Evaluating Cloud Properties at Scott Base: Comparing Ceilometer Observations with ERA5, JRA55, and MERRA2 Reanalyses Using an Instrument Simulator

A. J. McDonald[1], P. Kuma[2], M. Panell[1], O. K. L. Petterson[1], G. E. Plank[1], M. A. H. Rozliaiani[1] and L. E. Whitehead[3]

[1]School of Physical and Chemical Sciences, University of Canterbury, Christchurch, New Zealand
[2]Department of Meteorology (MISU) and Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[3]Department of Geosciences, University of Oslo, Oslo, Norway

Key Points:

- Cloud occurrence is underestimated below 3km in ERA5, JRA55 and MERRA2 reanalyses relative to observations, leading to cloud fraction biases
- Observed cloud occurrence is more strongly impacted by synoptic state than season, ERA5 simulates this pattern better than JRA55 and MERRA2
- Super-cooled liquid cloud derived from ceilometer data have higher occurrences than the three reanalyses, with MERRA2 having the least bias

Corresponding author: Adrian McDonald, `adrian.mcdonald@canterbury.ac.nz`

Abstract

This study compares CL51 ceilometer observations made at Scott Base, Antarctica, with statistics from the ERA5, JRA55, and MERRA2 reanalyses. To enhance the comparison we use a lidar instrument simulator to derive cloud statistics from the reanalyses which account for instrumental factors. The cloud occurrence in the three reanalyses is slightly overestimated above 3km, but displays a larger underestimation below 3 km relative to observations. Unlike previous studies, we see no relationship between relative humidity and cloud occurrence biases, suggesting that the cloud biases do not result from the representation of moisture. We also show that the seasonal variation of cloud occurrence and cloud fraction, defined as the vertically integrated cloud occurrence, are small in both the observations and the reanalyses. We also examine the quality of the cloud representation for a set of synoptic states derived from ERA5 surface winds. The variability associated with grouping cloud occurrence based on synoptic state is much larger than the seasonal variation, highlighting synoptic state is a strong control of cloud occurrence. All the reanalyses continue to display underestimates below 3km and overestimates above 3km for each synoptic state. But, the variability in ERA5 statistics matches the changes in the observations better than the other reanalyses. We also use a machine learning scheme to estimate the quantity of super-cooled liquid water cloud from the ceilometer observations. Ceilometer low-level super-cooled liquid water cloud occurrences are considerably larger than values derived from the reanalyses, further highlighting the poor representation of low-level clouds in the reanalyses.

Plain Language Summary

This study compares cloud observations from a CL51 ceilometer at Scott Base, Antarctica, with data from three weather reanalyses: ERA5, JRA55, and MERRA2. We used a lidar simulator to better match the reanalyses data with the ceilometer's measurements. The reanalyses slightly overestimate cloud presence above 3 km but significantly underestimate it below 3 km compared to the ceilometer data. Both the observations and reanalyses show only small seasonal changes in cloud presence. However, grouping the data by weather patterns shows that these patterns strongly influence cloud presence. The reanalyses still underestimated cloud presence below 3 km and overestimated it above 3 km for all weather patterns, but ERA5 data matched the observed changes better than the other reanalyses. We also used machine learning to estimate the amount of super-cooled liquid water clouds from the ceilometer data. The ceilometer detected many more low-level super-cooled liquid water clouds than the reanalyses simulations, highlighting that issues with the representation of low-level clouds in these models are widespread.

1 Introduction

Clouds are fundamental to the Earth's energy balance, influencing surface temperature by reflecting solar radiation, trapping and emitting infrared radiation. But, comparisons between observations and simulations reveal significant biases in the representation of clouds. In particular, large biases were identified over high latitudes in the Coupled Model Intercomparison Project phase 3 (CMIP3) models (Trenberth & Fasullo, 2010). Subsequent work has made improvements in the simulation of clouds and their properties, but biases are still large and can contain compensating errors which can hide biases (Schuddeboom & McDonald, 2021; Kuma et al., 2023). Identifying biases' sources is crucial, with previous studies identifying that both insufficient cloud cover and problems with the quantity of super-cooled water clouds simulated contribute to biases. The latter issue is a problem because liquid water cloud reflects more shortwave radiation than ice clouds containing the same amount of water (Vergara-Temprado et al., 2018). In particular, models often struggle to simulate super-cooled liquid water clouds accurately leading to significant shortwave radiation biases (Bodas-Salcedo et al., 2016; Kay et al., 2016;

Kuma et al., 2020). Unfortunately, these clouds which occur between the 0°C isotherm and -38°C isotherm, used to represent the homogeneous freezing level, are very common over the Southern Hemisphere (Hogan et al., 2004), the Southern Ocean (Bodas-Salcedo et al., 2016; Kuma et al., 2020) and Antarctica (Listowski et al., 2019). For example, Listowski et al. (2019) identified that the fraction of super-cooled liquid-water containing cloud (SLCC) was of the order of 0–35% over the Antarctic continent. These issues are important because Zelinka et al. (2020) highlighted that changes in the global Effective Climate Sensitivity (ECS) between CMIP phase 5 and 6 models could largely be attributed to changes in the representation of extra-tropical Southern hemisphere clouds.

Observational data on cloud properties at Southern high latitude sites is thus an important constraint on ECS and the representation of clouds. Satellite observations offer the most spatially complete constraints for models and also often provide the longest records above the Southern Ocean and Antarctica. They also have a relatively long history of usage as detailed in Lachlan-Cope (2010) and Bromwich et al. (2012). Satellite data has provided valuable insights on cloud cover, cloud phase, seasonality and the vertical distribution of clouds across the Antarctic continent (Verlinden et al., 2011; Bromwich et al., 2012; Adhikari et al., 2012). However, they do have a number of limitations. In particular, passive satellite sensors face challenges in cloud identification due to the similarity of the properties of snow- and ice-covered ground to low-level cloud (Frey et al., 2008). Additionally, low-level cloud layers and cloud base height observations by satellite instruments are severely limited by the presence of an almost continuous cloud cover in the Southern Ocean which acts to obscure these clouds. Additionally, passive satellite datasets, such as the Moderate Resolution Imaging Spectroradiometer (MODIS; (Platnick et al., 2003)) dataset and the data used in the International Satellite Cloud Climatology Project (ISCCP; (Rossow & Schiffer, 1999)) generally only observe radiation scattered or emitted from cloud top of optically thick clouds. Therefore, these satellites are not suitable for resolving the full vertical profile of clouds in some cases.

These issues are partially mitigated by active satellite instruments, such as the CloudSat Cloud Profiling Radar (CPR) (Stephens et al., 2008) and the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) instrument on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite (Winker et al., 2009). But, even these instruments have limitations. For example, the CPR is affected by ground clutter below 1.2 km (Marchand et al., 2008) while the CALIOP lidar signal is attenuated by optically thick cloud. Given the high occurrence of low-level cloud in the Southern Ocean (Haynes et al., 2011), this factor has been studied to examine the level of underestimation (Alexander & Protat, 2018; McErlich et al., 2021). McErlich et al. (2021) compared two sets of satellite derived cloud products, developed from a combination of CPR and CALIOP data, against ground-based observations made at McMurdo station, Antarctica, collected during the Atmospheric Radiation Measurement (ARM) West Antarctic Radiation Experiment (AWARE) campaign (Lubin et al., 2020). They highlighted that active satellite sensors underestimate low-level cloud relative to surface observations.

In particular, McErlich et al. (2021) showed that both the 2B-CLDCLASS-LIDAR R05 (2BCL5) (Sassen et al., 2008) and raDAR/liDAR (DARDAR) (Delanoe & Hogan, 2010) data products underestimate cloud occurrence below 1.5 km relative to surface observations, with both products distinguishing roughly one third of co-located cloud occurrences observed by AWARE at 0.5 km. Over the Arctic and Antarctic, Silber et al. (2021) also found that differences in instrument sensitivity and detection algorithms can reduce spaceborne estimates of cloud and surface precipitation occurrence frequency by more than 50% relative to surface measurements. More widely, Liu et al. (2016) identified that the CPR experiences contamination in the lowest 1 km due to ground clutter that hinders detection of low marine clouds, inducing an underestimation of up to 39% over the oceans. Other parameters are also known to be affected by signal attenuation due to low-level clouds and ground clutter, for example biases exist in satellite-based observations of radiation (Pei et al., 2023) when compared to Southern Ocean surface observations.

Surface and airborne observations over Antarctica and the Southern Ocean are thus of high value and provide a complement to satellite observations. But, observational campaigns in the Southern Ocean (Kremser et al., 2021; McFarquhar et al., 2021; Sellegri et al., 2023) and around Antarctica (Scott & Lubin, 2014, 2016; Lubin et al., 2020) are challenging, costly, and therefore rare (Lachlan-Cope, 2010). Surface observations of all types also have their own limitations. For example, the lidar signal from surface observations can be attenuated by optically thick low-level cloud which means that the occurrence of high level clouds will be underestimated relative to satellite observations (McErlich et al., 2021). This can also influence integrated quantities, such as cloud fraction, with Listowski et al. (2019) identifying that ceilometer observations of cloud fraction were significantly lower than corresponding values from the DARDAR product over Antarctica.

This study compares cloud data from a Vaisala Cl51 ceilometer at Scott Base, Antarctica, with sets of data from three reanalyses after the application of an instrument simulator (Kuma et al., 2021). By simulating cloud properties which account for instrumental sensitivies, instrument simulators allow a direct quantitative comparison of cloud properties across diverse numerical models with observations which allows a like for like comparison. The use of instrument simulators alleviates some of the issues detailed in Silber et al. (2021). This analysis complements previous work in the region which has directly compared model output with observations. For example, a comparative analysis between observational data from McMurdo Station, Antarctica, and the Community Atmosphere Model version 6 (CAM6) simulations was detailed in Yip et al. (2021). They found that the CAM6 simulation consistently overestimates (underestimates) cloud occurrence above (below) 3 km in every season of the year. However, the effect of instrument sensitives was not considered in that work. Previous work detailed in Kuma et al. (2020) compared ceilometer observations against nudged output from the Global Atmosphere (GA) version 7.1 of the HadGEM3 GCM and MERRA2 reanalysis output processed using an instrument simulator over the Southern Ocean collected across a set of 5 voyages and highlighted the value of using instrument simulators. Notably they found that both the GA7.1 and MERRA2 underestimate low cloud and fog occurrence relative to the ship observations by 4–9% for GA7.1 and 18% for MERRA2.

2 Data and Methodology

Observations from a Vaisala CL51 ceilometer operating at a wavelength of 910 nm (near infrared) deployed at Scott Base (77.8°S, 166.7°E) between February 2022 and December 2023 are used in this study. This wavelength is characterised by relatively low molecular backscattering, but is affected by water vapour absorption (Wiegner & Gasteiger, 2015). The maximum range of the instrument is 15.4 km, with a sampling rate of 6 s and a vertical resolution of 25 m. This instrument produces data files containing uncalibrated attenuated volume backscatter coefficients which are converted to NetCDF using the cl2nc software. These NetCDF files are then processed with the Automatic Lidar Ceilometer Framework software (Kuma et al., 2021) detailed in Section 2.1.

The present study uses outputs from three reanalyses, ECMWF Reanalysis 5 (ERA5) (Hersbach et al., 2020), Japanese 55-year Reanalysis (JRA55) (Kobayashi et al., 2015) and Modern-Era Retrospective analysis for Research and Applications, version 2 (MERRA2) (Gelaro et al., 2017). ERA5 is the fifth-generation ECMWF reanalysis model (Hersbach et al., 2020). 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 super-cooled liquid water and mixed-phase clouds relative to ERA-Interim.

We also use the Japanese 55-year Reanalysis (JRA55), this reanalysis extends for a 55 year period starting from 1958, when regular radiosonde observations became operational globally. Details about JRA55 are detailed in Kobayashi et al. (2015). This study also uses data from the Modern-Era Retrospective analysis for Research and Applications (MERRA2) reanalysis (Gelaro et al., 2017). We used the 3-hourly instantaneous assimilated meteorological fields (inst3_3d_asm_Nv (M2I3NVASM)), to generate simulated ceilometer profiles using ALCF. The four-dimensional MERRA2 fields were provided on pressure and model levels. The analysed time period of all three reanalyses datasets was between 14th February 2022 and 31st December 2023 unless otherwise stated.

We also used data from the Antarctic Mesoscale Prediction System (AMPS), which is an operational forecasting system which uses a version of the Weather Research and Forecasting (WRF) model modified for polar regions (Powers et al., 2012). This study examines forecast output from the Polar WRF version 24 documented online and examines output from AMPS Domain 3.

That domain covers a 1802 × 2766 km area at 2.67 km horizontal resolution and spans the Ross Sea, the Ross Ice Shelf, and the South Pole. AMPS archive data only includes 17 vertical levels with forecasts issued at midday and midnight UTC. Hourly forecasts are utilised in this study. For reference, AMPS obtains initial and boundary conditions from NCEP GFS model output. While near-real-time ice and snow extent (NISE) data provide input sea ice concentration (SIC) values (Brodzik & Stewart., 2016). Unfortunately, the polar WRF model configuration used in the AMPS operational system is changed as improvements become available and these changes are not logged. The Polar WRF output available from the AMPS operational system is therefore not a strong focus of this study.

Instead, we focus on examining the cloud representation in the three reanalyses available. An intercomparison of the wind field for each of these reanalyses over the Ross Ice Shelf region is detailed in McDonald and Cairns (2020) and highlighted that these products were broadly consistent with each other during the satellite period. However, to our knowledge no study examining the quality of multiple reanalyses cloud representation has occurred over Antarctica. The present study aims to fill this gap by comparing reanalyses output with ceilometer data. Ceilometers can provide valuable information on cloud and aerosols, but have not been widely used in the evaluation of climate models, reanalyses and numerical weather prediction models. This is partially related to the wide range of ceilometer instruments, a lack of standardised calibration and the difficulty in directly comparing observations with model outputs. The ALCF software allows the calibration of ceilometer data and the application of its instrument simulator to model outputs removes much of this uncertainty.

2.1 ALCF

This study uses the Automatic Lidar and Ceilometer Framework (ALCF) tool which was first used in Kuma et al. (2020) and was subsequently described in more detail in Kuma et al. (2021). ALCF provides a framework for converting ceilometer data from different manufacturers into a common format, calibrates the backscatter data, resamples data, and also completes a noise removal and cloud detection process.

ALCF also includes a ground-based lidar simulator, which calculates the radiative transfer of laser radiation and allows one-to-one comparison between models and observations. The ALCF ground-based lidar simulator is a development of the CFMIP Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011), a set of instrument simulators developed by the Cloud Feedback Model Intercomparison Project (CFMIP). COSP was originally developed as a satellite simulator package whose aim is to produce virtual satellite (and more recently ground-based) observations from atmospheric model fields in order to improve comparisons of model output with observations (Bodas-Salcedo et al., 2011). This approach is required because physical quantities derived from satellite observations generally do not directly correspond to model fields. ALCF developed a ground-based lidar simulator by modifying the COSP Active Remote Sensing Simulator (Chiriaco et al., 2006). This extension produces virtual backscatter measurements from model fields. Resampling, noise reduction and cloud detection were also performed on observational and derived model lidar output in a consistent way to reduce structural uncertainty. We used the ALCF software to create calibration coefficients for the CL51 ceilometer using the methodology detailed in O'Connor et al. (2004) rather than using the default CL51 calibration available within the package. ALCF developments required reversing the vertical layers, as the surface lidar looks from the surface up rather than down from space to the surface, and changing the radiation wavelength affecting Mie scattering by cloud droplets and Rayleigh scattering by air molecules. We only present a brief description of the surface lidar simulator and instead encourage interested readers to examine Kuma et al. (2021).

The recently introduced COSP version 2 (Swales et al., 2018) added support for a surface lidar simulator, although we believe that ALCF, developed before COSPv2 was available, is more complete in the present context due to its treatment of Mie scattering at wavelengths other than 532 nm (the wavelength of the CALIOP lidar). It also adds a more detailed simulation of ice crystal optical properties. The surface lidar simulator takes model cloud liquid and ice mixing ratios, cloud fraction and thermodynamic profiles as the input, and calculates vertical profiles of attenuated backscatter.

2.2 Super-cooled cloud detection

Guyot et al. (2022) developed an algorithm to detect super-cooled liquid water containing clouds (SLCC) based on the co-polarization backscatter measured by ceilometers using observations from a training dataset collected at Davis station, Antarctica. This classification model used an extreme gradient boosting (XGBoost) framework ingesting backscatter data with an accuracy as high as 0.91. More recently the same framework has also been applied with modifications over mid-latitudes by Whitehead et al. (2023), the modifications being necessary because regions which also include warm liquid cloud impact the accuracy of the Guyot et al. (2022) scheme outside the polar environment. This study applies the Guyot et al. (2022) classification scheme to ceilometer backscatter measurements made at Scott Base. We note that a validation of the Guyot et al. (2022) scheme is not possible without reference data. But, visual inspection initially identified poor classification results when the Guyot et al. (2022) scheme was applied to Scott Base data using the default ALCF calibration coefficient for a CL51 ceilometer. However, after using the O'Connor et al. (2004) methodology to calibrate the ceilometer the scheme worked well based on visual inspection, with perhaps some periods where SLCC is under reported. We thus detail the results of the application of the Guyot et al. (2022) classification scheme to the CL51 backscatter data in this paper.

2.3 Synoptic typing

Jolly et al. (2018) has suggested that cloud occurrence over the Ross Ice Shelf is strongly impacted by synoptic state and this variation is significantly larger than the observed seasonal cycle. We therefore create a set of synoptic states using a similar methodology to that used in McDonald and Cairns (2020). In particular, our synoptic types are derived using the Self-Organizing Map (SOM) technique applied to ERA5 10m wind speeds between 1979 and 2023 to derive representative surface wind patterns.

SOMs are an iterative unsupervised learning scheme commonly used in clustering (Kohonen, 1990). The learning process adjusts a set of reference vectors based on the differences between the reference vector and each input record. A learning rate determines how the adjustment is related to the difference between the reference vector and the input data measured by the Euclidean distance metric. Training then entails adjusting reference vectors iterativly until a set of stable values are reached. The learning rate and width of the kernel are reduced as a function of time such that the SOM evolves rapidly initially. The Euclidean distance is used to identify reference vectors within a certain range of the best matching vector. The vectors that fall within this neighborhood are then updated which produces the coherent organization of output. During each iteration, the reference vector that best matches the input record is identified and then modified to better reflect the input data. The training process ultimately produces reference vectors that represent the multidimensional input space.

Rather than apply the SOM technique directly to all the ERA5 output, we reduce the quantity of input into the SOM by applying an Empirical Orthogonal Function analysis to the space-time cube of the surface winds (both zonal and meridional winds) and then apply the SOM technique to the largest Principal Components (PCs) only. In this study, we truncated the set of PCs when the explained variance was 90% of the total variance of the dataset. In this study, we used the implementation of the SOM methodology available in the mini-SOM python package (Vettigli, 2018).

The usage of the Empirical Orthogonal Functions (EOF) analysis requires anomalies as inputs and the climatological mean from each latitude/longitude point for the 1979–2023 reference period was used to derive anomalies. Our analysis focuses on the geographic domain (60–90°S, 140–220°E) used previously in McDonald and Cairns (2020). We also derived a daily average to reduce the processing requirements for the study. Previous work detailed in Tastula et al. (2013) identified that near-surface wind speed displays low diurnal variability in both observations and in reanalyses products over Antarctica and thus our choice to use daily averages should not impact our results.

3 Results

Figure 1 displays the mean cloud occurrence as a function of altitude derived from the CL51 ceilometer observations, and predictions of cloud occurrence derived with the ALCF surface lidar simulator from input from Polar WRF, ERA5, JRA55 and MERRA2 models. These mean values are derived for the period 14th February 2022 to 31st December 2023 where both ceilometer and model outputs are available. The maximum in cloud occurrence for the CL51 observations peaks at the surface, but this peak is potentially contaminated by low-level fog and wind blown snow trappped below the commonly observed low-level inversion layer (Hofer et al., 2021). The backscatter near the surface is also more uncertain because of the overlap function used. Thus, the secondary peak with a value of just over 25% cloud occurrence at approximately 800m above the surface is likely the true maxima observed by this system. These values of cloud occurrence at this peak are roughly 10% lower than previous surface observations from McMurdo station detailed in Silber et al. (2018). Though, the general form of the vertical profile of cloud occurrence is very similar. The difference may be partially connected to the greater attenuation of the ceilometer signal due to obscuring optically thick clouds compared to those detailed in Silber et al. (2018) which used a more powerful lidar instrument and also included information from a Ka-Band cloud radar. The High Spectral Resolution Lidar (HSRL) is also more sensitive to tenuous cloud. Additionally, the variability from day to day, seasonally and with synoptic types is large based on the ground-based observations discussed in Silber et al. (2018) and therefore interannual variability could also partially explain this difference. We also note that the cloud thresholding scheme available in the ALCF software likely provides a conservative estimate of cloud occurrence.

Satellite observations averaged over the Ross Ice Shelf detailed in Jolly et al. (2018) show peak cloud occurrences between 20 and 30% at approximately 2 km for all seasons which are larger than the peak values observed by the CL51 ceilometer in Figure 1. However, the satellite observations have lower cloud occurrences than the CL51 ceilometer

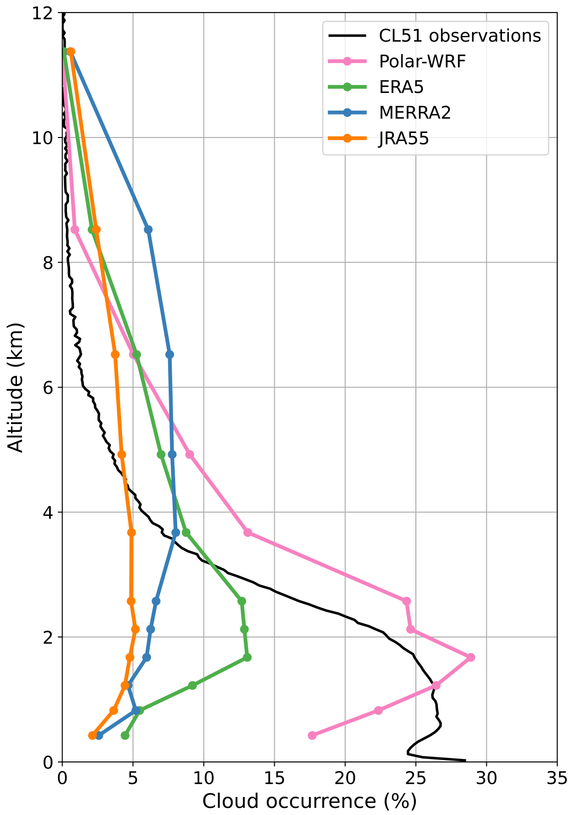

Figure 1. Mean vertical profiles of cloud occurrence for the period 14th February 2022 to 31st December 2023 derived from CL51 ceilometer observations (black line) and the AMPS (pink line), ERA5 (green line), JRA55 (orange line) and MERRA2 (blue line) model fields after processing using the ALCF ground-based lidar simulator.

values at altitudes below approximately 2 km. This difference is likely due to different instrument sensitivities. In particular, satellite observations of low-level cloud will likely be underestimates, while ground-based observations of upper-level cloud occurrence will be underestimates (McErlich et al., 2021). This comparison highlights the important of different instrument sensitivities. Comparison between model properties and observations which do not account for instrument sensitivities can thus bias model evaluations.

Comparison between the CL51 observations and the ERA5, JRA55 and MERRA2 cloud occurrence profiles derived using ALCF displayed in Figure 1 show low biased values relative to the CL51 observations of cloud occurrence for altitudes below 3 km and high biased cloud occurrences in ERA5, JRA55, and MERRA2 above that altitude. The Polar WRF cloud occurrence values derived from ALCF are slightly higher than the CL51 observations above 1 km and lower than the CL51 observations below 1 km. But, compared to the three reanalyses display very good correspondence with the CL51 ceilomter observations. Unfortunately, the Polar WRF simulations used are derived from the operational Antarctic Mesoscale Prediction System and the configuration of the Polar WRF changes multiple times during this study. We therefore limit the use of this dataset in later analysis and focus on the three reanalyses. However, the fact that a numerical model which includes tuning for the polar environment displays such a significant improvement relative to the reanalyses is notable.

Work detailed in Yip et al. (2021) compared the same AWARE observations as used in Silber et al. (2018) with CAM6 model data. The CAM6 simulations examined in Yip et al. (2021) were nudged toward MERRA2 reanalyses temperature and wind fields. Similar to our results, they identified sizable overestimates (underestimates) of cloud occurrence above (below) 3 km in the model. We also note that the general form of the vertical profile of cloud occurrence in MERRA2 displayed in Figure 1 is rather similar to the CAM6 equivalent, though the CAM6 cloud occurrence is roughly 15% greater than the corresponding MERRA2 values at the same altitude. This is likely associated with changes in the comparison process due to the use of an instrument simulator in this study.

Yip et al. (2021) also identified that cloud occurrence biases were closely associated with concurrent biases in relative humidity in the CAM6 model. With high relative humidity biases between the CAM6 data and observations above 2 km and low relative humidity biases below 2 km. To test whether this may also be a controlling factor for the three reanalyses, we compare the relative humidity from the reanalyses with radiosonde observations launched from the nearby (less than 3 km separation) McMurdo station. Figure 2 displays the median and interquartile ranges of the difference between radiosonde observations and the three reanalyses (model-observation). The difference between the McMurdo radiosonde relative humidity and the ERA5 values, shown in Figure 2 (a), display overestimates of the relative humidity below 2 km, a region of underestimates between 2 and 4.5 km and larger overestimates above this altitude. A similar pattern of bias between the observations and the JRA55 reanalyses relative humidity is displayed in Figure 2 (b), though the biases are larger than those from the ERA5 dataset apart from at around 9 km. The MERRA2 observations display the same structure of bias as ERA5 and JRA55 relative to the radiosonde observations.

Figure 3 (a)-(d) displays vertical profiles of the median and interquartile range cloud occurrence for each season from the ceilometer observations and the three reanalyses. Examination of the CL51 observations shows the largest cloud occurrences and the largest range of values occurs in austral autumn (MAM) and winter (JJA) and lower values o fthe cloud occurrence and interquartile range in austral spring (SON) and summer (DJF). In particular, the median cloud occurrence is up to 25% in austral autumn and winter, but below 20% in austral spring and summer. However, in every season the largest cloud occurrences are observed in the 2 km directly above the surface in the Cl51 observations and cloud occurrence reduces relatively quickly between 2 and 4 km to only a few percent above 4 km in all seasons. The reanalyses results display marginally higher values

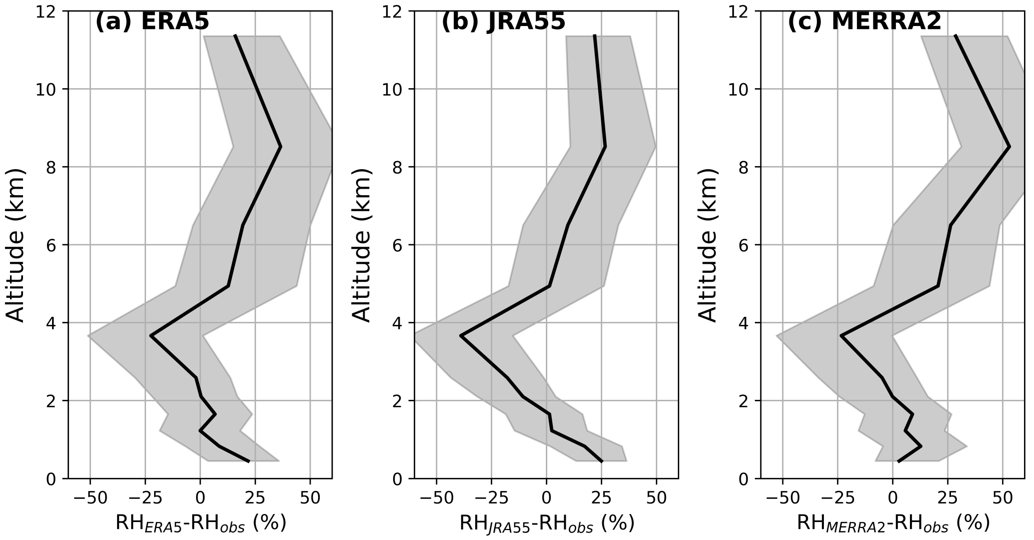

Figure 2. Median and interquartile ranges of the difference between radiosonde observations and reanalyses values of relative humidity for the ERA5 (a), JRA55 (b) and MERRA2 (c) reanalyses.

in austral spring, but small seasonal variations. It is also relatively clear that cloud occurrences are lower in the JRA55 reanalysis than the other two reanalyses which show more comparable results. Though, as previously seen in Figure 1 the MERRA2 model output has higher cloud occurrences at higher altitudes than the ERA5 values.

The biases between the reanalyses and the CL51 ceilometer observations (model - observations) are shown in Figure 3 (e)-(h) for each season. In each season, a large underestimate (15-30%) in cloud occurrence is observed for all three reanalyses below 3 km with a smaller overestimate in cloud occurrence in the reanalyses above 3 km. Interestingly, the bias at low altitudes is comparable to the values identified by Yip et al. (2021) and Kuma et al. (2021). The ERA5 reanalysis displays the smallest biases of the three reanalyses at most altitudes in most seasons. The JRA55 reanalysis displays the largest underestimates below 3 km in all seasons and the MERRA2 reanalysis has the largest overestimates relative to the CL51 observations above 3 km in all seasons. Closer inspection of Figure 3 (e)-(h) shows variations in the bias with season, with the largest underestimates below 3 km in austral autumn and winter and the smallest underestimates in austral spring and summer. An examination of the altitudes where the reanalyses overestimate cloud occurrence also shows that the austral spring displays the largest overestimates, which reach 10% in MERRA2. Notably the magnitudes of the underestimated and overestimated values are more similar in Yip et al. (2021) than in this study. This difference can likely be explained by the use of an instrument simulator in this study which allows a more robust comparison between the observations and the model output. Effectively, the low cloud occurrences in the observations at higher altitudes are likely impacted by instrument sensitivity which means that they are likely low biased estimates.

The median and interquartile ranges for cloud fraction, which is defined in this study as the temporal average cloud occurrence independent of altitude, as a function of month are displayed in Figure 4. A quite small variation in the median cloud fraction between months is observed for both the CL51 observations and the three reanalyses. In particular, for the CL51 observations and for the three reanalyses the interquartile range in

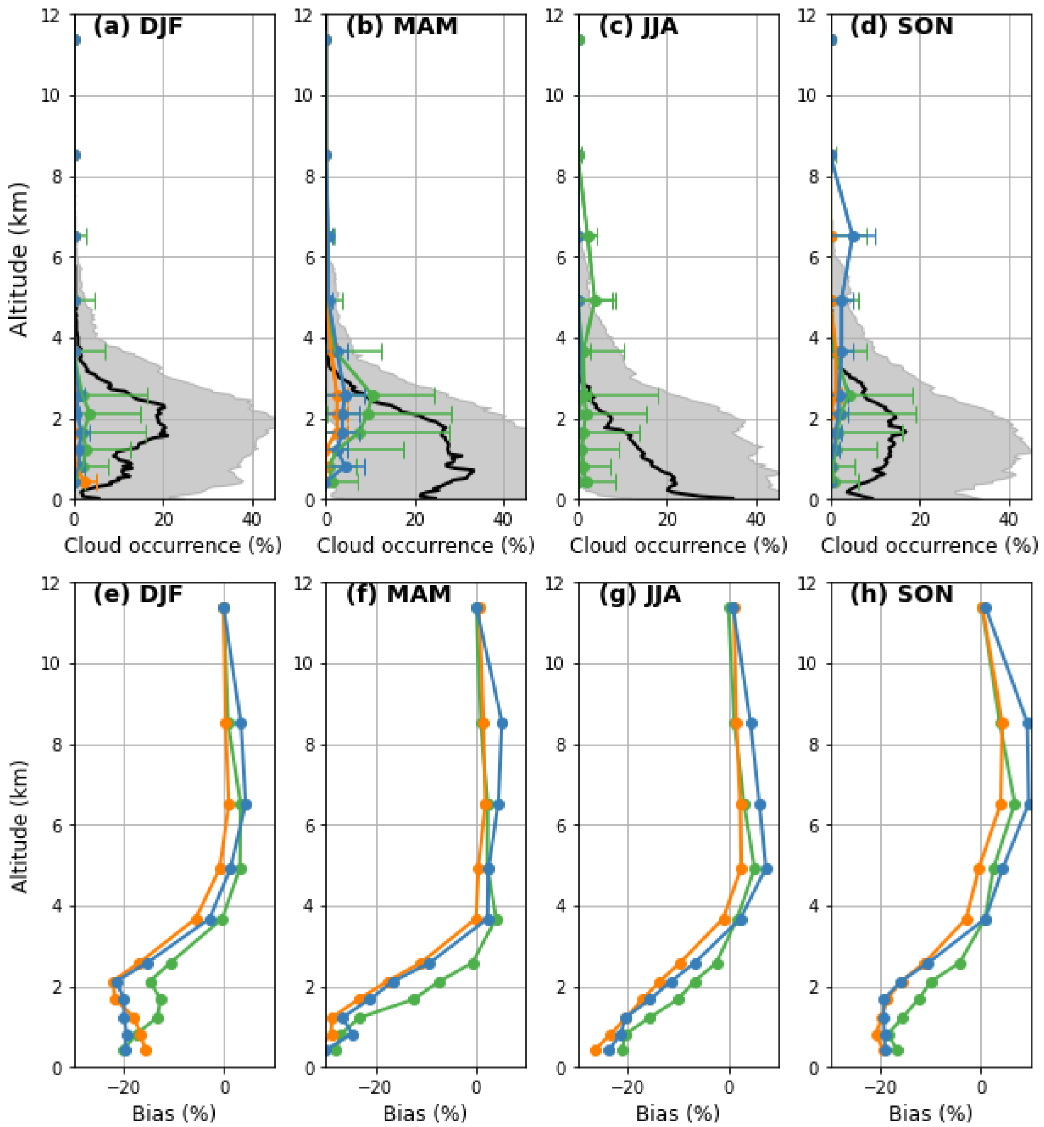

Figure 3. Median and interquartile ranges of cloud occurrence vertical profiles derived from CL51 ceilometer observations (black line), ERA5 (green line), JRA55 (orange line) and MERRA2 (blue line) model fields after processing using the ALCF ground-based lidar simulator are displayed for austral summer (a), autumn (b), winter (c) and spring (d). The model bias (model minus observation means) is also shown for austral summer (e), autumn (f), winter (g) and spring (h).

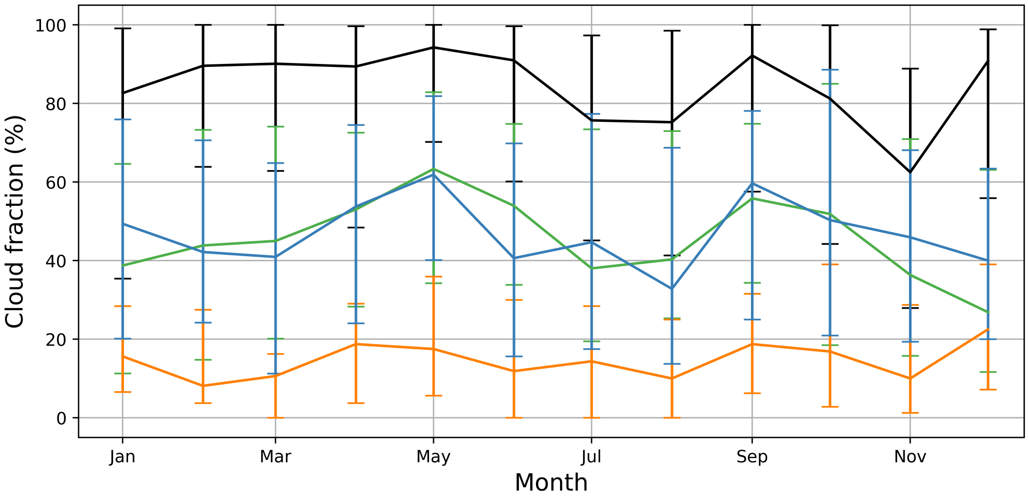

Figure 4. Median and interquartile ranges of cloud fraction each month derived from CL51 ceilometer observations (black line), ERA5 (green line), JRA55 (orange line) and MERRA2 (blue line) model fields after processing using the ALCF ground-based lidar simulator.

any month usually contains the range of median values for all months. However, comparison of the median cloud fraction between the CL51 observations and the three reanalyses very clearly shows large offsets. In particular, the median values of cloud fraction for the CL51 observations are between 62-92%, the ERA5 values are between 27-63%, the JRA55 values are between 10-22% and the MERRA2 values are 32-62%. It is thus very clear that all three reanalyses underestimate cloud fraction, though the underestimate is particularly large for the JRA55 reanalyses.

3.1 Synoptic classification

The relatively small variation between the different seasons observed in Figure 3 and 4 has previously been identified in other studies. In particular, Jolly et al. (2018) and Silber et al. (2018) identified that the synoptic situation has a much larger impact on vertical cloud distributions in this region than seasonal variability. We therefore complete a synoptic classification over the Ross Sea region, this allows the CL51 ceilometer data collected from Scott Base between February 2022 and January 2024 to be categorised based on synoptic state. The method used to complete this synoptic classification is detailed in Section 2.3. The surface horizontal wind vectors and wind speeds associated with each synoptic state are shown Figure 5 over the Ross Sea region. This classification is used to group the corresponding data from the CL51 observations and the output of the ALCF lidar simulator output derived from the three reanalyses in Figure 6.

A 3 × 3 SOM (3 columns and 3 rows) was selected for our classification because it minimized quantization error and represented a good balance in terms of representation of the wind patterns over the region. The set of nodes from this reference period are used throughout this study for grouping the ceilometer data collected and also to group the corresponding data from the ALCF lidar simulator (Kuma et al., 2021) derived from the ERA5, JRA55 and MERRA2 models.

The different wind patterns in Figure 5 are dominated by southerly winds in all of the nodes derived, except for node 2. But, the magnitude of the wind changes significantly. For example, nodes 2 and 5 display weak winds directly over Ross Island, the

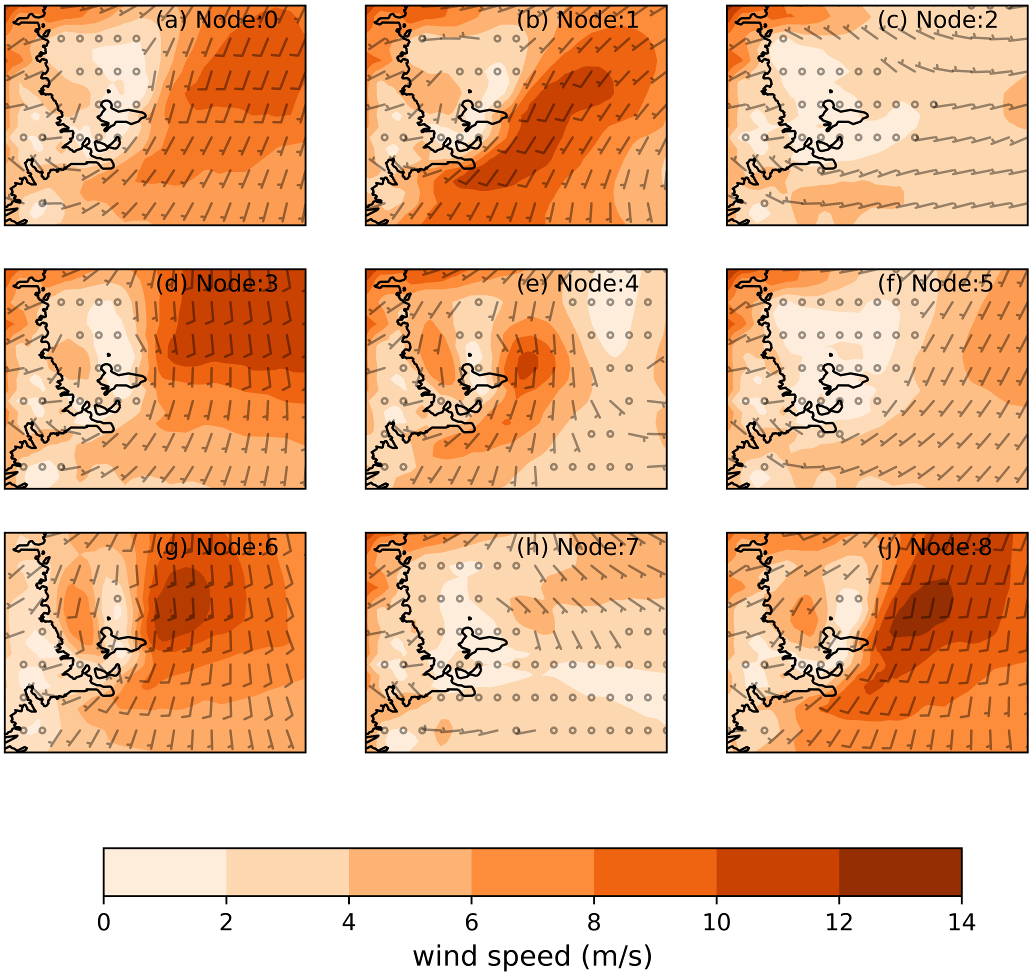

Figure 5. Near-surface (10m) horizontal wind speeds and directions for each of the 9 nodes in the SOM derived from ERA5 reanalysis output for the period 1980–2024.

site of Scott Base and the ceilometer, while node 1 displays rather strong winds directly to the east of Ross Island. We also note that the nodes at the opposite corners of the SOM (node 2 and 6) display the largest difference in terms of wind magnitudes. Node 2 is also dominated by westerly winds.

Figure 6 displays vertical profiles of cloud occurrence for the CL51 observations and the results of the application of the ALCF instrument simulator to the three reanalyses grouped based on the synoptic conditions displayed in Figure 5. Examination of the CL51 cloud occurrence patterns shows significantly larger variability between synoptic states (Figure 6) than for different seasons (Figure 3). In particular, nodes 1 and 4 display maximum cloud occurrences above 40% at altitudes below 2 km. Though, the cloud occurrence begins to fall from around 50% from 1 km. While the lowest CL51 cloud occurrences are observed in node 2 and 5. We also note that high cloud occurrences are very close to the surface in nodes 6 and 7 which potentially suggests the presence of wind blown snow, fog or low-level temperature inversions in these synoptic situations.

Vertical profiles of the cloud occurrence derived from the ERA5 reanalysis show higher values for node 1 and 4 close to 2 km and lower values in node 2 and 5 in Figure 6. These patterns match closely with the CL51 ceilometer observations for these synoptic situations above 2 km. However, the cloud occurrence is underestimated for all nodes below 2 km. Additionally, nodes 6, 7 and 8 display substantial overestimates in a relative sense for cloud occurrence above 2 km.

Vertical profiles of the cloud occurrence derived from the JRA55 and MERRA2 reanalyses also show higher values for node 1 and 4 and lower values in node 2 and 5 in Figure 6. However, these patterns match much less closely with the CL51 ceilometer observations than the ERA5 values. The patterns are quite consistent between the JRA55 and MERRA2 simulation results, though notably the MERRA2 cloud occurrences are higher at nearly every altitude in every node than the corresponding JRA55 values. Additionally, for nodes 3, 6, 7 and 8 the JRA55 and MERRA2 values display substantially overestimated cloud occurrence above 2 km relative to the CL51 ceilometer observations. These overestimates are also significantly larger than those observed between ERA5 and the CL51 ceilometer observations. However, the cloud occurrence is underestimated for all nodes below 2 km.

3.2 Cloud phase analysis

To obtain more information from the ceilometer observations we apply the XGBoost algorithm detailed in Guyot et al. (2022) to derive the fraction of cloud that is associated with super-cooled liquid water. Figure 7 (a) displays the attenuated volume backscatter coefficient data over Scott Base for the 27th March 2022. On this day, a narrow band of low-altitude (< 1 km) multi-layer cloud occurs between 00:00 and 09:00 UTC, while a thicker band of cloud is present at altitudes ranging from 1-3 km between 15:00 and 23:00 UTC. The cloud classification displayed in Figure 7 displays thin ice cloud layers close to the surface between 00:00 and 09:00 UTC capped by strongly attenuating super-cooled cloud layers. The thicker cloud layer between 15:00 and 23:00 UTC either has an ice cloud or undefined classification. The presence of a small amount of super-cooled liquid cloud within that thicker layer may suggest the presence of mixed phase cloud or may be a classification error. The lack of depolarisation data from the CL51 ceilometer means that we can not validate the Guyot et al. (2022) scheme. But, visual inspection does suggest that the strongly attenuating cloud layer between 00:00 and 09:00 UTC is correctly classified as super-cooled liquid cloud. Though, the lack of continuity in the backscatter coefficient data for that layer suggests that any estimates super-cooled liquid water cloud pressence are likely to be conservative. This also tallies with results from Guyot et al. (2022) which came to the same conclusion. We also note the relatively high proportion of undefined clouds in this case, which means that these peaks in the attenuated volume backscatter coefficient can not be classified as either ice or super-cooled liquid water. This is potentially due to the presence of mixed phase cloud in our observations. We also reiterate that the classification scheme detailed in Guyot et al. (2022) does appear to be sensitive to the calibration factors applied to the data. We thus advise future users to complete calibration using the O'Connor et al. (2004) or Hopkin et al. (2019) scheme rather than using default values from ALCF.

The application of the Guyot et al. (2022) XGBoost scheme allows us to derive the fraction of cloud peaks classified as super-cooled liquid water relative to other classes (the combination of ice and undefined). The mean fraction of super-cooled liquid water cloud as a function of altitude can then be derived from the CL51 ceilometer observations between 14th February 2022 to 31st December 2023. Figure 8 (a) displays vertical profiles of the occurrence of super-cooled liquid water cloud and all cloud. We note that the super-cooled liquid water cloud fraction remains relatively constant between 0.5 to 2.5 km at an occurrence rate of 5%, then rapidly declines to near zero values above 4 km. The form of this vertical profile is very similar to that previously displayed in Silber et al. (2018) with near constant values between 0.5 and 3 km. Though, we note that the cloud occurrence is again lower than that identified in Silber et al. (2018). We again be-

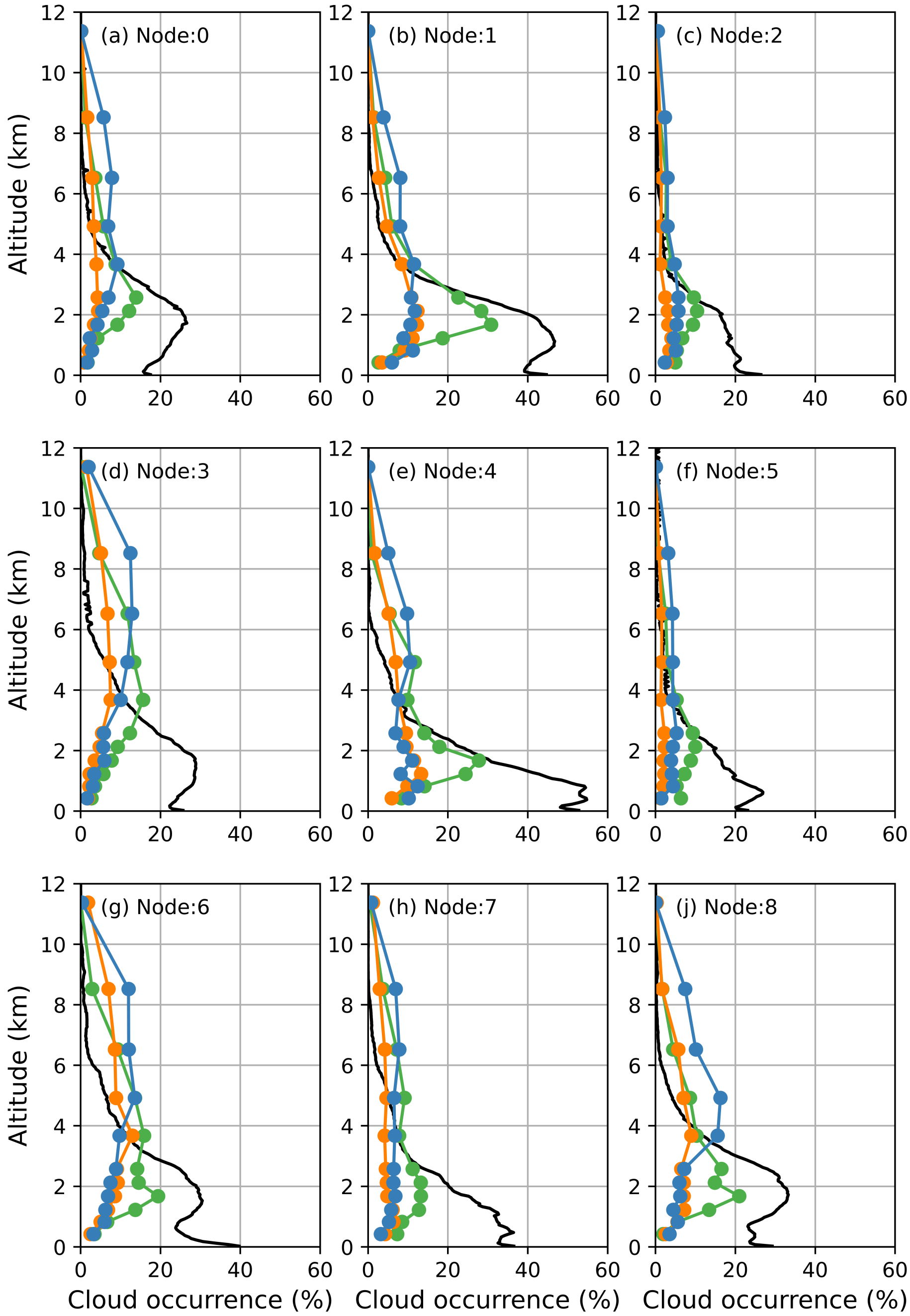

Figure 6. Mean vertical profiles of cloud occurrence for each synoptic state for CL51 ceilometer observations (black), ERA5 (green), JRA55 (orange), and MERRA2 (blue) model fields.

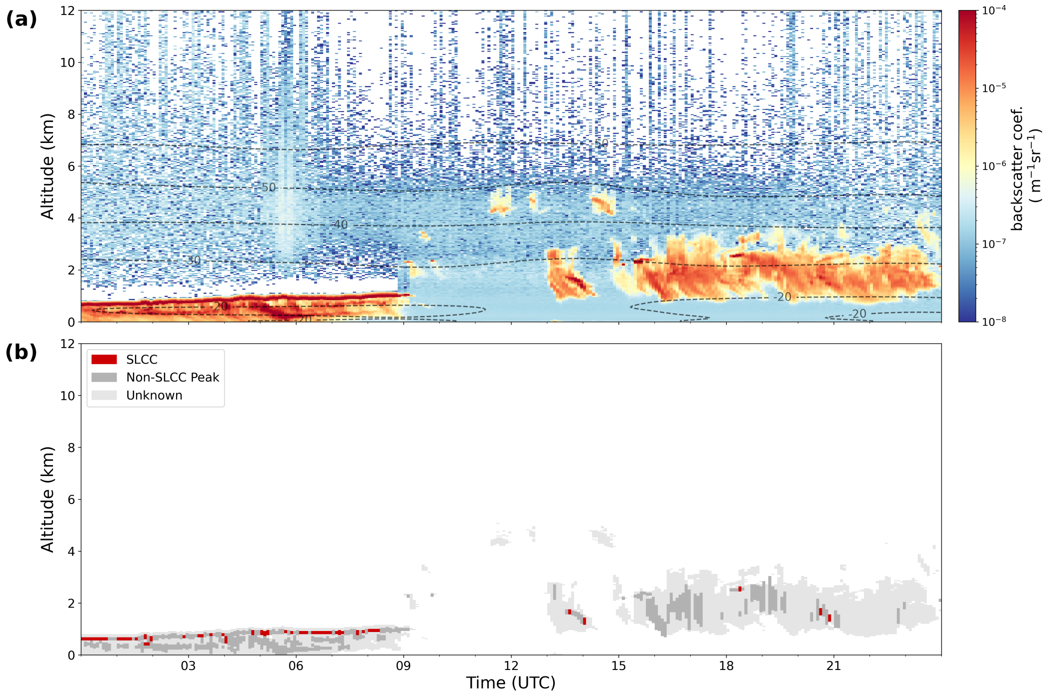

Figure 7. Cl51 ceilometer attenuated volume backscatter coefficient data over (a), and the (Guyot et al., 2022) cloud mask (b) for 27th March 2022. AMPS air temperature contours are overlaid in (a) for reference.

lieve that this is associated with differences in the capability of the HSRL used in Silber et al. (2018) and the CL51 ceilometer observations used in the present study.

Figure 8 (b)-(e) display the mean cloud occurrence and the super-cooled liquid water cloud fraction for each season. The mean cloud occurrence displays similar patterns to the median values previously shown in Figure 3 as expected. Comparison of the super-cooled liquid water cloud occurrences between the seasons shows the highest super-cooled liquid water fractions in austral summer (Figure 8 (b)) and the lowest values in the austral winter (Figure 8 (d)). Thus, while the vertical profile of cloud occurrence is strongly defined by synoptic state (see Figure 6), cloud phase is strongly controlled by season. This likely reflects variations in the the occurrence of temperatures between the 0°C isotherm and the homogeneous freezing level (-38°C) with season.

The Guyot et al. (2022) XGBoost scheme requires information on the width of cloud peaks. In particular, super-cooled liquid water cloud is partially identified by narrow peaks in the vertical profiles of attenuated volume backscatter coefficient. The low vertical resolution of the reanalysis and their varying vertical resolution with altitude precludes the use of the scheme as derived in Guyot et al. (2022) on this model output. Instead we apply the simple scheme detailed in Desai et al. (2023) in which cloud phase is defined using the ice mass fraction (μ_{ice}). The ice mass fraction is shown in Equation 1 and is obtained by taking the ratio of the ice water content (IWC) to the total water content. Desai et al. (2023) classified grid points where $\mu_{ice} > 0.9$ as ice, $0.1 \leq \mu_{ice} \leq 0.9$ as mixed phase, and $\mu_{ice} < 0.1$ as liquid phase.

$$\mu_{ice} = \frac{\mathrm{IWC}}{\mathrm{IWC} + \mathrm{LWC}} \tag{1}$$

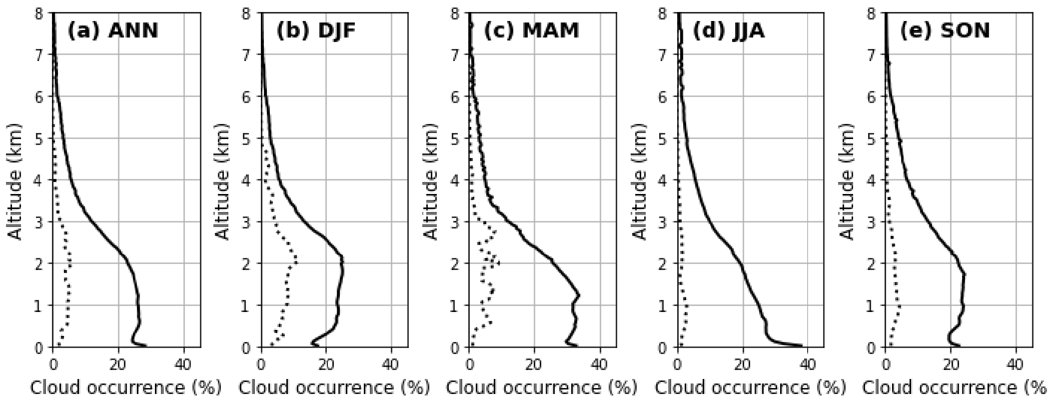

Figure 8. Mean vertical profiles of cloud occurrence averaged over the observational period (a) and austral summer (b), spring (c), winter (d) and autumn (e) derived from CL51 ceilometer observations (black line) and the corresponding super-cooled liquid water occurrence (black dotted line).

Figure 9 (a)-(c) displays vertical profiles of the mean ice water fraction for the ERA5, JRA55 and MERRA2 reanalysis data at Scott Base derived between 2022 and 2023 inclusive. The ice water fraction is near one for ERA5 everywhere apart from the lowest 2 km of the atmosphere (see Figure 9 (a)) which suggests that nearly all the cloud identified would be ice cloud based on the Desai et al. (2023) scheme. The ice water fraction is even larger for the JRA55 reanalysis with only the lowest altitude displaying a value which would be connected to mixed phase cloud. Interestingly, the MERRA2 reanalysis shows much smaller ice cloud fraction values than ERA5 and JRA55, with values between 0.1 and 1.0 between the surface and 6 km, above which the mean value is one.

Figure 9 (d)-(f) display the total cloud occurrence taken directly from the reanalyses, and the cloud occurrence associated with ice and liquid water derived using the Desai et al. (2023) scheme. Comparison between the total cloud occurrence in Figure 9 (d)-(f) and Figure 1 allows the effect of the instrument simulator to be examined. Comparison between Figure 9 (d) and Figure 1 shows that the raw ERA5 cloud occurrences are around 4-5% higher than those derived after the application of the instrument simulator. This difference likely represents the impact of attenuation by low-level cloud and instrument sensitivity affects meaning that tenuous clouds will not meet the backscatter threshold used in the cloud detection scheme. Comparison between Figure 9 (e) and Figure 1 shows that the raw JRA55 cloud occurrences have marginally higher values than those derived via the instrument simulator, the small difference may be associated with the small quantity of low-level cloud in the JRA55 simulation. Finally, inspection of Figure 9 (f) and Figure 1 shows a sizable difference between the raw MERRA2 cloud occurrences and those derived from the instrument simulator. The difference is particularly large above 2 km, again likely due to instrument sensitivity factors and the simulation of the effect of attenuating low-level clouds. This comparison demonstrates the value of the use of instrument simulators in the evaluation of model output.

We now focus on the occurrence of ice and liquid water cloud, as identified by the Desai et al. (2023) scheme, in Figure 9 (d)-(f). Notably ERA5 and JRA55 display such small quantities of liquid water that the occurrence of that cloud is barely visible relative to the zero occurrence line. Interestingly, MERRA2 displays liquid water occurrences up to approximately 2.5% below 2 km. However, these values are still considerably smaller

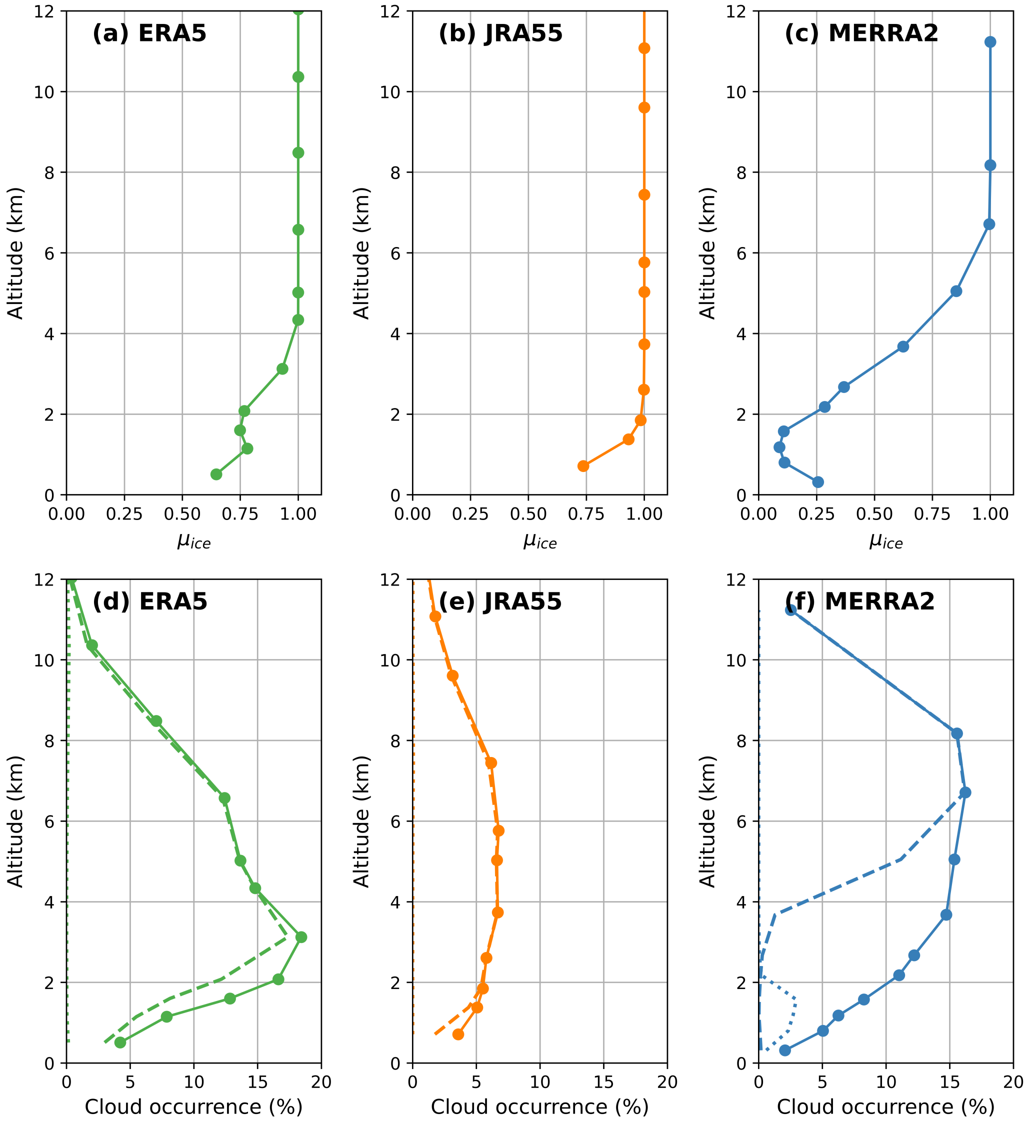

Figure 9. Mean vertical profiles of the ice water faction (μ_{ice}) for the ERA5 (a), JRA55 (b) and MERRA2 (c) reanalysis. Mean vertical profiles of the cloud occurrence (full line) from the ERA5 (d), JRA55 (e) and MERRA2 (f) reanalysis. The cloud occurrence associated with ice (dashed line), liquid (dotted line) water derived from the analysis of μ_{ice} using the methodology detailed in Desai et al. (2023), as well at the total cloud occurrence (full line) are presented in (d) through (f). Note that cloud occurrences have not been derived from data processed using the ALCF lidar simulator.

than the occurrence of super-cooled liquid cloud in Figure 8, though the fraction of liquid water to the total cloud occurrence is larger than that in Figure 8. Comparison of the ice water cloud occurrence (dashed line) and total cloud occurrence in Figure 9 (d) show that some mixed phase cloud, as identified by the Desai et al. (2023) scheme, exists at altitudes below 4 km in the ERA5 dataset. Similar comparison for Figure 9 (e) shows a very small of mixed phase cloud exists below 2 km in the JRA55 reanalysis. Finally, the difference between the ice cloud occurrence line and the total cloud occurrence line in Figure 9 shows that mixed phase cloud makes up the majority of the cloud observed between approximately 2 and 5 km in the MERRA2 reanalysis.

4 Conclusions and Discussion

This paper has principally detailed an analysis of CL51 ceilometer observations relative to ERA5, JRA55 and MERRA2 model output that has been processed using an instrument simulator. The application of the instrument simulator to the reanalyses output allows the derivation of pseudo-backscatter profiles, which in turn can be processed using the same cloud mask algorithm. This processing therefore allows a like-for-like comparison to be performed between the ceilometer and reanalyses output which accounts for instrumental sensitivities and differences in the way that the models represent cloud. Comaprison between cloud occurrences derived from the instrument simulator (Figure 1) and those taken directly from the reanalyses (Figure 9) highlight the value of this methodology. However, it must be borne in mind that the nature of the radiative transfer calculations used in the lidar simulator mean that the impact of both cloud phase and cloud fraction are convolved.

Comparison of the CL51 ceilometer vertical profiles of cloud occurrence relative to previous observations made during the AWARE campaign (Lubin et al., 2020) suggest that low-level cloud may be underestimated because of differences in instrument sensitivity. However, comparison with previous CALIOP-CloudSat climatologies over the Ross Ice Shelf (Jolly et al., 2018) suggest that these observations observe significantly more cloud below 2 km than the satellite observations, this result further supports the conclusions made in McErlich et al. (2021).

Critically, we find that the vertical profile of cloud occurrence for all three reanalyses shows significant underestimation below 3km and a smaller overestimation above that altitude relative to the CL51 observations. This result compares qualitatively with a comparison between CAM6 simulations and the AWARE dataset detailed in Yip et al. (2021) which was partially attributed to low biases in humidity relative to observations.

Recent work detailed in Zhang et al. (2023) has identified that output from the Energy Exascale Earth System Model version 2 (EAMv2) tends to overestimate cloud frequency of occurrence throughout the year in Antarctica which differs from our results. However, they also find that cloud base height and cloud top height are much higher than observations across the year. This would suggest underestimates of cloud occurrence at low altitudes and overestimates at higher altitudes which match with the results observed in the present study. They also identify that EAMv2 tends to simulate stratiform mixed-phase clouds with significantly underestimated liquid water paths at McMurdo station. This matches with results form the ERA5 and JRA55 reanalysis which show smaller fractions of liquid water clouds than identified in the ceilometer observations using the XG-Boost scheme (Guyot et al., 2022).

Furthermore, Yip et al. (2021) highlighted a strong positive relationship between biases in cloud occurrence and relative humidity between CAM6 model output and observations made at McMurdo station. Examination of Figure 2 and Figure 3 (e)-(h) shows that this relationship is not identified when looking at the three reanalyses relative to the observations used in this study. In particular, the relative humidity is overestimated in all of the reanalyses relative to radiosonde observations made at McMurdo station in the bottom 2 km of the atmosphere, while cloud occurrences are underestimated. This difference between the present study and the result in Yip et al. (2021) could be partially explained by our use of an instrument simulator which allows a more robust comparison between the cloud occurrence observations and the model output. The similar magnitudes of the underestimated and overestimated cloud occurrences below and above 3km in Yip et al. (2021) are potentially caused by the a lack of consideration of instrumental factors relative to our analysis which shows much larger biases at low altitudes. The lack of correlation between relative humidity and cloud occurrence biases at low-levels for the three reanalyses suggests that the cloud occurrence biases are likely due to parameterisation errors.

Further support for the robustness of the present analysis comes from results in Kuma et al. (2020), which compared ceilometer observations against nudged HadGEM3 general circulation model and MERRA2 reanalysis output processed using the ALCF instrument simulator. In particular, the biases between the MERRA2 cloud occurrences and the ceilometer observations over the Southern Ocean were quite similar to those observed in the present study at low altitudes.

Our results also show that the CL51 ceilometer seasonal cloud occurrence and cloud fraction shows little variation, similar to results in Jolly et al. (2018) and Silber et al. (2018). Notably, we find that there is a lack of a strong seasonal cycle in cloud fraction in both the CL51 ceilometer observations and the reanalyses. However, the cloud fraction is underestimated by around 25% in ERA5 and MERRA2 and by 70% in JRA55 relative to the CL51 ceilometer observations. This work thus further demonstrates the value of instrument simulators model evaluation.

Given that previous work has highlighted the importance of synoptic state on cloud properties, we derived a synoptic classification using a similar methodology to that detailed in McDonald and Cairns (2020). As expected, when grouping cloud occurrence vertical profiles by synoptic state mean values display much larger variability than that observed for different seasons. All three reanalyses continue to display underestimates of cloud occurrence above 3km and overestimates above 3km relative to the ceilometer observations for all the different nodes in our synoptic classification. However, the ERA5 reanalyses variability in cloud occurrence matches the changes observed in the CL51 observations for different synoptic state much better than the other two reanalyses. In particular, higher cloud occurrences are observed for node 1 and 4 close to 2 km and lower values in node 2 and 5. Given that much of the higher altitude cloud in this region is associated with large scale synoptic features, such as extra-tropical cyclones, this suggests that ERA5 represents these controlling factors better in these situations than either JRA55 or MERRA2.

Additionally, we note that the cloud occurrence is underestimated for all nodes below 2 km in all three reanalyses. Positively, vertical profiles of the cloud occurrence derived from the JRA55 and MERRA2 reanalyses do show variations in cloud occurrence which correspond with the CL51 observations, though the correspondence is much poorer than that between the CL51 observations and the ERA5 reanalyse above 2 km. Wile the patterns are quite consistent between the JRA55 and MERRA2 simulation results in general, the MERRA2 cloud occurrences are higher at nearly every altitude in every node than the corresponding JRA55 values. These results likely represent differences between the underlying cloud parameterisations in the different reanalyses.

Finally, we apply a machine learning scheme developed for the classification of cloud phase from attenuated volume backscatter coefficient data. This scheme has been developed and validated previously for polar conditions as discussed in detail in Guyot et al. (2022). While we can not validate this algorithm at Scott Base because of a lack of polarisation data, visual inspection of attenuated volume backscatter coefficient data and cloud classifications appears to confirm that this scheme works well (see Figure 7), though may provide a conservative estimate of super-cooled liquid cloud. Classification of the climatological attenuated volume backscatter coefficient data from the CL51 observations at Scott Base allows the mean occurrence of super-cooled liquid water cloud to be derived. The super-cooled liquid water cloud fraction remains relatively constant between 0.5 to 2.5 km at an occurrence rate of 5% and rapidly declines above that level. This pattern matches with vertical profile identified in Silber et al. (2018), though the cloud occurrence is again lower. This suggests that these relatively inexpensive vertically pointing lidars which can be left unattended for long periods can be a valuable source of data on cloud properties in the Antarctic environment which complements satellite observations. Application of a simple classification of reanalyses output, see details in (Desai et al., 2023), shows that ERA5 and JRA55 appear to significantly underestimate liquid-water cloud and mixed phase cloud relative to the values derived from the Cl51 observations. While liquid water and mixed phase cloud makes up the majority of the cloud observed in the MERRA2 reanalysis below 5 km, possibly explaining the large difference between the raw cloud occurrence and the cloud occurrence derived from the instrument simulator for this reanalyses.

In summary, our results highlight that the vertical profile of cloud occurrence for all three reanalyses shows significant underestimation below 3km and a smaller overestimation above that altitude relative to the CL51 observations. The low-level biases are largest for the JRA55 reanalysis in terms of cloud occurrence and cloud phase. The MERRA2 reanalysis displays the largest cloud occurrence biases at higher altitudes relative to the CL51 observations and appears to overestimate the proportion of super-cooled liquid and mixed phase cloud at low levels. The larger bias at higher altitudes likely offsets the low-level cloud occurrence biases in MERRA2 when cloud fraction is examined. Finally, the ERA5 cloud occurrence is significantly under-estimated relative to the ceilometer observations at low-levels, but displays small biases elsewhere. In particular, the ERA5 reanalysis displays an improved representation of cloud occurrence when data is grouped based on synoptic state relative to the other two reanalyses.

Further work will apply the machine learning scheme detailed in Guyot et al. (2022) to a set of ceilometer observations made across the Antarctic continent. This will provide a set of surface observations distributed over a wide geographic region for comparison with satellite observations and reanalyses for the first time. It will also allow us to determine whether Scott Base can be considered to be a representative site for further cloud property analyses.

5 Open Research

The ERA5 reanalyses data used in this study are available for download from the Climate Data Store at `https://doi.org/10.24381/cds.143582cf`. The JRA-55: Japanese 55-year Reanalysis 3-hourly data is available from the Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory at `https://doi.org/10.5065/D6HH6H41`. The MERRA2 data is available for downalod from the GES-DISC download site at `https://doi.org//10.5067/WWQSXQ8IVFW8`. The AMPS archive data used in this study can be downloaded from `https://www.earthsystemgrid.org/dataset/ucar.mmm.amps.html`. The McMurdo Station Radiosonde Observations are available from `https://doi.org/10.48567/ka0n-n046`.

All of the University of Canterbury ceilometer data proccesed using ALCF and the output from the ALCF lidar simulator derived from the various model archives (AMPS, ERA5, JRA55 and MERRA2) used in this study are accessible at Zenodo, along with code for creating all figures (`https://doi.org/10.5281/zenodo.11458722`, McDonald and Plank, 2024). The Automatic Lidar Ceilometer Framework software package is available at `https://doi.org/10.5281/zenodo.3764287` (Kuma et al., 2021).

Acknowledgments

AJM acknowledges the support provided by the Deep South National Science Challenge (Grant C01X1901). PK acknowledges support by the NextGEMS project funded by the European Union's Horizon 2020 research and innovation program (Grant 101003470). We also acknowledge the JRA55 data set used in this study which is provided by the Japanese 55-year Reanalysis project carried out by the Japan Meteorological Agency (JMA). We would also like to acknowledge the ERA5 reanalysis provided by the European Centre for Medium Range Weather Forecasting. We also recognise the efforts of MDISC, managed by the NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC), who provide access to the MERRA2 data set. We would also like to thank Lee Welhouse from the Antarctic Meteorological Research and Data Center (AMRDC) who provided us with access to the McMurdo station radiosonde data used in this paper.

References

Adhikari, L., Wang, Z. E., & Deng, M. (2012). Seasonal variations of Antarctic clouds observed by CloudSat and CALIPSO satellites. *Journal of Geophysical Research-Atmospheres*, *117*, 17. doi: 10.1029/2011jd016719

Alexander, S. P., & Protat, A. (2018). Cloud Properties Observed From the Surface and by Satellite at the Northern Edge of the Southern Ocean. *Journal of Geophysical Research-Atmospheres*, *123*(1), 443-456. doi: 10.1002/2017jd026552

Bodas-Salcedo, A., Andrews, T., Karmalkar, A. V., & Ringer, M. A. (2016). Cloud liquid water path and radiative feedbacks over the Southern Ocean. *Geophysical Research Letters*, *43*(20), 10938-10946. doi: 10.1002/2016gl070770

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., ... John, V. O. (2011). COSP Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, *92*(8), 1023-1043. doi: 10.1175/2011bams2856.1

Brodzik, M. J., & Stewart., J. S. (2016). *Near-Real-Time SSM/I-SSMIS EASE-Grid Daily Global Ice Concentration and Snow Extent, Version 5.* NASA National Snow and Ice Data Center Distributed Active Archive Center. Retrieved from `https://nsidc.org/data/NISE/versions/5` doi: 10.5067/3KB2JPLFPK3R

Bromwich, D. H., Nicolas, J. P., Hines, K. M., Kay, J. E., Key, E. L., Lazzara, M. A., ... van Lipzig, N. P. M. (2012). Tropospheric Clouds in Antarctica. *Reviews of Geophysics*, *50*, 40. doi: 10.1029/2011rg000363

Chiriaco, M., Vautard, R., Chepfer, H., Haeffelin, M., Dudhia, J., Wanherdrick, Y., ... Protat, A. (2006). The Ability of MM5 to Simulate Ice Clouds: Systematic Comparison between Simulated and Measured Fluxes and Lidar/Radar Profiles at the SIRTA Atmospheric Observatory. *Monthly Weather Review*, *134*(3), 897-918. doi: https://doi.org/10.1175/MWR3102.1

Delanoe, J., & Hogan, R. J. (2010). Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research-Atmospheres*, *115*, 17. doi: 10.1029/2009jd012346

Desai, N., Diao, M., Shi, Y., Liu, X., & Silber, I. (2023). Ship-Based Observations and Climate Model Simulations of Cloud Phase Over the Southern Ocean. *Journal of Geophysical Research: Atmospheres*, *128*(11), e2023JD038581. doi: https://doi.org/10.1029/2023JD038581

Frey, R. A., Ackerman, S. A., Liu, Y. H., Strabala, K. I., Zhang, H., Key, J. R., & Wang, X. G. (2008). Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5. *Journal of Atmospheric and Oceanic*

Technology, *25*(7), 1057-1072. doi: 10.1175/2008jtecha1052.1

Gelaro, R., McCarty, W., Suarez, M. J., Todling, R., Molod, A., Takacs, L., ... Zhao, B. (2017). The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate*, *30*(14), 5419-5454. doi: 10.1175/jcli-d-16-0758.1

Guyot, A., Protat, A., Alexander, S. P., Klekociuk, A. R., Kuma, P., & McDonald, A. (2022). Detection of supercooled liquid water containing clouds with ceilometers: Development and evaluation of deterministic and data-driven retrievals. *Atmos. Meas. Tech.*, *15*(12), 3663-3681. doi: 10.5194/amt-15-3663-2022

Haynes, J. M., Jakob, C., Rossow, W. B., Tselioudis, G., & Brown, J. (2011). Major Characteristics of Southern Ocean Cloud Regimes and Their Effects on the Energy Budget. *Journal of Climate*, *24*(19), 5061-5080. doi: 10.1175/2011jcli4052.1

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. doi: https://doi.org/10.1002/qj.3803

Hofer, S., Amory, C., Kittel, C., Carlsen, T., Le Toumelin, L., & Storelvmo, T. (2021). The Contribution of Drifting Snow to Cloud Properties and the Atmospheric Radiative Budget Over Antarctica. *Geophysical Research Letters*, *48*(22). doi: 10.1029/2021GL094967

Hogan, R. J., Behera, M. D., O'Connor, E. J., & Illingworth, A. J. (2004). Estimate of the global distribution of stratiform supercooled liquid water clouds using the lite lidar. *Geophysical Research Letters*, *31*(5).

Hopkin, E., Illingworth, A. J., Charlton-Perez, C., Westbrook, C. D., & Ballard, S. (2019). A robust automated technique for operational calibration of ceilometers using the integrated backscatter from totally attenuating liquid clouds. *Atmos. Meas. Tech.*, *12*(7), 4131-4147. doi: 10.5194/amt-12-4131-2019

Jolly, B., Kuma, P., McDonald, A., & Parsons, S. (2018). An analysis of the cloud environment over the Ross Sea and Ross Ice Shelf using CloudSat/CALIPSO satellite observations: the importance of synoptic forcing. *Atmos. Chem. Phys.*, *18*(13), 9723-9739. doi: 10.5194/acp-18-9723-2018

Kay, J. E., Bourdages, L., Miller, N. B., Morrison, A., Yettella, V., Chepfer, H., & Eaton, B. (2016). Evaluating and improving cloud phase in the Community Atmosphere Model version 5 using spaceborne lidar observations. *Journal of Geophysical Research-Atmospheres*, *121*(8), 4162-4176. doi: 10.1002/2015jd024699

Kobayashi, S., Ota, Y., Harada, Y., Ebita, A., Moriya, M., Onoda, H., ... Takahashi, K. (2015). The JRA-55 Reanalysis: General Specifications and Basic Characteristics. *Journal of the Meteorological Society of Japan*, *93*(1), 5-48. doi: 10.2151/jmsj.2015-001

Kohonen, T. (1990). The Self-Organizing Map. *Proceedings of the IEEE*, *78*(9), 1464-1480. doi: 10.1109/5.58325

Kremser, S., Harvey, M., Kuma, P., Hartery, S., Saint-Macary, A., McGregor, J., ... Parsons, S. (2021). Southern Ocean cloud and aerosol data: a compilation of measurements from the 2018 Southern Ocean Ross Sea Marine Ecosystems and Environment voyage. *Earth Syst. Sci. Data*, *13*(7), 3115-3153. doi: 10.5194/essd-13-3115-2021

Kuma, P., Bender, F. A. M., Schuddeboom, A., McDonald, A. J., & Seland, O. (2023). Machine learning of cloud types in satellite observations and climate models. *Atmos. Chem. Phys.*, *23*(1), 523-549. doi: 10.5194/acp-23-523-2023

Kuma, P., McDonald, A. J., Morgenstern, O., Alexander, S. P., Cassano, J. J., Garrett, S., ... Williams, J. (2020). Evaluation of Southern Ocean cloud in the HadGEM3 general circulation model and MERRA-2 reanalysis using ship-based observations. *Atmospheric Chemistry and Physics*, *20*(11), 6607-6630. doi: 10.5194/acp-20-6607-2020

Kuma, P., McDonald, A. J., Morgenstern, O., Querel, R., Silber, I., & Flynn, C. J. (2021). Ground-based lidar processing and simulator framework for comparing models and observations (ALCF 1.0). *Geosci. Model Dev.*, *14*(1), 43-72. doi: 10.5194/gmd-14-43-2021

Lachlan-Cope, T. (2010). Antarctic Clouds. *Polar Research*, *29*(2), 150-158. doi: 10.1111/j.1751-8369.2010.00148.x

Listowski, C., Delanoë, J., Kirchgaessner, A., Lachlan-Cope, T., & King, J. (2019). Antarctic clouds, supercooled liquid water and mixed phase, investigated with DARDAR: geographical and seasonal variations. *Atmos. Chem. Phys.*, *19*(10), 6771-6808. doi: 10.5194/acp-19-6771-2019

Liu, D., Liu, Q., Qi, L., & Fu, Y. (2016). Oceanic single-layer warm clouds missed by the Cloud Profiling Radar as inferred from MODIS and CALIOP measurements. *Journal of Geophysical Research: Atmospheres*, *121*(21), 12,947-12,965. doi: https://doi.org/10.1002/2016JD025485

Lubin, D., Zhang, D., Silber, I., Scott, R. C., Kalogeras, P., Battaglia, A., … Vogelmann, A. M. (2020). AWARE: The Atmospheric Radiation Measurement (ARM) West Antarctic Radiation Experiment. *Bulletin of the American Meteorological Society*, *101*(7), E1069-E1091. doi: 10.1175/BAMS-D-18-0278.1

Marchand, R., Mace, G. G., Ackerman, T., & Stephens, G. (2008). Hydrometeor detection using Cloudsat - An earth-orbiting 94-GHz cloud radar. *Journal of Atmospheric and Oceanic Technology*, *25*(4), 519-533. doi: 10.1175/2007jtecha1006.1

McDonald, A. J., & Cairns, L. H. (2020). A New Method to Evaluate Reanalyses Using Synoptic Patterns: An Example Application in the Ross Sea/Ross Ice Shelf Region. *Earth and Space Science*, *7*(1). doi: 10.1029/2019EA000794

McErlich, C., McDonald, A., Schuddeboom, A., & Silber, I. (2021). Comparing Satellite- and Ground-Based Observations of Cloud Occurrence Over High Southern Latitudes. *Journal of Geophysical Research: Atmospheres*, *126*(6), e2020JD033607. doi: https://doi.org/10.1029/2020JD033607

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., … McDonald, A. (2021). Observations of Clouds, Aerosols, Precipitation, and Surface Radiation over the Southern Ocean: An Overview of CAPRICORN, MARCUS, MICRE, and SOCRATES. *Bulletin of the American Meteorological Society*, *102*(4), E894-E928. doi: https://doi.org/10.1175/BAMS-D-20-0132.1

O'Connor, E. J., Illingworth, A. J., & Hogan, R. J. (2004). A technique for autocalibration of cloud lidar. *Journal of Atmospheric and Oceanic Technology*, *21*(5), 777-786. doi: 10.1175/1520-0426(2004)021⟨0777:atfaoc⟩2.0.co;2

Pei, Z., Fiddes, S. L., French, W. J. R., Alexander, S. P., Mallet, M. D., Kuma, P., & McDonald, A. (2023). Assessing the cloud radiative bias at Macquarie Island in the ACCESS-AM2 model. *Atmos. Chem. Phys.*, *23*(23), 14691-14714. doi: 10.5194/acp-23-14691-2023

Platnick, S., King, M. D., Ackerman, S. A., Menzel, W. P., Baum, B. A., Riedi, J. C., & Frey, R. A. (2003). The MODIS cloud products: Algorithms and examples from Terra. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 459-473. doi: 10.1109/tgrs.2002.808301

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*, *93*(11), 1699-1712. doi: 10.1175/bams-d-11-00186.1

Rossow, W. B., & Schiffer, R. A. (1999). Advances in understanding clouds from ISCCP. *Bulletin of the American Meteorological Society*, *80*(11), 2261-2287. doi: 10.1175/1520-0477(1999)080⟨2261:aiucfi⟩2.0.co;2

Sassen, K., Wang, Z., & Liu, D. (2008). Global distribution of cirrus clouds from CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) measurements. *Journal of Geophysical Research-Atmospheres*, *113*(D20), 12. doi: 10.1029/2008jd009972

Schuddeboom, A. J., & McDonald, A. J. (2021). The Southern Ocean Radiative Bias, Cloud Compensating Errors, and Equilibrium Climate Sensitivity in CMIP6 Models. *Journal of Geophysical Research: Atmospheres*, *126*(22), e2021JD035310. doi: https://doi.org/10.1029/2021JD035310

Scott, R. C., & Lubin, D. (2014). Mixed-phase cloud radiative properties over Ross Island, Antarctica: The influence of various synoptic-scale atmospheric circulation regimes. *Journal of Geophysical Research-Atmospheres*, *119*(11), 6702-6723. doi: 10.1002/2013jd021132

Scott, R. C., & Lubin, D. (2016). Unique manifestations of mixed-phase cloud microphysics over Ross Island and the Ross Ice Shelf, Antarctica. *Geophysical Research Letters*, *43*(6), 2936-2945. doi: 10.1002/2015gl067246

Sellegri, K., Harvey, M., Peltola, M., Saint-Macary, A., Barthelmeß, T., Rocco, M., ... Law, C. S. (2023). Sea2Cloud: From Biogenic Emission Fluxes to Cloud Properties in the Southwest Pacific. *Bulletin of the American Meteorological Society*, *104*(5), E1017-E1043. doi: https://doi.org/10.1175/BAMS-D-21-0063.1

Silber, I., Fridlind, A. M., Verlinde, J., Ackerman, A. S., Cesana, G. V., & Knopf, D. A. (2021). The prevalence of precipitation from polar supercooled clouds. *Atmos. Chem. Phys.*, *21*(5), 3949-3971. doi: 10.5194/acp-21-3949-2021

Silber, I., Verlinde, J., Eloranta, E. W., & Cadeddu, M. (2018). Antarctic Cloud Macrophysical, Thermodynamic Phase, and Atmospheric Inversion Coupling Properties at McMurdo Station: I. Principal Data Processing and Climatology. *Journal of Geophysical Research: Atmospheres*, *123*(11), 6099-6121. doi: 10.1029/2018JD028279

Stephens, G. L., Vane, D. G., Tanelli, S., Im, E., Durden, S., Rokey, M., ... Marchand, R. (2008). CloudSat mission: Performance and early science after the first year of operation. *Journal of Geophysical Research-Atmospheres*, *113*(D23), 18. doi: 10.1029/2008jd009982

Swales, D. J., Pincus, R., & Bodas-Salcedo, A. (2018). The Cloud Feedback Model Intercomparison Project Observational Simulator Package: Version 2. *Geosci. Model Dev.*, *11*(1), 77-81. doi: 10.5194/gmd-11-77-2018

Tastula, E. M., Vihma, T., Andreas, E. L., & Galperin, B. (2013). Validation of the diurnal cycles in atmospheric reanalyses over Antarctic sea ice. *Journal of Geophysical Research-Atmospheres*, *118*(10), 4194-4204. doi: 10.1002/jgrd.50336

Trenberth, K. E., & Fasullo, J. T. (2010). Simulation of Present-Day and Twenty-First-Century Energy Budgets of the Southern Oceans. *Journal of Climate*, *23*(2), 440-454. doi: https://doi.org/10.1175/2009JCLI3152.1

Vergara-Temprado, J., Miltenberger, A. K., Furtado, K., Grosvenor, D. P., Shipway, B. J., Hill, A. A., ... Carslaw, K. S. (2018). Strong control of Southern Ocean cloud reflectivity by ice-nucleating particles. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(11), 2687-2692. doi: 10.1073/pnas.1721627115

Verlinden, K. L., Thompson, D. W. J., & Stephens, G. L. (2011). The Three-Dimensional Distribution of Clouds over the Southern Hemisphere High Latitudes. *Journal of Climate*, *24*(22), 5799-5811. doi: 10.1175/2011jcli3922.1

Vettigli, G. (2018). *Minisom: minimalistic and numpy-based implementation of the self organizing map.* Retrieved from `https://github.com/JustGlowing/minisom/`

Whitehead, L. E., McDonald, A. J., & Guyot, A. (2023). Supercooled liquid water cloud classification using lidar backscatter peak properties. *EGUsphere*, *2023*, 1-30. doi: 10.5194/egusphere-2023-1085

Wiegner, M., & Gasteiger, J. (2015). Correction of water vapor absorption for aerosol remote sensing with ceilometers. *Atmos. Meas. Tech.*, *8*(9), 3971-3984. doi: 10.5194/amt-8-3971-2015

Winker, D. M., Vaughan, M. A., Omar, A., Hu, Y. X., Powell, K. A., Liu, Z. Y., ... Young, S. A. (2009). Overview of the CALIPSO Mission and CALIOP Data Processing Algorithms. *Journal of Atmospheric and Oceanic Technology*, *26*(11), 2310-2323. doi: 10.1175/2009jtecha1281.1

Yip, J., Diao, M., Barone, T., Silber, I., & Gettelman, A. (2021). Evaluation of the CAM6 Climate Model Using Cloud Observations at McMurdo Station, Antarctica. *Journal of Geophysical Research: Atmospheres*, *126*(16). doi: 10.1029/2021JD034653

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., ... Taylor, K. E. (2020). Causes of Higher Climate Sensitivity in CMIP6 Models. *Geophysical Research Letters*, *47*(1). doi: 10.1029/2019GL085782

Zhang, M., Xie, S., Liu, X., Zhang, D., Lin, W., Zhang, K., ... Zhang, Y. (2023). Evaluating EAMv2 Simulated High Latitude Clouds Using ARM Measurements in the Northern and Southern Hemispheres. *Journal of Geophysical Research: Atmospheres*, *128*(15). doi: 10.1029/2022JD038364