Cloud Influence on ERA5 and AMPS Surface Downwelling Longwave Radiation Biases in West Antarctica

Israel Silber¹, Johannes Verlinde¹
Sheng-Hung Wang², David H. Bromwich²,³
Ann M. Fridlind⁴
Maria Cadeddu⁵
Edwin W. Eloranta⁶
Connor J. Flynn⁷

[1] Department of Meteorology and Atmospheric Science, Pennsylvania State University, University Park, Pennsylvania, USA
[2] Polar Meteorology Group, Byrd Polar & Climate Research Center, The Ohio State University, Columbus, Ohio, USA
[3] Atmospheric Sciences Program, Department of Geography, The Ohio State University, Columbus, Ohio, USA
[5] Argonne National Laboratory, Argonne, Illinois, USA
[6] Space Science and Engineering Center, University of Wisconsin, Madison, Wisconsin, USA
[7] Pacific Northwest National Laboratory, Richland, Washington, USA

Correspondence to: Israel Silber (israel0silber@gmail.com)

August 2019
Abstract

The surface downwelling longwave radiation component ($LW_\downarrow$) is crucial for the determination of the surface energy budget and has significant implications for the resilience of ice surfaces in the polar regions. Accurate model evaluation of this radiation component requires knowledge about the phase, vertical distribution, and associated temperature of water in the atmosphere, all of which control the $LW_\downarrow$ signal measured at the surface. In this study, we examine the $LW_\downarrow$ model errors found in the Antarctic Mesoscale Prediction System (AMPS) operational forecast model and the ERA5 reanalysis model relative to observations from the AWARE campaign at McMurdo Station and the West Antarctic Ice Sheet (WAIS) Divide. The errors are calculated separately for observed clear-sky conditions, ice-cloud occurrences, and liquid-bearing cloud layer (LBCL) occurrences. The analysis results show a tendency in both models at each site to underestimate the $LW_\downarrow$ during clear sky conditions, high error variability (standard deviations > 20 W/m$^2$) during any type of cloud occurrence, and negative $LW_\downarrow$ biases when LBCLs are observed (bias magnitudes > 15 W/m$^2$ in tenuous LBCL cases; > 43 W/m$^2$ in optically thick/opaque LBCLs instances). We suggest that a generally dry and liquid-deficient atmosphere responsible for the identified $LW_\downarrow$ biases in both models is the result of excessive ice formation and growth, which could stem from model initial and lateral boundary conditions, microphysics scheme, aerosol representation, and/or limited vertical resolution.
1. Introduction

The surface energy budget in the polar regions is controlled by the amount, phase, and distribution of water in the atmosphere. Water is often advected from open-water sources, located mainly at lower latitudes, into the polar regions by meso and synoptic-scale systems (Carrasco et al., 2003; Nigro & Cassano, 2014; Pithan et al., 2018). Some of these moist and warm air intrusions accompanied by cloud formation have been shown to result in extensive ice sheet melting events (Bennartz et al., 2013; Nicolas et al., 2017), enhancement of meltwater runoff (Van Tricht et al., 2016), and sea-ice decline (Francis & Hunter, 2006).

The synoptic-scale patterns of moisture advection into the polar regions are generally well-captured by models (e.g., Coggins & McDonald, 2015; Nicolas & Bromwich, 2011; Scott et al., 2019; Scott & Lubin, 2014; Sedlar & Tjernström, 2017; Woods et al., 2013). However, discrepancies between observations and large-scale model representation of polar-cloud properties and occurrence are often reported (e.g., Cesana et al., 2012; Klein et al., 2009). These modeled cloud deviations from the measurements likely result from incomplete understanding and difficulties in the model representation of mixed-phase cloud microphysical processes, such as ice-formation mechanisms and ice-crystal growth (e.g., van Diedenhoven et al., 2009; Fridlind & Ackerman, 2018; Harrington et al., 2013). The outcome of such discrepancies is poorly modeled atmospheric profiles of water-phase partitioning that, together with an inaccurate prediction of hydrometeor properties, can generate biases in the model estimations of the surface energy budget. These surface radiation biases are largely dominated in the polar regions by the surface downwelling longwave (LW↓) component, which is highly sensitive to the water distribution in the atmosphere and its properties (e.g., phase, temperature; see Bennartz et al., 2013; Ohmura, 2001; Persson et al., 2017; Shupe & Intrieri, 2004).

Large-scale model intercomparisons of the LW↓ in Antarctica, performed at various temporal resolutions, from 3 h to full seasons, were reported by several studies. These intercomparisons were
made with climate models (e.g., Lenaerts et al., 2017), regional climate and forecast models (e.g., Bromwich et al., 2013; Hines et al., 2019; King et al., 2015), and reanalysis models (e.g., Välisuo et al., 2014; Wang & Dickinson, 2013), which were evaluated using satellite measurement data products (e.g., Lenaerts et al., 2017; van Wessem et al., 2018) and ground-based measurements (e.g., Välisuo et al., 2014; Valkonen et al., 2014; Wang & Dickinson, 2013). These studies showed predominantly large mean model relative bias and standard deviation (SD) that occasionally exceeded 10 W/m² and 30 W/m², respectively, and moderate Pearson’s correlation coefficient (r) values within the 0.5-0.8 range (see Bromwich et al., 2013; King et al., 2015; Listowski & Lachlan-Cope, 2017; Välisuo et al., 2014; van Wessem et al., 2018). The sources for these high model biases and moderate r values were consistently attributed to the poor model prediction of cloud physical properties (e.g., Bromwich et al., 2013; King et al., 2015; Lenaerts et al., 2017), although these conclusions were made without separately quantifying the model biases for cloudy periods. For instance, Listowski and Lachlan-Cope (2017) supported their conclusion by pointing to the reduction in the model bias and SD when more advanced cloud microphysical schemes were implemented in a regional forecast model.

In this study, we examine the $LW_{\downarrow}$ bias of the Antarctic Mesoscale Prediction System (AMPS) regional forecast model (Powers et al., 2012) and the latest, fifth version of the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis model, ERA5 (Hersbach & Dee, 2016). The biases are calculated relative to comprehensive surface observations gathered at McMurdo Station (77.85ºS, 166.72ºE; 70 m above mean sea level) and the West Antarctic Ice Sheet (WAIS) Divide (79.47ºS, 112.08ºW; 1803 m above mean sea level) during the U.S. Department of Energy (DOE) Atmospheric Radiation Measurement (ARM) West Antarctic Radiation Experiment (AWARE) field campaign (Lubin et al., 2015; Witze, 2016). We isolate and compare the biases generated during the occurrence of tenuous and opaque liquid-bearing clouds, ice clouds, and clear-sky periods. Finally, we discuss the results and potential bias sources.
2. Methodology

2.1. Observational Data

We follow the methodology of Silber et al. (2018a) and Silber et al. (2019) for the detection of clouds and phase classification at McMurdo Station and the WAIS Divide, respectively. These methods are based mainly on the analysis of the Ka-Band ARM Zenith Radar (KAZR; Widener et al., 2012) and the High Spectral Resolution Lidar (HSRL; Eloranta, 2005) measurements at McMurdo, and the polarization-sensitive micropulse lidar (MPL; Flynn et al., 2007) at the WAIS Divide, the raw measurements of which were processed by the authors (Silber et al., 2018b). The Vaisala ceilometer (Morris, 2016) and microwave radiometer (MWR; Morris, 2006) measurements are also used in the processing of the data from both sites.

Cloud occurrence is determined on an hourly basis to match the temporal resolution of the models (1 h, except for AMPS at the WAIS Divide, which has a temporal resolution of 3 h). An hour is considered cloudy if the hourly cloud occurrence fraction is above 25% (15 min), and clear otherwise. Similarly, a cloudy hour is considered as containing liquid-bearing cloud layers (LBCLs; likely mixed-phase clouds) if the detected liquid occurrence fraction is above 25%, and containing only ice clouds otherwise. The LBCLs are classified as opaque when the hourly mean liquid water path (LWP) is higher than 25 g/m², and tenuous otherwise. Clouds with LWP exceeding this value act similar to blackbody emitters in the longwave (LW; Shupe & Intrieri, 2004, Figure 6; Turner et al., 2007b, Figure SB1). The LWP is retrieved using the MWR measurements (Cadeddu et al., 2007; Turner et al., 2007a), or with the combined data product of the MWR and the G-Band Vapor Radiometer Profiler (GVRP; Cadeddu, 2010; Cadeddu et al., 2009), when measurements from both instruments are available. Due to the dependence on the LWP retrievals in the determination of the atmospheric regime (clear sky, tenuous LBCLs, etc.), only periods with MWR data are considered for
the $LW_\downarrow$ radiation comparison; that is, February to December 2016 (11 months) at McMurdo Station, and December 19, 2015, to January 17, 2016, at the WAIS Divide.

The $LW_\downarrow$ data are based on the sky radiation radiometers (SKYRAD; Andreas et al., 2018) measurements, the instruments of which were deployed at both sites throughout the campaign. Finally, we also utilize the twice-daily and four-times daily radiosonde measurements at McMurdo (from the entire year of 2016) and the WAIS Divide (from December 6, 2015, to January 17, 2016), respectively, to compare various atmospheric parameter profiles with the output from the models. Both the sounding and model data are linearly interpolated up to 6 km in 100 m intervals for all statistics presented in this study.

2.2. AMPS

AMPS is an operational Antarctic weather prediction model that is based on the Polar Weather Research and Forecasting (WRF) model (Hines & Bromwich, 2008), and is the result of a joint collaboration between the National Center for Atmospheric Research (NCAR) and the Polar Meteorology Group at The Ohio State University’s Byrd Polar and Climate Research Center. The Polar-WRF version on which AMPS is based on was upgraded from version 3.3.1 to version 3.7.1 on January 19, 2016; hence, the entire WAIS deployment period is covered by the earlier version, while most of the McMurdo deployment (including the full $LW_\downarrow$ dataset used here) is supported by the newer version of Polar WRF. AMPS currently utilizes WRF’s computationally-efficient single-moment 5-class microphysics scheme (WSM5), which considers cloud water that is allowed to be supercooled, cloud ice, rain, and snow, and operates on the resolved wind fields (see Hong et al., 2004).

AMPS forecasts are initiated twice-daily at 00 UTC and 12 UTC, using initial and lateral boundary conditions from the National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS). We use the AMPS data after a model spin-up of 12 h. The data for McMurdo Station and the WAIS Divide are bi-linearly interpolated from AMPS domains 5 and 2 (see Powers et al.,
which have a spatial resolution of ~1.1 km and ~10 km, and are provided in intervals of 1 h and 3 h, respectively.

2.3. ERA5

ERA5 is the fifth generation ECMWF reanalysis model and is the successor of the ERA-Interim reanalysis model (Dee et al., 2011). The cloud and large-scale precipitation processes are described in ERA5 by prognostic equations for cloud water and ice, rain, snow, and cloud fraction. The model considers various sources and sinks of all cloud variables, and provides better physical representation of supercooled liquid water and mixed-phase clouds relative to ERA-Interim (see ECMWF, 2016, Chapter 7; Tiedtke, 1993). The full list of improvements in ERA5 with respect to ERA-Interim can be found in Hersbach and Dee (2016). While ERA5 integrated atmospheric water amount and $LW_\downarrow$ prediction performance over Antarctica have not been reported to our knowledge, previous studies did find a better agreement of ERA-Interim with Antarctic measurements relative to other reanalysis models (e.g., Lenaerts et al., 2017; Wang & Dickinson, 2013).

ERA5 provides hourly output data at a spatial resolution of up to 0.25° (used here), which corresponds to ~23.4 km in the north-south direction and ~5.8 km and ~5.1 km in the east-west direction at the latitudes of McMurdo Station and the WAIS Divide, respectively. The ERA5 $LW_\downarrow$ data are taken from the model forecast fields (initialized at 06 UTC and 18 UTC) while allowing 3 hours of model spin-up. All the ERA5 data in this study are bi-linearly interpolated to the coordinates of McMurdo Station and the WAIS Divide. While the adjacent ERA5 grid cells used for the bi-linear interpolation over McMurdo are clear of the two volcanoes to the north-northeast (no surface-altitude biases in the data), the model performance is strongly influenced by the nearby topography. This is evident by the better $LW_\downarrow$ agreement with the observations at McMurdo Station (but not at the WAIS Divide) when the bi-linear interpolation is used (not shown). This improvement results from the
inclusion of grid cells located further away from the topographic features of Ross Island into the estimate.

3. Results

3.1. Model LW↓ Bias

Figure 1 shows the model error (model minus observation) frequency of occurrence in both AMPS and ERA5 at McMurdo Station and the WAIS Divide during 4 different observed regimes, namely, the occurrence of cloud-free periods (clear sky), ice clouds (no liquid), tenuous LBCLs, and opaque LBCLs. This figure is complemented by Table 1, which provides error statistics for each of the models, sites, and regimes. These statistics include the error percentiles, mean, SD, and \( r \) between the observations and the models. \( r \) is also calculated for the deseasonalized data at McMurdo, that is, the residual data after the removal of the annual cycle of the \( LW↓ \), which is estimated using a non-linear least square fit of sine with a periodicity of 365.25 days. This calculation is not performed over the WAIS Divide due to the short dataset (30 days).

Both models show a negative bias (\( LW↓ \) less than that observed) at both sites when the full datasets are considered, in agreement with several previous reports (cf., Bromwich et al., 2013; Fréville et al., 2014; Lenaerts et al., 2017). The relative occurrence of each atmospheric regime during the two deployment periods controls the magnitude of these total biases. The occurrence of these atmospheric regimes is influenced by the regional synoptic conditions, which were fairly representative of their long-term patterns for the periods corresponding with the AWARE deployment dates at each site (Silber et al., 2019). However, other factors also influence the locally-observed atmospheric regimes, for example, mesoscale forcing, sea-ice extent, and atmospheric rivers (e.g., Carrasco & Bromwich, 1993; Scott & Lubin, 2014, Figure 17; Silber et al., 2019). Therefore, the magnitude of the total-mean biases for the AWARE-deployment periods may vary from the values reported in Table 1.
The full dataset SD > 26 W/m² (in both models at both sites) and the r values for the full McMurdo datasets are comparable with previous reports (see Section 1). However, the r values of the model output for the WAIS Divide are much lower, partially as a result of the smaller number of samples. The smaller differences in r during cloudy periods between the raw and the deseasonalized data in both models relative to the clear-sky instances suggest that the observed biases are largely due to cloud properties, and do not show a significant dependence on the annual temperature cycle. This conclusion is supported by the fact that results from a seasonal analysis (based on months; not shown) both correspond with the annual statistics discussed below and are similar for different seasons. The observed seasonal variabilities are related to the observed cloud properties. For example, both models exhibit exceptionally high SD values in spring season (September-October-November) during observed clear-sky and LBCL periods (higher than 24 W/m² and 36 W/m², respectively). These higher values likely ensue from the combination of the low occurrence and predominantly shallower clouds relative to the other seasons (see Figures 4, 6, and 9 in Silber et al., 2018a), which could impact the model performance (see Section 4).

AMPS predominantly underestimates the LW↓ at both sites and in all regimes. The AMPS histograms in Figure 1 (all a and c panels) display a strong tendency for negative values. Some LW↓ overestimation instances do occur, mainly in clear-sky and ice-cloud periods, but these are hardly noticeable when the full datasets are considered. For example, the LW↓ is underestimated at least up to the 75th percentile in all the regimes (Table 1).

AMPS LW↓ in clear-sky periods is underestimated half of the time by more than 5 W/m² and has a mean error of -3.6 W/m² and -6.4 W/m² at McMurdo and the WAIS, respectively. However, r > 0.86 at both sites, and has a reasonable value even with the annual cycle removed, which suggests that moisture advection patterns to both sites are well captured but underestimated in amplitude (supported by precipitable water vapor comparison; not shown).
During cloudy periods, AMPS shows poorer performance in $LW_\downarrow$ prediction (SD > 20 W/m$^2$), with greater underestimation values. The tendency for $LW_\downarrow$ underestimation is emphasized during LBCL occurrences (Figure 1a4,a5,c4,c5), especially in the opaque cases, with mean biases below -51 W/m$^2$ at both sites.

ERA5 shows in general a tendency for model $LW_\downarrow$ underestimation as well, but with more overestimation occurrences than AMPS (all b and d panels in Figure 1). Clear-sky periods are rather variable, with SD values above 14 W/m$^2$ at both sites. The $LW_\downarrow$ during these periods is mostly underestimated (negative medians in Table 1), while the mean $LW_\downarrow$ is negative at McMurdo and positive at the WAIS. The different mean and median signs suggest that the overestimation error magnitudes at the WAIS (on the order of a few tens of W/m$^2$) are larger than the underestimation error magnitudes (see Figure 1d2 and the matching percentiles in Table 1).

ERA5 also tends to underestimate the $LW_\downarrow$ when ice clouds are observed at McMurdo (median $LW_\downarrow$ of -6.9 W/m$^2$; Figure 1b3), and overestimate at the WAIS (median $LW_\downarrow$ of 4.8 W/m$^2$; Figure 1d3). Similar to the AMPS model comparison, the ERA5 $LW_\downarrow$ biases are largest (at both sites) during the occurrence of LBCLs, with mean and median biases below -43 W/m$^2$ in opaque LBCL cases.

In multiple robustness tests using the McMurdo dataset, in which the hourly thresholds for the determination of cloud occurrence and type are modified, there are merely small changes in the clear-sky and cloudy period statistics presented in Table 1, except for the ice-cloud and tenuous-LBCL percentiles and mean biases (see Table S1 in the Supporting Material). These ice-cloud and tenuous-LBCL mean biases may increase by up to ~20% and 60%, respectively, when higher hourly thresholds are used, an increase that is attributed to the omission of short-lived, less “healthy”, and optically-thinner clouds from the analysis.

The notable magnitudes of the $LW_\downarrow$ biases during LBCL occurrences (in both models), as well as the scale of the lower percentile errors in ice cloud instances (Table 1), are comparable to the
instantaneous cloud radiative forcing of polar LBCL and ice clouds, respectively (e.g., Miller et al., 2015; Silber et al., 2019). Disparities of the modeled water quantities (and atmospheric temperature profiles) and lack of correspondence between the modeled and observed atmospheric regime could bring about these cloudy period biases and the large clear-sky variability discussed above, via changes in the atmospheric optical-depth profile and the effective LW brightness temperature of the sky. These disparities are examined below.

3.2. Model Atmospheric Regime and Hydrometeor Amount Errors

Figure 2 presents scatter plots of the modeled versus observed LWP and ice water path (IWP) during modeled and observed cloud occurrence, together with $r$ and linear fits between the datasets. The thresholds used to determine model LBCL and ice-cloud occurrences are based on annual (2016) statistics from the AWARE campaign (not shown). The threshold values are set to the 1st percentile of the observed LWP (1 g/m$^2$) and IWP (10$^{-3}$ g/m$^2$) values. The IWP comparison is only made for McMurdo Station because the IWP retrievals (following Hogan et al., 2006) are based on radar measurements, which were not performed at the WAIS deployment.

Table 2 shows the atmospheric regime correspondence between the model and the observations at both sites. The modeled atmospheric regimes are calculated with the same thresholds that are mentioned above. This table is divided into two sections; the left section gives the correspondence percentages relative to the number of observed cases per atmospheric regime. For example, the observed clear-sky periods at McMurdo were determined by AMPS as clear sky, ice cloud, and LBCL occurrence in 45.72%, 54.15%, and 0.13% of the time, respectively (together, 100% of observed clear-sky occurrences). The percentages in the right section are calculated relative to the number of modeled cases per atmospheric regime. For example, the AMPS modeled clear-sky periods at McMurdo were actually clear sky, ice cloud, and LBCL occurrence periods 55.63%, 23.18%, and 21.20% of the time, respectively (together, 100% of the clear-sky periods based on AMPS).
The small number of samples in the AMPS LWP plots (Figure 2a,e) are the result of the model producing LBCLs for less than 2.5% of the cases LBCLs are observed (Table 2, left section). Instead, the model tends to generate only ice clouds on at least 80% of the LBCL occurrences, which may occasionally compensate for the lack of liquid in the resolved \( L W'_↓ \) (not shown). On the other hand, the model does not tend to produce LBCLs when they are not observed. The AMPS tendency to over-produce ice clouds is also seen during clear-sky periods, with the model producing ice clouds during \(~54\%\) and \(40\%\) of the observed clear-sky hours at McMurdo and the WAIS, respectively. The lower percentage at the WAIS may explain the higher \( r \) during clear-sky periods (see Table 1). The periods when the model generates only ice clouds without being observed account for more than \(57\%\) and \(87\%\) of the total model-produced “pure” ice cloud cases at McMurdo and the WAIS, respectively (Table 2, right section).

Given AMPS tendency to over-produce clouds, it comes as no surprise that “pure” ice cloud occurrence is well captured by AMPS. The modeled ice clouds are present in more than \(84\%\) of the time these clouds are observed and have some IWP correspondence with the observations, especially at higher values (Figure 2c). Some tendency for \( L W'_↓ \) underestimation is found when the model predicts IWPs that are too small relative to the observations. Conversely, the \( L W'_↓ \) tends to be overestimated when the model exaggerates the IWP (with values predominantly higher than \(10 \text{ g/m}^2\); see Figure 2c).

ERA5 lacks a LWP correlation with the observations at both sites (Figure 2b,f), and has a tendency to underestimate the observed LWP by up to 2 orders of magnitude (most markers are below the 1:1 correspondence line). The model LWP values rarely exceed \(~100 \text{ g/m}^2\), despite numerous observations of higher values (at McMurdo).

Similar to AMPS, ERA5 tends to generate ice clouds instead of LBCLs more than \(58\%\) of the time (Table 2, left section), which can reduce the \( L W'_↓ \) errors (not shown). Relative to the total model
production of liquid, liquid is incorrectly produced (during ice cloud or clear sky conditions) in more than 56% of the time at McMurdo, but only 20% of the time at the WAIS (Table 2, right section).

Ice clouds are frequently produced by the ERA5 model and have a reasonable IWP agreement with the observations (Figure 2d). Similar to AMPS, ERA5 tends to underestimate (overestimate) the $LW_{\downarrow}$ when its IWP values are too low (high) relative to the observations. While ERA5 captures ~76% and ~94% of the observed ice cloud occurrences at McMurdo and the WAIS, these cases only account for ~45% and ~12% of the model “pure” ice-clouds periods, respectively.

ERA5 also tends to generate hydrometeors (mainly ice) when clear-sky conditions are observed (more than 80% of the time, at both sites). These instances govern the larger magnitude of the positive (relative to the negative) $LW_{\downarrow}$ errors during observed clear-sky conditions (see Section 3.1). We see the same result, although to a much smaller extent (Figure 1a2), with AMPS at McMurdo.

The $LW_{\downarrow}$ when both the models and the observations describe clear-sky conditions is consistently underestimated (at least up to the 88th percentile; mean magnitudes > 4.8 W/m² and SDs smaller than the mean magnitudes; see Table S2 in the Supporting Material). This conspicuous clear-sky bias, the rarely-produced LBCLs (with deficient LWP; see Figure 2a,b,e,f and Table 2, left section), and the excess generation of ice clouds relative to the observations (Table 2, right section) suggest that the models inaccurately capture the atmospheric temperature and water vapor profiles.

3.3. Atmospheric Temperature and Water Vapor Errors

Figure 3 depicts the profiles of the mean (solid blue) and median (solid black) model errors of the temperature ($T$), specific humidity ($q$) and relative humidity (RH). The model RH and the observed $q$ are calculated using the saturation vapor pressure equation provided by Murphy and Koop (2005). The markers represent the root-mean-square errors (RMSEs) of the under and overestimated cases, where the model underestimation RMSEs are multiplied by -1 to distinguish them from model
overestimation cases. The marker colors designate the percentage of the under and overestimated cases at each altitude. AMPS tends to form an atmosphere which is too cold and dry relative to the observations (Figure 3a1,a2,c1,c2; see also Hines et al., 2019), which may also explain the tendency to produce smaller amounts of ice (Figure 2c). The largest $T$ bias at McMurdo (Figure 3a1) is observed near the surface (mean and median values of -1.4 K), in agreement with Bromwich et al. (2013, their Figure 9). This near-surface bias suggests that the model produces too frequent and/or too strong surface $T$ inversions, which correspond with potentially more and/or stronger surface $q$ inversions (Figure 3a2). These low-level $T$ and $q$ biases in AMPS (seen both in clear sky and cloudy conditions; see Figure S1 in the Supporting Material) are likely related to high near-surface wind speed intensities originated from katabatic flow along the Ross Ice Shelf (see Wille et al., 2016, 2017).

The $q$ profiles in the model show a negative bias that is most pronounced below 2.5 km (Figure 3a2). This bias is reflected in underestimation of RH by averages of at least 4% (with a maximum negative RMSE of ~22%) up to ~2 km (Figure 3a3). The $q$ impact on the RH is moderated above 2 km by the large magnitude of the negative $T$ bias (~0.7-0.9 K on average, up to ~4.5 km). However, RH continues to be underestimated up to ~8 km (not shown). It should be noted that the generally decreasing (with altitude) $q$ biases and RMSEs (as seen in the other $q$ panels as well; Figure 3b2,c2,d2) are the result of the decreasing $T$ with altitude (e.g., Figure 5 in Silber et al., 2018a), following the Clausius-Clapeyron Equation.

At the WAIS Divide, the $T$ bias (> 1 K) at the altitude range of ~3.7-5.3 km is the result of consistent (> 75% of the time) $T$ underestimation with negative RMSEs of more than 1.4 K (Figure 3c1). Higher negative RMSEs are observed near the surface, but relatively higher $T$ overestimation occurrences with high RMSEs (> 1.5 K) mitigate the mean $T$ bias. Similar to McMurdo, the $q$ bias peaks at lower altitudes than the $T$ bias (maximum mean and median magnitudes of 0.31 g/kg and 0.21 g/kg at ~2.5 km, respectively; Figure 3c2). The combined negative $q$ and $T$ biases are projected in the
mean RH bias, which exceeds 10% from ~2.5 km to ~4 km, and stays higher than 5% up to an altitude of ~6 km (Figure 3c3).

The lower resolution of ERA5 raised some initial concerns regarding the model’s ability to capture the surface and atmospheric variability at McMurdo Station. That is, because of the complex local topography that results in a strong influence of mesoscale flow patterns in that region (Carrasco et al., 2003; Monaghan et al., 2005; Silber et al., 2019). However, examination of the ERA5 results in Figure 3 (as well as in Figure 1 and 2 and Table 1 and 2) reveals no clear signs of inferior model performance at McMurdo Station relative to the WAIS Divide.

ERA5 exhibits a negative T bias from ~1.1 km up to ~5 km at McMurdo. This bias is emphasized from 2 km up to nearly 4 km, with mean values > 0.4 K (Figure 3b1). This altitude range (where the T bias peaks) has a large overlap with that of AMPS (Figure 3a1). An overlap with AMPS is also observed in the altitude range where the magnitude of the q bias is the highest (Figure 3a2,b2). The q bias in ERA5 reaches a maximum mean magnitude of ~0.13 g/kg close to the surface, at 0.5 km (Figure 3b2). At these near-surface altitudes (below 1 km) T is overestimated by the model and maximizes near the surface with a mean bias of ~1.1 K (Figure 3b1). The combination of the negative q bias and positive T bias yields negative RH bias magnitudes > 7% between the surface and ~1 km (Figure 3b3). The RH continues to be underestimated up to ~3 km and is slightly overestimated (by ~2-3%) above 3 km.

The ERA5 T profiles at the WAIS Divide display a negative bias that maximizes near the surface (mean of -0.7 K), and overall decreases in magnitude with height (Figure 3d1). A similar bias pattern of a near-surface maximum (-0.15 g/kg) and a decrease with height is observed in q (Figure 3d2). The RH values produced by the model are predominantly underestimated up to 4.5 km (magnitudes > 4% below 4 km), and slightly overestimated above that altitude (Figure 3d3).
Generally speaking, both models show the largest $T$, $q$, and RH RMSEs closer to the surface. Separation of the error calculation based on the atmospheric regime results in very similar error profiles (see Figures S1 and S2 in the Supporting Material), which emphasizes the consistency of these model errors. The negative biases stem both from the larger (in magnitude) and more frequent underestimation errors. The consistent negative RH errors below 3 km at McMurdo Station and 4 km at the WAIS Divide (Figure 3, right panels), correspond with the altitudes where LBCLs are most frequently observed at these sites (Hines et al., 2019; Scott & Lubin, 2016; Silber et al., 2018a). These RH errors, which are accentuated by their consistency and larger amplitudes in observed liquid-bearing vertical grid cells (see Figures S1 and S2 in the Supporting Material), are able to tip the scales between LBCL and clear-sky occurrence conditions, and can ultimately result in the large $LW_{\downarrow}$ biases (Section 3.1).

To qualitatively examine the likelihood of liquid condensation in the models relative to the observations, we inspect the RH with respect to ice ($RH_{ice}$), calculated using the formulation in Murphy and Koop (2005). Figure 4a-f illustrates the cumulative distribution function (CDF) of the observed and modeled $RH_{ice}$ up to 6 km (bin width of 2.5%), together with the corresponding highest $T$ that is observed/modelled in every bin, which is designated by the bin color. The bins are colored in white when the total bin counts are equal to 0. Figure 4g shows $RH_{ice}$ at water saturation.

The observations at both sites (Figure 4a,d) exceed ice saturation in more than 14% and 23% of the data grid cells at McMurdo Station and the WAIS Divide, respectively, and exhibit high ice supersaturation values that extend slightly beyond 40%, generally consistent with the observed occurrence of supercooled liquid. The highest temperatures at which these supersaturations are observed consistently correspond with the $RH_{ice}$ at water saturation (Figure 4g), and the cases where $RH_{ice}$ is higher than $RH_{ice}$ at water saturation fall within the 5% uncertainty range of the radiosonde RH measurement (Holdridge et al., 2011).
The RH\textsubscript{ice} observations connote the limited ice nucleating particle (INP) concentrations in the West Antarctic atmosphere (e.g., Belosi et al., 2014), especially deposition-mode INP concentrations capable to nucleate ice and deplete atmospheric water vapor prior to liquid water condensation (e.g., de Boer et al., 2011). The models clearly do not exhibit the “ice-starved atmosphere” behavior apparent in observations. AMPS does show RH\textsubscript{ice} values close to water saturation at supercooled temperatures, but only where RH\textsubscript{ice} is < 110% (weakly supercooled) and these cases account for less than 1% of the data grid cells (Figure 4b,e). ERA5 displays RH\textsubscript{ice} that are significantly below RH\textsubscript{ice} at water saturation (Figure 4c,f). The presence of some liquid in the model can be attributed to the model’s liquid generation schemes (ECMWF, 2016, Chapters 6 and 7.2.1).

Based on the results presented here, we suggest that an overly dry atmosphere (see Figure 3 and the higher CDF values at lower RH\textsubscript{ice} in Figure 4) leads to insufficient condensation of liquid, and evaporates existing liquid hydrometeors too quickly. The near-absence of ice supersaturation in the simulations provides further evidence that overproduction of ice in the models is driving water vapor concentrations to near-ice-saturated values, thereby desiccating the atmosphere via sedimentation. All of these processes may result in the large negative \( LW\downarrow \) biases found in the models (which are emphasized during LBCL occurrence).

4. Conclusions

In this study, we examined the surface downwelling longwave radiation (\( LW\downarrow \)) biases in the AMPS operational regional forecast model and the ERA5 reanalysis model at two West Antarctic sites (McMurdo Station and the WAIS Divide). The better performance of ERA5 relative to AMPS is probably associated with the different model types, and not necessarily a function of the model parameterizations. In a forecast model, the simulations are free running after initialization, whereas in a reanalysis model, the simulations are constrained by frequent observation assimilation. Therefore, one should expect a less skillful performance by AMPS relative to ERA5. An evaluation of a climate-
model’s performance over these two sites is also possible, given the spatial extent of high correlation of atmospheric parameters over McMurdo Station and the WAIS Divide with the regions surrounding the two sites (see Figures S3-S5 in the Supporting Material), and would be able to complement this study.

From the analysis presented above, we deduce the following:

- Both ERA5 and AMPS exhibit mean negative $LW_{\downarrow}$ biases of more than 14 $W/m^2$, with high SDs (> 26 $W/m^2$).

- The $LW_{\downarrow}$ during clear-sky periods are mostly underestimated in both models and at both sites by a few $W/m^2$ (negative medians). However, ERA5 slightly overestimates (on average) the $LW_{\downarrow}$ at the WAIS Divide. The low bias values near 0 $W/m^2$ in ERA5 (at both sites) and AMPS (at McMurdo) are the results of the clear-sky underestimation instances due to the dryer atmosphere, balanced by overestimation periods induced by excess production of ice clouds when they are not observed.

- The $LW_{\downarrow}$ during ice cloud occurrence has a large variability (SD ≥ 21 $W/m^2$), although the models manage to produce ice clouds on most of these times. During these periods, AMPS has a negative $LW_{\downarrow}$ bias with a mean magnitude larger than 21 $W/m^2$, while ERA5 shows a mixed behavior with $LW_{\downarrow}$ underestimation and overestimation (of several $W/m^2$) at McMurdo and the WAIS, respectively.

- Liquid-bearing cloud layer (LBCL) occurrences produce the largest (negative) $LW_{\downarrow}$ bias in both models. The biases are accentuated during the occurrence of “opaque” LBCLs (LWP ≥ 25 $g/m^2$), with mean magnitudes larger than 51 $W/m^2$ and 43 $W/m^2$ in AMPS and ERA5, respectively. These large biases are mainly the result of deficient/lack of condensate production in the models.
• Both models produce too much ice and too little liquid at the West Antarctic sites. We hypothesize that excessive ice formation and growth in the models may ultimately result in the dry and liquid-deficient atmosphere.

• The model $LW_\downarrow$ errors during clear-sky periods are strongly dependent on the annual cycle, while the errors during cloud occurrence are more dependent on the representation and prediction of cloud processes.

• AMPS potentially produces more and/or stronger surface temperature inversions, and both models tend to produce additional and/or more intense surface moisture inversions.

Diminution of the large model $LW_\downarrow$ biases by treatment of their sources could increase the accuracy of the surface energy budget estimation, and potentially, also improve the credibility of climate projections. Different choices of the initial and lateral boundary conditions applied in the models have been shown in previous studies to be capable of reducing these biases (e.g., Bromwich et al., 2011, 2013; Hines et al., 2019). Small changes in the AMPS $LW_\downarrow$ bias at McMurdo Station as a function of the model runtime (not shown) suggest that the initial conditions could play a significant role in the disagreement of the model with the observations. The initial boundary conditions might also have a significant influence on the ERA5 biases because the model diagnoses convection (ECMWF, 2016, Chapter 6), which (rarely) may take place in the model representation of the Antarctic atmosphere. The impact of the initial and boundary conditions on the $LW_\downarrow$ model output will be the topic of future studies.

The $LW_\downarrow$ biases could also be influenced by the model microphysical scheme. Previous reports have demonstrated better model performance in predicting the surface radiation when more advanced microphysical parametrization schemes are used (e.g., Hines et al., 2019; Listowski & Lachlan-Cope, 2017). Recently, Barrett et al. (2017a, 2017b) concluded, based on examination of large-scale models
and mid-latitude measurements, that the main cause for liquid deficiency in climate models is the results of the coarse resolution in the models combined with the parametrization of the ice-growth processes. Exploring this potential bias source (e.g., by implementing a higher vertical resolution or the parametrization introduced by Barrett et al., 2017b) might improve the large-scale model representation of Antarctic mixed-phase clouds, and consequently, the $\text{LW}_\downarrow$ biases as well.

Finally, cloud properties and processes are influenced by limited-aerosol regimes (e.g., Fridlind & Ackerman, 2018; Mauritsen et al., 2011), such as those found in Antarctica (e.g., Shaw, 1980, 1988). The impact of these Antarctic regimes and the sharp contrast in aerosol-particle concentrations relative to other regions (e.g., Holben et al., 2001; Spracklen et al., 2008) on the model representation of clouds warrant further investigation. Different methods for the determination of aerosol concentrations in large-scale models might reduce the $\text{LW}_\downarrow$ and other model biases that are related to the aerosol indirect effect in the high-latitude regions (e.g., cloud reflectivity; see McCoy et al., 2015; Vergara-Temprado et al., 2018). However, the exact impact of the aerosol indirect effect and its representation in models are still highly active fields of research (e.g., Christensen et al., 2016; Fridlind & Ackerman, 2018; Garrett & Zhao, 2006; Jackson et al., 2012; Lubin & Vogelmann, 2006, 2007; McFarquhar et al., 2011) and debate (in a global context; cf., Rosenfeld et al., 2019; Sato et al., 2018), which necessitates further study.

The quantification of the impact of each of these four plausible sources and the mitigation of the $\text{LW}_\downarrow$ biases via treatment of these sources remain two major model-specific problems. Because of the interplay between the sources, for example, the vertical resolution impact on the aerosol representation, which controls the cloud microphysical scheme, bulk analysis would likely overlook some of the feedbacks in the representation of clouds in different models. Therefore, we suggest that a case-study approach is more suitable for this type of research, as it allows for independent investigation of the different sources of error, for example, with the use of model sensitivity tests.
Additional comprehensive Antarctic field campaigns such as AWARE could benefit this type of studies as well, by providing valuable datasets for model intercomparisons. Ongoing and future studies will focus on this problem in forecast, reanalysis, and climate model perspectives, and hopefully, assist in addressing the representation of the LW surface energy budget in these models.

Acknowledgments

AWARE is supported by the DOE ARM Climate Research Facility and NSF Division of Polar Programs. The data used in this study are available in the ARM data archive (http://www.archive.arm.gov). The HSRL data can be obtained from the ARM archive or from the University of Wisconsin-Madison HSRL Lidar Group (http://lidar.ssec.wisc.edu). AMPS data can be requested from the Ohio State University Polar Meteorology Group (http://polarmet.osu.edu/AMPS/). ERA5 reanalysis data are accessible via the Copernicus Climate Change Service (C3S) Climate Data Store (CDS; https://cds.climate.copernicus.eu). The authors wish to thank Fuqing Zhang for the fruitful discussions. I.S. and J.V. are supported by the National Science Foundation grant PLR-1443495 and by the DOE grant DE-SC0017981. D.H.B. and S.H.W. are supported by National Science Foundation grant PLR-1443443 and by DOE grant DE-SC0017981. Contribution 1583 of Byrd Polar & Climate Research Center. A.F. is supported by the NASA Radiation Science and Modeling, Analysis and Prediction programs.

5. References


https://doi.org/10.1029/2006GL028750


https://doi.org/10.5194/acp-11-165-2011


List of Tables:

Table 1: AMPS and ERA5 surface downwelling longwave radiation ($LW_{\downarrow}$) model-error statistics during different atmospheric regimes at McMurdo Station and the WAIS Divide (LBCL stands for liquid-bearing cloud layer). The Pearson’s correlation coefficient ($r$) is calculated for the raw data at both sites, as well as for the residual data (at McMurdo), after the removal of the annual cycle using a non-linear least-square sine fit.

Table 2: AMPS and ERA5 atmospheric regime correspondence with the observations at McMurdo Station and the WAIS Divide. The percentages in the left section are given relative to the total number of observed cases per atmospheric regime (i.e., the sum of vertical cell triplets equals 100%), whereas in the right section, the percentages are relative to the total number of modeled cases per atmospheric regimes (i.e., the sum of horizontal cell triplets equals 100%). The model thresholds for ice-cloud and LBCL hourly occurrences are LWP $\geq 1$ g/m$^2$ and IWP $\geq 10^{-3}$ g/m$^2$, respectively (see text for details).
List of Figures:

Figure 1: AMPS and ERA5 surface downwelling longwave radiation ($LW_{↓}$) model error (model minus observation) at McMurdo Station (top two rows) and the WAIS Divide (bottom two rows). The left panels show the error occurrence frequency for the full datasets, partitioned into the occurrence of cloud-free periods (clear sky), ice clouds, tenuous liquid-bearing cloud layers (LBCLs), and opaque LBCLs (see legend). The other four panels in each row depict the occurrence frequency for clear sky and the three cloud types separately to make them easier to evaluate.

Figure 2: AMPS and ERA5 liquid water path (LWP) and ice water path (IWP) model output versus the observations at McMurdo Station and the WAIS Divide. The illustrated LWP and IWP data are based only on periods when both the observations and the model indicate the occurrence of LBCLs and ice clouds, respectively. The marker colors represent the $LW_{↓}$ model errors. The dotted-gray lines designate the 1:1 correspondence lines. The dashed black lines denote least-squares linear fits of the modeled as a function of the observed parameters. Linear fit coefficients and the Pearson’s correlation coefficient ($r$; statistically significant for all IWP panels, and the ERA5 LWP at McMurdo) are given at the top of each panel. The IWP is compared only at McMurdo Station because the retrieval of this parameter is based on radar measurements, which were not performed at the WAIS Divide.

Figure 3: Modeled profile biases of (from left to right panels) the temperature, water vapor mixing ratio ($q$), and relative humidity (RH). The black and blue curves designate the median and mean biases, respectively. The square and circle markers to the left and right of the mean curves denote the root-mean-square-error (RMSE) of the model under and overestimation cases, respectively, where the model underestimation RMSEs are multiplied by -1 to distinguish them from model overestimation cases. The colors represent the percentage of the data samples that are under or overestimated, such that the percentage sum of the model under and overestimated cases at a given altitude equals to 100%. The thick horizontal black line designates the surface altitude. This comparison is only based on hours where radiosonde measurements were made.

Figure 4: (a-f) Observed and modeled cumulative distribution function (CDF) of RH with respect to ice ($RH_{ice}$) between the surface and 6 km altitude (same altitude range that is shown in Figure 3; bin width of 2.5%). The bar colors designate the highest observed-modeled temperature in each $RH_{ice}$ bin, following the temperature color representation depicted in panel g. White bars indicate zero bin counts. The dashed-black lines denote $RH_{ice} = 100\%$. (g) $RH_{ice}$ at water saturation.
Table 1: AMPS and ERA5 surface downwelling longwave radiation ($LW_\downarrow$) model-error statistics during different atmospheric regimes at McMurdo Station and the WAIS Divide (LBCL stands for liquid-bearing cloud layer). The Pearson’s correlation coefficient ($r$) is calculated for the raw data at both sites, as well as for the residual data (at McMurdo), after the removal of the annual cycle using a non-linear least-square sine fit.

<table>
<thead>
<tr>
<th>Site</th>
<th>Model</th>
<th>Atmospheric Regime</th>
<th>Percentiles [W/m²]</th>
<th>Mean [W/m²]</th>
<th>SD [W/m²]</th>
<th>$r$ (no annual)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>25</td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td>McMurdo Station</td>
<td>AMPS</td>
<td>Clear sky</td>
<td>-14.5</td>
<td>-9.0</td>
<td>-5.0</td>
<td>-1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice</td>
<td>-65.5</td>
<td>-36.0</td>
<td>-16.2</td>
<td>-5.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenuous LBCL</td>
<td>-75.7</td>
<td>-56.8</td>
<td>-40.3</td>
<td>-21.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opaque LBCL</td>
<td>-86.0</td>
<td>-75.6</td>
<td>-61.1</td>
<td>-31.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>-74.8</td>
<td>-41.5</td>
<td>-13.5</td>
<td>-4.4</td>
</tr>
<tr>
<td></td>
<td>ERA5</td>
<td>Clear sky</td>
<td>-18.9</td>
<td>-11.0</td>
<td>-3.9</td>
<td>3.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice</td>
<td>-43.4</td>
<td>-18.7</td>
<td>-6.9</td>
<td>6.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenuous LBCL</td>
<td>-70.3</td>
<td>-49.3</td>
<td>-29.9</td>
<td>-12.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opaque LBCL</td>
<td>-85.5</td>
<td>-66.4</td>
<td>-43.6</td>
<td>-21.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>-65.9</td>
<td>-25.6</td>
<td>-10.5</td>
<td>0.8</td>
</tr>
<tr>
<td>WAIS Divide</td>
<td>AMPS</td>
<td>Clear sky</td>
<td>-16.4</td>
<td>-10.4</td>
<td>-5.8</td>
<td>-2.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice</td>
<td>-77.4</td>
<td>-38.1</td>
<td>-20.0</td>
<td>-7.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenuous LBCL</td>
<td>-90.4</td>
<td>-71.4</td>
<td>-45.9</td>
<td>-11.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opaque LBCL</td>
<td>-110.4</td>
<td>-104.0</td>
<td>-91.8</td>
<td>-77.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>-104.0</td>
<td>-75.9</td>
<td>-37.9</td>
<td>-9.0</td>
</tr>
<tr>
<td></td>
<td>ERA5</td>
<td>Clear sky</td>
<td>-10.8</td>
<td>-5.2</td>
<td>-1.2</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice</td>
<td>-22.4</td>
<td>-9.1</td>
<td>4.8</td>
<td>22.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tenuous LBCL</td>
<td>-64.2</td>
<td>-36.2</td>
<td>-13.1</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Opaque LBCL</td>
<td>-97.1</td>
<td>-81.5</td>
<td>-68.8</td>
<td>-54.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>-83.7</td>
<td>-43.0</td>
<td>-10.8</td>
<td>4.2</td>
</tr>
</tbody>
</table>
Table 2: AMPS and ERA5 atmospheric regime correspondence with the observations at McMurdo
Station and the WAIS Divide. The percentages in the left section are given relative to the total number
of observed cases per atmospheric regime (i.e., the sum of vertical cell triplets equals 100%), whereas
in the right section, the percentages are relative to the total number of modeled cases per atmospheric
regimes (i.e., the sum of horizontal cell triplets equals 100%). The model thresholds for ice-cloud and
LBCL hourly occurrences are LWP ≥ 1 g/m² and IWP ≥ 10⁻³ g/m², respectively (see text for details).

<table>
<thead>
<tr>
<th>Site</th>
<th>Model</th>
<th>Atmospheric Regime (rows – model, columns – observations)</th>
<th>% relative to the total number of observed cases per atmospheric regime</th>
<th>% relative to the total number of modeled cases per atmospheric regime</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Clear sky</td>
<td>Ice clouds</td>
<td>LBCL</td>
</tr>
<tr>
<td>McMurdo Station</td>
<td>AMPS</td>
<td>Clear sky</td>
<td>45.72%</td>
<td>15.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice cloud</td>
<td>54.15%</td>
<td>84.70%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LBCL</td>
<td>0.13%</td>
<td>0.14%</td>
</tr>
<tr>
<td></td>
<td>ERA5</td>
<td>Clear sky</td>
<td>18.19%</td>
<td>2.39%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice cloud</td>
<td>54.97%</td>
<td>76.01%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LBCL</td>
<td>26.84%</td>
<td>21.60%</td>
</tr>
<tr>
<td>WAIS Divide</td>
<td>AMPS</td>
<td>Clear sky</td>
<td>60.00%</td>
<td>15.79%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice cloud</td>
<td>40.00%</td>
<td>84.21%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LBCL</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td>ERA5</td>
<td>Clear sky</td>
<td>20.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ice cloud</td>
<td>65.56%</td>
<td>94.34%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>LBCL</td>
<td>14.44%</td>
<td>5.66%</td>
</tr>
</tbody>
</table>
Figure 1: AMPS and ERA5 surface downwelling longwave radiation ($LW_\downarrow$) model error (model minus observation) at McMurdo Station (top two rows) and the WAIS Divide (bottom two rows). The left panels show the error occurrence frequency for the full datasets, partitioned into the occurrence of cloud-free periods (clear sky), ice clouds, tenuous liquid-bearing cloud layers (LBCLs), and opaque LBCLs (see legend). The other four panels in each row depict the occurrence frequency for clear sky and the three cloud types separately to make them easier to evaluate.
AMPS

(a) $y = 0.30x + 0.60, r = 0.12$

(b) $y = 0.30x + 0.31, r = 0.27$

ERA5

(c) $y = 0.52x + 0.09, r = 0.47$

(d) $y = 0.56x + 0.53, r = 0.65$

WAIS Divide

(e) $y = -0.19x + 0.24, r = -0.98$

(f) $y = 0.23x + 0.19, r = 0.28$

LW↓ model error [W/m²]
Figure 2: AMPS and ERA5 liquid water path (LWP) and ice water path (IWP) model output versus the observations at McMurdo Station and the WAIS Divide. The illustrated LWP and IWP data are based only on periods when both the observations and the model indicate the occurrence of LBCLs and ice clouds, respectively. The marker colors represent the $LW_{\downarrow}$ model errors. The dotted-gray lines designate the 1:1 correspondence lines. The dashed black lines denote least-squares linear fits of the modeled as a function of the observed parameters. Linear fit coefficients and the Pearson’s correlation coefficient ($r$; statistically significant for all IWP panels, and the ERA5 LWP at McMurdo) are given at the top of each panel. The IWP is compared only at McMurdo Station because the retrieval of this parameter is based on radar measurements, which were not performed at the WAIS Divide.
Figure 3: Modeled profile biases of (from left to right panels) the temperature, water vapor mixing ratio ($q$), and relative humidity (RH). The black and blue curves designate the median and mean biases, respectively. The square and circle markers to the left and right of the mean curves denote the root-
mean-square-error (RMSE) of the model under and overestimation cases, respectively, where the model underestimation RMSEs are multiplied by -1 to distinguish them from model overestimation cases. The colors represent the percentage of the data samples that are under or overestimated, such that the percentage sum of the model under and overestimated cases at a given altitude equals to 100%. The thick horizontal black line designates the surface altitude. This comparison is only based on hours where radiosonde measurements were made.
Figure 4: (a-f) Observed and modeled cumulative distribution function (CDF) of RH with respect to ice (RH$_{\text{ice}}$) between the surface and 6 km altitude (same altitude range that is shown in Figure 3; bin width of 2.5%). The bar colors designate the highest observed-modeled temperature in each RH$_{\text{ice}}$ bin, following the temperature color representation depicted in panel g. White bars indicate zero bin counts.

The dashed-black lines denote RH$_{\text{ice}}$ = 100%. (g) RH$_{\text{ice}}$ at water saturation.