smaller domain-averaged RMSE of PSFC than the *PWstatic*.

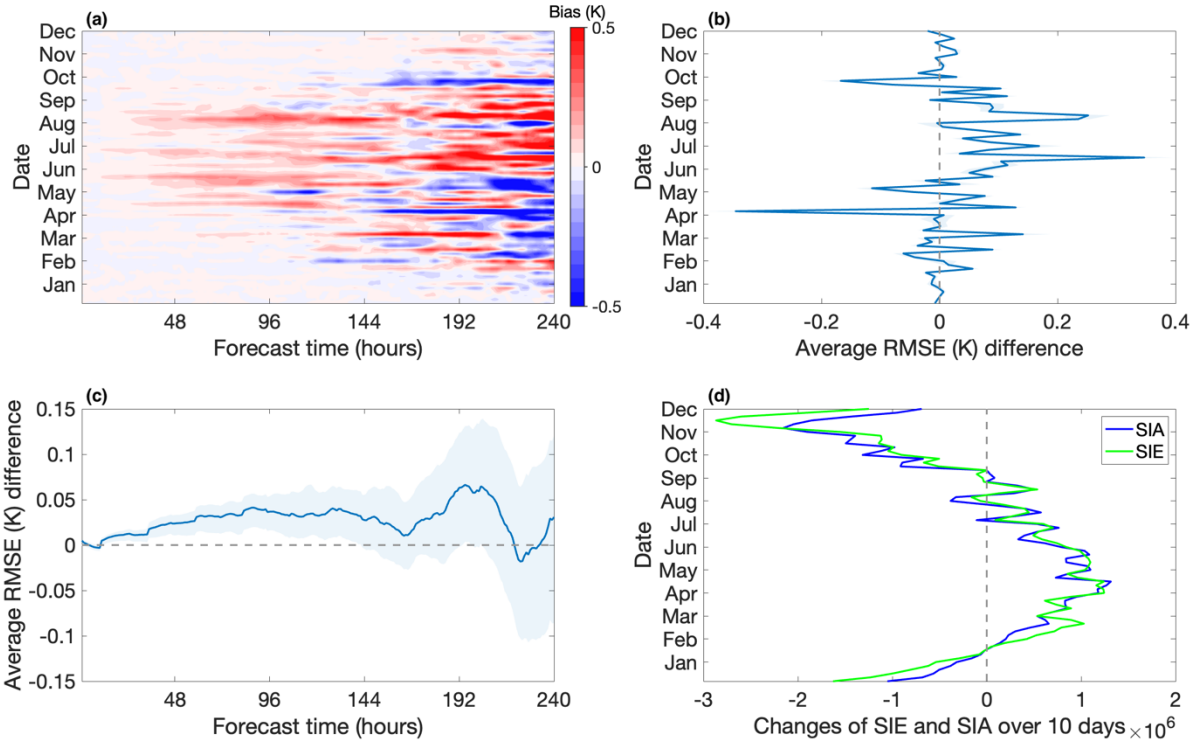


Figure 5: (a) Hovmöller diagram of domain-averaged RMSE difference (*PWstatic* minus *PWupdate*) south of 60° S for T2m between *PWstatic* and *PWupdate* compared with ERA5. Here red indicates that *PWupdate* outperforms the simulation with *PWstatic*. The x-axis represents the forecast time of each experiment and the y-axis represents the initiation date of each forecast experiment. (b) Time series (across the year 2018) of the mean RMSE difference of each forecast period from the Hovmöller diagram. Positive values indicate that the simulation with *PWupdate* outperforms *PWstatic*. (c) Time series (across the forecast period) of the mean RMSE difference (*PWstatic* minus *PWupdate*). The blue shaded area represents the 95% confidence interval using bootstrap sampling from the 72 model runs. Panel (d) shows the changes of sea-ice extent (SIE) and sea-ice area (SIA) over each 10-day forecast period.

The near-surface wind forecast shows slight improvement in the *PWupdate* experiment (Table 2). Table 2 indicates that U10m and V10m have smaller or similar domain-averaged RMSE values (south of 60° S) for nearly every day in the first 10 days of forecasts, except the zonal wind speed in the day 7 to day 8 average, where *PWstatic* shows a 0.01 m/s smaller RMSE.

	Day 1-2 mean				Day 3-4 mean				Day 5-6 mean				Day 7-8 mean				Day 9-10 mean			
	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR
2 m air temperature (K)																				
ERA5	-16.61				-16.58				-16.61				-16.61				-16.62			
updated sea ice	-16.59	0.02	3.22	0.96	-16.21	0.37	4.18	0.911	-16.08	0.54	5.13	0.849	-15.94	0.67	5.89	0.784	-16	0.62	6.31	0.738
static sea ice	-16.56	0.05	3.24	0.959	-16.17	0.41	4.22	0.907	-16.01	0.61	5.17	0.842	-15.86	0.75	5.93	0.776	-15.85	0.77	6.34	0.726
2 m dewpoint temperature (K)																				
ERA5	-19.98				-19.94				-19.98				-19.97				-19.98			
updated sea ice	-19.77	0.21	3.22	0.951	-19.36	0.57	4.36	0.89	-19.22	0.76	5.45	0.816	-19.08	0.89	6.28	0.744	-19.12	0.86	6.8	0.69
static sea ice	-19.75	0.23	3.24	0.95	-19.32	0.61	4.4	0.887	-19.16	0.82	5.48	0.811	-19.01	0.97	6.34	0.736	-18.98	1	6.83	0.68
Surface pressure (hPa for average, Pa for others)																				
ERA5	898.35				898.1				898.3				898.2				898.01			
updated sea ice	898.4	5.16	344.94	0.984	897.65	-44.93	536.04	0.927	897.17	-113.24	799.14	0.802	896.78	-141.88	1031.21	0.666	896.97	-104.28	1215.01	0.513
static sea ice	898.39	4.5	345.25	0.984	897.63	-46.46	537.12	0.926	897.15	-115.39	801.82	0.801	896.69	-150.38	1036.3	0.663	896.88	-112.53	1210.46	0.509
Zonal wind speed (m/s) u component																				
ERA5	-0.09				0.06				-0.08				-0.03				0.05			
updated sea ice	-0.36	-0.27	2.7	0.893	-0.08	-0.14	4.16	0.733	-0.05	0.03	5.35	0.547	-0.03	0	6.17	0.403	-0.15	-0.2	6.63	0.315
static sea ice	-0.36	-0.27	2.7	0.893	-0.08	-0.14	4.18	0.732	-0.05	0.03	5.37	0.545	0.02	0.05	6.16	0.405	-0.11	-0.17	6.64	0.312
Meridional wind speed (m/s) v component																				
ERA5	0.92				0.85				0.89				0.89				0.85			
updated sea ice	1.01	0.09	2.49	0.888	0.85	0	3.82	0.727	0.82	-0.07	4.95	0.537	0.82	-0.07	5.7	0.392	0.85	0	6.12	0.285
static sea ice	1.01	0.09	2.49	0.888	0.85	0	3.83	0.726	0.82	-0.07	4.96	0.536	0.81	-0.08	5.72	0.392	0.85	0	6.14	0.281

Table 2: The domain-averaged mean state, bias, RMSE and CORR of T2m, TD2m, PSFC and U10m/V10m, against their ERA5 counterparts. The metrics were calculated from the average of 72 model runs during 2018 (without spin-up) for every 48-hour forecasts.

Since T2m shows the largest improvement, a more comprehensive error analysis south of 60° S is provided. The Hovmöller diagram of the domain-averaged RMSE difference between *PWstatic* and *PWupdate* sea ice (Figure 5 a) shows that using updated sea ice can improve the forecast skill of T2m for most forecast dates and forecast time periods of the year. The period from July to September shows the largest improvement and fastest response to the more realistic updating of sea-ice concentration. In these three months, the surface temperature shows improvement after only one or two days forecast time, while it needs twice as long to demonstrate improvement in other months. The maximum mean RMSE reduction can exceed 0.5 K at +144 hours forecast, and 0.3 K at +96 hours in June and September. Figure 5 (b) shows that the updated sea-ice distribution has positive effects in most months of 2018, especially in the months of sea-ice advance. Figure 5 (c) indicates that the updated sea ice outperforms static sea ice at nearly every forecast time period (on average). However, PWupdate shows a remarkable reduction in the advantage of forecast skill on the last two days of the 10-day forecast period. This may be due to the reduction in accuracy and reliability in the final two days. Overall, the contribution of updated sea ice generally increases with the passing of the forecast time in the first eight days of forecast. Figure 5 (d) shows the 10-day change in sea-ice extent and area initiated from each forecast period. The periods of strongest T2m improvement roughly correspond to periods of sea-ice advance, indicating the updated sea ice gives the largest NWP skill improvement in the sea-ice formation season.

The mechanism that can be posited for underlying the T2m forecast skill increase, mainly during sea-ice advance, relies on the large temperature difference between the atmosphere (the top of the snow on the sea ice is close to the temperature of the atmosphere) and the ocean. By comparing Figures 5 (b) and (d), it is evident that there is not an exact correspondence between the seasonality of T2m improvement and the seasonality of sea-ice advance, i.e., the T2m improvement peaks in July to September (Figure 5 b) while the rate of sea-ice advance peaks in April (Figure 5 d). It can be considered that the magnitude of NWP forecast skill improvement can be partitioned into three phenological regimes: (a) no significant improvement to the modelled T2m during the period of sea-ice retreat, since the ice temperature is similar to the ocean temperature (at the sea ice melting point -- so addition

of a more realistic sea-ice field does not change the surface temperature, which strongly controls T2m); (b) minor improvement to modelled T2m during early sea-ice advance (i.e., when the air temperature is still cooling down -- e.g., April, when the heat flux from ocean to atmosphere (versus sea ice to atmosphere) is not remarkable); and (c) significant and rapid skill increase during sea-ice advance in the presence of a cold, near-surface air temperature (e.g., July - September) - even though the advance is not as rapid as in April, the stronger heat flux contrast gives a much more robust T2m forecast skill improvement.

c. Influence on surface heat balance

To diagnose the mechanism of T2m improvement, the influence of updated sea ice on the surface energy balance is compared with the static sea-ice experiment in Polar WRF. During times of minimal heat flux through sea ice, the near-surface air temperature is mainly controlled by the surface heat energy balance (Valkonen et al. 2014). As T2m shows a statistically significant improvement in this series of experiments, it is reasonable to suppose the updated (realistic) sea ice makes a positive contribution to surface heat balance modification. Table 3 lists the terms contributing to the net surface heat flux balance south of 60° S in ERA5, and the statistics for *PWupdate and PWstatic*. Figure 6 spatially indicates the four surface heat flux terms and the surface net heat flux for the updated sea-ice assimilation.

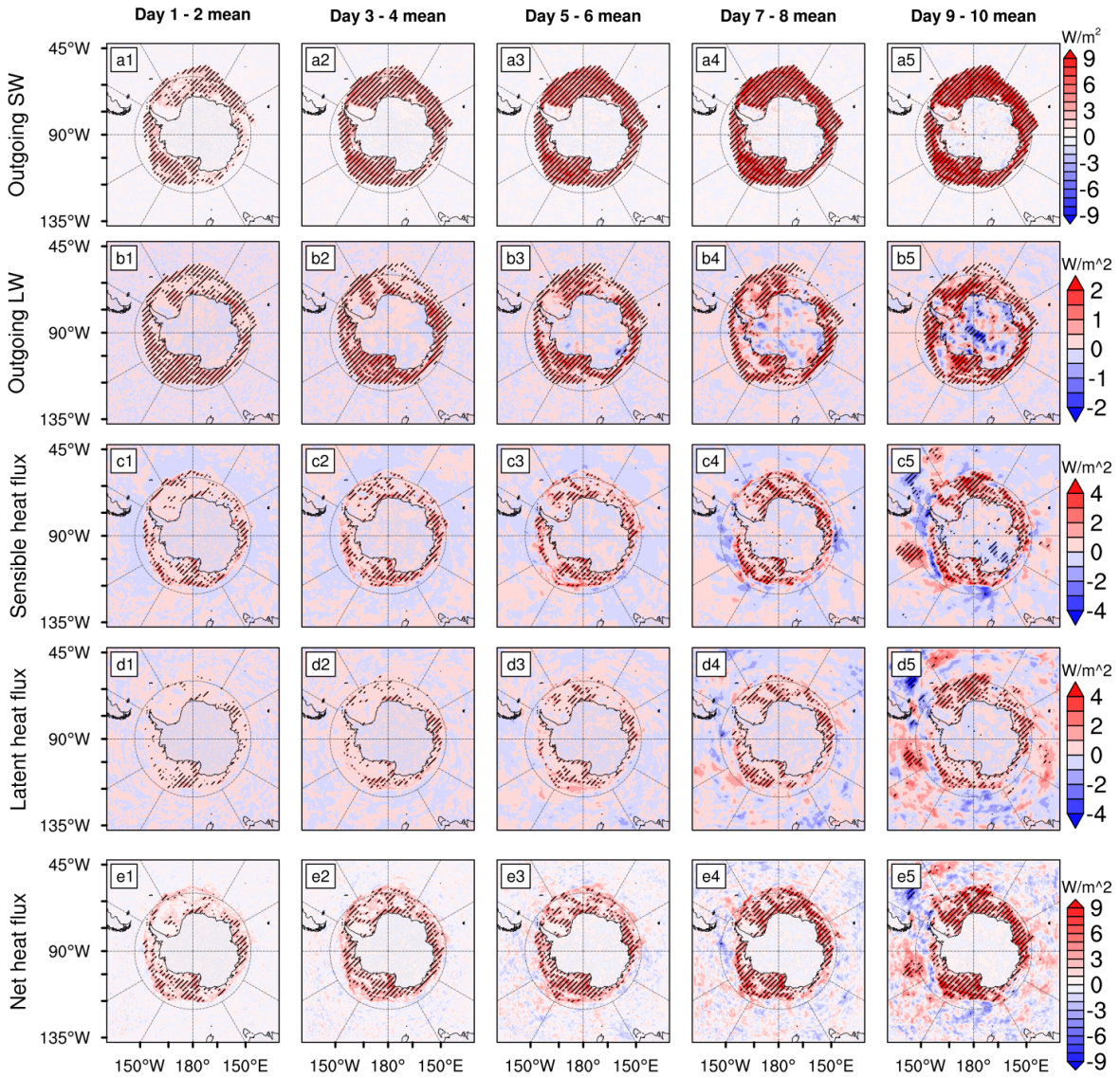


Figure 6: Difference (*PWstatic* minus *PWupdate*) of annual-averaged RMSE in outgoing shortwave/longwave radiation fluxes, latent heat flux, sensible heat flux and net heat flux at the surface between *PWstatic* and *PWupdate* when compared with ERA5 reanalysis data. The metrics were calculated from the average of 72 model runs during 2018 (without spin-up) for every 48-hour forecasts. Here red shading indicates that the updated sea-ice experiment outperforms the static sea-ice experiment. Stippling indicates differences that are significant at the 99% level.

Table 3 gives a comprehensive assessment and comparison for area-weighted surface energy balance for each two-day forecast average (south of 60º S). For radiative heat fluxes in each upward and downward direction, the bias, RMSE and CORR show improvements in all five two-day averaging periods. The outgoing shortwave radiation has the best improvement (when shortwave is present), where the area-weighted average bias reduces from 3.05 to 0.79 W/m^2, RMSE reduces from 36.8 to 29.3 W/m^2 and CORR increases from 0.93 to 0.96 for the

day 9 to 10 averaging period, despite the considerable variability in T2m forecast performance over these days (Figure 5 c). The outgoing longwave radiation CORR increases from 0.80 to 0.82, the bias reduces from 4.02 to 0.33 W/m² while the RMSE reduction is not so noticeable (22.9 to 22.4 W/m^2). For sensible heat flux, the WRF forecast value has large biases around Antarctica compared with ERA5. Previous research also showed poor correlation of sensible heat flux from WRF compared to observations (Tastula et al. 2012; Valkonen et al. 2014). However, *PWupdate* reduces the RMSE from 36.0 to 35.1 W/m^2 and increases CORR from 0.44 to 0.47 for the day 9-to-10 averaging period, again despite the relatively poor forecast performance of T2m. For latent heat flux, the bias of *PWupdate* is always larger than that of *PWstatic* i.e., *PWupdate* tends to underestimate the latent heat flux. However, CORR in *PWupdate* is always larger than in *PWstatic*. The lower RMSE of *PWupdate* shows the *PWupdate*-forecast latent heat flux is closer to the reanalysis field value.

	Day 1-2 mean				Day 3-4 mean				Day 5-6 mean				Day 7-8 mean				Day 9-10 mean			
	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR	AVE	Bias	RMSE	CORR
LW↓ (W/m²)																				
ERA5	205.82				206.18				205.67				206.15				205.89			
Updated sea ice	212.64	6.82	24.94	0.814	213.94	7.77	32.10	0.686	214.23	8.56	38.01	0.559	214.85	8.69	42.04	0.459	214.87	8.98	44.77	0.378
Static sea ice	212.75	6.93	24.91	0.814	214.13	7.96	32.15	0.684	214.58	8.91	37.98	0.558	215.23	9.08	42.08	0.457	215.49	9.60	44.90	0.372
LW↑ (W/m²)																				
ERA5	250.46				250.55				250.44				250.47				250.43			
Updated sea ice	251.95	1.49	12.71	0.963	253.08	2.53	15.75	0.931	253.49	3.06	18.65	0.890	253.89	3.43	20.95	0.848	253.76	3.33	22.40	0.819
Static sea ice	252.12	1.66	12.85	0.959	253.35	2.80	16.06	0.921	253.88	3.44	19.01	0.876	254.36	3.89	21.34	0.830	254.45	4.02	22.88	0.796
SW↓ (W/m²)																				
ERA5	120.17				120.16				120.30				119.78				120.38			
Updated sea ice	125.09	4.92	43.34	0.973	125.16	5.00	50.01	0.965	125.09	4.79	55.63	0.956	124.72	4.93	59.36	0.949	124.91	4.53	61.61	0.946
Static sea ice	125.26	5.09	43.45	0.973	125.47	5.31	50.16	0.965	125.47	5.17	55.77	0.956	125.22	5.44	59.55	0.949	125.53	5.14	61.56	0.947
SW↑ (W/m²)																				
ERA5	67.88				67.64				67.78				67.65				67.69			
Updated sea ice	69.08	1.20	22.52	0.976	68.65	1.00	25.22	0.970	68.61	0.82	27.38	0.963	68.54	0.89	28.49	0.960	68.49	0.79	29.33	0.956
Static sea ice	69.60	1.72	24.13	0.971	69.62	1.97	28.39	0.957	70.01	2.22	32.03	0.943	70.36	2.71	34.62	0.933	70.74	3.05	36.82	0.925
SH (W/m²)																				
ERA5	-3.56				-3.76				-3.53				-3.59				-3.60			
Updated sea ice	4.40	7.96	21.17	0.815	3.53	7.29	25.17	0.714	3.13	6.66	29.70	0.614	2.95	6.55	32.72	0.523	2.89	6.50	35.11	0.468
Static sea ice	4.52	8.08	21.68	0.808	3.72	7.48	25.94	0.700	3.37	6.90	30.68	0.595	3.23	6.83	33.62	0.499	3.23	6.84	36.01	0.444
LH (W/m²)																				
ERA5	14.12				13.99				13.97				14.04				14.00			
Updated sea ice	14.06	-0.06	11.19	0.797	13.76	-0.22	15.76	0.710	13.53	-0.44	19.17	0.618	13.56	-0.48	21.92	0.552	13.49	-0.52	23.81	0.500

Static sea ice	14.19	0.07	11.42	0.793	13.98	0.00	16.15	0.701	13.81	-0.16	19.60	0.606	13.93	-0.11	22.45	0.535	13.90	-0.11	24.37	0.480
Qnet (W/m²)																				
ERA5	-2.92				-2.08				-2.68				-2.63				-2.26			
Updated sea ice	-1.77	1.15	42.10	0.639	0.09	2.17	51.52	0.610	0.56	3.24	60.64	0.583	0.62	3.25	67.07	0.563	1.15	3.41	71.87	0.548
Static sea ice	-2.43	0.49	43.10	0.637	-1.07	1.02	53.10	0.605	-1.01	1.67	62.65	0.574	-1.42	1.21	69.44	0.551	-1.31	0.95	74.51	0.532

Table 3: The domain-averaged mean state, bias, RMSE and CORR for each term of surface energy balance, along with their ERA5 counterparts. The metrics were calculated from the average of 72 model runs during 2018 (without spin-up) for every 48-hour forecasts.

Figure 6 indicates that the outgoing radiative surface fluxes, including both shortwave and longwave radiation, has clear improvement over the domain with *PWupdate*. The turbulent surface fluxes, namely the latent and sensible heat fluxes, also show an improvement when using updated sea ice. The outgoing shortwave radiation flux has the most significant reduction in RMSE, and it dominates the improvement in the net heat flux (the net surface heat flux is the summation of net longwave and shortwave radiation and turbulent fluxes). The reduction in outgoing longwave radiation RMSE is likely driven by a more realistic surface temperature simulation when including an updated sea-ice field. The improvement appears mainly in the sea-ice region, showing that a better sea-ice description improves forecast skill of surface heat balance. The reduction in RMSE for outgoing radiative heat flux appears to only affect the sea-ice region on a relatively long timescale (10 days) with limited lower latitude influence.

By checking the seasonality of the surface heat balance (Figure B1 in Appendix B), the upward shortwave radiation shows a significant improvement in the Antarctic summer, while the longwave radiation shows a reasonable improvement for the sea-ice advance season (May-September). The improved longwave radiation during the period of sea-ice advance further indicates that it is due to the surface temperature improvement in *PWupdate* which has more realistic sea-ice dynamic representation. Sensible and latent heat fluxes also gained forecast skill in austral winter where the largest temperature difference occurs between ocean and atmosphere.

As expected from the *PWupdate* experiment, both downward shortwave and longwave radiation improvements are not as strong as those in the upward direction (Table 3). The

upward longwave and shortwave radiation terms will have differences directly driven by the specifications in the two experiments, whereas the downward longwave and shortwave radiation should not be directly impacted by the prescribed differences between experiments. However, the downward longwave and shortwave radiation can be indirectly influenced by the different cloud simulations in *PWstatic* and *PWupdate.* The downward shortwave radiation shows an amount of noise when comparing the static and updated sea-ice experiments (Figure B2 in Appendix B). This is mainly due to the modifications in the upper-level atmosphere and cloud simulation. Accurate Southern Ocean cloud simulation remains a challenge in atmospheric NWP models (Hines et al. 2019). The cloud fraction products in the WRF model are unreliable for use by forecasters due to their considerable bias with observations (Hines et al. 2019). The unreliable cloud and upper air simulation in the model limits improvement in downward radiative fluxes, although in the downward longwave radiation, improvements are still found in the Antarctic coastal region and the region covered by sea ice, while pockets of decrease in model forecast skill still exist at the ice edge.

Both latent and sensible heat fluxes show improvements over the domain (Figure 6) with a statistically significant RMSE reduction over the sea-ice region. The sensible heat flux has a larger RMSE reduction than that of latent heat flux. The region showing the strongest improvement occurs at the sea-ice edge, indicating that the surface turbulent heat fluxes are more sensitive to the sea-ice advance/retreat than the sea-ice concentration change within consolidated ice.

d. Influence on the vertical structure of the atmosphere

To examine the influence of the updated sea-ice field through the troposphere, three-hourly vertical profiles of three Polar WRF model variables (air temperature, relative humidity and geopotential height) were interpolated to the 37 ERA5 pressure levels to facilitate comparison. This comparison is shown in Figure 7.

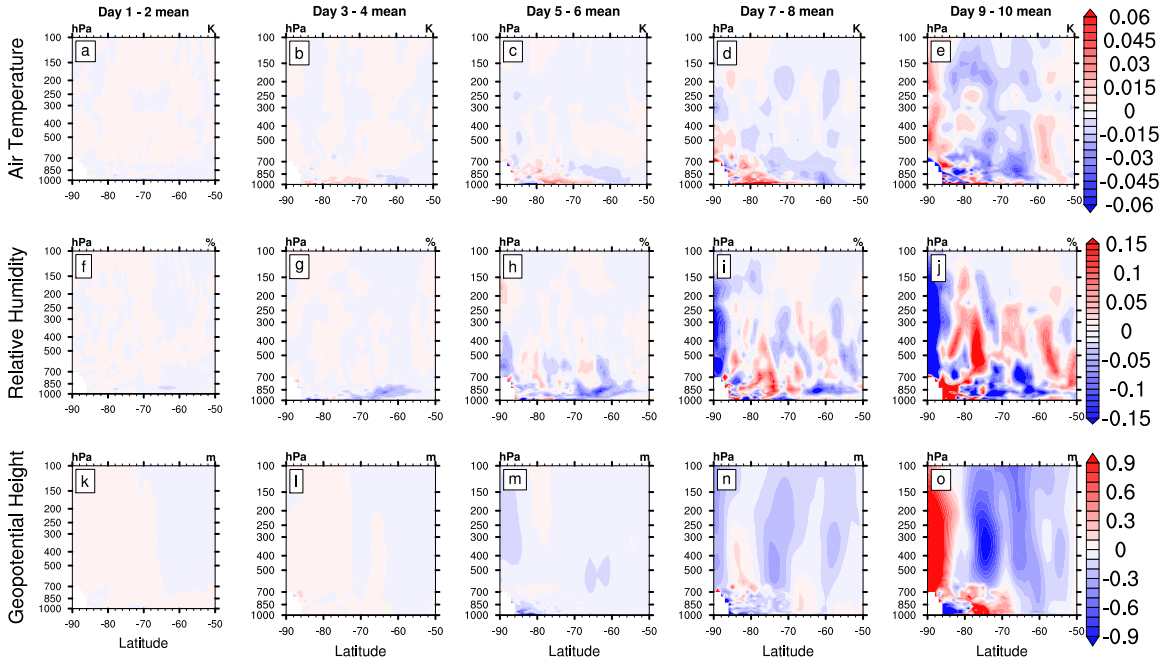


Figure 7: Vertical profiles of differences in annual average RMSE in air temperature, relative humidity and geopotential height between *PWstatic* and *PWupdate* when compared with ERA5 reanalysis data. The metrics were calculated from the average of 72 model runs during 2018 (without spin-up) for every 48-hour forecasts. Here red indicates that the updated sea ice outperforms static sea ice.

For the air temperature vertical profile (Figure 7, a - e), there is a clear improvement in sea ice-covered latitudes (65° S to 80° S) from the surface to 900 hPa after +4 days, the improvement can reach further inland (75° S to 82° S), and propagates to the 600 hPa level after +6 days. However, the *PWupdate* experiment tends to show an air temperature forecast skill decrease originating from the Southern Ocean (60° S) and propagating vertically to 600 hPa at around 75° S after +8 days.

Humidity is the most challenging variable to model in the upper troposphere, especially in the polar regions (Elliott and Gaffen 1991; Wilson et al. 2011). The relative humidity result here shows a large RMSE difference between WRF output and ERA5, with the value of 9 to 27% between 900 and 500 hPa, even in the first 48 hours, and the RMSE difference can reach 45% after 10 days (figure not shown). Since both the WRF output and ERA5 reanalysis values are subject to their own independent biases, a relatively poor correspondence may be expected. The upper troposphere relative humidity analyses are presented here for reference (Figure 7 f - j), and results above 500 hPa will not be discussed. Relative humidity shows a forecast skill

decrease in the *PWupdate* experiment for the first 6 days (Figure 7 f - h), while an improvement occurs from day 7 to 8 at around 950 hPa from 68° S to 82° S (Figure 7 i - j).

The RMSE differences of geopotential height between *PWupdate* and *PWstatic* are not large (Figure 7 k - o). Even though the *PWupdate* does not depict a perceptible improvement in the first 8 days forecast (Figure 7 k - n), the negative effects are not significant, especially in the first 6 days (Figure 7 k - m). Considering the surface pressure is also not strongly changed, this indicates that the atmospheric pressure may not be significantly impacted by including updated sea ice.

4. Conclusion

Daily-updated sea-ice concentrations have been assimilated into the Polar WRF model to compare with model runs using static sea ice throughout a 10-day forecast period across Antarctica and the Southern Ocean. ERA5 was used to force the initial and boundary conditions and compared with the model output as the "real world" reference. These experiments were repeated for a total of 72 times throughout 2018.

The first 48 hours forecast of near-surface variables were used to evaluate the model performance. Near-surface variable forecasts from Polar WRF with updated sea ice compared well with the ERA5 reanalysis, showing a small bias and very high correlation. The bias, MAE and RMSE errors were in good agreement with previous studies (Wilson et al. 2011; Hines et al. 2019; Valkonen et al. 2014). The bias and errors are in a controlled and expected range, especially in regions away from steep topography, indicating that the Polar WRF model is suitable for use in our study to capture the effects of changes in sea-ice distribution.

For near-surface variables, the T2m and TD2m have the most statistically-significant improvements with a domain-averaged value of 9% - 15% bias reduction after 48 hours forecast. The surface pressure appears to be relatively insensitive to updated sea ice within a 10-day forecast. Near-surface winds, both in the meridional and zonal directions, show an improvement but the magnitudes are relatively small.

Although limited improvement for indirectly-influenced variables is found, such as for winds and surface pressure, the improvements are not as remarkable as the directly-influenced variables, namely, the near-surface temperature and humidity. Temperature and humidity are

positively impacted due to the better simulation of surface heat fluxes, while improvements in the winds and surface pressure may be due to the modification of temperature and humidity. Based on these results, the improvement in pressure and winds may have a greater lag and muted response, which is hard to capture in a relatively short forecast time (10-day forecast), but may be more obvious in a seasonal or climate context.

Adding updated sea ice into the Polar WRF model significantly modified the surface heat balance, e.g., upward shortwave radiation has a domain-averaged value of 30% - 75% bias reduction immediately after spin-up. The improvement for upward shortwave radiation plays a dominant role in the modification. The upward longwave radiation, latent and sensible heat flux were also dramatically improved over the sea-ice region. These contribute to the energy balance, which results in the modification of surface temperature.

Even though this study found a statistically significant improvement at the surface and near-surface when adding updated realistic sea-ice concentrations into the NWP model, propagation of this improvement towards the upper troposphere is limited. Improvement appears to be constrained within the planetary boundary layer (PBL), where atmospheric variables can be directly influenced by the surface changes. Above the PBL, the propagation of the improving signal weakens steadily since, particularly over much of the Antarctic plateau, the turbulence and vertical mixing are rare. Compared with the Arctic, the polar vortex over Antarctica is stronger and more resistant to blocking air mass exchange with mid-latitudes (Waugh and Randel 1999; Qian et al. 2021). The zonally-dominated circulation of the atmosphere at mid to high southern latitudes seems to provide a dynamic boundary, limiting the propagation of improvements northward. In summary, the improvement of using updated sea ice in the NWP model is most distinctive in +120 to +192 hour forecasts, between June and September, south of 60° S, and below 700 hPa.

Further studies should include the investigation of the diurnal cycle of error propagation and the relationships between each impacted variable. The impact of updating sea ice on Antarctic NWP is a complicated process, and so the annual-average analyses presented here may ignore some important fine-scale responses, such as atmosphere and ocean interactions with the sea ice that would be naturally included in a fully coupled regional model. A series of case studies showing maximum impacts on atmospheric factors is necessary to provide more insight on those processes contributing to the forecast improvement. Furthermore, the

effect of a more realistic prescription of sea-ice thickness and snow depth, the values of which are fixed within these model runs in Antarctic NWP, should be investigated as well. In addition, the NWP model implemented for this study is an atmospheric-only model, albeit with enhancements to increase the realism of sea ice. Coupling such a model to a computationally-efficient ice/ocean model is the next logical step for use in operational forecasting and requires further investigation.

Acknowledgments

This research was supported by scholarships from the Australian Government and the Australian Research Council's Special Research Initiative for the Antarctic Gateway Partnership (SRI40300001). This project received grant funding from the Australian Government as part of the Antarctic Science Collaboration Initiative program. We would like to acknowledge the use of the high-performance computing facilities provided by the Tasmanian Partnership for Advanced Computing (TPAC) funded and hosted by the University of Tasmania. This research was also supported by use of the Nectar Research Cloud; a collaborative Australian research platform supported by the NCRIS-funded Australian Research Data Commons (ARDC). We thank David Bromwich, Keith Hines and Lesheng Bai from The Ohio State University PMG group for providing the Polar WRF model and the constructive and helpful discussion. We thank Marcus Thatcher and Jack Katzfey from The Commonwealth Scientific and Industrial Research Organisation (CSIRO) for their helpful comments and discussion in relation to nudging. The authors thank NCAR Mesoscale and Microscale Meteorology for providing the standard release of the WRF model, and the Copernicus Climate Change Service (C3S) Climate Data Store (CDS) for providing the ERA5 reanalysis data.

Data Availability Statement.

ERA5 data used in this study can be acquired through the Copernicus Climate Change Service (C3S) Climate Data Store (CDS; https://cds.climate.copernicus.eu). Polar WRF model can be downloaded from The Ohio State University Polar Meteorology Group (http://polarmet.osu.edu/PWRF/) via registration. Model output used in this study can be acquired from the corresponding author upon request.

APPENDICES

Appendix A: The influence of nudging on error propagation in Polar WRF

Nudging is an important technique to constrain model simulations from drifting away from the real conditions (Glisan et al. 2013). However, whether this technique should be used depends on the purpose of the experiment. Here, using 10-day forecasts implemented in this study, the influences of nudging on error propagation and on experiment purpose are investigated.

The results from a single nudging versus no-nudging experiment with Polar WRF are presented here. The model runs from 1-12 September 2018. One test is run with spectral nudging techniques, and includes nudging of winds, temperature, and geopotential height from ERA5 every six hours, with daily-updated- and static- sea-ice concentration. Nudging is only applied to model layer 40 (around 400 hPa) and above following the suggestion of Zou et al. (2021). A control test is run with the same configuration but does not use the nudging technique. Both model runs use daily-updated sea-ice concentrations.

The difference in error propagation between nudging and non-nudging experiments is shown in Figure A1. The error propagation is well-constrained by spectral nudging, especially at the Antarctic continental margin and sea-ice regions (Figure A1 a1 – e1). For the non-nudging experiment (Figure A1 a2 – e2), the RMSE develops and propagates after 48 hours forecast period to spread across the Antarctic continent along the Transantarctic Mountains and Dronning Maud Land, as well as from the Weddell Sea and Amundsen Sea to spread to other sea-ice-covered regions of primary interest in this study. This indicates that the nudging technique may interfere with (or confound) the results presented here.

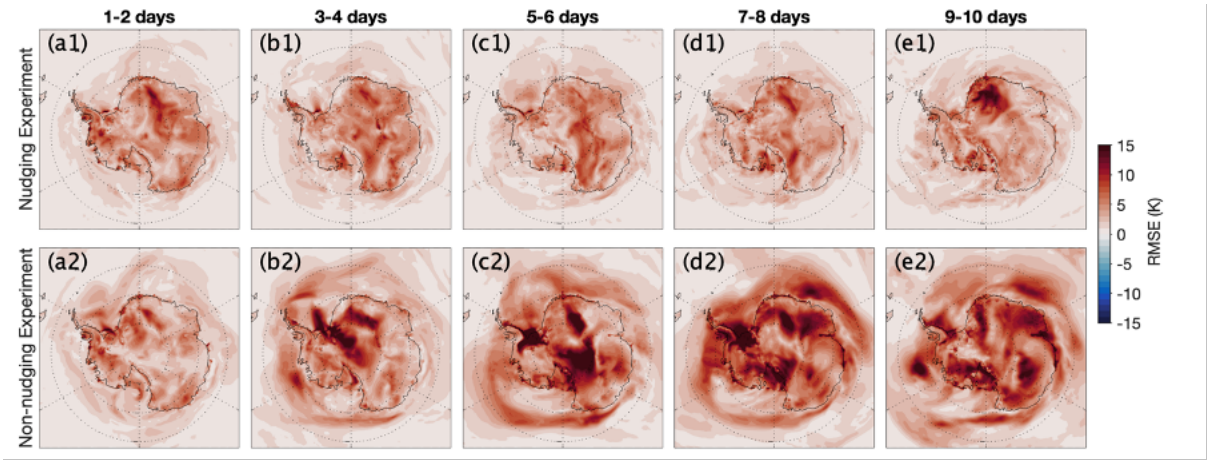


Figure A1: RMSE of T2m in the daily-updated sea-ice experiment using ERA5 nudging (upper panel), and without nudging experiment (bottom panel).

As the error propagation has been well constrained by the nudging technique, the differences caused by daily-updated-sea ice are much smaller in the nudging experiment than in the non-nudging experiment (Figure A2). It is noted that the nudged T2m RMSE difference for the 7 - 8 day average is similar in magnitude to the 5 - 6 day time period, which indicates the nudging technique may have an excessive interference on error propagation. This may imply that nudging has a strong effect on constraining the error propagation that may overwhelm the magnitude of differences between static- and daily-updated-sea ice.

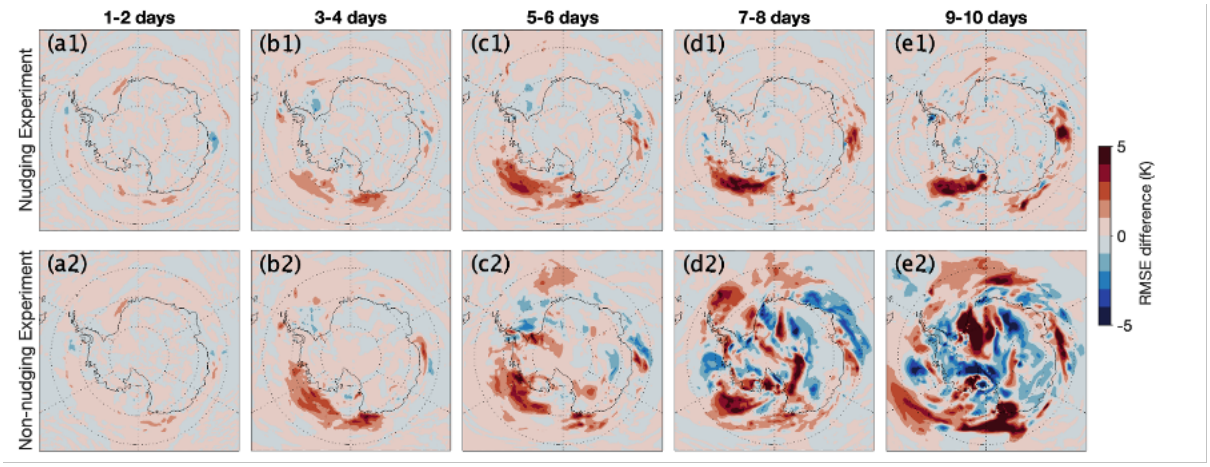


Figure A2: T2m RMSE difference (static minus daily-updated) in the nudging experiment (upper panels) and non-nudging experiment (lower panels). Each column represents a two-day average of forecasts. Red shading indicates regions where Polar WRF with daily-updated-sea ice outperforms Polar WRF with static-sea ice. The first 24 hours are discarded as a spin-up period.

The nudging technique has relatively little influence on the first 48 hours of forecast period (Figure A3), but the influence is stronger after that. The nudging technique keeps the error tightly bound throughout the duration of the 10-day forecast.

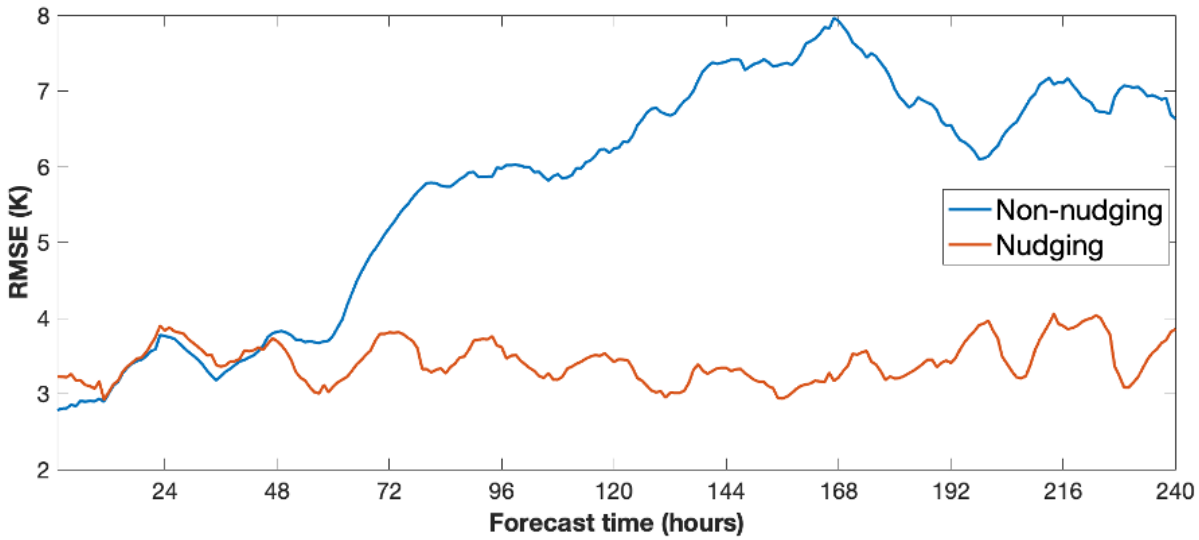


Figure A3: Domain-averaged T2m RMSE in daily-updated sea-ice run for the nudging and non-nudging experiments.

The RMSE difference between static and updated runs with nudging is much less than the runs without nudging (Figure A4). It appears that the nudging technique has little influence on the RMSE difference pattern in the first 144 hours (i.e., nudging just influences the RMSE magnitude), however, the error pattern also changes after 144 hours.

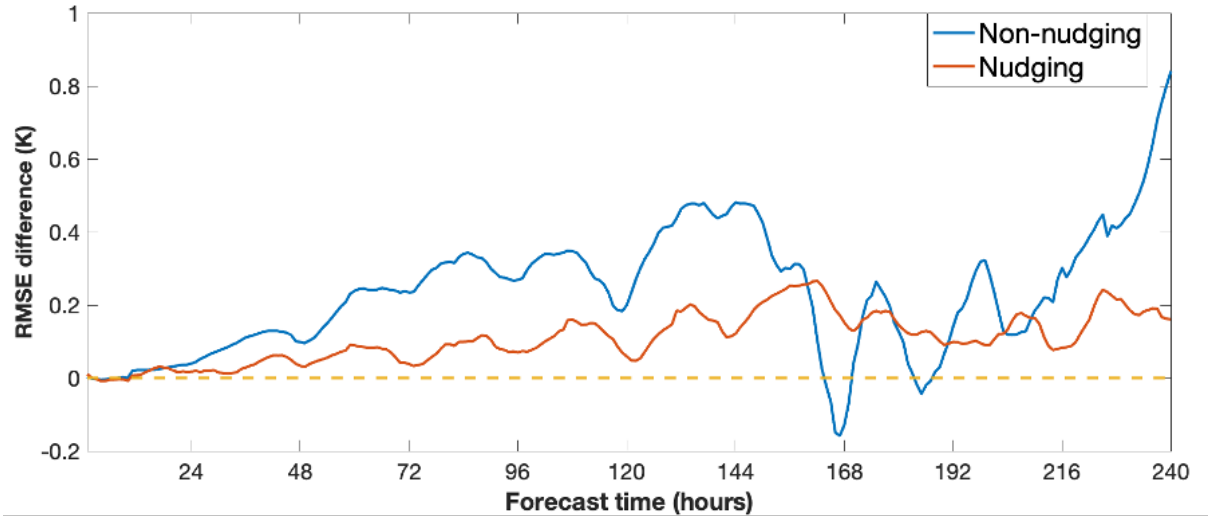


Figure A4: The difference of domain-averaged T2m RMSE between static and updated runs in the nudging and non-nudging experiments, respectively.

Hence, although nudging can effectively constrain model error, it also strongly influences the differences in error propagation between static and daily-updated runs. The ultimate aim of this study is not to provide an extremely accurate forecast, rather to investigate how static sea ice impacts the forecast accuracy, and to provide an "upper bound" of forecast skill improvement when coupling a dynamic sea-ice component into NWP. As the nudging technique can cause interference in the interpretation of results, nudging is not used in this study. In addition, nudging from reanalysis is impossible in the case of an operational short-term forecast which is what this study is aiming to simulate.

Appendix B: The influence of non-static sea ice on surface heat balance

Figure B1 shows the seasonality of the surface heat balance, including (a) outgoing shortwave radiation, (b) outgoing longwave radiation, (c) latent heat flux, and (d) sensible heat flux. Difference (*PWstatic* minus *PWupdate*) of annual-averaged RMSE in incoming shortwave/longwave radiation fluxes at the surface between *PWstatic* and *PWupdate* when compared with ERA5 reanalysis data are shown in Figure B2.

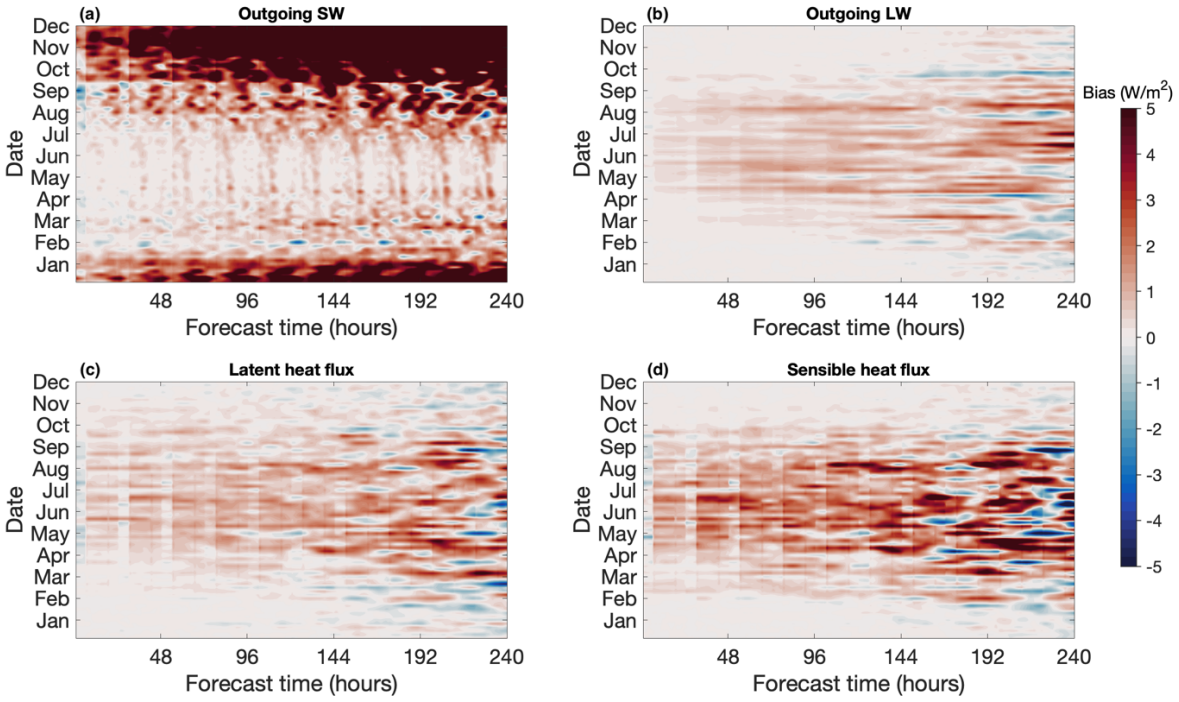


Figure B1: Hovmöller diagram of domain-averaged RMSE difference (PWstatic minus PWupdate) south of 60° S for outgoing (a) shortwave radiation, (b) longwave radiation, (c) latent heat flux and (d) sensible heat flux between PWstatic and PWupdate compared with ERA5. Here red indicates that PWupdate outperforms the simulation with PWstatic. The x-axis represents the forecast time of each experiment and the y-axis represents the initiation date of each forecast experiment.

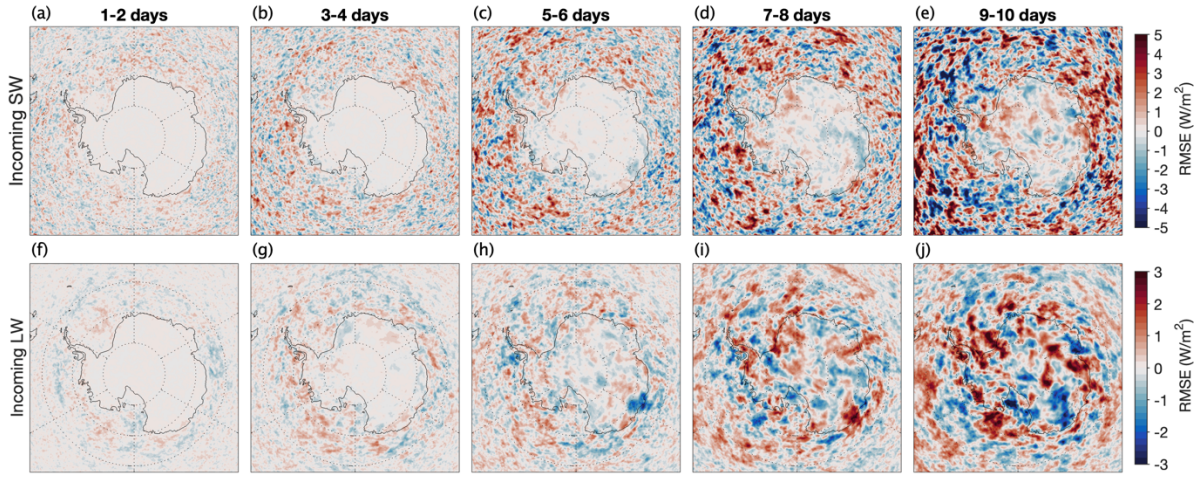


Figure B2: As in Figure 6, but for (a) – (e) incoming shortwave radiation and (f) – (g) incoming longwave radiation.

REFERENCES

Al-Yahyai, S., Y. Charabi, and A. Gastli, 2010: Review of the use of Numerical Weather Prediction (NWP) Models for wind energy assessment. *Renew. Sustain. Energy Rev.*, **14**, 3192–3198, https://doi.org/10.1016/j.rser.2010.07.001.

Bromwich, D. H., L. Bai, and G. G. Bjarnason, 2005: High-Resolution Regional Climate Simulations over Iceland Using Polar MM5*. *Mon. Weather Rev.*, **133**, 3527–3547, https://doi.org/10.1175/MWR3049.1.

——, K. M. Hines, and L. Bai, 2009: Development and testing of Polar Weather Research and Forecasting model: 2. Arctic Ocean. *J. Geophys. Res.*, **114**, D08122, https://doi.org/10.1029/2008JD010300.

——, F. O. Otieno, K. M. Hines, K. W. Manning, and E. Shilo, 2013: Comprehensive evaluation of polar weather research and forecasting model performance in the Antarctic. *J. Geophys. Res. Atmos.*, **118**, 274–292, https://doi.org/10.1029/2012JD018139.

Cassano, J. J., M. W. Seefeldt, S. Palo, S. L. Knuth, A. C. Bradley, P. D. Herrman, P. A. Kernebone, and N. J. Logan, 2016: Observations of the atmosphere and surface state over Terra Nova Bay, Antarctica, using unmanned aerial systems. *Earth Syst. Sci. Data*, **8**, 115–126, https://doi.org/10.5194/essd-8-115-2016.

Chen, F., and J. Dudhia, 2001: Coupling an Advanced Land Surface–Hydrology Model with

the Penn State–NCAR MM5 Modeling System. Part I: Model Implementation and Sensitivity. *Mon. Weather Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Clough, S. A., M. W. Shephard, E. J. Mlawer, J. S. Delamere, M. J. Iacono, K. Cady-Pereira, S. Boukabara, and P. D. Brown, 2005: Atmospheric radiative transfer modeling: a summary of the AER codes. *J. Quant. Spectrosc. Radiat. Transf.*, **91**, 233–244, https://doi.org/10.1016/j.jqsrt.2004.05.058.

Day, J. J., S. Keeley, G. Arduini, L. Magnusson, K. Mogensen, M. Rodwell, I. Sandu, and S. Tietsche, 2022: Benefits and challenges of dynamic sea ice for weather forecasts. *Weather Clim. Dyn.*, **3**, 713–731, https://doi.org/10.5194/wcd-3-713-2022.

Dong, X., Y. Wang, S. Hou, M. Ding, B. Yin, and Y. Zhang, 2020: Robustness of the Recent Global Atmospheric Reanalyses for Antarctic Near-Surface Wind Speed Climatology. *J. Clim.*, **33**, 4027–4043, https://doi.org/10.1175/JCLI-D-19-0648.1.

Eayrs, C., D. Holland, D. Francis, T. Wagner, R. Kumar, and X. Li, 2019: Understanding the Seasonal Cycle of Antarctic Sea Ice Extent in the Context of Longer-Term Variability. *Rev. Geophys.*, **57**, 1037–1064, https://doi.org/10.1029/2018RG000631.

Ebert, E., and Coauthors, 2013: Progress and challenges in forecast verification. *Meteorol. Appl.*, **20**, 130–139, https://doi.org/10.1002/met.1392.

Edwards, J. M., A. C. M. Beljaars, A. A. M. Holtslag, and A. P. Lock, 2020: Representation of Boundary-Layer Processes in Numerical Weather Prediction and Climate Models. *Boundary-Layer Meteorol.*, **177**, 511–539, https://doi.org/10.1007/s10546-020-00530-z.

Eerola, K., 2013: Twenty-One Years of Verification from the HIRLAM NWP System. *Weather Forecast.*, **28**, 270–285, https://doi.org/10.1175/WAF-D-12-00068.1.

Elliott, W. P., and D. J. Gaffen, 1991: On the Utility of Radiosonde Humidity Archives for Climate Studies. *Bull. Am. Meteorol. Soc.*, **72**, 1507–1520, https://doi.org/10.1175/1520-0477(1991)072<1507:OTUORH>2.0.CO;2.

Glisan, J. M., W. J. Gutowski, J. J. Cassano, and M. E. Higgins, 2013: Effects of Spectral Nudging in WRF on Arctic Temperature and Precipitation Simulations. *J. Clim.*, **26**, 3985–3999, https://doi.org/10.1175/JCLI-D-12-00318.1.

Hersbach, H., and Coauthors, 2020: The ERA5 global reanalysis. *Q. J. R. Meteorol. Soc.*, **146**, 1999–2049, https://doi.org/10.1002/qj.3803.

Hines, K. M., and D. H. Bromwich, 2008: Development and Testing of Polar Weather Research and Forecasting (WRF) Model. Part I: Greenland Ice Sheet Meteorology*. *Mon. Weather Rev.*, **136**, 1971–1989, https://doi.org/10.1175/2007MWR2112.1.

——, ——, L. Bai, C. M. Bitz, J. G. Powers, and K. W. Manning, 2015: Sea Ice Enhancements to Polar WRF*. *Mon. Weather Rev.*, **143**, 2363–2385, https://doi.org/10.1175/MWR-D-14-00344.1.

——, ——, S.-H. Wang, I. Silber, J. Verlinde, and D. Lubin, 2019: Microphysics of summer clouds in central West Antarctica simulated by the Polar Weather Research and Forecasting Model (WRF) and the Antarctic Mesoscale Prediction System (AMPS). *Atmos. Chem. Phys.*, **19**, 12431–12454, https://doi.org/10.5194/acp-19-12431-2019.

Jung, T., and M. Matsueda, 2016: Verification of global numerical weather forecasting systems in polar regions using TIGGE data. *Q. J. R. Meteorol. Soc.*, **142**, 574–582, https://doi.org/10.1002/qj.2437.

——, and Coauthors, 2016: Advancing Polar Prediction Capabilities on Daily to Seasonal Time Scales. *Bull. Am. Meteorol. Soc.*, **97**, 1631–1647, https://doi.org/10.1175/BAMS-D-14-00246.1.

Kain, J. S., 2004: The Kain–Fritsch Convective Parameterization: An Update. *J. Appl. Meteorol.*, **43**, 170–181, https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2.

Liu, H., K. Jezek, B. Li, and Z. Zhao, 2001: Radarsat Antarctic Mapping Project digital elevation model version 2. *Natl. Snow Ice Data Cent.*,.

Mass, C. F., and Y.-H. Kuo, 1998: Regional Real-Time Numerical Weather Prediction: Current Status and Future Potential. *Bull. Am. Meteorol. Soc.*, **79**, 253–263, https://doi.org/10.1175/1520-0477(1998)079<0253:RRTNWP>2.0.CO;2.

Massom, R., and D. Lubin, 2006: *Polar Remote Sensing - Volume I: Atmosphere and Oceans*.

Massom, R. A., and S. E. Stammerjohn, 2010: Antarctic sea ice change and variability - Physical and ecological implications. *Polar Sci.*, **4**, 149–186, https://doi.org/10.1016/j.polar.2010.05.001.

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of Cloud Microphysics on the Development of Trailing Stratiform Precipitation in a Simulated Squall Line: Comparison of One- and Two-Moment Schemes. *Mon. Weather Rev.*, **137**, 991–1007, https://doi.org/10.1175/2008MWR2556.1.

Nakanishi, M., and H. Niino, 2006: An Improved Mellor–Yamada Level-3 Model: Its Numerical Stability and Application to a Regional Prediction of Advection Fog. *Boundary-Layer Meteorol.*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

Nigro, M. A., and J. J. Cassano, 2014: Identification of Surface Wind Patterns over the Ross Ice Shelf, Antarctica, Using Self-Organizing Maps. *Mon. Weather Rev.*, **142**, 2361–

826 2378, https://doi.org/10.1175/MWR-D-13-00382.1.
827 Parish, T. R., and J. J. Cassano, 2003: The Role of Katabatic Winds on the Antarctic Surface
828 Wind Regime. *Mon. Weather Rev.*, **131**, 317–333, https://doi.org/10.1175/1520-
829 0493(2003)131<0317:TROKWO>2.0.CO;2.
830 Parkinson, C. L., 2019: A 40-y record reveals gradual Antarctic sea ice increases followed by
831 decreases at rates far exceeding the rates seen in the Arctic. *Proc. Natl. Acad. Sci.*, **116**,
832 14414–14423, https://doi.org/10.1073/pnas.1906556116.
833 Pezza, A., K. Sadler, P. Uotila, T. Vihma, M. D. S. Mesquita, and P. Reid, 2016: Southern
834 Hemisphere strong polar mesoscale cyclones in high-resolution datasets. *Clim. Dyn.*, **47**,
835 1647–1660, https://doi.org/10.1007/s00382-015-2925-2.
836 Phillips, N. A., 1971: Numerical weather prediction. *Eos, Trans. Am. Geophys. Union*,
837 https://doi.org/10.1029/EO052i006pIU420.
838 Powers, J. G., K. W. Manning, D. H. Bromwich, J. J. Cassano, and A. M. Cayette, 2012: A
839 Decade of Antarctic Science Support Through Amps. *Bull. Am. Meteorol. Soc.*, **93**,
840 1699–1712, https://doi.org/10.1175/BAMS-D-11-00186.1.
841 Qian, Y., Y. Luo, F. Si, T. Yang, and D. Yang, 2021: Three-Year Observations of Ozone
842 Columns over Polar Vortex Edge Area above West Antarctica. *Adv. Atmos. Sci.*, **38**,
843 1197–1208, https://doi.org/10.1007/s00376-021-0243-7.
844 Roussel, M.-L., F. Lemonnier, C. Genthon, and G. Krinner, 2020: Brief communication:
845 Evaluating Antarctic precipitation in ERA5 and CMIP6 against CloudSat observations.
846 *Cryosph.*, **14**, 2715–2727, https://doi.org/10.5194/tc-14-2715-2020.
847 Schroeter, B. J. E., P. Reid, N. L. Bindoff, and K. Michael, 2019: Antarctic Verification of
848 the Australian Numerical Weather Prediction Model. *Weather Forecast.*, **34**, 1081–
849 1096, https://doi.org/10.1175/WAF-D-18-0171.1.
850 Seefeldt, M. W., and J. J. Cassano, 2012: A description of the Ross Ice Shelf air stream
851 (RAS) through the use of self-organizing maps (SOMs). *J. Geophys. Res. Atmos.*, **117**,
852 https://doi.org/10.1029/2011JD016857.
853 Simpkins, G. R., L. M. Ciasto, D. W. J. Thompson, and M. H. England, 2012: Seasonal
854 Relationships between Large-Scale Climate Variability and Antarctic Sea Ice
855 Concentration. *J. Clim.*, **25**, 5451–5469, https://doi.org/10.1175/JCLI-D-11-00367.1.
856 Skamarock, W. C., and Coauthors, 2008: A Description of the Advanced Research WRF
857 Version 3. *A Descr. Adv. Res. WRF Version 3*, 113, https://doi.org/10.5065/D6DZ069T.
858 Skamarock, W. C., and Coauthors, 2019: A description of the advanced research WRF model
859 version 4. *Natl. Cent. Atmos. Res. Boulder, CO, USA*, **145**, 145.

860 Smith, G. C., F. Roy, and B. Brasnett, 2013: Evaluation of an operational ice-ocean analysis
861 and forecasting system for the Gulf of St Lawrence. *Q. J. R. Meteorol. Soc.*, **139**, 419–
862 433, https://doi.org/10.1002/qj.1982.
863 ——, and Coauthors, 2018: Impact of Coupling with an Ice–Ocean Model on Global
864 Medium-Range NWP Forecast Skill. *Mon. Weather Rev.*, **146**, 1157–1180,
865 https://doi.org/10.1175/MWR-D-17-0157.1.
866 Smith, G. C., and Coauthors, 2021: The Regional Ice Ocean Prediction System v2: a pan-
867 Canadian ocean analysis system using an online tidal harmonic analysis. *Geosci. Model*
868 *Dev.*, **14**, 1445–1467, https://doi.org/10.5194/gmd-14-1445-2021.
869 Tastula, E.-M., T. Vihma, and E. L. Andreas, 2012: Evaluation of Polar WRF from Modeling
870 the Atmospheric Boundary Layer over Antarctic Sea Ice in Autumn and Winter. *Mon.*
871 *Weather Rev.*, **140**, 3919–3935, https://doi.org/10.1175/MWR-D-12-00016.1.
872 Tetzner, D., E. Thomas, and C. Allen, 2019: A Validation of ERA5 Reanalysis Data in the
873 Southern Antarctic Peninsula—Ellsworth Land Region, and Its Implications for Ice Core
874 Studies. *Geosciences*, **9**, 289, https://doi.org/10.3390/geosciences9070289.
875 Thorndike, A. S., D. A. Rothrock, G. A. Maykut, and R. Colony, 1975: The thickness
876 distribution of sea ice. *J. Geophys. Res.*, **80**, 4501–4513,
877 https://doi.org/10.1029/JC080i033p04501.
878 Uotila, P., T. Vihma, A. B. Pezza, I. Simmonds, K. Keay, and A. H. Lynch, 2011:
879 Relationships between Antarctic cyclones and surface conditions as derived from high-
880 resolution numerical weather prediction data. *J. Geophys. Res.*, **116**, D07109,
881 https://doi.org/10.1029/2010JD015358.
882 Valkonen, T., T. Vihma, M. M. Johansson, and J. Launiainen, 2014: Atmosphere-sea ice
883 interaction in early summer in the Antarctic: evaluation and challenges of a regional
884 atmospheric model. *Q. J. R. Meteorol. Soc.*, **140**, 1536–1551,
885 https://doi.org/10.1002/qj.2237.
886 Vignon, E., F. Hourdin, C. Genthon, B. J. H. Van de Wiel, H. Gallée, J. Madeleine, and J.
887 Beaumet, 2018: Modeling the Dynamics of the Atmospheric Boundary Layer Over the
888 Antarctic Plateau With a General Circulation Model. *J. Adv. Model. Earth Syst.*, **10**, 98–
889 125, https://doi.org/10.1002/2017MS001184.
890 Waugh, D. W., and W. J. Randel, 1999: Climatology of Arctic and Antarctic Polar Vortices
891 Using Elliptical Diagnostics. *J. Atmos. Sci.*, **56**, 1594–1613,
892 https://doi.org/10.1175/1520-0469(1999)056<1594:COAAAP>2.0.CO;2.
893 Wilby, R. L., and T. M. L. Wigley, 1997: Downscaling general circulation model output: a

review of methods and limitations. *Prog. Phys. Geogr. Earth Environ.*, **21**, 530–548, https://doi.org/10.1177/030913339702100403.

Wille, J. D., D. H. Bromwich, J. J. Cassano, M. A. Nigro, M. E. Mateling, and M. A. Lazzara, 2017: Evaluation of the AMPS Boundary Layer Simulations on the Ross Ice Shelf, Antarctica, with Unmanned Aircraft Observations. *J. Appl. Meteorol. Climatol.*, **56**, 2239–2258, https://doi.org/10.1175/JAMC-D-16-0339.1.

Wilson, A. B., D. H. Bromwich, and K. M. Hines, 2011: Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain: Surface and upper air analysis. *J. Geophys. Res.*, **116**, D11112, https://doi.org/10.1029/2010JD015013.

Xue, J., Z. Xiao, D. H. Bromwich, and L. Bai, 2022: Polar WRF V4.1.1 simulation and evaluation for the Antarctic and Southern Ocean. *Front. Earth Sci.*, 1–20, https://doi.org/10.1007/s11707-022-0971-8.

Zhu, J., A. Xie, X. Qin, Y. Wang, B. Xu, and Y. Wang, 2021: An Assessment of ERA5 Reanalysis for Antarctic Near-Surface Air Temperature. *Atmosphere (Basel).*, **12**, 217, https://doi.org/10.3390/atmos12020217.

Zou, X., D. H. Bromwich, A. Montenegro, S. Wang, and L. Bai, 2021: Major surface melting over the Ross Ice Shelf part I: Foehn effect. *Q. J. R. Meteorol. Soc.*, **147**, 2874–2894, https://doi.org/10.1002/qj.4104.

Zwally, H. J., J. C. Comiso, C. L. Parkinson, D. J. Cavalieri, and P. Gloersen, 2002: Variability of Antarctic sea ice 1979–1998. *J. Geophys. Res.*, **107**, 3041, https://doi.org/10.1029/2000JC000733.