NAISS Activity Report 2023/22-202

Peter Kuma*, Frida A.-M. Bender and Thorsten Mauritsen

Department of Meteorology (MISU) and Bolin Centre for Climate Research, Stockholm University, Stockholm SE-106 91, Sweden

*E-mail: peter.kuma@misu.su.se, Web: https://peterkuma.net

30 May 2024

Introduction

We applied for the NAISS Small Compute allocation 2023/22-202 to support our research as part of the NextGEMS project (https://nextgems-h2020.eu), funded through the European Union's Horizon 2020 research and innovation programme under the grant agreement number 101003470. NextGEMS is focused on the research and development of high resolution (km-scale) storm-resolving climate models, which are going to be the next generation of production-ready climate models. Because of the high resolution, some previously parameterised physical processes can be resolved on the model grid. Thus, such parameterisations and related uncertainties can be eliminated.

The main aim of our research is to evaluate a high-resolution version of the Icosahedral Nonhydrostatic Weather and Climate Model (ICON), developed by NextGEMS jointly with Deutscher Wetterdienst (DWD) and the Max-Planck-Institute for Meteorology (MPI-M). Previous studies have identified substantial large-scale biases in climate model clouds over the Southern Ocean, affecting sea surface temperature and the Earth's albedo overall. Our aim is to quantify how well the high-resolution ICON model is simulating clouds in this region, particularly in light of the fact that subgrid-scale clouds are not parameterised in this model. This region is mostly dominated by boundary layer clouds generated by shallow convection, and these are problematic to observe by spaceborne lidar and radars, which are affected by attenuation by overlapping and thick clouds and ground clutter, respectively. Therefore, we chose to use a large set of ship-based observations conducted with ceilometers and lidars on board of the RV *Polarstern* and other voyages and stations.

Altogether, we analysed over 1500 days of data from 31 voyages and 1 sub-antarctic station covering diverse longitudes of the Southern Ocean. To achieve a like-for-like comparison with the model, we used a ground-based lidar simulator called the Automatic Lidar and Ceilometer Framework (ALCF) (Kuma et al., 2021). We contrasted the results with the European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis 5 (ERA5) (ECMWF, 2019) and the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Gelaro et al., 2017).

The outlined research is still ongoing and we have not finalised our analysis. Intermediate results were presented at two conferences in the form of posters: the 15th Bolin Days in Stockholm on 29–30 November 2023 (Kuma and Bender, 2023) and the 2nd Swedish Climate Symposium in Norrköping on 15–17 May 2024 (Kuma and Bender, 2024a). We also presented some of the intermediate results at the NextGEMS hackathon in Hamburg on 4–8 March 2024. A manuscript covering this work is currently in preparation and will be submitted later in 2024 (Kuma and Bender, 2024b).

We used the NAISS Small Compute allocation for compiling and test-running ICON in an atmosphere-only configuration.
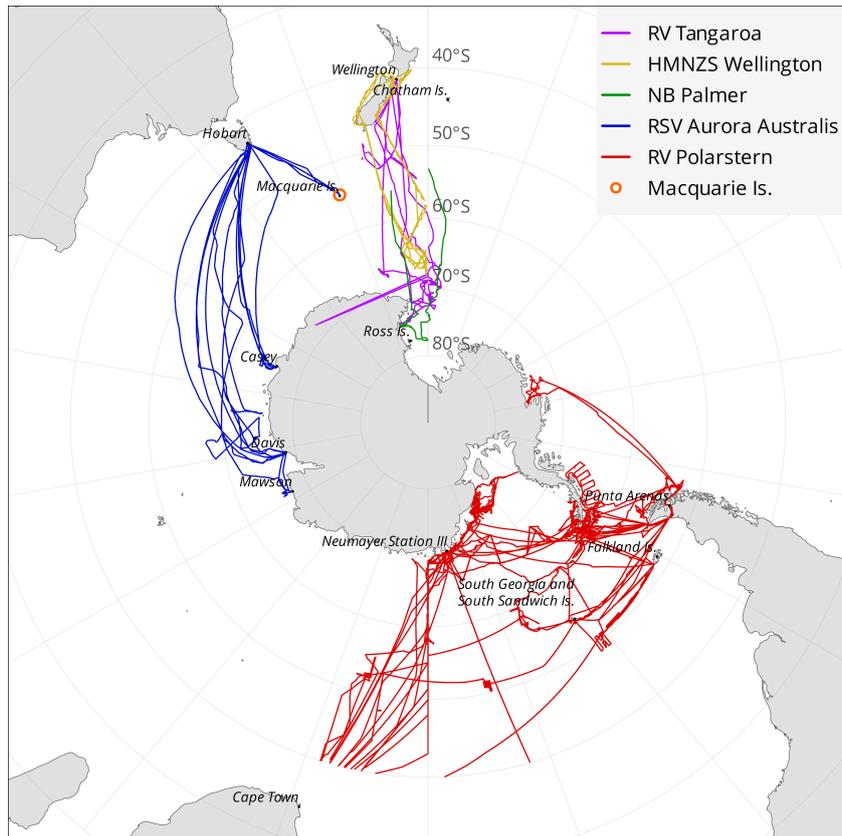
Figure 1 | Map showing voyage tracks and a station analysed.

Sections 1 and 2 briefly outline the methods used and the main scientific results. These are presented in more detail in the referenced posters (available online). The results will be described fully in the manuscript in preparation.

1 Methods

1.1 Voyage and station data

Together, we analysed data from 31 voyages of RV *Polarstern*, RSV *Aurora Australis*, RV *Tangaroa*, RV *Nathaniel B. Palmer*, HMNZS *Wellington* and one sub-Antarctic station (Macquarie Island) in the Southern Ocean south of 40°S between year 2010 and 2021 (Fig. 1).

These data sources contained ceilometer (lidar) observations with the Vaisala CL51 operating at 910 nm and the Lufft CHM-15k operating at 1064 nm. A ceilometer is a low-power vertically pointing lidar with the purpose of measuring the cloud base, but it also measures the full vertical structure of clouds as long as the laser signal is not attenuated by thick clouds. A total of about 1500 days of lidar observations were included in our analysis.

Apart from lidar observations, radiosondes were launched on weather balloons at regular synoptic times on the RV *Polarstern* voyages, and surface meteorological quantities were measured continuously.

1.2 ICON

The storm-resolving version of the ICON model is in development by the NextGEMS project (Hohenegger et al., 2023). NextGEMS has so far produced four cycles of model runs. In our

research, we used the Cycle 3 run due to limited data availability in Cycle 4. The horizontal resolution is about 5 km. Unlike current general circulation models (GCMs), it does not use convective and cloud parameterisation but relies on explicit simulation of convection and clouds on the model grid. While this makes the code development simpler without having to rely on uncertain parameterisations, it can miss smaller-scale clouds below the grid resolution. Turbulence and cloud microphysics are still parameterised in this model. In our analysis, we used 4 years of coupled simulation output between years 2021 and 2024 (inclusive). Because the model is free-running, weather and climate oscillations are not equivalent to reality at the same time and place. To compare with observations taken during different years (2010–2021), we compared the model output with observations at the same time of year and geographical location.

1.3 ALCF

The Automatic Lidar and Ceilometer Framework (ALCF) is a ground-based lidar simulator. It performs radiative transfer calculations to derive equivalent lidar backscattered radiation in an atmospheric model, which can then be compared with observed lidar backscatter (Kuma et al., 2021). For this purpose, it takes atmospheric fields of cloud fraction, liquid and ice mass mixing ratio, temperature and pressure fields as an input and can be run offline (on the model output rather than inside the model code). The lidar simulator in the ALCF is based on the instrument simulator Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP) (Bodas-Salcedo et al., 2011). After the backscattered radiation is calculated, a cloud mask, cloud occurrence by height and the total cloud fraction are determined.

1.4 Precipitation identification using machine learning

It was not possible to include precipitation simulation in the ALCF due to the absence of required fields in our model output and reanalysis data (liquid and ice precipitation mass mixing ratios) because the required radiation calculations are not implemented in the ALCF. At the same time, it is relatively difficult to distinguish precipitation backscatter from cloud backscatter in lidar observations, especially when only one wavelength channel and no polarisation channel are available.

In order to get a fair comparison of observations and models, one needs to exclude profiles with precipitation either manually or using an automated method. In models, this can be accomplished relatively easily by excluding profiles with an amount of surface precipitation flux exceeding a threshold. In observations, this can be more complicated, especially on voyages, because rain gauges are known to be unreliable on ships.

On the RV *Polarstern* voyages, regular synoptic observations were available. These included precipitation observations by an observer. We used this dataset to train a convolutional artificial neural network (ANN) of the U-Net type (Ronneberger et al., 2015) to recognise precipitation in the lidar data. Samples of short time intervals (10 min) of near-surface backscatter (0–250 m) were classified as clear, rain, snow and fog, using the synoptic observations as a training set. From these, a binary classification of profiles as wet or dry was derived.

The ANN achieved 65% sensitivity and 87% specificity when the true positive rate (26%) was made to match observations. We considered these success rates satisfactory for the purpose of filtering precipitation profiles.

2 Results

2.1 Cloud occurrence

We used the ALCF to derive cloud occurrence by height and the total cloud fraction from observations, ICON and two reanalyses MERRA-2 and ERA5 (Fig. 2). We aggregated the sources by calculating the averages and percentiles of all individual profiles (Fig. 3). The analysis shows that the total cloud fraction (determined as the fraction of profiles with clouds at any height in the lidar cloud mask) is underestimated in ICON and reanalyses by about 10% and 20%, respectively.

In particular, ICON overestimates cloud occurrence below 1 km and underestimates it above, MERRA-2 underestimates cloud occurrence at all heights, especially near the surface, and ERA5 simulates cloud occurrence relatively well above 1 km, but strongly underestimates it near the surface. We note that fog or near-surface clouds are strongly lacking in the reanalyses. As shown in Fig. 2, the biases are relatively consistent across voyages and longitudes. The ICON results are overall better matching the observations than the reanalyses.

The ICON comparison is limited by the fact that the model is free-running. Thus, the vertical profiles are not expected to represent the same weather conditions as in observations, but long-term statistical comparison is still possible. Only profiles with the same sea ice conditions (present or not) were included to avoid comparing sea ice with open sea conditions.

2.2 Thermodynamic profiles

We compared about 2000 radiosonde profiles from 24 RV *Polarstern* voyages between the observations and the model. Profiles in the model were taken at the same geographical location and time relative to the start of the year. Only profiles for which the sea ice conditions (sea ice present or absent) are the same in the observations and the model were included. We derived potential temperature profiles, lifting condensation level (LCL), and aggregated the profiles by sea ice conditions (Fig. 4).

We found that the variability of potential temperature in the model to be smaller than in the observations. This indicates that the model does not represent the entire natural variability. The LCL peaks at the surface in the observations, but the peak in the model is higher (about 200 m). This probably relates to the greater occurrence of fog and the peak of cloud occurrence at the surface in observations, whereas in the model the peak is higher.

3 Code developments

3.1 ALCF

As part of this research, many code improvements were made to the open source software ALCF. The following non-exhaustive list highlights some of the improvements:

- adding support for the ICON model
- adding support for the Atmospheric Radiation Measurement (ARM) format of Vaisala CL51 data
- adding support for the Antarctic Mesoscale Prediction System (AMPS) GRIB format
- implementing faster vertical interpolation

Southern Ocean cloud ALC comparison
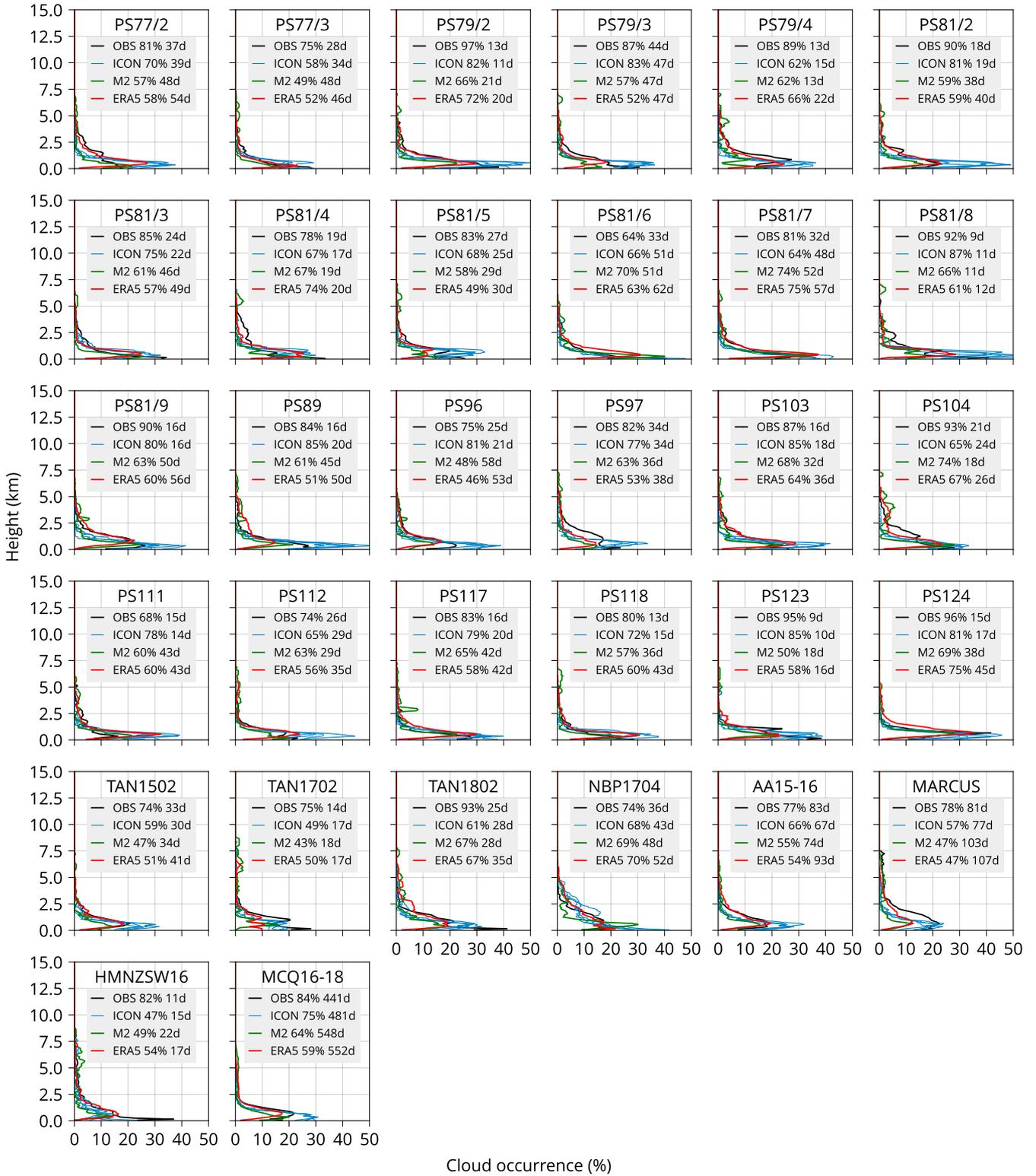
Figure 2 | Cloud occurrence by height for 32 voyages and stations in observations (OBS), ICON, MERRA-2 (M2) and ERA5. Total cloud fraction is shown in the legend.

5
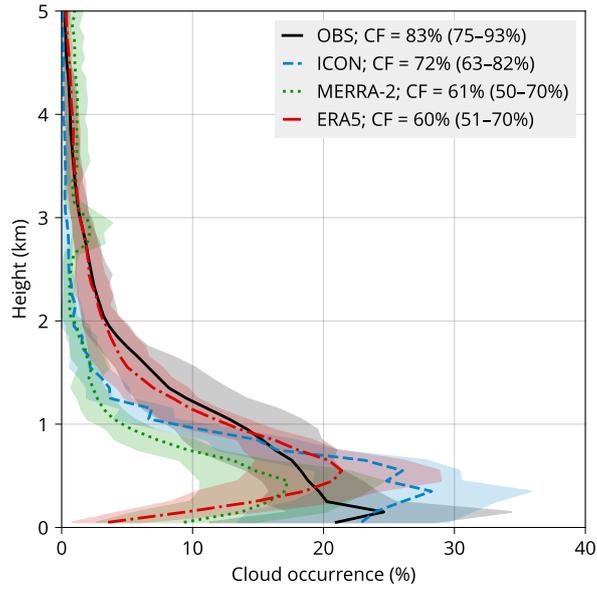
Figure 3 | Cloud occurrence by height aggregated for all voyages and stations. The total cloud fraction is shown in the legend. The ranges are from the 16th to the 84th percentile.
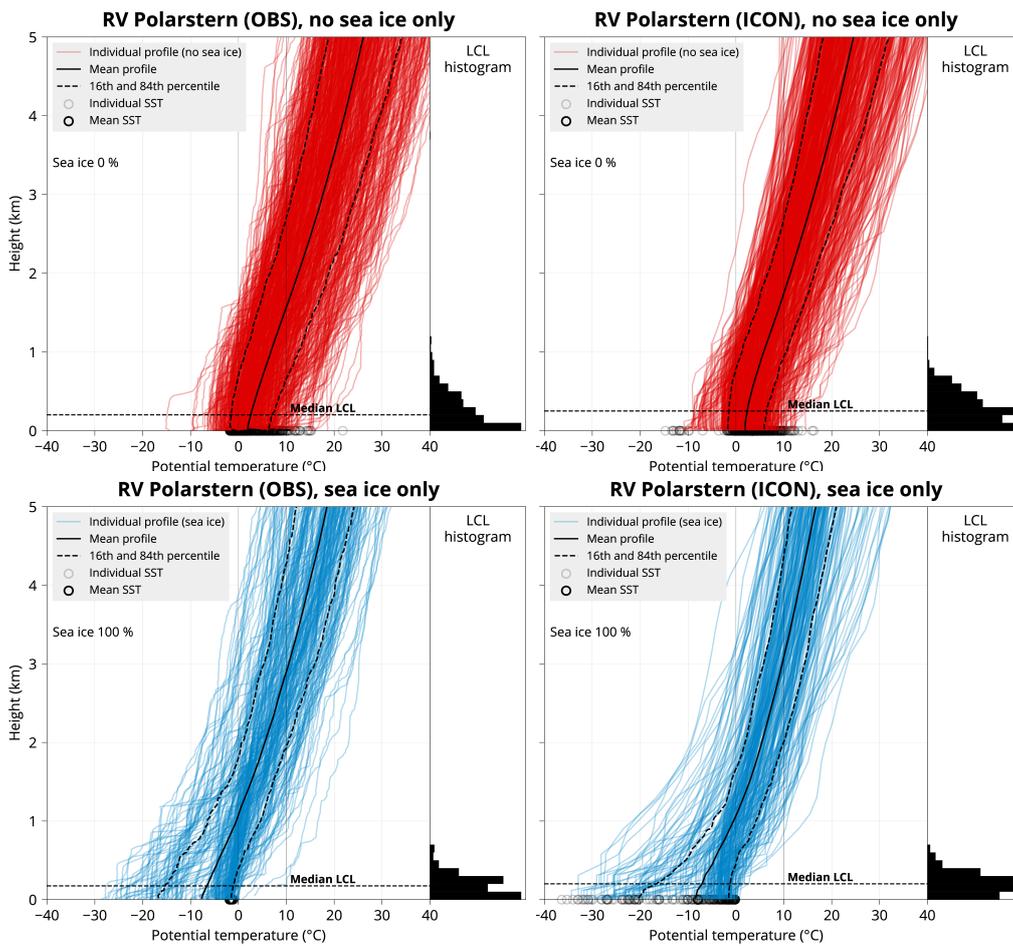
Figure 4 | Radiosonde profiles of potential temperature and lifting condensation level (LCL) on all RV *Polarstern* voyages.

- implementing new vertical interpolation algorithms
- parallel processing of model input
- automatic downloading of MERRA-2 and ERA5 data for a specified station or voyage track
- various improvements in handling of voyage tracks and time intervals
- various improvements in the documentation

The improvements are listed in more detail at https://alcf.peterkuma.net/installation/#releases.

3.2   Data processing and comparison code and the ANN

A substantial amount of new code has been developed for voyage and station data processing and comparison with the models, as well as the ANN described in Section 1.4. All of the developed code will be published under an open source license together with our upcoming manuscript.

Conclusions

We show that the model underestimates the total cloud fraction by about 10%, with an overestimation of clouds below 2 km, and an underestimation of clouds above 2 km. The reanalyses also underestimate the total cloud fraction by about 20%. ERA5 overestimates cloud below 1 km but underestimates near-surface cloud or fog. In addition to lidar data, we compare radiosonde profiles acquired on the RV *Polarstern* voyages with ICON. Notably, the model exhibits smaller natural variability than observations, and its lifting condensation level tends to be higher. This might explain why cloud occurrence is peaking higher in the model (at 500 m) than in observations (at the surface).

We compared radiosonde profiles acquired on the RV *Polarstern* voyages with ICON. The ICON model exhibits smaller internal variability than observations, and its LCL level tends to be higher, which might explain why a peak of cloud occurrence in observations is located at the surface, while in ICON it is located higher, corresponding to the LCL.

The results imply that Southern Ocean cloud biases are still a significant issue in a km-scale resolution model, even though an improvement over the lower-resolution reanalyses is notable. More effort is needed to improve model cloud simulations in this fast-changing and understudied region. The advancement from convection and cloud parameterisation to cloud-resolving models might not solve this bias without additional effort.

Further analysis will focus on investigating the role of cyclones in the identified biases by using cyclone tracking and the role of local thermodynamic stability using lower tropospheric stability and estimated inversion strength.

As part of our research we also continued the development of the open-source ground-based lidar simulator ALCF, which has already been used by several research teams for model and reanalysis evaluation (Kuma et al., 2020; Kremser et al., 2021; Guyot et al., 2022; Pei et al., 2023; Whitehead et al., 2023).

References

A. Bodas-Salcedo, M. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. Klein, Y. Zhang, R. Marchand, J. Haynes, R. Pincus, et al. COSP: Satellite simulation software for model assessment. *Bulletin of the American Meteorological Society*, 92(8):1023–1043, 2011. doi:10.1175/2011BAMS2856.1.

ECMWF. Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), 11 2019. URL https://cds.climate.copernicus.eu/cdsapp#!/home.

R. Gelaro, W. McCarty, M. J. Suárez, R. Todling, A. Molod, L. Takacs, C. A. Randles, A. Darmenov, M. G. Bosilovich, R. Reichle, et al. The modern-era retrospective analysis for research and applications, version 2 (MERRA-2). *Journal of Climate*, 30(14):5419–5454, 2017. doi:10.1175/JCLI-D-16-0758.1.

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

C. Hohenegger, P. Korn, L. Linardakis, R. Redler, R. Schnur, P. Adamidis, J. Bao, S. Bastin, M. Behravesh, M. Bergemann, J. Biercamp, H. Bockelmann, R. Brokopf, N. Brüggemann, L. Casaroli, F. Chegini, G. Datseris, M. Esch, G. George, M. Giorgetta, O. Gutjahr, H. Haak, M. Hanke, T. Ilyina, T. Jahns, J. Jungclaus, M. Kern, D. Klocke, L. Kluft, T. Kölling, L. Kornblueh, S. Kosukhin, C. Kroll, J. Lee, T. Mauritsen, C. Mehlmann, T. Mieslinger, A. K. Naumann, L. Paccini, A. Peinado, D. S. Praturi, D. Putrasahan, S. Rast, T. Riddick, N. Roeber, H. Schmidt, U. Schulzweida, F. Schütte, H. Segura, R. Shevchenko, V. Singh, M. Specht, C. C. Stephan, J.-S. von Storch, R. Vogel, C. Wengel, M. Winkler, F. Ziemen, J. Marotzke, and B. Stevens. ICON-Sapphire: simulating the components of the Earth system and their interactions at kilometer and subkilometer scales. *Geoscientific Model Development*, 16(2):779–811, 2023. doi:10.5194/gmd-16-779-2023.

S. Kremser, M. Harvey, P. Kuma, S. Hartery, A. Saint-Macary, J. McGregor, A. Schuddeboom, M. von Hobe, S. T. Lennartz, A. Geddes, R. Querel, A. McDonald, M. Peltola, K. Sellegri, I. Silber, C. S. Law, C. J. Flynn, A. Marriner, T. C. J. Hill, P. J. DeMott, C. C. Hume, G. Plank, G. Graham, and S. Parsons. Southern Ocean cloud and aerosol data: a compilation of measurements from the 2018 Southern Ocean Ross Sea Marine Ecosystems and Environment voyage. *Earth System Science Data*, 13(7):3115–3153, 2021. doi:10.5194/essd-13-3115-2021.

P. Kuma and F. A.-M. Bender. Using ship observations to assess Southern Ocean clouds in a storm-resolving general circulation model ICON, 2023. URL https://doi.org/10.5281/zenodo.10956026. The 15th Bolin Days, Stockholm, Sweden, 29–30 November 2023.

P. Kuma and F. A.-M. Bender. Using ship observations to assess Southern Ocean clouds in a storm-resolving general circulation model ICON, 2024a. URL https://doi.org/10.5281/zenodo.11204414. Swedish Climate Symposium, Norrköping, Sweden, 15–17 May 2024.

P. Kuma and F. A.-M. Bender. Using ship observations to assess Southern Ocean clouds in a storm-resolving general circulation model ICON. 2024b. Manuscript in preparation.

P. Kuma, A. J. McDonald, O. Morgenstern, S. P. Alexander, J. J. Cassano, S. Garrett, J. Halla, S. Hartery, M. J. Harvey, S. Parsons, G. Plank, V. Varma, and J. Williams. 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, 2020. doi:10.5194/acp-20-6607-2020.

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

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

O. Ronneberger, P. Fischer, and T. Brox. U-Net: Convolutional networks for biomedical image segmentation, 2015.

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