# 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[1,3], 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)

**Key Points:**

- Improvements in forecasting skill are evaluated by assessing the relative performance of static versus daily-updated sea ice experiments using a numerical weather prediction model
- The study presents the improvements in near surface variables and heat fluxes by using daily-updated Antarctic sea ice concentrations in the circumpolar high-latitude Southern Ocean region.
- The study emphasizes the importance of incorporating realistic Antarctic sea ice distribution in near-real-time weather forecasting.

# Abstract

Although operational weather forecasting centers are increasingly using coupled atmosphere-ocean-ice models to replace atmosphere-only models for short-term (10 days) 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 (Polar WRF) is implemented within a circumpolar Antarctic domain to investigate the impact of daily-updates of sea ice concentrations on short-term weather forecasting. Apart from some steep plateau regions adjacent to the Antarctic continental margin, Polar WRF shows good forecast skill in Antarctic surface variables. A statistically significant improvement in near-surface temperature and humidity is shown from +96 hours to +192 hours when assimilating daily sea ice concentration into the model. Improvements in model performance are enhanced during July through September, which is a period of late sea ice advance. Regionally, model improvements are shown to encompass almost all sea ice regions, although marked in the Ross and Weddell seas sectors. The surface heat budget balance also shows remarkable improvement in outgoing radiative heat fluxes and both sensible and latent heat fluxes after 48 hours. Our results demonstrate the non-negligible effect of including daily-updates of sea ice concentrations in numerical weather forecasting, and endorsing the necessity of a fully coupled atmosphere-ocean-ice model in operational high-latitude Southern Hemisphere weather forecasting.

# 1 Introduction

Atmospheric numerical weather prediction (NWP) is the primary tool used for real-time forecasting of the short-term weather conditions (generally out to +10 days). NWP is achieved, most frequently, using an atmospheric model which employs a number of dynamic/thermodynamic governing equations, numerical computing methods and appropriate parameterizations of some physical processes (Phillips, 1971). By forcing the initial conditions with the current meteorological state at the forecast initialization, plus boundary conditions of atmospheric state at the edge of the model domain (in the case of a regional model), NWP provides estimates of the fundamental atmospheric variables such as temperature, winds, surface pressure and precipitation for the next several days (Mass & Kuo, 1998).

Weather forecasting in the polar regions presents extra complexities in comparison to mid- and lower-latitudes (Jung & Matsueda, 2016). Forecast guidance obtained from NWP models is generally better at lower latitudes rather than at the poles, with strong interannual variability of their performance over the higher latitudes (Jung et al., 2016). This is particularly so in the Antarctic region. Model initialization over Antarctica and the Southern Ocean is problematic, given the sparsity of 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 &

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 best represented in a regional (nested) model, thus insulating changes to the model's physics from impacting upon forecast performance at lower latitudes (Tastula et al., 2012).

An example of regional modelling over Antarctica is the Antarctic Mesoscale Prediction System (AMPS) which is used in support of the U.S. Antarctic Program. AMPS is based on Polar WRF and provides a real-time atmosphere forecast product in six Antarctic domains, ranging from the entire Antarctic region with a 30 km grid size, to the McMurdo regional area with ~1 km grid resolution (Powers et al., 2012). The AMPS project shows good performance in providing information for transportation and navigation purposes in the Antarctic region, suggesting that Polar WRF is a useful tool for providing NWP forecast guidance operationally within the Antarctic domain (Powers et al., 2012).

Another important process, distinctive to the polar regions is cryosphere interactions – in particular, interactions between sea ice and the atmosphere. Sea ice is an essential part of our Earth system. It plays a key role in the Antarctic weather and climate system as a key modulator of atmospheric processes and ocean-ice-atmosphere interaction (R. Massom & Lubin, 2006; Simpkins et al., 2012). Antarctic sea ice is characterized by strong seasonality, changing dramatically from ~ 19 million square kilometers at maximum in late September (austral winter peak) to ~ 3 million square kilometers at minimum in late February (austral summer peak) every year (Eayrs et al., 2019; Parkinson, 2019). Sea ice is highly dynamic in Antarctica over the NWP time period, which can change by up to ~ 3 million $km^2$ over 10 days (Figure 5). From the perspective of the atmosphere, sea ice has an insulative effect between the relatively warm ocean and cold atmosphere, which highly modifies heat and momentum exchange and water vapor transport (Cassano et al., 2016; R. A. Massom & Stammerjohn, 2010), especially when established snow cover, an even more effective thermal insulator, is present. 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 ocean (Thorndike et al., 1975).

The parameterizations of small scale or complicated Earth system processes play a crucial role in determining the forecast accuracy (Bauer et al., 2015). In doing so, accurate representation of sea ice is a key point to improve predictive skill in polar atmospheric forecasts. In recent decades, having a finer (temporal and/or spatial) resolution and more comprehensive representation of sea ice has become increasingly important in both global and regional model simulations, as demonstrated in the Arctic (Hines et al., 2015; Puri et al., 2013; Rinke et al., 2006; Smith et al., 2021; Yao et al., 2016). One popular regional model designed for high latitude applications is the polar-optimized Weather Research and Forecasting model (Polar WRF). Polar WRF has been found to have a good forecast performance in polar regions because of appropriate polar modifications allowing more realistic sea ice representation (Bromwich et al., 2009), including more accurate representations of thermal properties of sea ice and snow on sea ice (Hines & Bromwich, 2008). Modifications within the Noah Land Surface model (Noah LSM) (e.g., allowing prescription of fractional sea ice (Bromwich et al., 2009) specification of spatially-varying sea ice and snow thickness) are among the most important polar optimizations to the standard WRF model (Hines et al., 2015).

Global weather forecast models also provide real-time forecast products covering the Antarctic region. The Australian Community Climate and Earth-System Simulator (ACCESS;

Puri et al., 2013) is a global atmosphere-only NWP model operated by the Australian Bureau of Meteorology (BoM), and is based on the UK Met Office Unified Model (UM; Cullen, 1993). Since there is no polar-optimized version of ACCESS currently operational, the global variant version (ACCESS-G) is used by BoM to provide real-time weather forecast guidance in polar regions (Schroeter et al., 2019). In the ACCESS-G Australian Parallel Suite (APS2) configuration (Puri et al., 2013), however, *static* sea ice is used for the lower boundary throughout the entire forecast period, and is initialized with the National Centers for Environmental Prediction (NCEP)-derived sea ice extent, i.e., the sea ice extent in the model does not update throughout each simulation. This fixing of the sea ice field throughout the forecast period is expected to have detrimental effects on forecast accuracy especially around times of maximum rate of sea ice retreat and advance, though the magnitude and spatial extents of these effects have yet to be determined. These are the primary aims of this study.

Here we show the impact of *non-static* (i.e., daily-updating) Antarctic sea ice concentration distribution on the synoptic-scale performance of NWP over a circum-Antarctic domain. In this study, our aim is to investigate how the prescription of *static* sea ice in the NWP models negatively impacts the short-term weather forecast in Antarctica and the Southern Ocean. By comparing these experiments, we will: a) explore whether or not forecast accuracy is increased when *non-static* sea ice is implemented in the Polar WRF model; b) characterize 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 NWP; d) spatially characterize the forecast improvement when daily-updated sea ice concentrations are prescribed; and e) determine why these improvements happen, i.e. could we ascribe the improvements to radiation, heat flux, or other parameters? By answering these questions, we will give a detailed analysis of the effects of *non-static* sea ice in NWP. Our research results provide the impetus to move towards operational regional coupled modelling and provide a baseline against which to compare future model performance.

# 2 Methods

Polar WRF version 4.1.1 was implemented in this study. Our domain covers the entire Antarctic region, including the region of maximum sea ice extent, using a polar stereographic projection. Domain corners reach 30° S while the latitude at the middle of domain reaches 45° S. We compare the Polar WRF output using two model experiments over a 10 day forecast period: (a) *static* sea ice (denoted *PWstatic*); and (b) daily *updating* sea ice (denoted *PWupdate*). Daily sea ice concentration distributions are taken from ERA-5 reanalysis data with 0.25° horizontal resolution (Hersbach & Dee, 2016) itself based on the EUMETSAT Data Center (EDC) OSI SAF operational dataset. Sea surface temperature (SST) is updated from ERA-5 (Hersbach et al., 2020) every six hours in both experiments. Each experiment is initiated approximately every five days in 2018 (the most complete calendar year at the start of this experiment), for the whole year, i.e., runs initiated on the 1st, 6th, 11th, 16th, 21st, 26th of each month, representing systematic seasonal sea ice phenological events in a year. We compare our model outputs with ERA-5 global reanalysis data to give a comprehensive evaluation of each experiment for the spatial and temporal impacts of *updating* sea ice concentration. More details of model configuration and experimental design can be found in the section Polar WRF Model Configuration.

## 2.1 Model Description

Polar WRF is a modification of the Weather Research and Forecasting (WRF) model aimed to better represent polar processes (Hines & Bromwich, 2008; Skamarock et al., 2008). The standard WRF model is a non-hydrostatic, mesoscale numerical forecast model designed for numerical weather prediction and atmospheric system simulation (Skamarock et al., 2008). The WRF model is developed and maintained by the USA National Center for Atmospheric Research (NCAR) and other collaborative organizations. WRF is commonly used for atmospheric research ranging from large-eddy resolving to global scale and has wide applications, such as weather forecasting, regional climate simulation, and air quality monitoring. Polar WRF is developed and maintained by Ohio State University's Polar Meteorology Group (PMG) as a code supplement to the standard WRF model (Hines et al., 2015). The model framework of Polar WRF is mainly based on the standard WRF developed by NCAR and others. Both WRF and Polar-WRF use the fully compressible and Euler non-hydrostatic equations for their dynamic components in the atmosphere scheme. The Arakawa C-grid is used as the grid staggering in the horizontal coordinate and terrain-following (TF) or hybrid vertical coordinate (HVC) are alternative options as the vertical hydrostatic pressure coordinate (NCAR and MMM, 2012).

Polar WRF modifications to WRF 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, etc (Hines et al., 2015)). The main modification of Polar WRF is optimizing the Noah LSM for better representation of snow and sea ice processes and heat transfer in polar regions (Bromwich et al., 2009). Through the implementation of fractional (i.e., non-binary) sea ice concentration (SIC), the net quantity of model parameters, such as surface temperature, heat fluxes and humidity, are calculated by the mosaic method (Hines et al., 2015):

$$A = SIC \times A_i + (1 - SIC) \times A_w$$

Where $SIC$ is the fraction of ice-covered surface within each grid cell, $A_i$ is the net quantity of each parameter of ice, and $A_w$ is the net quantity of each parameter of open water.

The surface layer scheme is first called for calculating the surface conditions assuming a 100% sea ice fraction, then for a 100% open water fraction. The surface values are the results which sum up with the weight by the actual sea ice fraction. Furthermore, the surface temperature and specific humidity in the sea ice grid are extracted before the lower surface parameters are applied to the land surface model and the selected planetary boundary layer scheme (PBL). After the computations with LSM and PBL, the net values of each parameter in the actual grid cell are reassembled. The values for the open water grid are calculated by the surface boundary layer scheme and the LSM is no longer used (Hines et al., 2015).

## 2.2 Polar WRF Model Configuration

The choice of physics parameterizations used within the study are mainly based on the latest practice in the Ohio PMG group which has been tested as a mature and appropriate physical scheme combination for 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 parameterization and updated every time step. For radiation schemes, we choose the Rapid Radiative Transfer Model for GCMs (RRTMG; Clough et al., 2005) as the parameterization for both shortwave and longwave radiation. This parameterization shows an improved radiation performance in polar regions than the prior version (Hines et al., 2015).  The shortwave and longwave radiation updates every 30 minutes. We use the Mellor-Yamada Nakanishi and Niino Level 2.5 PBL scheme (MYNN; Nakanishi & Niino, 2006) for the planetary boundary layer and update every time step. The Nakanishi and Niino PBL's surface layer scheme (Nakanishi & Niino, 2006) is used as the corresponding atmospheric surface layer. The land surface scheme is the Unified Noah Land Surface Model (Chen & Dudhia, 2001) with Polar optimization modified by the Ohio State University's Polar Meteorology Group (Hines & Bromwich, 2008). Previous research has shown that a higher model pressure layer top (i.e., representing a higher altitude) gives a better representation of gravity wave propagation (Bromwich et al., 2005), so we choose an upper model top pressure value of 3 hPa. We set up a staggered vertical grid on 71 full-η levels on WRF hybrid vertical coordinate from the sea surface to 3 hPa with vertical velocity damping within the top 8 km of the model to gain a better vertical stability. Sea ice albedo and thickness are set to uniform, circumpolar values of 0.8 and 1 m respectively. Snow depth was initialized to 5 cm. It can be increased or decreased by precipitation or melting during the simulation but is always at least 5 cm. Table 1 describes the main schemes and parameters that we implemented in the Polar WRF model configuration.

Polar WRF model configuration overview

| | |
|---|---|
| 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 | ERA-5 0.25-degree sea ice fraction |
| Initial and boundary conditions | ERA-5 0.25-degree 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 and 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 optimization |
| Microphysics | Morrison double-moment scheme |
| Cumulus Parameterization | Kain-Fritsch (KF) |
| Spin Up | First 24 h used as model spin-up time |
| Time Step | 60 S; not adaptive |

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

## 2.3 Model Domain and Input Data

Figure 1 shows the model domain used for this study. We design a single domain with a polar stereographic projection, centered at the south pole. The model domain has a 30 km horizontal resolution with 330x349 grid points, which covers 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).

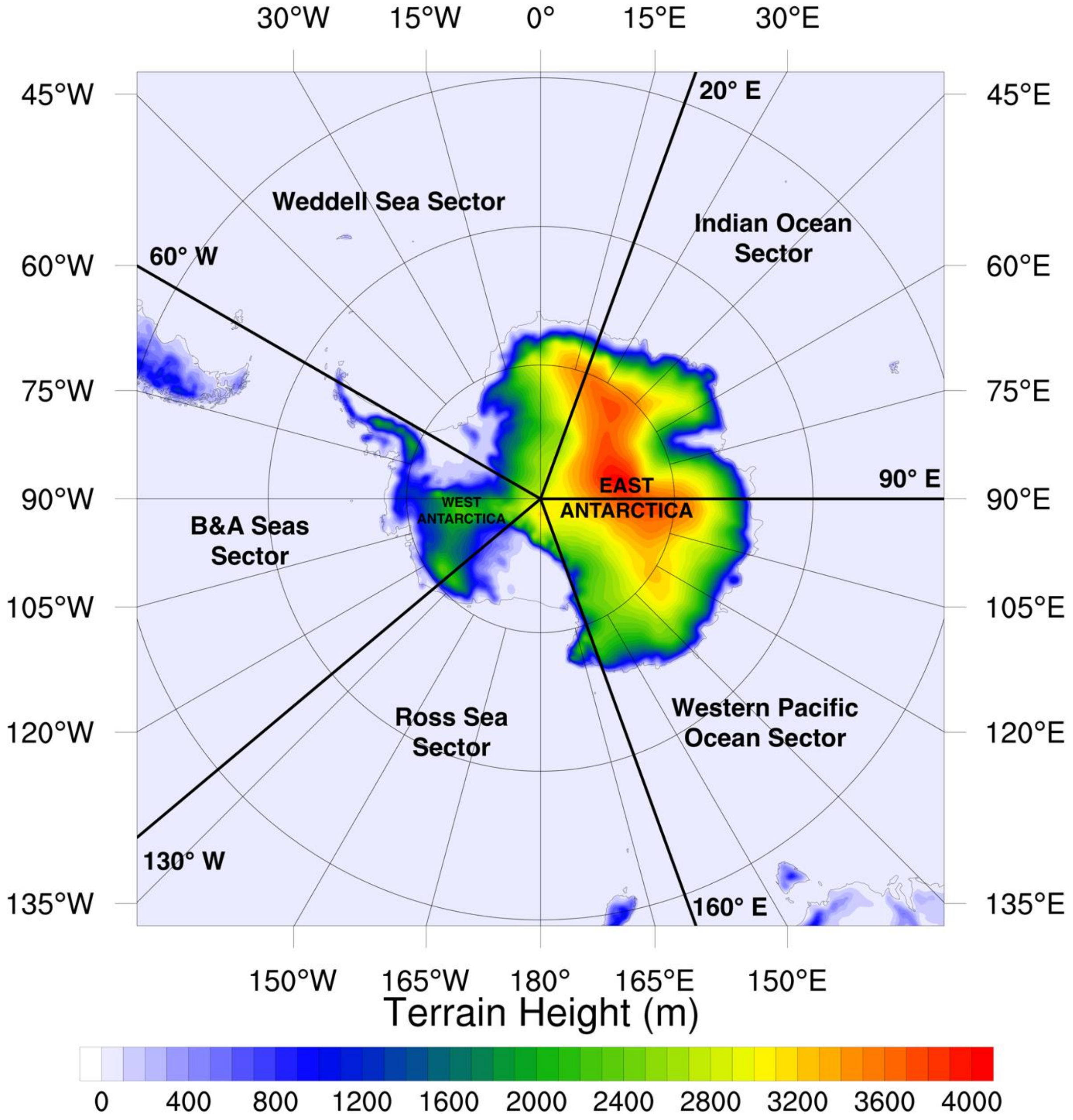


**Figure1**. The Polar WRF domain used for this study.

WRF requires meteorological variables as both initial and boundary conditions. The model's initial and boundary conditions are obtained from ECMWF ERA-5 0.25-degree global reanalysis data (Hersbach & Dee, 2016). 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 1 km resolution. The SST and SIC are also obtained from ERA-5 with the same grid size and time interval. Meteorological input data updates occur every 6 hours, so the WRF Preprocessing System (WPS) will linearly assimilate the SIC and SST data into the lower boundary with the same frequency if the SST update option is used.

## 2.4 Experimental Design

In order to investigate the influence of updating sea ice concentrations on the accuracy of NWP, Polar WRF is configured to run a suite of 10-day forecasts. Prior to each 10-day forecast, we perform a spin-up for 24 hours 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 & Cassano, 2003), and Hines and Bromwich (2008) found minimal difference between a 12 hour and 24 hour spin-up for Polar WRF. Considering that we perform a relatively long forecast (10 days), we choose a 24-hour spin-up time, i.e., a total of 11 days model running period in each case. In order to ensure all the meteorological conditions are identical between the two experiments after spin-up, all the parameterizations and initial/boundary conditions stay the same, including the sea ice concentration and SST updates, throughout the first 24 hours.

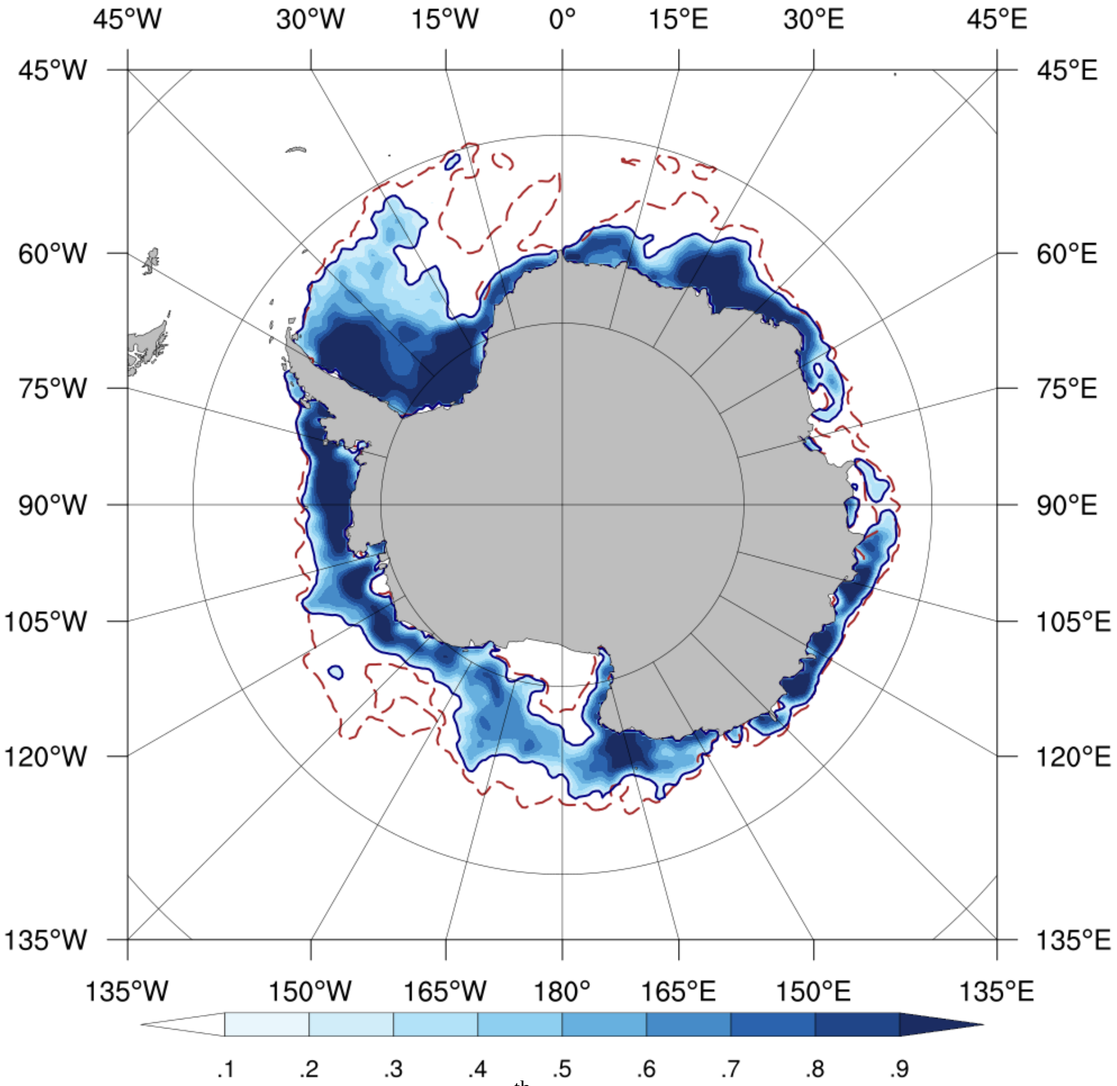


**Figure 2**. Sea ice concentration on the 26th of December 2018 and the sea ice edge on the 17th (red dashed line) and 26th (solid blue line) of December. The magnitude of retreat of the sea ice extent over this period is 1.40 million km$^2$.

## 2.5 Evaluation Data and Validation Methods

ERA-5 global reanalysis data (Hersbach & Dee, 2016) were used to evaluate the model performance of each experiment. ERA-5 is the fifth generation of European Centre for Medium-Range Weather Forecasts (ECMWF) atmospheric reanalyzes of the global climate. ERA-5 provides hourly atmospheric reanalysis data at a high spatial resolution (0.25°x0.25°). Although using reanalysis data as the reference to evaluate model performance is not ideal due to the difficulties associated with validation of re-analysis itself in the region characterized by a lack of observational data, this method is commonly used in operational verification (e.g., Eerola, 2013; Schroeter et al., 2019). While comparison with the 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, we treat ERA-5 as the “real world” reference when comparing to Polar WRF model output.

Our two experiments are evaluated by comparing hourly forecast output from Polar WRF against hourly ERA-5 data for several surface and near-surface variables, as well on pressure levels throughout the atmosphere. Key parameters at the near-surface level, such as the 10 m winds (U10/V10) and 2 m temperature (T2m) and dewpoint (TD2m), are compared each hour throughout the 10-day forecast period. On pressure levels, geopotential height, u and v winds, temperature and relative humidity at 37 levels are selected to investigate the upper-level model performance.  The ERA-5 data are interpolated to the same grid as the Polar WRF model output using spline interpolation. We choose four commonly-used validation metrics: mean error (ME), mean absolute error (MAE), root-mean-squared error (RMSE) and the Pearson correlation coefficient (CORR).

ME is expressed as:

$$ME = \frac{\sum_{i=1}^{n} y_i - x_i}{n} \quad (1)$$

Where $y_i$is a particular variable from the WRF output, $x_i$ is the counterpart from ERA-5 reanalysis, and $n$ represents the sample size for the variable of interest.
This metric is used to observe whether the estimated value is over- or under-estimated since the sign of the bias is taken into consideration. Since the magnitude of the ME metric tends toward zero, we also use the mean absolute error (MAE) to represent the error magnitude:

$$MAE = \frac{\sum_{i=1}^{n} |y_i - x_i|}{n} \quad (2)$$

We also assess model performance using root-mean-squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - x_i)^2}{n}} \quad (3)$$

RMSE is more appropriate than MAE when the error distribution is Gaussian (Chai & Draxler, 2014). However, this metric is sensitive to outliers. Hence, a comprehensive assessment of errors based on a variety of metrics is necessary in our study. The grid cell area-weighted ME, MAE and RMSE were also used for computing the domain-averaged statistics.

We use a two-tailed paired t-test to examine whether the RMSE between *PWstatic and PWupdate* are significantly different. As the hourly time series of atmosphere properties exhibit strong autocorrelation in each grid cell (Su et al., 2021), we take account of the effective number of degrees of freedom, following the methods in Davis (1976) when applying the t-test.

The linear dependence of WRF output and ERA-5 is measured by CORR. We did not detrend the data because the time periods are too short to include climate trends. We also did not remove the diurnal cycle because the cycle is also an important assessment point in this study.

Bootstrapped confidence intervals are estimated in order to provide a measure of uncertainty in RMSE (Davison & Hinkley, 1997). We create 2,000 bootstrap samples by randomly sampling 100 times with replacement from the domain-averaged time series of 72 model runs, and calculate the 95% confidence interval for the sample average.

# 3 Results

## 3.1 Model performance and quality control

The annual average model forecast performance (*PWupdate*) for five near-surface variables and three assessment metrics is shown in Figure 3. We calculated the RMSE, MAE and ME of the variables in the first 48 hours (after spin-up) of each forecast period against ERA-5, then averaged these for the 72 forecast periods to provide an annual-average performance. For T2m, the performance is very good over the Southern Ocean (as expected due to regularly-updating SST), with <1 K MAE and RMSE. Relatively larger errors are observed over the Antarctic continent. Our model tends to overestimate (positive ME) 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 (Hines et al., 2019; Valkonen et al., 2014) where the RMSE in T2m are 1.7 ~ 2.8 K and 2.0 ~ 2.7 K, respectively. The errors obtained within the first 48 hour forecast 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 ERA-5 on the Antarctic Continent, even though there is noise present, with the magnitudes of the order of the difference between WRF and ERA-5 surface pressure around the Antarctic continent coastlines and the region with steep orography. This noise appears to be partly caused by the spline interpolation method, which interpolates the ERA-5 pressure to the WRF grid. The good simulation of PSFC displayed here indicates that Polar WRF has the ability to accurately model synoptic-scale pressure systems. U10m and V10m also show good agreement with ERA-5. U10m shows an underestimation (negative ME) along the coastal region of 1 ~ 2 $m.s^{-1}$ while V10m shows a slight overestimate with a similar magnitude. The magnitude of errors (MAE and RMSE) of near-surface winds are smaller than 3 $m.s^{-1}$ throughout almost the entire domain.

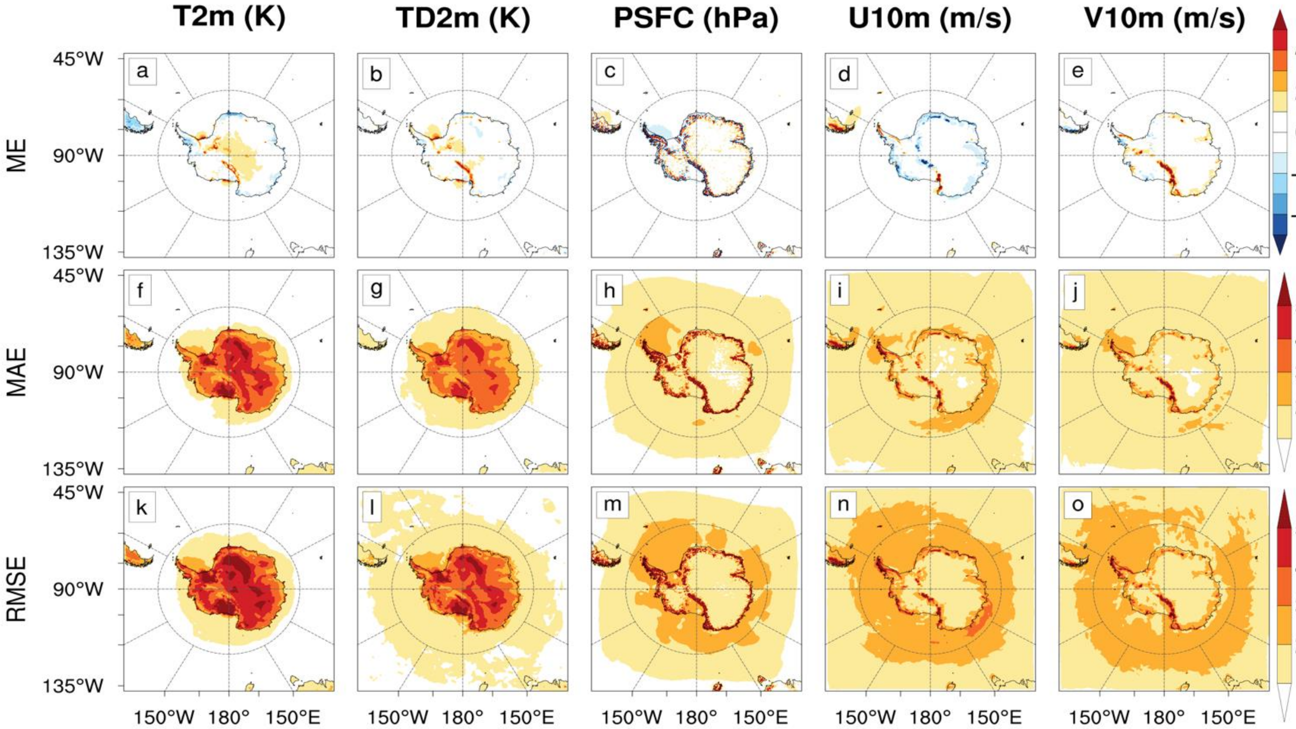


**Figure 3.** Model validation figure showing the annual average ME, MAE and RMSE in T2m, TD2m, PSFC and U10m/V10m in the zonal and meridional direction. These metrics are calculated using the *PWupdate* experiment output against ERA-5 reanalysis. The metrics were calculated using 72 model runs for the first 48 hours of each forecast period (after spin-up).

## 3.2 Influence of updating sea ice on the near-surface variables

Following validation, which indicated that this configuration of Polar WRF is appropriate for addressing the aims of this project, we now compare *PWupdate* and *PWstatic* experiments with ERA-5 reanalysis for the near-surface variables (Table 2). Of all the surface variables, T2m and TD2m show the most significant improvement when dynamic (updating) fractional sea ice is implemented. Figure 4 shows the difference of RMSE of T2m and TD2m between the simulation with *updated* sea ice and *static* sea ice, which were averaged from the 72 model runs. In almost all sea ice regions, the *updated* sea ice has a positive impact on the T2m and TD2m forecasts. For many regions, the difference becomes statistically significant (90% confidence level) after 2 days with marked improvement in the Ross Sea and Weddell Sea sectors. By checking the seasonal average of the difference for T2m (Figure 5), the improvement was mainly contributed between July and September, corresponding to the period of late sea ice advance. In September, the ice-covered ocean is more than 6 times larger than in late February. Updating sea ice concentration values appears to make a useful contribution to the model during this high heat flux period. TD2m also has significant improvements when using *updated* sea ice. The regional pattern of TD2m is very similar to that of T2m.

Figure 4 shows the influence of dynamic sea ice on PSFC. Unlike the temperature and humidity variables, the improvement for PSFC appears to be mixed. The impacts become statistically significant (90% confidence level) after 8 days of forecast while distinct influences can be seen in T2m and TD2m after only two days. From the day 7 to 8 averaged RMSE, we find a small positive modification at the Bellingshausen and Weddell Sea sectors (80% confidence level) and Antarctica continent in the Indian Ocean Sector (90% confidence level), and a small negative modification over the sea ice region (80% confidence level) and Antarctic continent of the Western Pacific Ocean Sector (90% confidence level). The impacts become significant after 8 days. A more than 0.5 hPa RMSE reduction was detected in the Amundsen Sea Sector when using updated sea ice while a more than 0.5 hPa RMSE increase was detected in the Bellingshausen Sea Sector as well. Table 2 shows the *PWupdate* generally has a slightly smaller domain-averaged RMSE than the *PWstatic*.

The near-surface wind forecast shows reasonable improvement in the *PWupdate* experiment. Table 2 indicates that U10m and V10m have smaller domain averaged RMSE values (south of 60° S) in nearly every day in the first 10 day forecasts, except the zonal wind speed in the day 9 to day 10 average where they become similar. Spatially, both U10m and V10m show an improvement in West Antarctica, mainly resulting from the rapid sea ice advance and retreat in these areas. Temporarily, by averaging the U10m and V10m every three months (figure not shown), seasonality is not easily detectable but the impact tends to appear more quickly in winter. Namely, the influence becomes detectable after 4 days in winter around the sea ice region while it needs 6 days in summer.

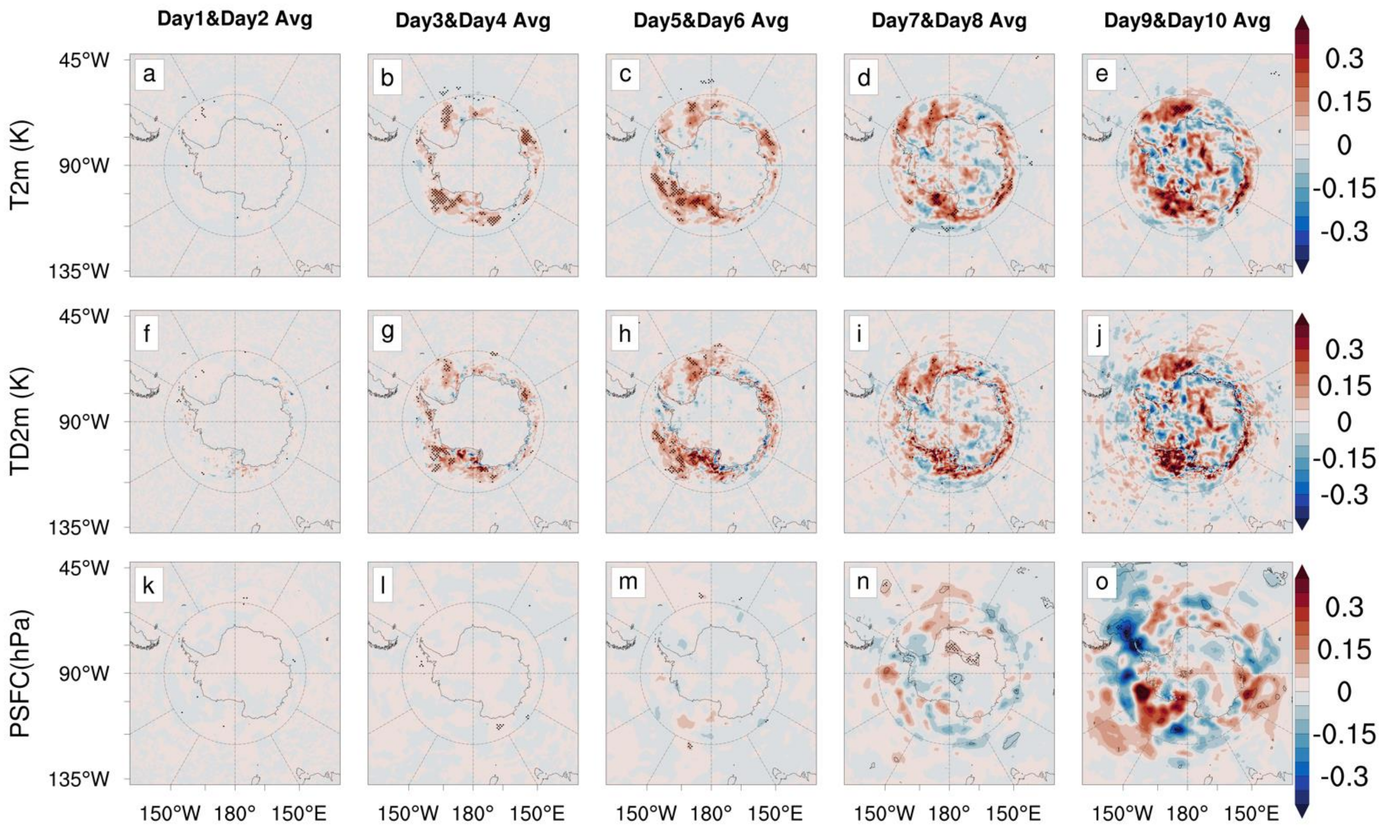

**Figure 4.** The difference of annual average RMSE in T2m, TD2m and PSFC between Polar WRF with static sea ice and updated sea ice when compared with ERA-5 reanalysis data. The metrics were calculated by the average of 72 model runs during 2018 using every 2 day average of each forecast period. Here red shading indicates that updated sea ice outperforms static sea ice. Stippled areas indicate differences are significant at the 90% level.

| Domain-Averaged Surface Air Statistics for Every 2 day Forecast Average (south of 60 degree) | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | day1-day2 average | | | | day3-day4 average | | | | day5-day6 average | | | | day7-day8 average | | | | day9-day10 average | | | |
| | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR |
| 2m Temperature (K) | | | | | | | | | | | | | | | | | | | | |
| ERA5 | -16.61 | | | | -16.58 | | | | -16.61 | | | | -16.61 | | | | -16.62 | | | |
| updated sea ice | -16.59 | 0.02 | 3.22 | 0.960 | -16.21 | 0.37 | 4.18 | 0.911 | -16.08 | 0.54 | 5.13 | 0.849 | -15.94 | 0.67 | 5.89 | 0.784 | -16.00 | 0.62 | 6.31 | 0.738 |
| static sea ice | -16.58 | 0.03 | 3.22 | 0.960 | -16.18 | 0.40 | 4.20 | 0.909 | -16.02 | 0.59 | 5.17 | 0.844 | -15.87 | 0.74 | 5.93 | 0.777 | -15.90 | 0.72 | 6.36 | 0.728 |
| 2m 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 | 0.890 | -19.22 | 0.76 | 5.45 | 0.816 | -19.08 | 0.89 | 6.28 | 0.744 | -19.12 | 0.86 | 6.80 | 0.690 |
| static sea ice | -19.77 | 0.21 | 3.22 | 0.951 | -19.33 | 0.60 | 4.38 | 0.889 | -19.17 | 0.81 | 5.48 | 0.812 | -19.02 | 0.95 | 6.32 | 0.738 | -19.03 | 0.94 | 6.84 | 0.683 |
| Surface Pressure (hPa for average, Pa for others) | | | | | | | | | | | | | | | | | | | | |
| ERA5 | 898.35 | | | | 898.10 | | | | 898.30 | | | | 898.20 | | | | 898.01 | | | |
| updated sea ice | 898.40 | 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 | 898.40 | 5.10 | 344.97 | 0.984 | 897.64 | -45.60 | 536.39 | 0.927 | 897.16 | -114.41 | 798.97 | 0.802 | 896.76 | -143.84 | 1031.28 | 0.665 | 896.95 | -106.27 | 1215.33 | 0.514 |
| 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 | 2.70 | 0.893 | -0.08 | -0.14 | 4.16 | 0.733 | -0.05 | 0.03 | 5.35 | 0.547 | -0.03 | 0.00 | 6.17 | 0.403 | -0.15 | -0.20 | 6.63 | 0.315 |
| static sea ice | -0.36 | -0.27 | 2.70 | 0.893 | -0.08 | -0.14 | 4.17 | 0.733 | -0.05 | 0.03 | 5.36 | 0.547 | -0.02 | 0.01 | 6.18 | 0.402 | -0.14 | -0.20 | 6.63 | 0.314 |
| 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 | 0.00 | 3.82 | 0.727 | 0.82 | -0.07 | 4.95 | 0.537 | 0.82 | -0.07 | 5.70 | 0.392 | 0.85 | 0.00 | 6.12 | 0.285 |
| static sea ice | 1.01 | 0.09 | 2.49 | 0.888 | 0.85 | 0.00 | 3.83 | 0.726 | 0.82 | -0.07 | 4.95 | 0.537 | 0.83 | -0.06 | 5.71 | 0.392 | 0.86 | 0.01 | 6.13 | 0.284 |

**Table 2.** The domain-averaged mean state, ME, RMSE and Pearson's correlation of T2m, TD2m, PSFC and U10m/V10m with their ERA-5 counterparts.

Since T2m shows the largest improvement, we provide a more comprehensive error analysis south of 60° S. The Hovmöller diagram of the domain averaged RMSE difference between *static* and *updated* sea ice (Figure 5a) shows that using *updated* sea ice can improve the forecast skill of T2m for most forecast dates and forecast time periods of the year. June, July and August show the largest improvement and fastest response to the realistic sea ice updating. In these three months, the surface temperature shows positive modification after only three days forecast time, while it needs twice as long in other months. The maximum mean RMSE reduction can reach 0.5 K in +216 hours forecast and thereafter, and 0.3 K in +144 hours forecast in June and September. Figure 5b shows the updated sea ice has positive effects in most months of 2018, especially in the months of sea ice advance. Figure 5c shows that the updated sea ice outperforms static sea ice at nearly every forecast time period on average. As a whole, the contribution of updated sea ice is generally increased with the passing of the forecast time. Figure 5d shows the 10-day change in sea ice extent and area initiated from each forecast period. We find 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.

We posit that the mechanism underlying the T2m forecast skill increase, mainly during sea ice advance, relies on a 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 5b and d, we find 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 August while the rate of sea ice advance peaks in April. We consider 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 (vs sea ice to atmosphere) is not remarkable). (c) significant and rapid skill increase during sea ice advance in the presence of a cold near-surface air temperature (e.g., July - August) - 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.

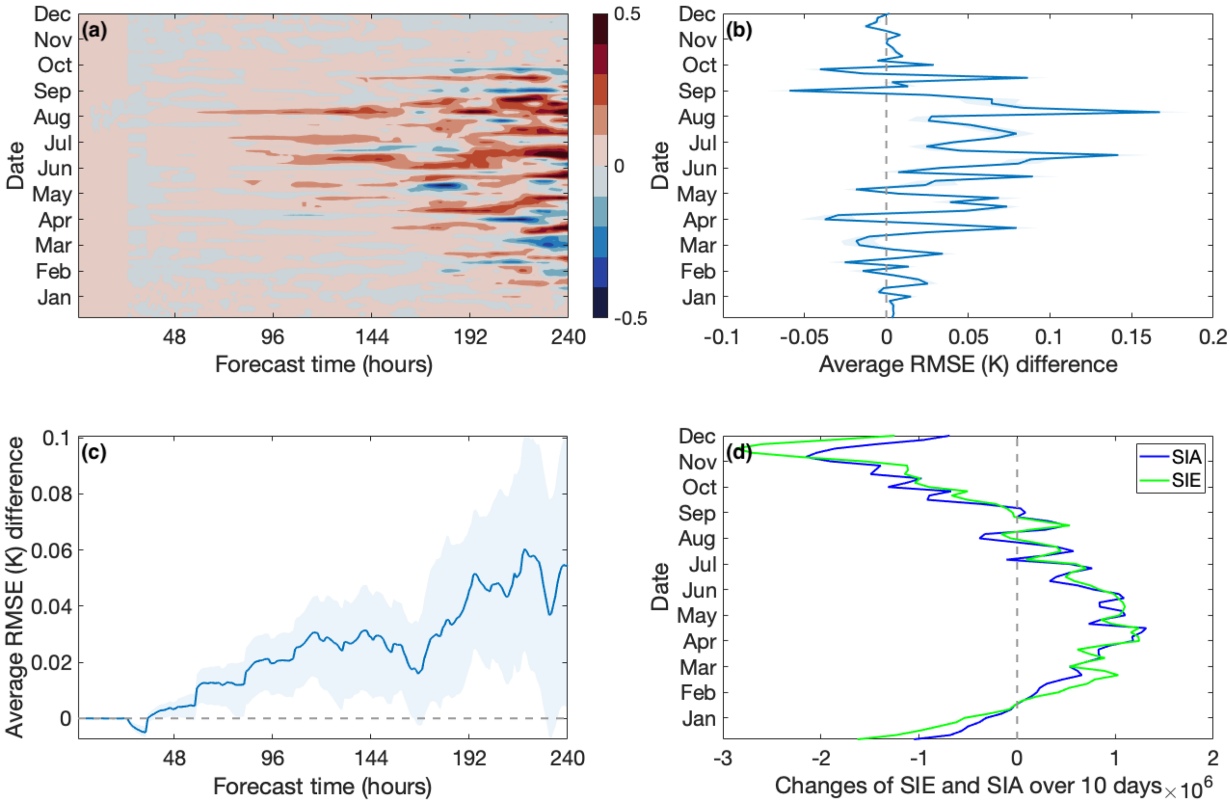


**Figure 5.** (a) Hovmöller diagram of domain-averaged RMSE difference south of 60° S for T2m between static and updated sea ice experiments compared with ERA-5. Here red indicates that *PWupdate* outperforms the simulation with static sea ice. The x-axis represents the forecast time of each experiment.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

## 3.3 Influence on Surface Heat Budget Balance

In order to diagnose the mechanism of T2m improvement, we now compare the influence of *updated* sea ice on the surface energy balance. 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 the T2m has a remarkable improvement, it is reasonable to suppose the updated realistic sea ice makes a positive contribution to surface heat budget balance modification. Table 3 shows the terms contributing to the net surface heat flux balance south of 60° S in ERA-5, and the statistics for *PWupdate and PWstatic*. Figure 6 spatially shows the four surface heat flux terms and the surface net heat flux with statistically-significant responses (90% confidence level) for the *updated* sea ice assimilation.

Table 3 gives a comprehensive assessment and comparison for area-weighted surface energy balance for every 2-day forecast average (south of 60º S). For radiative heat fluxes in each upward and downward direction, the ME, RMSE and CORR show improvements in all five 2-day averaging periods. The outgoing shortwave radiation has the best improvement (when shortwave is present), where the area-weighted average RMSE reduced from 35.2 to 29.3 $W.m^{-2}$ and the CORR increased from 0.93 to 0.96 for the day 9 to 10 averaging period. The outgoing longwave radiation CORR increased from 0.80 to 0.82 while the RMSE reduction is insignificant (22.9 to 22.4 $W.m^{-2}$). For sensible heat flux, the WRF forecasted value has large biases around Antarctica compared with ERA-5. Previous research also shows poor correlation of sensible heat flux from WRF against observations (Tastula et al., 2012; Valkonen et al., 2014). However, we still find *PWupdate* reduces the RMSE from 35.7 to 35.1 $W.m^{-2}$ and increases the CORR from 0.45 to 0.47 for the day 9 to 10 averaging period. For latent heat flux, the ME of *PWupdate* is always larger than that of *PWstatic* i.e., *PWupdate* tends to underestimate the latent heat flux. However, the CORR in *PWupdate* is always larger than that in *PWstatic*. The lower RMSE of *PWupdate* shows the *PWupdate* forecasted latent heat flux is closer to the reanalysis field.

| Domain-Averaged Surface Energy Balance for Every 2 day Forecast Average (south of 60 degree) | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | day1-day2 average | | | | day3-day4 average | | | | day5-day6 average | | | | day7-day8 average | | | | day9-day10 average | | | |
| | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR | Ave | ME | RMSE | CORR |
| LW↓ (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 212.66 | 6.84 | 24.92 | 0.814 | 214.10 | 7.92 | 32.10 | 0.685 | 214.48 | 8.81 | 38.00 | 0.557 | 215.13 | 8.97 | 42.06 | 0.457 | 215.19 | 9.31 | 44.75 | 0.376 |
| LW↑ (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 251.99 | 1.53 | 12.72 | 0.963 | 253.29 | 2.74 | 15.89 | 0.926 | 253.80 | 3.37 | 18.93 | 0.879 | 254.26 | 3.79 | 21.27 | 0.833 | 254.24 | 3.81 | 22.85 | 0.799 |
| SW↓ (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 125.12 | 4.95 | 43.38 | 0.973 | 125.31 | 5.15 | 50.09 | 0.965 | 125.34 | 5.04 | 55.74 | 0.956 | 125.11 | 5.33 | 59.60 | 0.949 | 125.50 | 5.11 | 61.62 | 0.946 |
| SW↑(W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 69.18 | 1.30 | 22.77 | 0.976 | 69.17 | 1.52 | 26.82 | 0.964 | 69.56 | 1.77 | 30.45 | 0.949 | 69.93 | 2.28 | 33.10 | 0.939 | 70.34 | 2.65 | 35.22 | 0.931 |
| SH (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 4.44 | 8.00 | 21.20 | 0.815 | 3.70 | 7.46 | 25.58 | 0.708 | 3.34 | 6.87 | 30.37 | 0.601 | 3.18 | 6.78 | 33.31 | 0.508 | 3.16 | 6.76 | 35.68 | 0.451 |
| LH (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| 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 | 14.09 | -0.03 | 11.21 | 0.797 | 13.92 | -0.06 | 15.95 | 0.706 | 13.76 | -0.21 | 19.47 | 0.610 | 13.84 | -0.20 | 22.26 | 0.540 | 13.87 | -0.14 | 24.10 | 0.487 |
| Qnet (W/m^2) | | | | | | | | | | | | | | | | | | | | |
| ERA5 | -2.92 | | | | -2.08 | | | | -2.68 | | | | -2.63 | | | | -4.02 | | | |
| 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 | -1.92 | 0.99 | 42.19 | 0.639 | -0.67 | 1.41 | 52.29 | 0.608 | -0.64 | 2.05 | 61.94 | 0.577 | -0.98 | 1.66 | 68.58 | 0.555 | -0.92 | 1.34 | 73.44 | 0.537 |

**Table 3.** The domain-averaged mean state, ME, RMSE and Pearson's correlation of each term of surface energy budget with their ERA-5 counterparts.

From Figure 6, we find the outgoing radiative surface fluxes, including both shortwave and longwave, show clear improvement over the domain. 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 shows the most significant RMSE reduction, 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 outgoing longwave radiation RMSE reduction is likely driven by a more realistic surface temperature simulation when including an *updated* sea ice field. We also find the improvement appears mainly in the sea ice region showing the better sea ice description brings an increased forecast skill on surface heat budget. The RMSE reduction for outgoing radiative heat flux appears to only affect the sea ice region on a relatively long timescale (10 days) with limited outward spreading.

By checking the seasonality of surface heat budget (figure not shown), the upward shortwave radiation shows a statistically significant improvement in Antarctic summer, while the longwave radiation shows a reasonable improvement in sea ice advance season. The improved longwave radiation during the period of sea ice advance further indicates that it is due to the surface temperature improvement. Sensible and latent heat flux also gained forecast skill in winter where the larger temperature difference occurs between ocean and atmosphere.

As expected from an experiment modifying the sea ice concentration, both downward shortwave and longwave radiation improvement are not as strong as those in the upward direction. The downward shortwave radiation shows an amount of noise when comparing the static and updated sea ice experiments (figure not shown). This is mainly because of the modifications in the upper-level atmosphere and cloud simulation, which will be discussed later. 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 to be used 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, we still report an improvement in the Antarctic coastal region and the region covered by sea ice, while pockets of skill decrease still exist at the ice edge.

Both latent and sensible heat fluxes show improvements over the domain 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 and coastal polynyas.

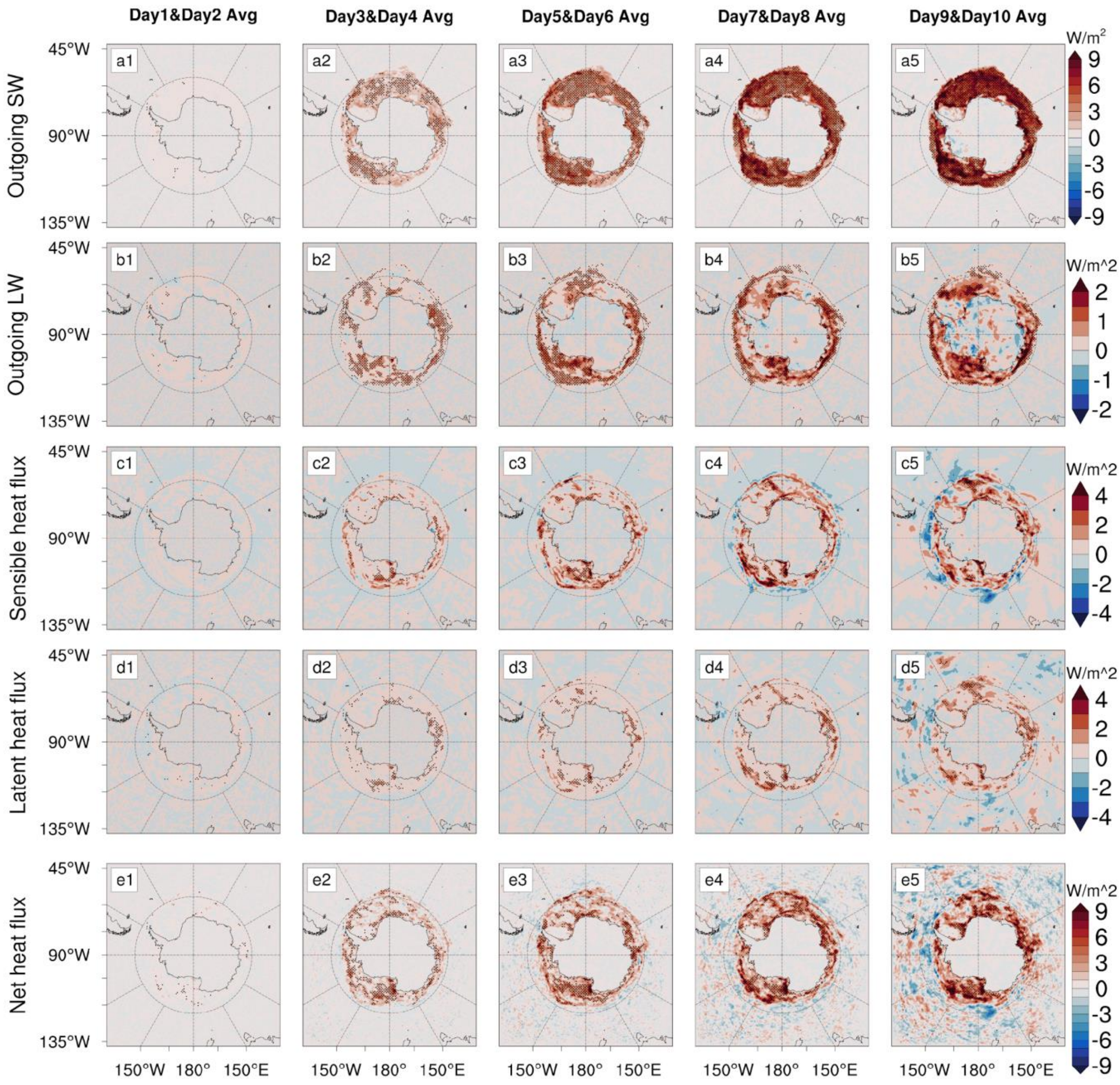


**Figure 6**. Difference of annual-averaged RMSE in outgoing shortwave/longwave radiation fluxes, latent heat flux, sensible heat flux and net heat flux at the surface between Polar WRF with static sea ice and updated sea ice when compared with ERA-5 reanalysis data. Here red shading indicates that updated sea ice outperforms static sea ice. Stippling indicates differences significant at the 90% level.

# 3.4 Influence on Vertical Structure of the Atmosphere

The 3-hourly vertical profile of three Polar WRF model variables (air temperature, relative humidity and geopotential height) were interpolated to 37 ERA-5 pressure levels to facilitate comparison. This is shown in Figure 7.

For the air temperature vertical profile (Figure 7, top row), there is a clear improvement in sea ice-covered latitudes (65° S to 75° S) from the surface to 900 hPa. After +8 days, the improvement can reach further inland (75° S to 82° S), and propagate to the 600 hPa level. However, the *PWupdate* experiment tends to have 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.

The humidity is the most challenging variable to model in the upper troposphere, especially in the polar regions(Elliott & Gaffen, 1991; Wilson et al., 2011). Our relative humidity results show a large error in RMSE between WRF output and ERA-5. Relative humidity shows an RMSE of 9 to 27% between 900 and 500 hPa, even in the first 48 hours, and the RMSE can reach 45% after 10 days. Since both the WRF output and ERA-5 reanalysis values are subject to biases, we expected a relatively poor correspondence, however we present the upper troposphere relative humidity analysis for reference, and we do not interpret results above 500 hPa. Relative humidity shows a forecast skill decrease in the *PWupdate* experiment for the first 6 days, while an improvement occurs from day 7 to 8 at around 950 hPa from 68° S to 82° S.

The RMSE of geopotential height between WRF and ERA-5 stays within a reasonable and acceptable range. Even though the *non-static* sea ice does not depict a perceptible improvement in the first 8 days forecast, the negative effects are not significant, especially in the first 6 days. Considering the surface pressure is also not strongly changed, this indicates that the pressure may not be significantly impacted by including *updated* sea ice.

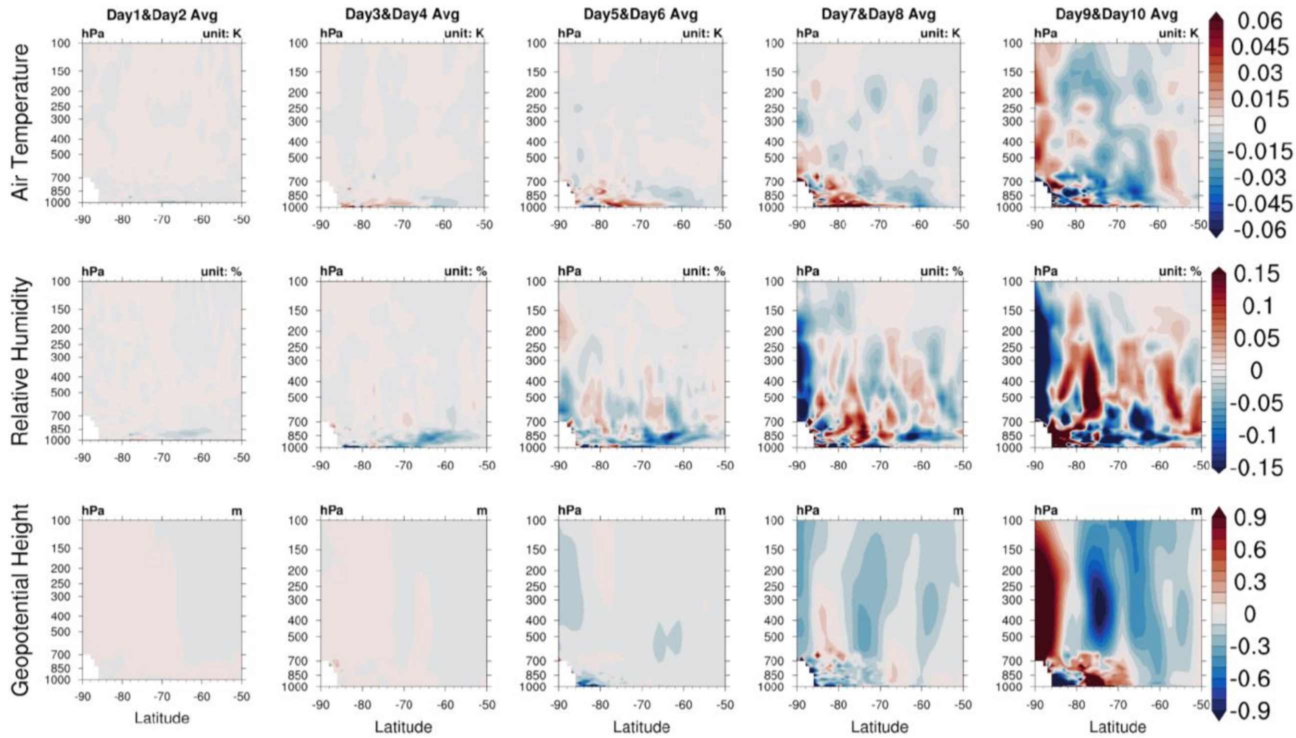


**Figure 7.** Vertical profiles of difference in annual average RMSE in air temperature, relative humidity and geopotential height between Polar WRF with static sea ice and non-static sea ice when compared with ERA-5 reanalysis data. Here red indicates that the update outperforms static sea ice.

# 4 Summary and 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. ERA-5 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. Polar WRF with *non-static* sea ice forecasted near-surface variables compared well with the ERA-5 reanalysis, showing a reasonable bias and very high correlation. The ME, MAE and RMSE errors had good agreement with previous studies (Hines et al., 2019; Valkonen et al., 2014; Wilson et al., 2011). The bias and error 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 research to capture the effects of sea ice.

For near-surface variables, the T2m and TD2m have the most statistically significant improvements. The surface pressure appears to be relatively insensitive to *non-static* sea ice in forecast time length. Near-surface winds, both in the meridional and zonal directions, show an improvement but the magnitudes are relatively small.

Although we found limited improvement in indirectly-influenced variables, such as winds and surface pressure, the improvements are not as remarkable as the directly-influenced variables, namely, the near-surface temperature and humidity. The temperature and humidity are directly impacted due to the better simulation of surface heat fluxes, while the winds and surface pressure improvement may owe to the modification of temperature and humidity. Based on this, 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 seasonal or climate context.

Adding *updated* sea ice into the Polar WRF model significantly modified the surface heat budget balance. 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 heat energy budget balance, which results in the modification of surface temperature.

Even though our research found a statistically significant improvement at the surface and near-surface when adding *updated* realistic sea ice concentration 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 error propagation is steadily weakening since the turbulence and vertical mixing are rare. Compared with the Arctic, the polar vortex over Antarctica is stronger and more resistant to block air mass exchange with mid-latitudes (Qian et al., 2021; Waugh & Randel, 1999). 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 statistically significant in +120 to +192 hour forecast, south of 60° S, and below 700 hPa.

Further research 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, so the annual average analyses presented here may ignore some important fine-scale responses. A series of case studies showing maximum impacts on atmospheric factors is necessary to provide more insight on processes contributing to the forecast improvement. Furthermore, the effect of a more realistic prescription of sea ice thickness and snow depth 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 computational-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 Ohio State University PMG group for providing the Polar WRF model and the constructive and helpful discussion. 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 ERA-5 reanalysis data.

# Reference

Al-Yahyai, S., Charabi, Y., & Gastli, A. (2010). Review of the use of numerical weather prediction (NWP) models for wind energy assessment. *Renewable and Sustainable Energy Reviews*, *14*(9), 3192–3198.

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. *Nature*, *525*(7567), 47–55.

Bromwich, D. H., Bai, L., & Bjarnason, G. G. (2005). High-resolution regional climate simulations over iceland using polar MM5. *Monthly Weather Review*, *133*(12), 3527–3547. https://doi.org/10.1175/MWR3049.1

Bromwich, D. H., Hines, K. M., & Bai, L. S. (2009). Development and testing of Polar Weather Research and Forecasting model: 2. Arctic Ocean. *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1029/2008JD010300

Bromwich, D. H., Otieno, F. O., Hines, K. M., Manning, K. W., & Shilo, E. (2013). Comprehensive evaluation of polar weather research and forecasting model performance in the antarctic. *Journal of Geophysical Research Atmospheres*, *118*(2), 274–292. https://doi.org/10.1029/2012JD018139

Cassano, J. J., Seefeldt, M. W., Palo, S., Knuth, S. L., Bradley, A. C., Herrman, P. D., et al. (2016). Observations of the atmosphere and surface state over Terra Nova Bay, Antarctica, using unmanned aerial systems. *Earth System Science Data*. https://doi.org/10.5194/essd-8-115-2016

Chai, T., & Draxler, R. R. (2014). Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, *7*(3), 1247–1250. https://doi.org/10.5194/gmd-7-1247-2014

Chen, F., & Dudhia, J. (2001). Coupling and advanced land surface-hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Monthly Weather Review*, *129*(4), 569–585. https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2

Clough, S. A., Shephard, M. W., Mlawer, E. J., Delamere, J. S., Iacono, M. J., Cady-Pereira, K., et al. (2005). Atmospheric radiative transfer modeling: A summary of the AER codes. *Journal of Quantitative Spectroscopy and Radiative Transfer*. https://doi.org/10.1016/j.jqsrt.2004.05.058

Cullen, M. J. P. (1993). The unified forecast/climate model. *Meteorological Magazine*, *122*(1449), 81–94.

Davis, R. E. (1976). Predictability of sea surface temperature and sea level pressure anomalies over the North Pacific Ocean. *J. Phys. Oceanogr.*, *6*, 249–266. https://doi.org/10.1175/1520-0485(1976)006<0249:POSSTA>2.0.CO;2

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge university press.

Eayrs, C., Holland, D., Francis, D., Wagner, T., Kumar, R., & Li, X. (2019). Understanding the Seasonal Cycle of Antarctic Sea Ice Extent in the Context of Longer-Term Variability. *Reviews of Geophysics*, *57*(3), 1037–1064. https://doi.org/10.1029/2018RG000631

Ebert, E., Wilson, L., Weigel, A., Mittermaier, M., Nurmi, P., Gill, P., et al. (2013). Progress and challenges in forecast verification. *Meteorological Applications*, *20*(2), 130–139. https://doi.org/10.1002/met.1392

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

Eerola, K. (2013). Twenty-one years of verification from the HIRLAM NWP system. *Weather and Forecasting*. https://doi.org/10.1175/WAF-D-12-00068.1

Elliott, W. P., & Gaffen, D. J. (1991). On the utility of radiosonde humidity archives for climate studies. *Bulletin - American Meteorological Society*. https://doi.org/10.1175/1520-0477(1991)072<1507:OTUORH>2.0.CO;2

Hersbach, H, & Dee, D. (2016). ERA5 reanalysis is in production. *ECMWF Newsletter*, *147*(7), 5–6.

Hersbach, Hans, Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., et al. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*. https://doi.org/10.1002/qj.3803

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

Hines, K. M., Bromwich, D. H., Bai, L., Bitz, C. M., Powers, J. G., & Manning, K. W. (2015). Sea Ice Enhancements to Polar WRF*. *Monthly Weather Review*, *143*(6), 2363–2385. https://doi.org/10.1175/MWR-D-14-00344.1

Hines, K. M., Bromwich, D. H., Wang, S. H., Silber, I., Verlinde, J., & Lubin, D. (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). *Atmospheric Chemistry and Physics*, *19*(19), 12431–12454. https://doi.org/10.5194/acp-19-12431-2019

Jung, T., & Matsueda, M. (2016). Verification of global numerical weather forecasting systems in polar regions using TIGGE data. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 574–582. https://doi.org/10.1002/qj.2437

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

Kain, J. S. (2004). The Kain - Fritsch convective parameterization: An update. *Journal of Applied Meteorology*. https://doi.org/10.1175/1520-0450(2004)043<0170:TKCPAU>2.0.CO;2

Liu, H., Jezek, K., Li, B., & Zhao, Z. (2001). Radarsat Antarctic Mapping Project digital elevation model version 2. *National Snow & Ice Data Center*.

Mass, C. F., & Kuo, Y. H. (1998). Regional Real-Time Numerical Weather Prediction: Current Status and Future Potential. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/1520-0477(1998)079<0253:RRTNWP>2.0.CO;2

Massom, R., & Lubin, D. (2006). *Polar Remote Sensing - Volume I: Atmosphere and Oceans*. *Chichester, UK : Springer*. https://doi.org/10.1007/3-540-30565-3

Massom, R. A., & Stammerjohn, S. E. (2010). Antarctic sea ice change and variability - Physical and ecological implications. *Polar Science*, *4*(2), 149–186. https://doi.org/10.1016/j.polar.2010.05.001

Morrison, H., Thompson, G., & Tatarskii, V. (2009). Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Monthly Weather Review*. https://doi.org/10.1175/2008MWR2556.1

Nakanishi, M., & Niino, H. (2006). An improved Mellor-Yamada Level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Boundary-Layer Meteorology*. https://doi.org/10.1007/s10546-005-9030-8

NCAR and MMM. (2012). ARW: Version 3 Modeling System User's Guide. In *WRF*.

Parish, T. R., & Cassano, J. J. (2003). The role of katabatic winds on the Antarctic surface wind regime. *Monthly Weather Review*, *131*(2), 317–333. https://doi.org/10.1175/1520-0493(2003)131<0317:TROKWO>2.0.CO;2

Parkinson, C. L. (2019). A 40-y record reveals gradual Antarctic sea ice increases followed by decreases at rates far exceeding the rates seen in the Arctic. *Proceedings of the National Academy of Sciences*, *116*(29), 14414–14423.

Phillips, N. A. (1971). Numerical weather prediction. *Eos, Transactions American Geophysical Union*. https://doi.org/10.1029/EO052i006pIU420

Powers, J. G., Manning, K. W., Bromwich, D. H., Cassano, J. J., & Cayette, A. M. (2012). A decade of antarctic science support through AMPS. *Bulletin of the American Meteorological Society*. https://doi.org/10.1175/BAMS-D-11-00186.1

Puri, K., Dietachmayer, G., Steinle, P., Dix, M., Rikus, L., Logan, L., et al. (2013). Implementation of the initial ACCESS numerical weather prediction system. *Australian Meteorological and Oceanographic Journal*. https://doi.org/10.22499/2.6302.001

Qian, Y., Luo, Y., Si, F., Yang, T., & Yang, D. (2021). Three-year observations of ozone columns over polar vortex edge area above West Antarctica. *Advances in Atmospheric Sciences*, 1–12.

Rinke, A., Maslowski, W., Dethloff, K., & Clement, J. (2006). Influence of sea ice on the atmosphere: A study with an Arctic atmospheric regional climate model. *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1029/2005JD006957

Schroeter, B. J. E., Reid, P., Bindoff, N. L., & Michael, K. (2019). Antarctic verification of the Australian Numerical Weather Prediction model. *Weather and Forecasting*. https://doi.org/10.1175/waf-d-18-0171.1

Simpkins, G. R., Ciasto, L. M., Thompson, D. W. J., & England, M. H. (2012). Seasonal relationships between large-scale climate variability and antarctic sea ice concentration. *Journal of Climate*. https://doi.org/10.1175/JCLI-D-11-00367.1

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., et al. (2008). A Description of the Advanced Research WRF Version 3. *A Description of the Advanced Research WRF Version 3*, (June), 113. https://doi.org/10.5065/D6DZ069T

Smith, G. C., Liu, Y., Benkiran, M., Chikhar, K., Surcel Colan, D., Gauthier, A.-A., et al. (2021). The Regional Ice Ocean Prediction System v2: a pan-Canadian ocean analysis system using an online tidal harmonic analysis. *Geoscientific Model Development*, *14*(3), 1445–1467. https://doi.org/10.5194/gmd-14-1445-2021

Su, Z., Pilo, G. S., Corney, S., Holbrook, N. J., Mori, M., & Ziegler, P. (2021). Characterizing Marine Heatwaves in the Kerguelen Plateau Region. *Frontiers in Marine Science*, *7*(January), 1–13. https://doi.org/10.3389/fmars.2020.531297

Tastula, E.-M., Vihma, T., & Andreas, E. L. (2012). Evaluation of Polar WRF from

Modeling the Atmospheric Boundary Layer over Antarctic Sea Ice in Autumn and Winter. *Monthly Weather Review*, *140*(12), 3919–3935. https://doi.org/10.1175/mwr-d-12-00016.1
Thorndike, A. S., Rothrock, D. A., Maykut, G. A., & Colony, R. (1975). The thickness distribution of sea ice. *Journal of Geophysical Research*. https://doi.org/10.1029/JC080i033p04501
Valkonen, T., Vihma, T., Johansson, M. M., & Launiainen, J. (2014). Atmosphere-sea ice interaction in early summer in the Antarctic: Evaluation and challenges of a regional atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, *140*(682), 1536–1551. https://doi.org/10.1002/qj.2237
Waugh, D. W., & Randel, W. J. (1999). Climatology of Arctic and Antarctic polar vortices using elliptical diagnostics. *Journal of the Atmospheric Sciences*, *56*(11), 1594–1613.
Wilby, R. L., & Wigley, T. M. L. (1997). Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, *21*(4), 530–548.
Wilson, A. B., Bromwich, D. H., & Hines, K. M. (2011). Evaluation of Polar WRF forecasts on the Arctic System Reanalysis domain: Surface and upper air analysis. *Journal of Geophysical Research Atmospheres*. https://doi.org/10.1029/2010JD015013
Yao, Y., Huang, J., Luo, Y., & Zhao, Z. (2016). Improving the WRF model's (version 3.6.1) simulation over sea ice surface through coupling with a complex thermodynamic sea ice model (HIGHTSI). *Geoscientific Model Development*, *9*(6), 2239–2254. https://doi.org/10.5194/gmd-9-2239-2016