Southern Ocean cloud and shortwave radiation biases in a nudged climate model simulation: does the model ever get it right?

Sonya L. Fiddes[1,2], Alain Protat[3,1], Marc D. Mallet[1], Simon P. Alexander[4,1], and Matthew T. Woodhouse[2,1]

[1]Australian Antarctic Program Partnership, Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Australia
[2]Climate Science Centre, Oceans and Atmosphere, Commonwealth Scientific and Industrial Research Organisation, Aspendale, Australia
[3]Bureau of Meteorology, Melbourne, Australia
[4]Australian Antarctic Division, Hobart, Australia

Correspondence: Sonya Fiddes (sonya.fiddes@utas.edu.au)

Abstract.

The Southern Ocean radiative bias continues to impact climate and weather models, including the Australian Community Climate and Earth System Simulator (ACCESS). The radiative bias, characterised by too much shortwave radiation reaching the surface, is attributed to the incorrect simulation of cloud properties, including frequency and phase. ~~In this work~~To identify cloud regimes important to the Southern Ocean, we use k-means cloud ~~clustering, combined with~~ histogram clustering, applied to a satellite product and then fitted to nudged simulations of the latest generation ACCESS atmosphere model~~, to evaluate~~. We identify instances when the model correctly or incorrectly simulates the same cloud type as the satellite product for any point in time or space. We then evaluate the cloud and radiation biases ~~when cloud types are correctly and incorrectly simulated~~in these instances.

We find that ~~even if~~ when the ACCESS model correctly simulates the cloud type, cloud property and radiation biases of equivalent, or in some cases greater, magnitude ~~then when they~~ remain compared to when cloud types are incorrectly simulated~~remain in the cloud and radiation fields examined~~. Furthermore, we find that even when radiative biases appear small on average, cloud property biases, such as liquid or ice water paths or cloud fractions remain large. Our results suggest that simply getting the right cloud type (or the cloud macrophysics) is not enough to reduce the Southern Ocean radiative bias. Furthermore, in instances where the radiative bias is small, it may be so for the wrong reasons. Considerable effort is still required to improve cloud microphysics, with a particular focus on cloud phase.

Key Figure

(used for advertising the paper)

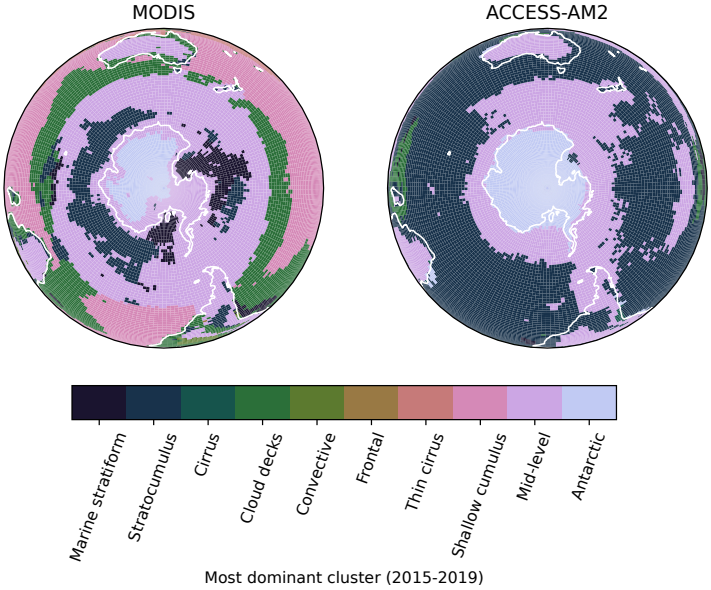

Figure 0. The most dominant cloud regime for each gridbox for the MODIS product (left) and the ACCESS-AM2 model (right)

1 Introduction

Global climate models, including coupled, Earth system and atmosphere-only models, have presented a significant radiative ~~biases~~ bias over the Southern Ocean (SO) for a number of generations, as documented by the Coupled Model Intercomparison Project (CMIP) phases 3, 5 and 6 (Bodas-Salcedo et al., 2014; Schuddeboom and McDonald, 2021). ~~Such biases are also~~ This bias has also been found in numerical weather prediction (NWP) models (Protat et al., 2017; McFarquhar et al., 2021). ~~These biases have~~ The radiative bias largely been found to be a result of poor simulation ~~by models~~ of cloud properties within models, in particular, the cloud phase in the cold sector of extra-tropical cyclones (Haynes et al., 2011; Bodas-Salcedo et al., 2016b). Over the SO, clouds are ubiquitously made up of supercooled liquid water droplets, setting them apart from clouds on other parts of the Earth (Huang et al., 2012; Chubb et al., 2013; Mace and Protat, 2018; Listowski et al., 2019). This distinction creates a key problem for the simulation of SO clouds given that the majority of cloud and aerosol microphysical parameterisations have been developed from observations collected in regions of the world where supercooled liquid water droplets do not occur frequently. The resulting positive surface radiative bias caused by the poor simulation of cloud characteristics has implications for numerous aspects of the global and regional climate system, including on the global energy budget (Haynes et al., 2011; Frey et al., 2017; Schuddeboom and McDonald, 2021), the strength and positioning of the Intertropical Convergence Zones (ITCZ) (eg. Hwang and Frierson (2013), debated in Hawcroft et al. (2017)) and the jet streams (Kay et al., 2016; Hawcroft et al., 2017) as well as SO sea surface temperatures (SSTs) (Hyder et al., 2018).

Within the cloud-radiation bias literature, model evaluation techniques now commonly involve both the use of an observational simulator (satellite or ground based) to ensure that model and observed products are comparable (eg. Bodas-Salcedo et al., 2011; Kuma et al., 2020), and a method to characterise cloud types and/or the associated synoptic conditions (eg. Williams and Webb, 2009; Field and Wood, 2007). Many of these cloud characterisation techniques have focused on a clustering-based approach using cloud histogram products from the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (COSP), utilising various satellites and models. Examples of both k-means clustering (Jakob and Tselioudis, 2003; Williams and Tselioudis, 2007; Williams and Webb, 2009; Tselioudis et al., 2013; Oreopoulos et al., 2014; Mason et al., 2015; Oreopoulos et al., 2016; Leinonen et al., 2016; Tselioudis et al., 2021; Cho et al., 2021) and self-organising maps (SOMs) (McDonald et al., 2016; Schuddeboom et al., 2018) are found in the literature, often presenting comparable cloud regime structures.

For example, Tselioudis et al. (2021), in analysing the CMIP6 model suite, find an improvement from CMIP5 in simulating global mean 'weather state' distributions (i.e. cloud regimes derived via k-means clustering from the updated International Satellite Cloud Climatology Project - ISCCP - cloud histograms). Shallow-cumulus clouds are found to be consistently underestimated by ~~the ensemble~~a selection of CMIP6 models, while for the other weather states described, a larger model spread is found (Tselioudis et al., 2021). However, they note, as does Schuddeboom et al. (2018), that averaging masks important regional biases, such as those over the SO. Schuddeboom and McDonald (2021) in fact find important differences in cluster frequency biases for the SO compared to the global values in CMIP6 models. Of particular interest, they find that stratocumulus clouds are occurring too frequently over the SO in the CMIP6 models, but compensate for the associated positive radiative

bias by not being bright enough, which is opposite to CMIP5 findings of too few and too bright clouds. Schuddeboom and McDonald (2021) also note that the models with the largest compensating errors tend to have the lowest radiation biases over the SO, indicating that updated parameterisations and tuning is causing the 'right' answer for the wrong reasons. This finding is particularly concerning, given the effort to reduce the SO radiative bias so far. Similar conclusions have been made by Gettelman et al. (2020). Additionally, Zelinka et al. (2020) find that CMIP6 generation models have a higher climate sensitivity, in part due to changes in how SO cloud microphysics are now treated.

Other studies have developed regimes dependant on cloud dynamic or thermodynamic properties, such as cyclone compositing techniques (Field and Wood, 2007) or mid-tropospheric large-scale vertical motion (Bony and Dufresne, 2005). Importantly, for many subsequent studies, the evaluation methods described in such papers as listed above are used to understand biases beyond that of the cloud histograms, (eg. radiative biases or super-cooled liquid water biases). For example, the Bodas-Salcedo et al. (2014) work relies on the methods presented by both Williams and Webb (2009) and Field and Wood (2007). ~~This workflow further demonstrates a need for consistent evaluation techniques across studies to help track and understand changes in model parameterisations.~~

~~Many~~ The majority of the aforementioned cloud-regime studies have compared ~~climate-scale runs~~free-running simulations, such as those performed for the CMIP experiments~~. This method limits~~ , including the Atmospheric Model Intercomparison Project (AMIP) where sea surface temperatures and sea ice concentrations are prescribed. However, with free-running simulations, the depth of analysis ~~applicable~~ is limited as the synoptic scale meteorology cannot be considered the same~~, despite often using prescribed SSTs and sea ice concentrations. To make up for this, many studies rely on~~ . Some studies have used synoptic compositing to alleviate this issue, where certain synoptic situations can be compared like-for-like, and location and timing is then considered irrelevant. However, these studies are often limited to one synoptic type, such as cyclone centers, ignoring a number of other synoptic situations relevant for the SO, as well as any compensating errors that may exist. Additionally, focusing on just one synoptic condition follows a pre-conceived idea of the error, which may or may not hold true for newer model generations. Alternatively, many studies use a decomposition of the radiative bias into three components: a term that quantifies the bias in the frequency of occurrence of cloud clusters, a term for the bias in the radiative balance itself, and a second order co-variation (Williams and Tselioudis, 2007; Williams and Webb, 2009; Mason et al., 2015; Schuddeboom et al., 2018; Schuddeboom and McDonald, 2021). This decomposition only considers the overall means or frequencies of occurrence, and so day-for-day comparison is not required. With this decomposition, both Mason et al. (2015) and Schuddeboom et al. (2018) found that the largest proportion of the radiative bias was explainable by biases in the cloud frequency of occurrence. This has led Schuddeboom et al. (2018) to further speculate that clouds themselves may be 'simulated well', but their distributions are wrong, leading to large errors in radiative biases. However ~~, due to the limiting nature of the model set-up,~~ this has not been explored further.

In this paper, we present an in-depth evaluation of the SO radiative bias of the Australian Community Climate and Earth System Simulator (ACCESS) - Atmosphere-only Model Version 2 (AM2). Importantly, we run a nudged simulation, so the model can be compared to satellite products directly in time and space. We then use cloud regime clustering to understand the underlying reasons for the observed cloud and radiative biases~~both for when clouds are correctly and incorrectly simulated.~~

. By using a nudged simulation, we are able to composite and evaluate days and locations when cloud regimes are correctly identified by the model (ie. are the same as what was seen by the satellite), as well as instances when the model incorrectly simulates the cloud regime. We aim to answer the following question: if the model simulates the correct cloud structure for the time and place, is the radiative bias improved? This work provides the essential first step towards a long-term goal of improving the ACCESS model (and possibly the Unified Model family) and the representation of clouds over the SO. The methods outlined in this work will be used to inform and further evaluate any developments made to the model to ensure that any resulting changes in the model are understood.

2 Data and methods

2.1 ACCESS-AM2 model setup

The ACCESS-AM2 model is an atmosphere-only configuration of the ACCESS-CM2 coupled climate model and is described in Bodman et al. (2020). The atmospheric component of ACCESS is the Unified Model (UM) at vn10.6, GA7.1, which is fully described in Walters et al. (2019). In brief, the radiation scheme used in ACCESS-AM2 is the Suite Of Community RAdiative Transfer codes based on Edwards and Slingo (SOCRATES) ~~Edwards and Slingo (1996)~~(Edwards and Slingo, 1996) , and is called hourly. The prognostic cloud fraction and condensate (PC2) cloud scheme is used (Wilson et al., 2008), which includes large-scale and convective clouds. The convective scheme, including downdraughts and momentum transport, is based on Gregory and Rowntree (1990). Further details on ACCESS-CM2 and GA7.1 configurations can be found in Bi et al. (2020) and Walters et al. (2019). In the atmosphere-only configuration, we prescribe sea surface temperature (SST) and sea ice concentration (SIC) following the CMIP6 AMIP ~~(Atmospheric Model Intercomparison Project)~~ protocol (Eyring et al., 2016). The input fields can be found at the input4MIPs webpage (https://esgf-node.llnl.gov/projects/input4mips/) and the data method is described in Hurrell et al. (2008). Solar forcing, greenhouse gases, volcanic aerosol optical depth, aerosol chemistry emissions and ozone are also prescribed according to CMIP6 (Eyring et al., 2016) and available at the input4MIPs site.

ACCESS-AM2 includes the GLOMAP-mode (GLObal Model of Aerosol Processes) aerosol microphysical scheme (Mann et al., 2010, 2012), including parameterised sulfur chemistry driven by prescribed oxidants. GLOMAP-mode is a two-moment, pseudo-modal microphysical aerosol scheme, representing four soluble (nucleation, Aitken, accumulation and coarse) and one insoluble (Aitken) aerosol modes. Each mode is internally mixed. GLOMAP uses a scaled ($1.7\times$, as recommended in Mulcahy et al., 2018) DMS surface water monthly climatology (Lana et al., 2011) combined with the Liss and Merlivat (1986) flux parameterisation (also see Fiddes et al., 2018). Sea salt emissions are calculated online and occur following the Gong (2003) parameterisation. Dust is treated separately according to Woodward (2001), using a six bin scheme.

Historical (pre-2014) anthropogenic aerosol emissions are provided by the Community Emissions Data System (Hoesly et al., 2018) and biomass burning by the Global Fire Emissions Database with small fires (GFED4s) (Van Marle et al., 2017). Post 2014, the shared socioeconomic pathway (SSP) 2-4.5, a 'middle of the road emissions pathway' (Fricko et al., 2017) is used, with emissions developed by the Integrated Assessment Models Consortium and described in Feng et al. (2020).

The land-surface scheme in ACCESS-AM2 is the Community Atmosphere Biosphere Land Exchange (CABLE) version 2.5 land surface model (also described in Bi et al., 2020). ACCESS-AM2 runs at a 1.25x1.875 degree horizontal resolution with 85 vertical levels and uses the 'ENDGame' dynamical core (Wood et al., 2014). Further information on the general model setup and preliminary model evaluation against standard climate fields, such as surface air temperatures, sea surface temperature, rainfall, mean sea level pressure (MSLP) and precipitation can be found in Bi et al. (2020) and Bodman et al. (2020).

The ~~simulations~~ simulation presented in this study ~~have~~ has been nudged to the European Centre for Medium-range Weather Forecasting (ECMWF) Reanalysis 5 (ERA5) product (Hersbach et al., 2020). Nudging occurs at every dynamical time step from reanalysis fields that are updated every three hours for the horizontal ~~wind~~ winds and temperature in the free troposphere and stratosphere. In this study, output is created at monthly and daily mean resolution for the years 2015-2019 (chosen to overlap with recent Southern Ocean field campaigns described in McFarquhar et al. (2021)).

2.2 Observational products

Simulated cloud fields are evaluated against the Moderate Resolution Imaging Spectroradiometer (MODIS) Combined Aqua/Terra, Level 3 daily, 1x1 degree grid, Collection 6.1, COSP product (MCD06COSP_D3_MODIS) (Platnick et al., 2017). This product has been derived specifically for use in model evaluation using the COSP outputs for CMIP6 and is described in Hubanks et al. (2020). Available properties include cloud optical depth (τ) for total, ice and liquid clouds, cloud top pressure (CTP), cloud mask fraction (derived from pixel-level cloudiness assessments) for total, low, mid and high clouds, cloud retrieval fraction (derived from the successful retrieval of τ) for total, ice and liquid clouds, cloud effective particle radius (R_{eff}) for ice and liquid clouds, liquid water path (LWP), ice water path (IWP) and the joint histogram for CTP and τ for cloudy and partly cloud pixels. All cloud properties in the MCD06COSP_D3_MODIS product are for day-time only scenes. Full descriptions of these cloud properties can be found in Pincus et al. (2012); Platnick et al. (2017) and Hubanks et al. (2020).

Collection 6 now includes partly cloudy scenes ~~,~~ that represent heterogeneous broken cloudy or cloud edge pixels (Platnick et al., 2017). Successfully retrieved cloudy pixels are considered to be high quality, where as partly cloudy scenes have a higher rate of retrieval failure of 34%, and ~~a slighlty~~ are slightly less robust (Platnick et al., 2017). Collection 6 has shown a marked improvement upon Collection 5 in part due to the re-writing of the cloud optical property retrievals, resulting in an increase in cloud phase classification of 10%, and a 90% agreement in total cloud phases between MODIS and the CALIOP (Cloud Aerosol Lidar With Orthogonal Polarization) retrievals for cloudy pixels (Platnick et al., 2017; Marchant et al., 2016). The largest improvement in cloud phase detection was found for opaque clouds over ice or snow covered areas, whilst detection for thin cloud retrievals over warm or bright surfaces remain an issue. Improved optical retrievals have also reduced the biases in the R_{eff}, however, it has been noted that much of the evaluation performed has been for single layered clouds, with multi-layered clouds remaining un-assessed (Marchant et al., 2016).

Comparisons of MODIS retrievals to aircraft field campaigns over the north-east and south-east Pacific, as well as ground-based observations in Finland have indicated that liquid R_{eff} is overestimated, which impacts the LWP retrievals (Painemal and Zuidema, 2011; King et al., 2003; Sporre et al., 2016; Noble and Hudson, 2015; Min et al., 2012). However, a number of studies also noted that there is good agreement in variability between liquid R_{eff} and LWP (Min et al., 2012; Noble and

Hudson, 2015). The cloud retrieval fraction and cloud optical depth has been found to perform better and be strongly correlated with field observations (Sporre et al., 2016; Noble and Hudson, 2015). Ice cloud retrievals have been known to be more difficult for a number or reasons relating to crystal shape and scattering properties as well as lifetime and expanse. Evaluation of the MODIS Collection 5 ice retrievals found an overestimation compared to an infrared radiative closure method for determining ice τ (Holz et al., 2016). The updates to the optical retrieval method in Collection 6 has reduced the bias to a more satisfactory level. We note that the evaluations discussed above have not been performed for the specific COSP product we are using in this work, but give us an idea of the MODIS satellite retrievals overall performance.

In this work, we are using cloud retrieval fraction, which we subsequently refer to as cloud fraction (CF), for ice (CFI) and liquid clouds (CFL) ~~, CTP~~ and the joint histogram for CTP and τ for cloudy and partly cloudy pixels. Following the methods of Oreopoulos et al. (2016); Schuddeboom et al. (2018) and Saponaro et al. (2020) we use the combined joint histogram product (i.e. the sum of the cloudy and partly cloud products). Very large biases that we considered unrealistic were found for the modelled R_{eff} and hence these fields have not been used for this work. Similar biases were found for the LWP and IWP, and hence, the COSP-derived products (described in the next section) for these fields were not used, but replaced with the raw model output. While this adds a degree of uncertainty to this work, we believe such an analysis with the derived COSP fields would not have been useful.

For this analysis we have removed all instances of clear sky from both the model and the observed histograms. We have done this by finding the instances where the CTP-τ histograms summed to zero. While this allows us to focus on instances when the model is simulating cloud, either correctly or incorrectly, it also means we are not considering instances when either the model or satellite simulated cloud while the other simulates clear sky. In this work, we are considering grid boxes of 1.25x1.875 degrees and daily means, meaning that there are very few instances of clear sky occurring for a full day over a large domain. On average, there are less than 1.5% of grid boxes simulated as clear sky in ACCESS-AM2 over the Southern Ocean, and far fewer in the satellite (in part due to the addition of the partly cloudy optical depth retrievals as well). For these reasons, we believe this choice is robust.

To evaluate the radiative bias in ACCESS-AM2, we use the Clouds and the Earth's Radiant Energy System (CERES) Syn1Deg product (Doelling et al., 2013, 2016) for evaluation at daily timescales. The bias in the outgoing top of atmosphere (TOA) shortwave (SW) cloud radiative effect (CRE) ($SWCRE_{TOA}$) has been chosen for analysis in this work based on previous findings from the literature presenting the SWCRE as the most problematic aspect of the SO energy balance. The CRE presented in this work is the difference between the ~~all-sky~~ clear-sky radiation and the ~~clear-sky~~ all-sky radiation fields (for both the model and satellite products). Throughout this paper, a predominantly ~~negative CRE~~ positive $SWCRE_{TOA}$ is present, e.g. the ACCESS-AM2 model is allowing too much shortwave radiation to pass through the clouds and not reflecting enough shortwave radiation out to space via clouds, corresponding with too much short wave radiation reaching the surface.

2.3 COSP

To directly compare the ACCESS-AM2 cloud properties to that of satellite products, we use the COSP, described by (Bodas-Salcedo et al., 2011), as prescribed for simulations within the CFMIP activity of CMIP6 (Webb et al., 2017). For the time

period of interest to this study, the MODIS products provided the best coverage and, hence, only these fields are compared in this study. Simulated fields include the joint histograms of CTP and τ, as well as liquid and ice cloud fractions, water paths and τ. We refer to these cloud fields as cloud 'features' throughout this work. By using COSP output, comparison of model to satellite is appropriate as the assumptions and limitations applied to the satellite algorithms are also applied to the model, hence limiting the possibility that any biases found are due to processing issues. ~~However we note that for~~ For the LWP and IWP, we are considering the grid box mean (as opposed to the in-cloud mean). We note that in the IWP and LWP fields, significantly large and seemingly unrealistic biases between the model COSP product and the MODIS product were found~~, that were thought to be a function of propagating errors~~. IWP and LWP are reliant on R_{eff} retrievals, which as discussed above, are less well captured in satellite products. For this reason, as stated above, the actual simulated (i.e. direct model output) IWP and LWP are used for this analysis instead of the COSP product.

2.4 Cloud regime clustering

In this study, we use k-means clustering to derive cloud regimes (Anderberg, 1973). K-means clustering is a form of unsupervised machine learning that separates N points in k clusters by minimising the sum of squared distances within each cluster. In this case, the Euclidean distance is considered, which is equivalent to the minimisation of variance (or inertia) within each cluster. To perform this analysis, the SciKit Learn (Pedregosa et al., 2011) and Dask-Machine Learning python packages were used, where k-means clustering has been implemented for distributed computing.

We apply k-means clustering to four years (2015-2019) of daily MODIS histogram data (CTP/τ), over the entire globe. ~~To perform this analysis, the SciKit Learn (Pedregosa et al., 2011) and Dask-Machine Learning python packages were used within a Pangeo environment, where k-means has been implemented for distributed computing~~We have normalised the CTP/τ histograms to one (as opposed to the cloud fraction) to limit the impact of cloud fraction biases within the ACCESS-AM2 model on the clustering results. Whilst this impacts our ability to compare to other studies, it allows the clustering to target cloud vertical extent and thickness regardless of the total fraction and how well it is captured by the ACCESS-AM2 model.

A complication of clustering methods, including k-means, is the choice in the number of clusters. While there is no 'right' or 'wrong' answer to the number of clusters to select, there are practical considerations and statistical metrics that can help guide this choice. Three statistical metrics were applied to this work in an attempt to aid the decision on the number of clusters to choose, including the 'Elbow' method (Wilks, 2011), the Calinski-Harabasz (CH) index (Calinski and Harabasz, 1974) and the Davies-Bouldin (DB) index (Davies and Bouldin, 1979). ~~Pedregosa et al. (2011) provide~~ The SciKit Learn application programming interface provides a detailed explanation of each metric in addition to their advantages and disadvantages. Unfortunately, the statistical guidance provided by these metrics was not useful in cluster number selection (suggesting 2, 4 and 17 clusters respectively). After consideration of a range of choices, we selected 12 clusters~~, consistent with previous studies and resulting in clear physical differences between clusters~~.

We then used the MODIS cluster regimes to predict the cloud regimes of ACCESS-AM2, by fitting each ACCESS-AM2 data point to one of the 12 cluster centres. We chose this method, as opposed to a hybrid approach as taken by Mason et al. (2015), so that we could apply the same cluster centres to multiple model simulations, allowing a direct comparison over a

number of simulations. A similar method is described in Williams and Webb (2009), where they also point out that this method eliminates further subjective choices.

We have chosen to perform clustering over the entire globe, though only results for the SO (defined for this work as the broad region from 30-69°S) are ~~shown~~ examined in this current study. This choice was made to allow us the option, after some change to the model has been made, to assess whether we have impacted other parts of the globe inadvertently. For this reason, while ~~all~~ the cloud-regime histograms are ~~shown initially, not all will be studied in detailif they are not important for the SO~~defined globally, only those important to the SO will be examined in detail.

2.5 Bias decomposition

To explore cloud property biases, we follow the decomposed bias metrics described in Williams and Webb (2009); Mason et al. (2015) and Schuddeboom et al. (2018) where the bias (Δ: ACCESS-AM2 $-$ satellite) of a simulated field of interest F for each cluster r can be summarised as the errors due to the RFO ($F_r^{sat} \cdot \Delta RFO_r$, referred to as RFO errors) plus the errors in the simulated field ($\Delta F_r \cdot RFO_r^{sat}$, referred to as field errors) plus a second order co-variation term ($\Delta F_r \cdot \Delta RFO_r$) otherwise referred to as the cross-term. ~~See Equation 1~~

$$\delta F_r = F_r^{sat} \cdot \Delta RFO_r + \Delta F_r \cdot RFO_r^{sat} + \Delta F_r \cdot \Delta RFO_r \quad (1)$$

3 The ACCESS-AM2 radiation and cloud biases

The annual and seasonal ACCESS-AM2 SWCRE$_{TOA}$ and cloud biases compared to the CERES-Syn1D or MODIS products are shown in Figure 1. The boundaries of our analysis, shown by the dashed lines, represent the mid-latitudes at 30-43°S, the sub-polar region 43-58°S, and the polar region 58-69°S of the SO. While seasonal expansion and contraction of the Ferrel and Polar atmospheric cells means these stationary boundaries may not capture the seasonal bias boundaries perfectly, for the purpose of this work they are satisfactory. These regions will be used throughout the rest of this work to explore how differently cloud types are contributing to the SWCRE$_{TOA}$ bias in each region.

Considering firstly the SWCRE$_{TOA}$ (row 1, Figure 1), a persistent ~~negative~~ positive bias in the polar region of the SO is found across all seasons. This bias has not been improved upon in this latest generation model, with a similar total SW$_{TOA}$ radiative bias (not shown) to that reported in a previous version of ACCESS presented in Fiddes et al. (2018). The summer (DJF), shown in Figure 1c1, continues to have the largest polar bias, while the winter (JJA) season has the smallest bias overall~~. A~~, in agreement with previous work (Bodas-Salcedo et al., 2012; Kuma et al., 2020). A positve/negative/positive ~~/negative~~ 'tripole' can broadly be seen in the annual mean biases (b1), where the SWCRE$_{TOA}$ is on average ~~underestimated~~ overestimated in the mid-latitudes, weakly ~~overestimated~~ underestimated in some parts of the sub-polar SO, and largely ~~underestimated~~ overestimated in the polar SO. MAM and JJA (Figure 1d1~~) shows this pattern more consistently, while it is shifted northwards in JJA with the expansion of the Polar Cell~~, e1) show very weak positive biases in the polar region, though with well defined transitions from negative to positive in the sub-polar to mid-latitude regions. In SON and DJF, the ~~negative~~ positive biases in

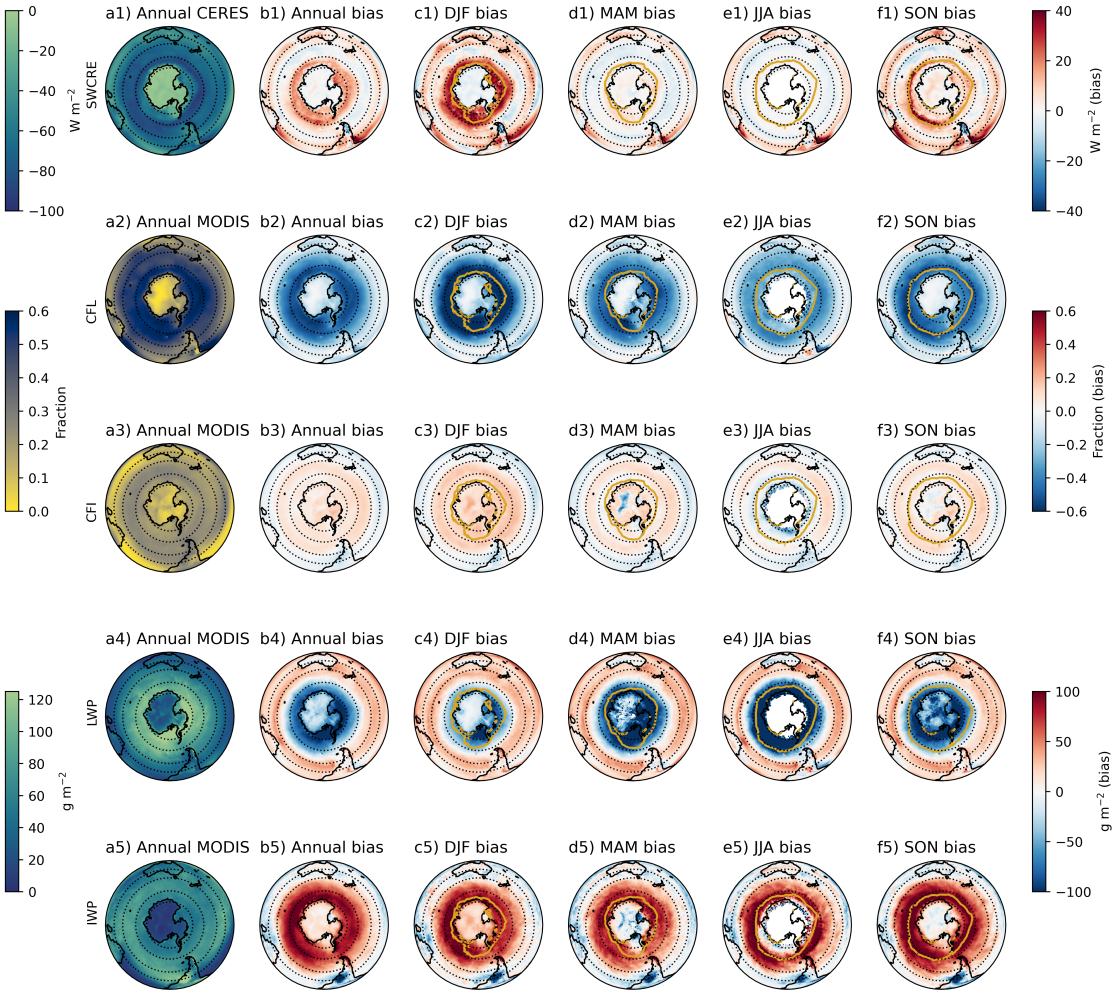

Figure 1. Column a) the annual mean satellite (CERES-Syn1D or MODIS) derived field for row 1: SWCRE$_{TOA}$, (W m^{-2}) 2: CFL 3: CFI , 4: LWP (g m^{-3}) and 5: IWP (g m^{-3}); column b) annual biases (ACCESS-AM2 - satellite) in the same respective fields; c-f) seasonal biases (DJF, JJA, MAM,SON). The yellow line in figure c-f) shows the mean seasonal extent of sea ice (calculated at 15% coverage), taken from the Hurrell et al. (2008) AMIP sea ice concentration data set. Dashed lines are located at 30°S, 43°S, 58°S, and 69°S

the polar and mid-latitude regions are well defined, while the ~~positive mid-latitude~~ negative sub-polar biases becomes far less ~~zonal~~zonally coherent. Some strongly ~~negative~~ positive regions in spring are also found along the edge of the sea-ice zone (yellow line). This feature could be due to some discrepancy between the sea-ice concentrations prescribed to the model and what was seen by the CERES satellite product, or could be issues with the satellite retrieving observations over ice-covered areas. Kuma et al. (2020) also noted a strong latitudinal dependence of radiative biases in a similar version of the UM and a reanalysis product. While much of the recent literature has focused on the strongly ~~negative~~ positive biases in the polar region

of the SO in DJF, we show here that the radiative biases for the Southern Ocean are complex and not always zonal. The summer time position of the positive/negative biases agrees broadly with recent observational work analysing airmass characteristics (eg. aerosol/airmass origin and impact on cloud condensation nuclei and cloud properties, Humphries et al., 2021; Simmons et al., 2021; Mace et al., 2021). This agreement highlights the importance of understanding not just the region where the biases are the most problematic, but the Southern Ocean system as a whole.

Rows 2-3 in Figure 1 show the respective biases for the CFL and CFIs compared to MODIS. Over all SO regions and seasons, negative CFL biases are found, in particular for the sub-polar and polar regions. The strongest bias is found in DJF for the polar region, though this extends into the sub-polar region, while the JJA appears to have the weakest biases. Interestingly, this bias in CFL is not compensated by a consistent overestimation in CFI for all regions (Figure 1, row 3). Annually, too much CFI is found for the polar and sub-polar regions, while this transitions to too little CFI in the mid-latitude region. In DJF, the positive sub-polar bias is very well defined and, as with the polar bias, at its largest. Over MAM, the sub-polar bias remains consistently constrained by the mid-latitude/sub-polar boundary, while the polar region bias begins to weaken and by JJA has become weakly negative. The clear distinction between the mid-latitude and sub-polar regions have also weakened in JJA, with the biases becoming less zonal. SON returns to predominantly positive polar and sub-polar regions, though with a less well defined boundary between the sub-polar and mid-latitude regions than what is seen for MAM or DJF.

Of note is the fact that the positive biases in the CFI are much weaker in magnitude than the negative biases in the CFL. Further, the spatial patterns of the biases are not the same. On average annually, the positive biases in the CFI over the polar and sub-polar regions only partially compensate for the negative biases in CFL. In the mid-latitude region however, both the cloud fractions are weakly negative, indicating too few clouds overall in this region, which can explain the ~~negative~~ positive SWCRE$_{TOA}$ bias. Too few liquid clouds which are instead simulated as ice clouds, will result in clouds that are more optically thin ~~and may precipitate out more easily, causing not enough~~ causing too little short wave radiation to be reflected out to space. This may help explain the average ~~negative~~ positive SWCRE$_{TOA}$ bias in the polar region, however does not explain the more ~~positive~~ negative bias in the sub-polar region. The lack of spatial correlation for the cloud fractions indicates that although the inaccurate partitioning of liquid and ice cloud is an important contributor to the radiative bias, it is not the whole story.

Rows 4-5 in Figure 1 show the respective biases for the LWP and IWPs compared to MODIS. Broadly, the LWP and IWPs tend to be of opposite signs to each other and similar in magnitude, with positive LWP biases in the mid-latitudes, a transition to negative in the sub-polar region, and negative LWP in the polar region. However, the change from positive to negative (or vice versa for IWP) appears to be more southwards for the LWPs, while more northwards for the IWPs. Too much ice, in place of water, potentially super cooled liquid water, again would produce an optically thinner cloud, which would cause too much sunlight to reach the surface and a ~~negative~~ positive SWCRE$_{TOA}$ bias. In the polar region, this process ~~may be~~ is likely contributing to the ~~negative~~ positive SWCRE$_{TOA}$ bias. In the sub-polar region, in DJF and to a lesser degree the other seasons, both water paths are in places positive, increasing the optical thickness of the cloud overall, which may contribute to a ~~positive radiation~~ negative SWCRE$_{TOA}$ bias. However, the regions of ~~positive~~ negative SWCRE$_{TOA}$ bias are not easily reconciled with the biases in the LWP and IWPs, indicating that this is a complex system~~to understand~~. Finally, in the mid-latitudes region, a positive LWP bias is found over all seasons and a positive IWP in winter and spring and weakly negative IWP in summer and

autumn. Positive water paths may have a compensating effect to the too little cloud fractions found in the region, resulting in the much weaker radiative bias. One reason for the different biases in the warmer mid-latitude regions compared to the other regions is likely the temperature dependence with respect to super cooled liquid water formation.

To summarise, in the polar region, the frequency and LWP of liquid clouds is largely underestimated, likely resulting in a very strong ~~negative~~ positive radiative bias. This is compounded by a slight overestimation of ice clouds containing too much ice, which are ~~likely to precipitate out too quickly~~ optically thinner adding to the ~~negative~~ positive radiative bias. This finding agrees well with the literature in that not enough liquid water exists below zero degrees Celsius, instead being simulated as ice (Bodas-Salcedo et al., 2016b). In the sub-polar region, the frequency of liquid clouds is weakly underestimated, but is compensated by too much LWP, creating too few, optically thicker clouds. Combining these biases with too much ice cloud, and too much IWP, results in a weak radiative bias that fluctuates between positive and negative. Finally, in the mid-latitudes, negative CFLs are again combined with positive LWPs, while the CFI and IWP are both weakly negative resulting in a weak ~~negative~~ positive radiation bias. These patterns are found to be generally consistent across the seasons, with some degree of variability in the strength of the biases.

It is clear from these results that while in the polar region the biases in cloud fraction and water paths can satisfactorily explain the SWCRE_{TOA} bias, for the other regions, it is not as clear cut. The role of cloud phase, and in particular that of supercooled liquid water, appears to have significant latitudinal dependence, likely influenced by a range of factors, including temperatures and ice nucleating particle (INP) availability. This is a significant issue for model development, as has been shown previously, where fixing one of these issues for the SO region (e.g. the ratio of liquid to ice clouds, or number of INP) may have detrimental effects on other parts of the system. For this reason, we suggest that we need a more in-depth analysis of the problem with respect to the ACCESS-AM2 model, before we can attempt to reduce this SWCRE_{TOA} bias. To do this, the next section presents results from the cloud regime-clustering, which will then be used to understand the radiative and cloud biases in more detail.

4 Cloud regimes

4.1 MODIS

The k-means clustering technique, ~~using 12 clusters for~~ spanning five years of daily-mean joint histograms ~~and applied to~~ over the entire globe, resulted in the 12 cluster centres shown in Figure 2. We have not shown the global relative frequency of occurrence (RFO), but have chosen to show the RFO for the Southern Hemisphere only in Figure 3, to focus on the Southern Ocean domain (30-69°S). The clusters have been approximately arranged from low to high CTP along the vertical and thin to thick τ in the horizontal.

Three of the resulting clusters are limited to the Antarctic region, and have similar (yet accordingly distinct) cloud characteristics. ~~While this is important to note, these clouds~~ These clusters have come about in part as a result of our normalisation by one, instead of cloud fraction and make up only a small fraction of overall cloud occurrence. Furthermore, as noted by Williams and Webb (2009), cloud retrievals over ice-covered regions with high albedo can be problematic. For these reasons, the three

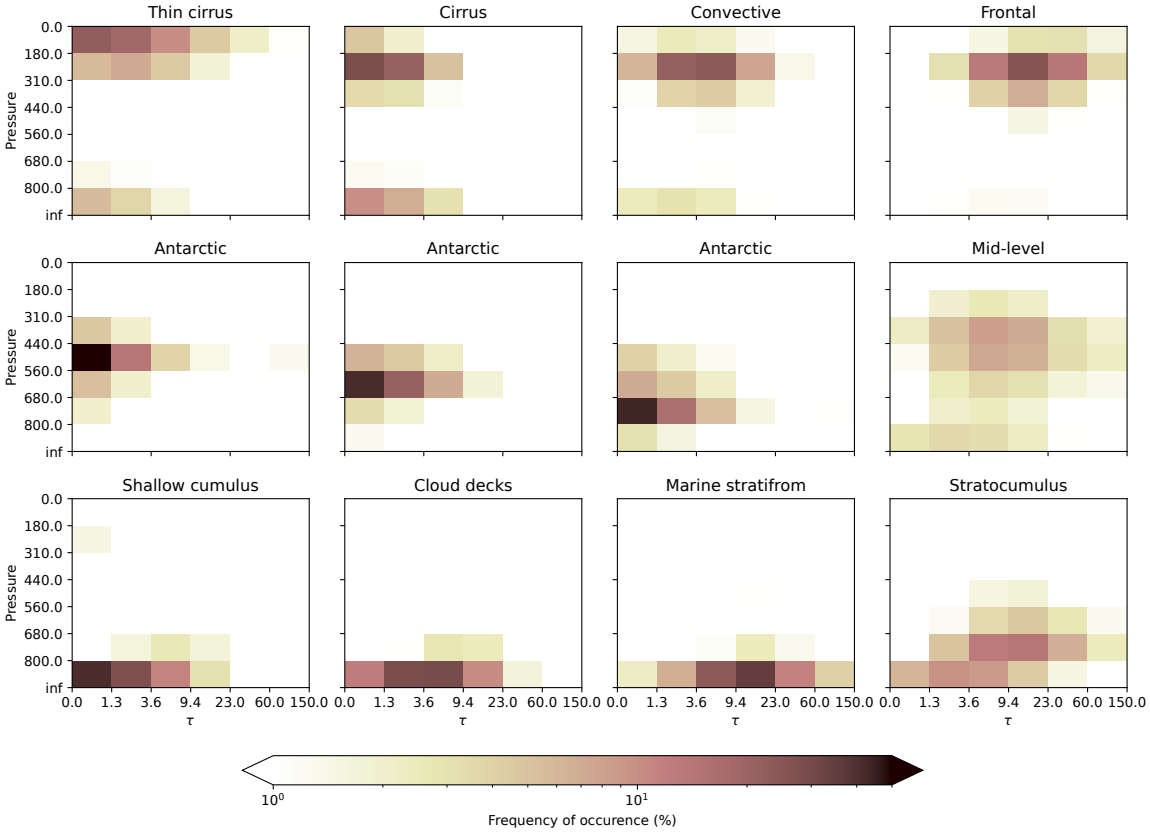

Figure 2. The mean CTP (y-axis, hPa) - τ (x-axis, optical depth unitless) histograms for each cloud regime found by k-means clustering on the MODIS product (daily means, 2015-2019). Each histogram has been assigned a 'cloud type'. The histograms have been loosely arranged by pressure and depth

Antarctic clusters are merged into one for the remainder of this analysis. We have labelled the remaining cloud regimes according to the cloud types that best reflect these profiles, however we acknowledge that the CTP-τ is often not enough to truly distinguish one cloud type from another (and that some quite different clouds may have similar CTP-τ profiles). Hence, these
330 labels should be considered just that: a way of easily differentiating the cloud regimes in this study.

Along the top row of Figures 2 and 3, high level clouds are shown. The thin cirrus cluster is restricted to tropical regions, particularly over the West Pacific/East Indian Ocean, and is characterised by clouds that are optically thin with very low CTP. Similarly, the cirrus cluster is characterised by very optically thin clouds, at a slightly lower altitude, and is found predominantly in the tropical and mid-latitude regions. Both the cirrus and thin cirrus show some thin, low level cloud too, likely associated

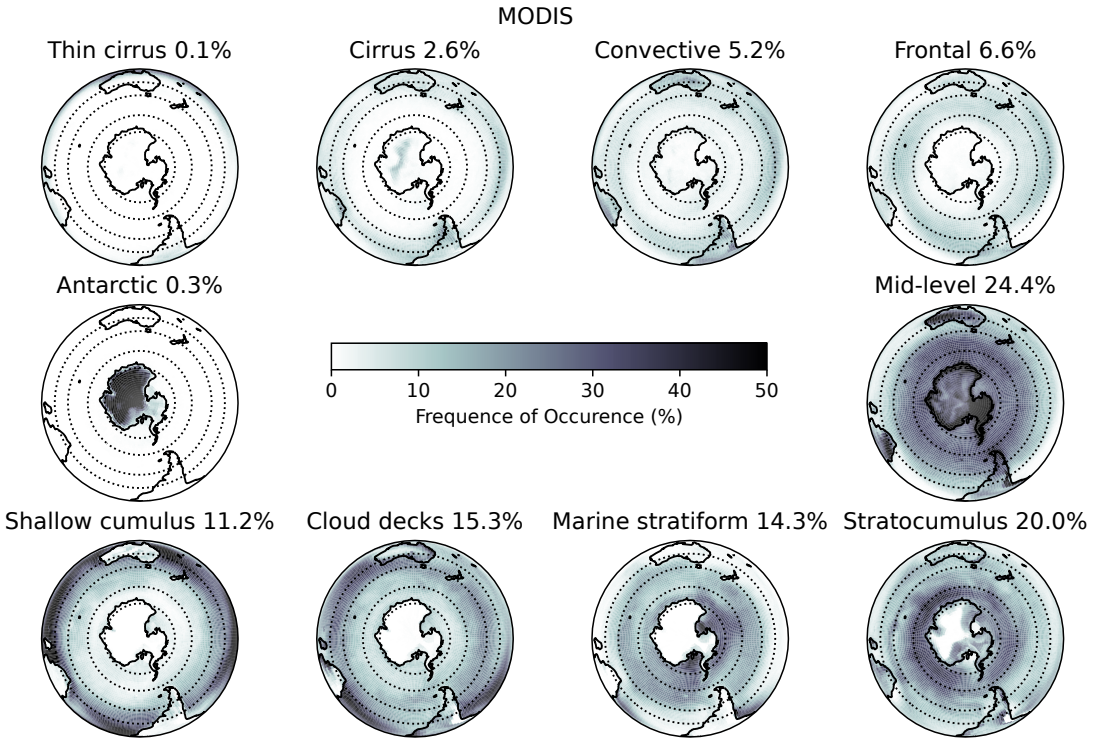

Figure 3. Spatial maps of the relative frequency of occurrence for each MODIS cloud cluster found by k-means clustering. Note that the three Antarctic clusters have been merged into one. The dashed lines represent the boundaries of the mid-latitude, sub polar and polar regions as defined in Section 3. The numbers in each title represent the mean ~~percentage~~ frequency of occurrence for each cloud regime (in time and space~~that each cloud makes up in the~~) over SO region

335 with shallow convective clouds. At a similar height, the convective cluster is more optically thick, and more strongly associated with the ITCZ and South Pacific convergence zone (SPCZ), although this is not shown by this map projection. Finally, along the top row, the frontal cluster, while similar in altitude as the convective cluster, is more optically thick, and is found more predominantly in the mid-latitudes, in particular around westerly storm tracks. We note that convective and frontal clouds can have similar CTP-τ profiles (Williams and Webb, 2009), and hence their location is also important to consider.

340 The mid-level cluster shows mid-range optical properties as well as mid-level CTPs and is without one or two clearly dominant bins as found in the other clusters (Figure 2). This mid-level cluster is most common over the SO, marine equatorial regions and terrestrial regions (in particular on the west coast of a number of continents). These varied geographic locations represent some significantly different cloud formation processes, raising an important point for this work: while the optical and

height characteristics within a cluster may be similar, each cloud's trigger mechanism is not necessarily the same. While further analysis to understand the formation processes for the cloud regimes is possible via meteorological fields such as proximity to cyclone centres or vertical motion, or physical fields such as proximity to topography would be of interest, such analysis is out of scope for this work.

Along the bottom row, the shallow cumulus cluster is found to have low, optically thin clouds that occur mostly over marine regions in the tropical and mid latitudes, away from any large-scale cloud formations such as the ITCZ/SPCZ (not shown) or stratiform cloud decks. This cluster has little vertical structure. By comparison, the cloud deck cluster is optically thicker and dominant in the eastern boundary of ocean basins, regions of extended stratiform cloud, again with little vertical structure. The marine stratiform cluster is the most optically thick cloud regime, with high CTPs and little vertical structure, and is found predominantly in higher latitude marine regions, as well as the very eastern boundaries of the cloud decks regions. Finally, the stratocumulus cluster characterised by mid-range optical thickness, but has lower ~~cloud top pressures~~ CTPs than those of the other low cloud clusters, suggesting a greater vertical extent associated with stratocumulus clouds. This cluster is geographically wide spread, although dominant in the polar region of the SO and absent in regions of deep tropical convection.

4.2 ACCESS-AM2

Using the cluster centres defined by the MODIS joint histograms described above, we fit the ACCESS-AM2 joint histograms to the same 12 cluster definitions (noting that the three Antarctic clusters were subsequently merged). Figure 4 shows the difference in frequency of occurrence between ACCESS-AM2 and MODIS. It is immediately clear that ACCESS-AM2 has significant problems with the mid-level and stratocumulus clouds, simulating them too frequently (by 13.4% and 19.8%) across all regions. The low-level cloud fields by contrast are all underestimated, including shallow cumulus (-9.9%), cloud decks (-9.5%) and marine stratiform clouds (-10.0%). The higher clouds (thin cirrus, cirrus, convective and frontal), are simulated comparatively well, although also with slight underestimation of frequency. Interestingly, for the SO~~region of interest~~, most of the RFO biases are spatially consistent in sign~~and for some, magnitude~~. This result could be interpreted as a consistent bias, with consistent causes, across the latitudes. We will explore this in more detail in the subsequent sections.

5 Understanding the biases

5.1 Bias decomposition

We can now use the biases in the frequency of occurrence of the cloud regimes to help gain a better understanding of the biases in radiation using the traditional decomposition method described in Section 2.5. Note that from this point on we only consider the broad SO region~~defined earlier~~, and the three sub-regions defined within. Figure 5 shows the cumulative components of the SWCRE_{TOA} biases over each region, for each cloud regime. The horizontal bar represents the total bias. Similar to previous studies, including Bodas-Salcedo et al. (2014), Mason et al. (2015) and, Schuddeboom et al. (2018), the largest component of the SWCRE_{TOA} bias is due to the errors in the RFO (blue bars). As previously discussed, DJF has the largest SWCRE_{TOA}

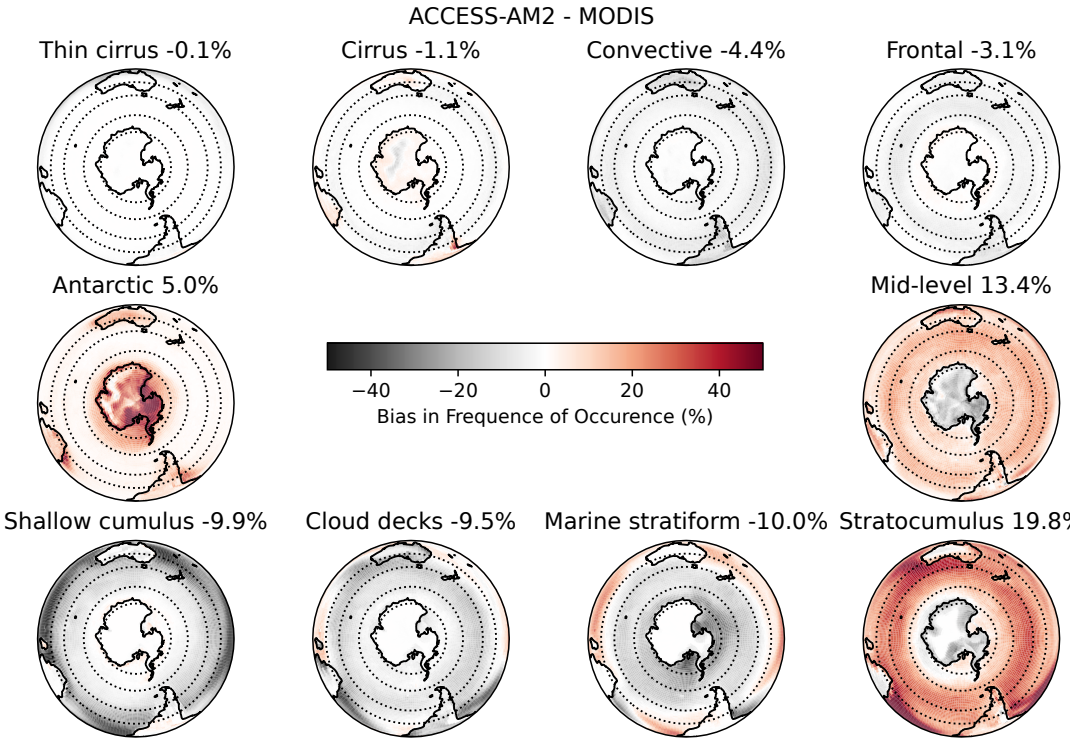

Figure 4. Spatial maps of the difference in the relative frequency of occurrence for ACCESS-AM2 - MODIS for each cloud cluster. The dashed lines represent the boundaries of the mid-latitude, sub polar and polar regions as defined in Section 3. The numbers in each title represent the ~~bias in the~~ mean ~~percentage that~~ frequency of occurrence for each cloud ~~makes up~~ regime (in ~~the~~ time and space) over SO region

375 biases, though the dominance of the RFO bias is true for all seasons. For DJF, clear ~~positive~~ negative RFO errors are found for the mid-level clouds and the stratocumulus clouds, which are both overestimated in frequency. This ~~positive~~ negative RFO error has a compensating effect, as the mid-level clouds are simulated in the place of other, lower cloud regimes (eg. marine stratiform, cloud decks and shallow cumulus), which are underestimated in frequency and hence have ~~negative~~ positive RFO errors. A much smaller component of the bias is made up by the error in the simulated field (pink bars), primarily in DJF,
380 while insignificant in the other seasons. Schuddeboom et al. (2018) hypothesised that the dominance of the RFO errors and comparatively smaller field errors, also found here, indicates that the clouds themselves, when simulated with the correct frequency, could in fact be simulated well. This hypothesis will be tested in the next section. The cross term (green bars) appears to contribute to the radiative biases again only in summer, and predominantly in the polar region. What this cross term

represents is not easy to evaluate, however it is clear that the field biases and RFO biases for this time and season can only explain part of the story.

We have also applied this bias decomposition technique to other cloud features, including biases in LWP, IWP and CFI and CFL (~~not shown~~see Appendix). The results for the LWP, IWP and CFL, for all regions, show large biases across all cloud clusters and decomposed terms. Interestingly, the field error and RFO error are of opposite signs and approximately cancel, leaving the cross term to make up the total radiative bias. For CFI, the RFO error was dominant, similar to the SWCRE_{TOA}, though with the field error and cross term being non-negligible, unlike for the SWCRE_{TOA}. The importance of the cross term in these results makes the respective biases much more difficult to understand, though it is clear that the decomposed errors in RFO and the fields themselves are playing a large role in the total bias.

The strength in this method is that it is time and space agnostic, i.e. it does not require the two 'climates' to be directly comparable, as it is using the mean RFOs and field values to make judgements about how each of them contribute to a particular bias. For this reason, this method has been popular in the literature which has often used CMIP simulations as the basis of their work. However, a limitation of this method is that it cannot compare like-for-like instances in cloud properties to gain a better understanding of what particular conditions are leading to large SWCRE_{TOA} biases. This work suggests that for all seasons and all regions, a ~~positive~~ negative bias occurs because the frequency of mid-level and stratocumulus clouds are overestimated, compensating for the underestimation of the lower level clouds. We ask, at what cost, specifically, were the mid-level and stratocumulous clouds overestimated? Is the compensating error due to the model predicting clouds that are too frequent, optically thick, high or with too much liquid? Applying the same technique (the bias decomposition) to such cloud fields was unable to satisfactorily give us these answers. Additionally, we would like to know, does the model ever actually do a good job, with the correct cloud type and a small (or wishfully - no) radiative bias?

5.2 Biases when clusters are correctly or incorrectly simulated

One strength of comparing a daily, nudged, simulation to daily MODIS fields is the ability to make direct comparisons in time and space. As the synoptic meteorology is considered to be the same in the model and the observed conditions, we therefore expect that the model microphysics, if accurate, would generate the same cloud types that the large-scale dynamics prescribes. With this assumption, we are able to isolate instances (exact points in time and space) where the model simulates the same cloud type as MODIS, which we define as 'correctly' simulating the cloud type. Similarly, we can also define the instances where the model simulates a different cloud type, which we consider an 'incorrect' cloud type assignment by the model. We can then use these instances to generate a statistical understanding of the conditions during these times and places. We demonstrate these definitions in Figure 6a and b.

In the previous section, we saw that the model generally tends to simulate the ~~wrong cloud frequency~~ cloud type RFOs incorrectly and that this ~~error~~ bias in RFOs dominates the radiative bias. This result demonstrates the limit of the decomposition method as we are not able to gain any further information about how the microphysical biases in cloud types affect the radiative properties, or the nuances of which cloud types are being simulated in place of others. In this Section, we take advantage of the model nudging and explore in detail the instances when ACCESS-AM2 correctly and incorrectly simulates the MODIS

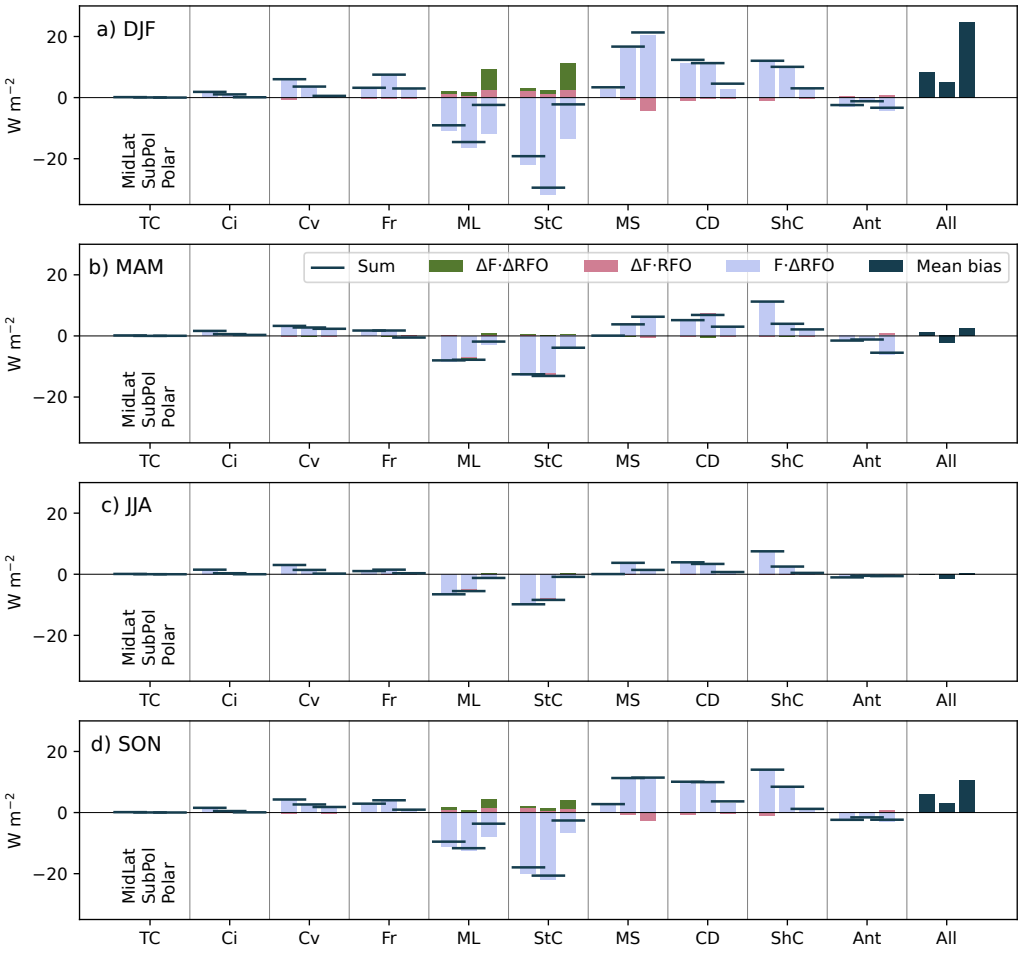

Figure 5. The decomposed mean biases in the SWCRE$_{TOA}$ ($W\,m^{-2}$) for each cloud regime (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant) over the three regions from left to right mid-latitudes, sub-polar and polar, for each season (DJF, plot a; MAM, b; JJA, c; SON, d respectively). The sum of the decomposed biases are shown by the horizontal bar, while the the terms of the bias decomposition (see eq. 1) are shown in blue ($F_r^{sat} \cdot \Delta RFO_r$), pink ($\Delta F_r \cdot RFO_r^{sat}$) and green ($\Delta F_r \cdot \Delta RFO_r$). On the far right of each plot is the mean radiative bias for each region/season as a reference

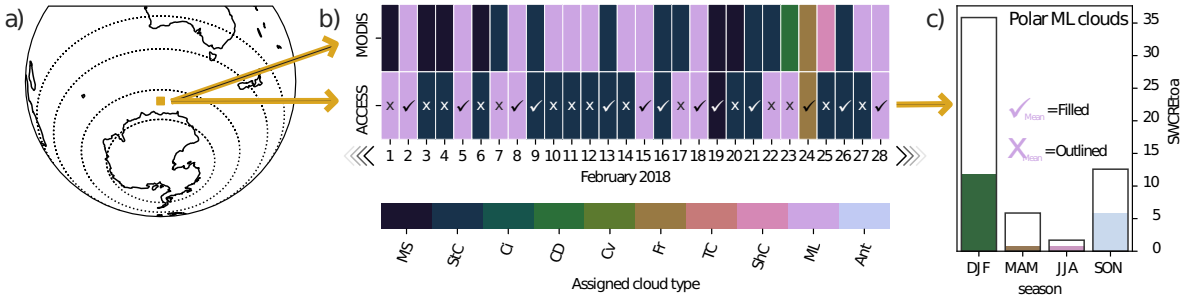

Figure 6. For each grid box (a), and for every point in time (b), the cloud type has been determined via k-means clustering for the MODIS product, and subsequently fitted to the ACCESS-AM2 output. The assignment of cloud type is shown for a one month period in b), where we can see where ACCESS-AM2 correctly (shown by a tick) or incorrectly (shown by a cross) simulates the same cloud type as the MODIS product. Using this information, we calculate the seasonal mean radiation or cloud property bias (c) over the instances where ACCESS correctly (filled bars) or incorrectly (outlined bars) simulated the cloud type for the three regions of the Southern Ocean. As an example, c) shows the mean radiation biases for polar mid-level clouds.

identified cloud regimes and begin to understand the associated radiative and cloud biases in much greater detail than has previously been achieved.

Each panel of Figure 7 shows for each MODIS cloud type (y-axis), the percent of time that each cloud type is assigned by ACCESS-AM2 (x-axis). The total number of instances that that cloud type is observed by MODIS for each region is shown on the right of each panel. If ACCESS-AM2 ~~were correctly simulating clusters~~simulated every cloud type the same as what the MODIS product did, we would expect a diagonal line through ~~Figure 7, which shows the percent of MODIS clusters correctly assigned by ACCESS-AM2 for the SO region of -30 to -69°S~~each panel in Figure 7 of 100%, with zeros elsewhere. What is shown however, reflects clearly the biases of the RFOs, where ACCESS incorrectly assigns the stratocumulus and mid-level clusters over most other clusters. For the low-level cloud regimes (shallow cumulus, cloud decks and marine stratiform), we see that a large number (~~> 50%~~ between 48% and 59%) of points have been wrongly assigned to the stratocumulus cluster, while the high level clouds (convective and frontal, as well as the less important cirrus and thin cirrus) tend to be assigned as the mid-level cluster (~~also > 50% for the more relevant clouds~~between 38% and 67%). Only ~~9%~~12%, 11% and 5% of the marine stratiform points have been correctly assigned ~~(this~~ by ACCESS-AM2 for each region (mid-latitudes, sub-polar and polar). This statistic is worse for some other, less important SO ~~clusters), whereas 67%and 61~~cloud types. In contrast, 63%, 68% and 68% of mid-level ~~and~~ clouds and 60%, 66% and 56% of stratocumulus clouds are simulated in the correct time and place for the ~~SO region~~three SO regions (mid-latitudes, sub-polar and polar).

Having identified exactly when ACCESS-AM2 correctly or incorrectly simulates cloud regimes, we can now begin to evaluate what the biases in radiation and other cloud fields are in these instances. Figures 8 (mid latitude), 9 (sub-polar) and 10 (polar), show seasonal mean biases when the ACCESS-AM2 correctly simulates each cloud cluster (coloured bars~~, grey outline~~) or incorrectly simulates each cloud cluster (black transparent bars)~~according to MODIS. The total number of instances~~

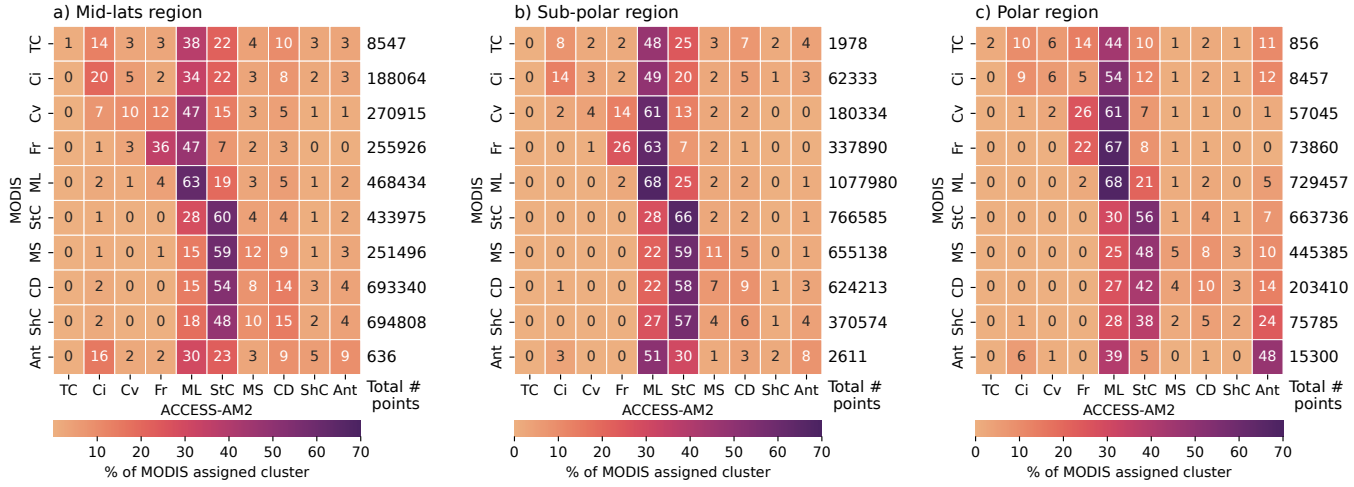

Figure 7. ~~A confusion matrix~~ Confusion matrices for each region, a) mid-latitudes, b) sub-polar and c) polar, showing the MODIS cluster assignment on the y-axis and the ACCESS-AM2 cluster assignment on the y-axis (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant). The colours (and inset text) represent the proportion of points assigned to each MODIS cluster (eg. sum to 100 along the x-axis). The text on the right hand side indicates the number of points (in time and space) that make up these statistics for each MODIS cluster

~~when this occurred are indicated by the size of the circle along the top (agree) and bottom (disagree) axes, with larger circles indicating a greater frequency of occurrence~~. An example of how one of these subsets (polar mid-level clouds) is derived is shown in Figure 6. We note that these results are not weighted by the frequency of occurrence, unlike the previous figure, as we are interested in understanding the microphysical properties of each cloud type, even if they occur infrequently. Figure 7 can provide an indication of the frequency of each cloud type and how often it is correctly or incorrectly identified. Furthermore, we emphasise that these plots show the impact in the radiation and cloud fields from the perspective of the 'right' cluster being assigned wrongly, unlike in Figure 5, which shows the radiative biases from the perspective of the accumulative ~~'~~'wrong' clusters. For this reason, some of the values between the figures have opposite signs, but support the same overall finding.

5.2.1 The mid-latitude region

Figure 8a shows the mid-latitude ~~regions~~ region's mean $SWCRE_{TOA}$ biases ~~. Here, it can be seen that when the cloud clusters are incorrectly defined by~~ when ACCESS-AM2 ~~, the biases in $SWCRE_{TOA}$ for most mid-low level cloud typesare worse than when the cluster is correctly identified (eg. black outlined bars are larger than grey coloured bars).~~ correctly and incorrectly simulates the cloud types. For the mid-level and marine stratiform clouds, the $SWCRE_{TOA}$ biases are larger when the cloud types are assigned incorrectly (outlined bars) than when they have been correctly simulated by ACCESS-AM2 (coloured bars). However, ~~we can see that even when the~~ when these cloud types are correctly assigned (which happens ~~67~~ 63% and 12% of the time respectively), the $SWCRE_{TOA}$ bias are ~~for most cases~~ also non-negligible in most

seasons. For ~~both these cloud types, the CFL is strongly underestimated when they are incorrectly simulated which is expected to produce negative SWCRE$_{TOA}$ biases, while the CFI is comparatively well captured. Interestingly, the LWP and IWPs seem to be relatively well captured. This suggests that the SWCRE$_{TOA}$ bias may be predominantly driven by an underestimated CFL, while the amount of water in them is somewhat correct. For the~~ the mid-level clouds, the SWCRE$_{TOA}$ biases are ~~of a similar nature, though~~ smaller in magnitude when the clusters are correctly identified by the model. Interestingly, this is not the case for the CFL and LWPs, which both have larger biases when the clusters agree, suggesting that the radiative effects ~~of too fewer~~ associated with too few liquid water clouds is partially compensated by them being too optically thick. This indicates that the lower SWCRE$_{TOA}$ bias when the mid-level clouds are correctly simulated may be occurring for the wrong reasons. For the marine stratiform clouds, the CFL is strongly underestimated when it is incorrectly simulated, which is expected to produce positive SWCRE$_{TOA}$ biases. The CFI is comparatively well simulated. Interestingly, the LWP and IWPs seem to be relatively well captured for marine stratiform clouds. This suggests that the SWCRE$_{TOA}$ bias may be predominantly driven by an underestimated CFL, while the amount of water in them is somewhat correct.

The stratocumulus cluster shows almost uniformly equal biases in all fields when the cluster assignments ~~agree and disagree. The negative~~ are correct and incorrect. The positive radiative biases are predominantly ~~driven by~~ associated with too few liquid water clouds, similar to that of the mid-level clouds. Considering that stratocumulus clouds are simulated correctly 60% of the time and are otherwise simulated as mid-level clouds 28% of the time, these biases make ~~some degree of~~ sense. These results suggest that we cannot focus solely on improving the lowest level clouds, but must also improve the mid-level cloud representation too.

~~Shallow cumulus clouds show a small negative SWCRE$_{TOA}$ bias when the clusters agree, which is also reflected in the cloud fraction and water paths, indicating that the model performs well when simulating these clouds correctly. However, the grey circles in Figure 8, plus data in Figures 4 and 7 show that the model infrequently gets this cloud type right (only 2% of the time), instead simulating it as stratocumulus 48% of the time in the mid-latitudes. The positive SWCRE$_{TOA}$ bias found in all seasons when shallow cumulus are incorrectly simulated is likely a reflection of the higher and more optically thick clouds found in the stratocumulus cloud type, indicated by overestimated LWP and to a lesser degree, IWPs. This result also agrees well with our previous plot (Figure 5), where too many stratocumulus clouds result in a compensating positive bias. A similar result can be observed for the less frequent cirrus clouds, which are too often simulated as lower, more optically thick stratocumulus or mid-level clouds (22 and 34% of the time) and cloud decks in winter and autumn.~~

Cloud decks present a notable exception of where the radiative biases are larger and of opposite sign when the cluster is correctly identified. Cloud decks are not simulated frequently enough in ACCESS-AM2 (correct only 14% of the time), instead assigned as stratocumulus clouds 54% of the time. ~~Similar to the cirrus and shallow cumulus clouds,~~ Figure 8 shows a weak ~~positive~~ negative SWCRE$_{TOA}$ bias for the MAM and JJA seasons when the ~~clusters disagree~~ACCESS-AM2 clusters are incorrect, while the bias in the other seasons is negligible. When cloud-decks are correctly assigned, the SWCRE$_{TOA}$ bias is ~~negative~~ positive in all seasons, indicating that for these low-lying clouds, even if we are able to simulate the height and optical thickness properties of the cloud ~~somewhat~~ correctly, issues remain. Looking at the cloud fields, most ~~are relatively well captured~~ show small biases (compared to the other cloud regimes), with the CFL having the largest negative bias. ~~Too~~

We suggest that too low cloud fraction would result in a ~~negative~~ positive SWCRE$_{TOA}$ bias, though how this seemingly minor cloud fraction bias is resulting in the more significant radiative bias is difficult to say. Regardless, this result demonstrates the complexity of this issue, where we not only need to get the right clouds ~~in the right place~~for the large-scale conditions, but also ensure the underlying microphysical properties are correctly simulated.

Shallow cumulus clouds show a small positive SWCRE$_{TOA}$ bias when the ACCESS-AM2 correctly simulates the cluster, which is also reflected in the cloud fraction and water paths, indicating that the model performs well in these instances. However, Figure 7a shows that the model infrequently gets this cloud type right (only 2% of the time), instead simulating it as stratocumulus 48% of the time in the mid-latitudes. The negative SWCRE$_{TOA}$ bias found in all seasons when shallow cumulus are incorrectly simulated is likely a reflection of the higher and more optically thick clouds found in the stratocumulus cloud type, indicated by overestimated LWP and to a lesser degree, overestimated IWP. This result also agrees well with our previous plot (Figure 5), where too many stratocumulus clouds result in a compensating negative bias. A similar result can be observed for the less frequent cirrus clouds, which are too often simulated as lower, more optically thick stratocumulus or mid-level clouds (22% and 34% of the time) and cloud decks in winter and autumn.

The frontal cloud type has a large ~~negative~~ positive SWCRE$_{TOA}$ bias when the model does not ~~agree with the MODIS assignment~~correctly simulate the MODIS cluster, which occurs 64% of the time in the mid-latitude region. The frontal clouds are a relatively high cloud (see Figure 2) and in these lower latitudes have a greater likelihood of being ice clouds. This bias, found across all seasons, could be due to a too large LWP, though the CFL is relatively well captured. For the frontal ice clouds, too few clouds are simulated with a negative IWP bias. In the case of correctly assigned frontal clouds, we can see that the SWCRE$_{TOA}$ bias is small, despite too little CFI and too high LWP. This result may again indicate that while the ACCESS-AM2 model is simulating correct SWCRE$_{TOA}$, it may be doing so for the wrong reasons. A similar result, i.e. strongly underestimated CFI causing a ~~negative~~ positive radiative bias, is found for the convective clouds, in instances when they both disagree and to a greater extent, agree. This result may indicate that for these higher clouds, getting the cloud fractions right, with the right phase partitioning, may go a long way to reducing the radiative biases~~right.~~. The thin cirrus clouds are shown to have large biases in the spring (SON) when correctly simulated for all fields. However, we note that these cloud types occur infrequently (see Figure 7a) and are only correctly simulated 1% of the time, and hence, we do draw any conclusions from this result.

5.2.2 The sub-polar region

~~Sub-polar total radiative biases~~The sub-polar total radiative bias, shown in the far right column of Figure ~~9a, are~~ 5, is in general quite small~~, with a few exceptions. Considering~~. If we consider just the zonal means, the model appears to perform well. However, after analysing Figure 1, we know that large spatial variability exists within this region that leads to a small overall bias. ~~By analysing each cluster, we expect to identify what type of clouds are contributing to each bias.~~ In this section we explore if some of that variability can be attributed to particular cloud types.

~~For the two cloud typesthat are simulated most accurately, the~~ The mid-level and stratocumulus clouds are the two most accurately simulated cloud types (68% and 66% of the time)~~, we can see that the~~. The cloud radiative biases in these cases

are relatively low, implying that the model may be doing a good job. We note that the biases are fairly similar for when they disagree too, and that is likely due to each one being predominately assigned as the other when they are incorrectly simulated. ~~For when the stratocumulus assignments do agree~~When the stratocumulus are correctly simulated, examination of the cloud fields indicates that despite the small radiative biases the apparent model skill may be misleading. In these cases, CFL and LWP are underestimated and the CFI and IWP are overestimated across most seasons. We suggest that too few, optically thinner (derived from too little LWP) liquid clouds, causing a strong ~~negative~~ positive SWCRE$_{TOA}$ bias may be partially compensated by too many, optically thicker (too much IWP) ice clouds. A similar result is found for the mid-level clouds, though with smaller biases in CFL and IWP, and negligible biases in CFI and LWP. ~~Considering now what may be contributing to the positive biases in this region, we~~

We can see that the majority of ~~this~~ the negative bias in the sub-polar region is driven by shallow cumulus clouds ~~,~~ that are incorrectly simulated. Shallow cumulus clouds are optically thin, and low lying (see Figure 2). They are only correctly simulated in this region 1% of the time. Instead they are simulated as stratocumulus clouds 57% of the time which are higher and more optically thick, both of which would lead to more sunlight being reflected back out to space. Again, this agrees well with the findings of Figure 5, where ~~positive~~ negative radiative biases are found to be due to too many stratocumulus and mid-level clouds. Looking at the shallow cumulus cloud features, we can see that when these clouds are incorrectly defined, we have too few liquid clouds, too many ice clouds and too high LWP and IWPs. When shallow cumulus are correctly simulated, the majority of these biases are much smaller, and the SWCRE$_{TOA}$ bias is weakly ~~negative~~positive, indicating again that the model is doing a better job.

The cloud deck clusters, when incorrectly assigned, also appear to contribute towards a ~~positive~~ negative SWCRE$_{TOA}$ bias. Cloud decks are only ~~correct~~ correctly simulated 9% of the time, and are also considered by the ACCESS-AM2 model to be stratocumulus clouds 57% of the time instead. Examining the biases in the cloud fields for these instances indicates a similar process ~~for the~~ as for the shallow cumulus clouds for the SWCRE$_{TOA}$ bias to occur. For the sake brevity, we will not discuss the remaining cloud types for this region. The majority of them (bar some of the less frequently occurring types) are generally contributing to a ~~negative~~ positive SWCRE$_{TOA}$ bias regardless of whether the clusters are correctly identified or not.

5.2.3 The polar region

Finally, we now consider the polar region in Figure 10. Here it is clear that the mid-level, stratocumulus and marine stratiform clouds are contributing the most to the ~~negative~~ positive SWCRE$_{TOA}$ bias, whether the clusters ~~agree or disagree~~are correctly or incorrectly assigned. Few compensating errors are found in cloud types with a meaningful impact on the region (e.g. that occur frequently), adding to the larger magnitude of the SWCRE$_{TOA}$ bias in this region. In each of the three aforementioned cloud types, CFL is strongly underestimated (by over 50% in most seasons) by the model. The LWP is predominantly underestimated (though interestingly, less so for summer) and the IWP is overestimated. Each of these cloud biases are expected to produce a ~~negative~~ positive SWCRE$_{TOA}$ bias. The CFI, for the marine stratiform and stratocumulus clouds is overestimated, while for the mid-level clouds, it is in most cases underestimated. This difference may be due to (in part) the higher ~~cloud top pressures~~ CTPs found in these mid-level clouds, meaning that ice formed in the model is done so correctly. The cloud biases found in these

plots are of similar magnitude throughout the year while the radiative bias in summer time is large and almost non-existent in winter. This is ~~likely~~ due the fact that despite consistently poorly simulated cloud features, they have much lesser impact on the radiative balance in winter due to the much smaller amount of solar radiation compared to the summer. Nevertheless, the fact that even when the cloud types are correctly simulated by the ACCESS-AM2 model the SWCRE_{TOA} bias are similarly, if not more, ~~negative~~ positive than when the clusters are incorrectly assigned is a cause for concern. These results strongly highlight the issues of cloud phase within the model, that exists even if the height and optical depth of the cloud is correctly simulated. Furthermore, these results show that the cloud property biases are not limited to the summer season, where they are able to have a large impact on the SWCRE_{TOA}.

5.2.4 Summary of findings

There are two broad groups of cloud types driving the SWCRE_{TOA} bias apparent from this analysis: the high clouds, including thin cirrus, cirrus, convective and frontal clouds; and mid-low level clouds, including the mid-level, stratocumulus, marine stratiform, cloud deck and shallow cumulus clouds. The high clouds consistently, across regions and seasons, contain too few ice clouds (negative CFI) with too much IWP. The IWP biases get larger at higher latitudes. While these cloud types do not make up a large fraction of clouds over the SO (see Figure 3), and hence have not been widely studied, it is clear that the different microphysical processes are also contributing to a SO ~~negative~~ positive SWCRE_{TOA} bias.

For the mid-low level clouds, too few liquid clouds are found consistently across the regions and seasons, which again, gets worse at higher latitudes. This ~~negative~~ contribution to the ~~negative~~ positive SWCRE_{TOA} bias is compounded in the optically thicker cloud types (mid-level, stratocumulus and marine stratiform) by negative biases in the LWP, also increasing in magnitude with higher latitudes. Positive biases in CFI and IWP are also found for these cloud types, which again are largest in magnitude in the polar region. These positive biases in ice clouds may have a compensating ~~positive~~ negative effect to the SWCRE_{TOA} bias if they are increasing the overall cloud and water path fractions, however, if they are simulated instead of super cooled liquid clouds, they may indeed have the opposite effect. The optically thinner shallow cumulus clouds are found to have a compensating negative effect when the optically thicker clouds are simulated in their place for all regions.

6 Conclusions

Recent evaluation of CMIP6 models has found continuing problems in the SO shortwave cloud radiation effect. While updates to model parameterisations and tuning have occurred since CMIP5, these have resulted in some models having too high climate sensitivity (Zelinka et al., 2020), particularly driven by the SO, that may be producing lower SO SWCRE_{TOA} biases for the wrong reasons (Schuddeboom and McDonald, 2021; Gettelman et al., 2020). In this work we take a detailed look at the SWCRE_{TOA} bias in one model, the ACCESS-AM2 model. By running ACCESS-AM2 with nudging, we are able to make day-for-day comparisons with satellite products (in this case CERES-Syn1D and MODIS COSP products) for the first time, allowing more in-depth analysis of this problem than what has previously been completed.

In this analysis, we use unsupervised k-means clustering on CTP-τ histograms, a method that has proven useful in numerous previous studies. We use 12 clusters for our analysis, though later merge three in to one 'Antarctic' cluster, leaving ten cloud regimes from which we can begin to understand under what conditions the SWCRE$_{TOA}$ bias occurs and the associated cloud properties. In particular, we are able to analyse the SWCRE$_{TOA}$ and cloud property biases in instances when the model correctly simulates (e.g. hits) a cloud regime and compare it against instances when the model incorrectly simulates a cloud regime (e.g. misses - assigned as something else) with respect to the MODIS product.

We find that ACCESS-AM2 strongly overestimates the occurrence of stratocumulus and mid-level clouds over the entire SO, in general agreement with that found for CMIP6 models in Schuddeboom and McDonald (2021). The over-prediction of these two cloud types comes predominantly at the cost of marine stratiform, cloud deck or shallow cumulus clouds (and others), while stratocumulus and mid-level clouds themselves are accurately predicted 60% and 67% of the time by the model, respectively. In particular, we find shallow cumulus clouds are simulated as the higher/thicker stratocumulus clouds by ACCESS-AM2 50% of the time, leading to a compensating ~~positive~~ negative SWCRE$_{TOA}$ bias over the entire SO for all seasons. A similar result is found in the sub-polar region for cloud decks in all seasons. Interestingly, when shallow cumulus clouds are correctly simulated by ACCESS-AM2, the radiative biases are much smaller, as are the biases in cloud properties. This result implies that correctly simulating the conditions for shallow cumulus to form may have a beneficial impact on the simulated radiative balance. The same cannot be said for the cloud deck clouds however, where errors in both the SWCRE$_{TOA}$ and cloud properties remain even if they are correctly simulated.

We find that when stratocumulus clouds are correctly simulated by the ACCESS-AM2 model, the SWCRE$_{TOA}$ bias in the mid-latitude and sub-polar parts of the SO is relatively small. However, examination of the respective biases in cloud properties indicates that the model is likely producing these smaller radiative biases for the wrong reasons, which differs by latitude. Concerningly, in the polar region summer time, stratocumulus clouds have a worse SWCRE$_{TOA}$ bias when they are correctly simulated by the ACCESS-AM2 model than when they are incorrectly assigned, indicating that significant issues remain within cloud microphysical properties, and in particular cloud phase, even if the macrophysics are somewhat correct. These findings provide ~~crucial~~ important knowledge to help guide future model development which must target not only simulating the correct cloud type but also the correct cloud microphysics.

~~In the region of the largest negative SWCRE$_{TOA}$ bias, the polar region of the SO, both correct and incorrect assignment of the marine-stratiform,~~ Marine-stratiform, stratocumulus and mid-level clouds contribute most to ~~this~~ the polar region's positive SWCRE$_{TOA}$ bias. Each of these ~~biases, again whether the cloud regime is correctly simulated or assigned as something else, is~~ cloud types is associated with large biases in cloud properties, including too few liquid clouds, too little LWP and too much IWP. This is the case both for when these cloud types are correctly and incorrectly assigned. In the case of the marine-stratiform and stratocumulus clouds, the CFI is overestimated, while for the mid-level clouds, it is underestimated in all seasons except summer, where the bias is small. These results indicate that the issue of cloud phase within the ACCESS-AM2 model is still causing significant problems for the polar radiative balance. This cloud phase problem in these three cloud regimes does not appear to be as large an issue for the same clouds in the mid-latitudes or even the sub-polar region. This finding implies that

any changes to the model parameterisations must be done with caution, taking into account latitudinal dependencies on things such as temperature, boundary layer coupling, ice nucleating particle presence, etc.

While we have in this work explored the $SWCRE_{TOA}$ biases between ACCESS-AM2 and the CERES-Syn1D product and the cloud feature and type biases between ACCESS-AM2 and MODIS product, some key caveats must be raised. Firstly, we must note that comparing across two different satellite products may not be ideal given different assumptions and requirements. While this is standard practice in the field, the two products are derived from different platforms with different limitations, pointing towards a greater need for combined 'earth system' satellite products. We note that other satellite products have not been used in this work due to inconsistencies with the time period of focus. Secondly, we must recognise that both satellite products have their own biases. While in this work we regard them as 'truth', this is not necessarily the case. Hinkelman and Marchand (2020) for example, compared CERES and CloudSat radiation fields to in-situ observations at Macquarie Island and found a $+10\,W\,m^{-2}$ bias, which they expect would exacerbate the biases found in model evaluation. As discussed in Section 2.2, significant issues remain with the retrievals of R_{eff}, and subsequently the water paths. The propagation of these errors through the COSP framework, meant that comparisons between the COSP product and MODIS product were found to be very unrealistic, and hence these fields were not used and the raw model field LWP and IWP were used instead, adding uncertainty to these results. Further work evaluating cloud properties derived from satellite products would be of upmost help to studies such as this.

~~Gettelman and Sherwood (2016)~~ Gettelman et al. (2020) also noted that the nudging choices made can have large, and in some cases detrimental effects on the ability to simulate cloud fields, in particular cloud phase partitioning and water content due to the control on temperatures Zhang et al. (2014). While a study such as this could not be performed without the use of nudging, it is possible that the effects found in Gettelman et al. (2020) may be present here. We must also consider the impact of a single moment cloud scheme used in ACCESS-AM2 (the PC2 scheme). A double moment cloud scheme, that includes prognostic equations for both the cloud droplet size and number is preferable for a study such as this. Even better would be a cloud scheme that is fully coupled to the aerosol scheme. While in ACCESS-AM2, the cloud droplet number concentration is resolved, ice nucleating particles are not explicitly resolved. We hope to be able to address both of these issues in the future within ACCESS.

The work presented in this study has provided a relatively qualitative view of the cloud and radiative biases associated with cloud types. Studies such as Bender et al. (2017) have used more quantitative methods to evaluate the role cloud fractions play in determining the radiative balance. Specifically, (Bender et al., 2017) examine the distributions of cloud albedo and the associated cloud fraction in CMIP5 models and their linear relationships. In further exploration of the data used in this work, we note that most of the relationships between cloud properties and radiative bias are in fact non-linear, which is also highlighted in this work by the differing relationships observed by latitude/cloud type. Similar results have been noted in Bodas-Salcedo et al. (2016a), where the LWP and radiative biases were not found to be as tightly coupled as expected. We will present findings of a quantitative analysis employing machine learning to understand the role cloud properties play in determining the radiative bias in an upcoming paper.

Furthermore, in this work we have only considered a few cloud properties that impact the radiation budget. Other important factors, such as the number and size of cloud droplets, precipitation phase and amount, or other thermodynamical properties are likely to impact the absorption and scattering properties of clouds and their lifetime (for example: Mülmenstädt et al., 2021) . Additionally, we have only considered these properties in isolation, eg. one field at a time, and not how they are operating together. This will also be addressed in our upcoming machine learning publication.

To summarise, in this work we have found that considerable radiative biases continue to exist within the ACCESS-AM2 model. By analysing this bias with respect to cloud types and their properties, we find that significant issues with respect to cloud phase remain in the mid-low level cloud regimes for the polar region of the SO, regardless of season, and that even if these cloud types are correctly simulated by the model, the large microphysical biases still persist. We also find compensating errors due to the underestimation of shallow cumulus clouds in favour of stratocumulus clouds. Our results show that significant effort must continue to reduce these SO cloud biases within models.

Code and data availability. Limited model data for this project is available at https://doi.org/10.5281/zenodo.6004062. The code to perform the analysis of this work can be found on Github (accessed 28 June 2022). The model set-up (under licence) can be found by UM users under suite bx400. CERES data can be downloaded from https://ceres.larc.nasa.gov/data/ (last accessed 25th March 2022). MODIS data can be downloaded from https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/61/MCD06COSP_D3_MODIS/ (last accessed 25th March 2022).

Author contributions. SF completed the model simulations, analysis and writing of this work. AP, MM, SA and MW contributed to the planning, ideas and revisions of this paper. MW provided guidance on the model setup

Competing interests. The authors have no competing interests to declare

Acknowledgements. This project received grant funding from the Australian Government as part of the Antarctic Science Collaboration Initiative program, under the Australian Antarctic Program Partnership, ASCI000002. This research was undertaken with the assistance of resources and services from the National Computational Infrastructure (Project jk72), which is supported by the Australian Government. S.F. would like to thank Scott Wales, and the ARC Centre of Excellence for Climate Extremes for their maintenance of virtual environments and code/model support. S.F. would also like to thank Robert Pincus and Alejandro Bodas-Salcedo for early advice on the use of MODIS/COSP products. The authors would like to acknowledge the teams at NASA CERES, NASA Earth data and the CFMIP project for making the tools and data used in this work available. The contribution of S.P.A. was supported by the Australian Antarctic Science project 4292

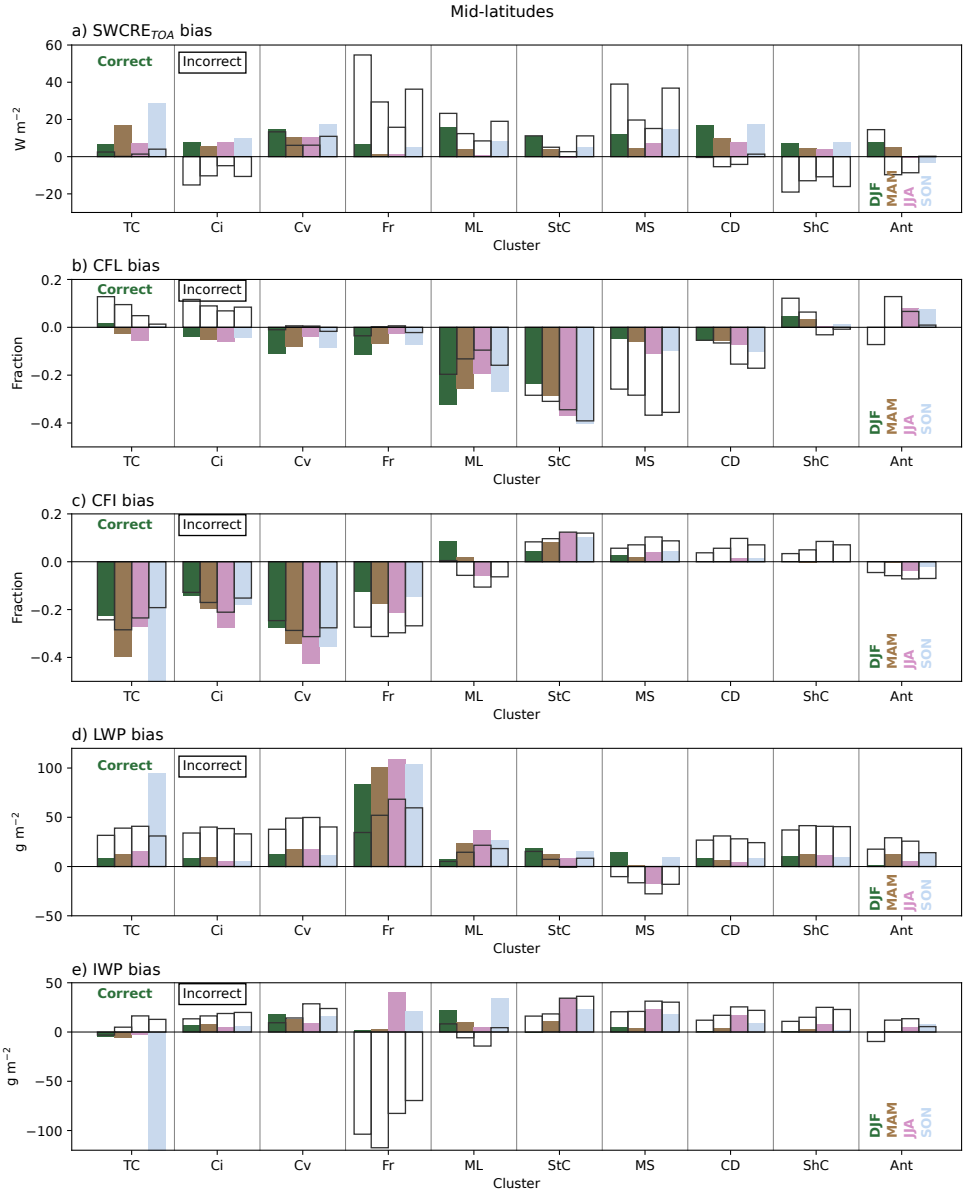

Figure 8. a) Mid-latitude SWCRE$_{TOA}$ biases ($W\,m^{-2}$) averaged over each cloud cluster (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant), seasonally (DJF, dark blue; MAM, green; JJA, pink; SON, light blue), for instances when the MODIS clusters are correctly assigned by ACCESS-AM2 (coloured bars~~with grey outline~~) and are incorrectly assigned (black outline) for the mid-latitude region. ~~The circles along the top and bottom axes (grey for agree, top, black for disagree, bottom) imply the relative importance of each cluster and its assignment with respect to the total number of points for that season. The final column, 'Total' (with hatching) shows the total mean bias over the region (for all instances) for context.~~ b-e) shows the same as plot a) but for the CFL (fraction), CFI (fraction), LWP ($g\,m^{-2}$) and IWP ($g\,m^{-2}$) biases respectively.

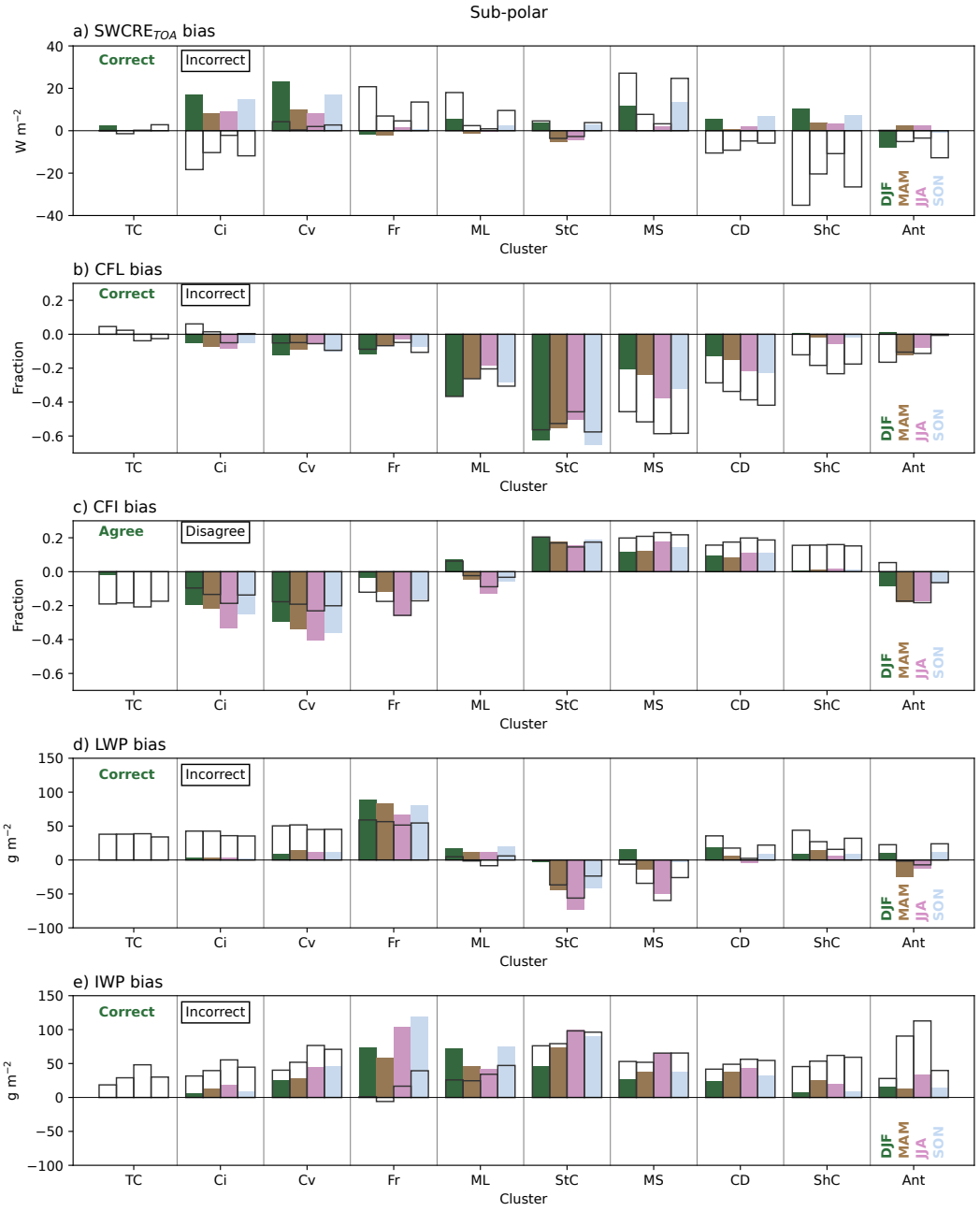

Figure 9. The same as Figure 8, but for the sub-polar region.

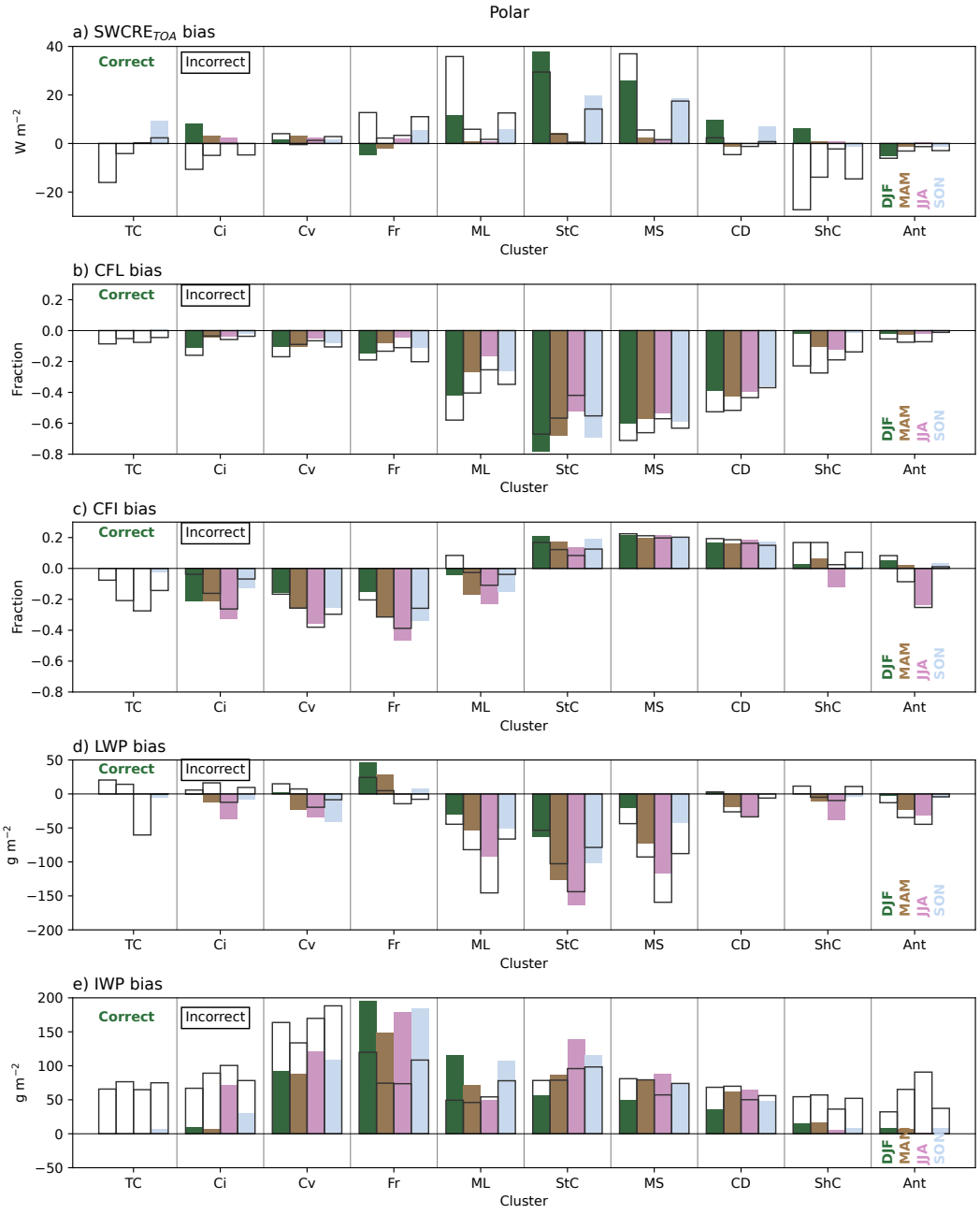

Figure 10. The same as Figure 8, but for the polar region.

Appendix A: Additional figures

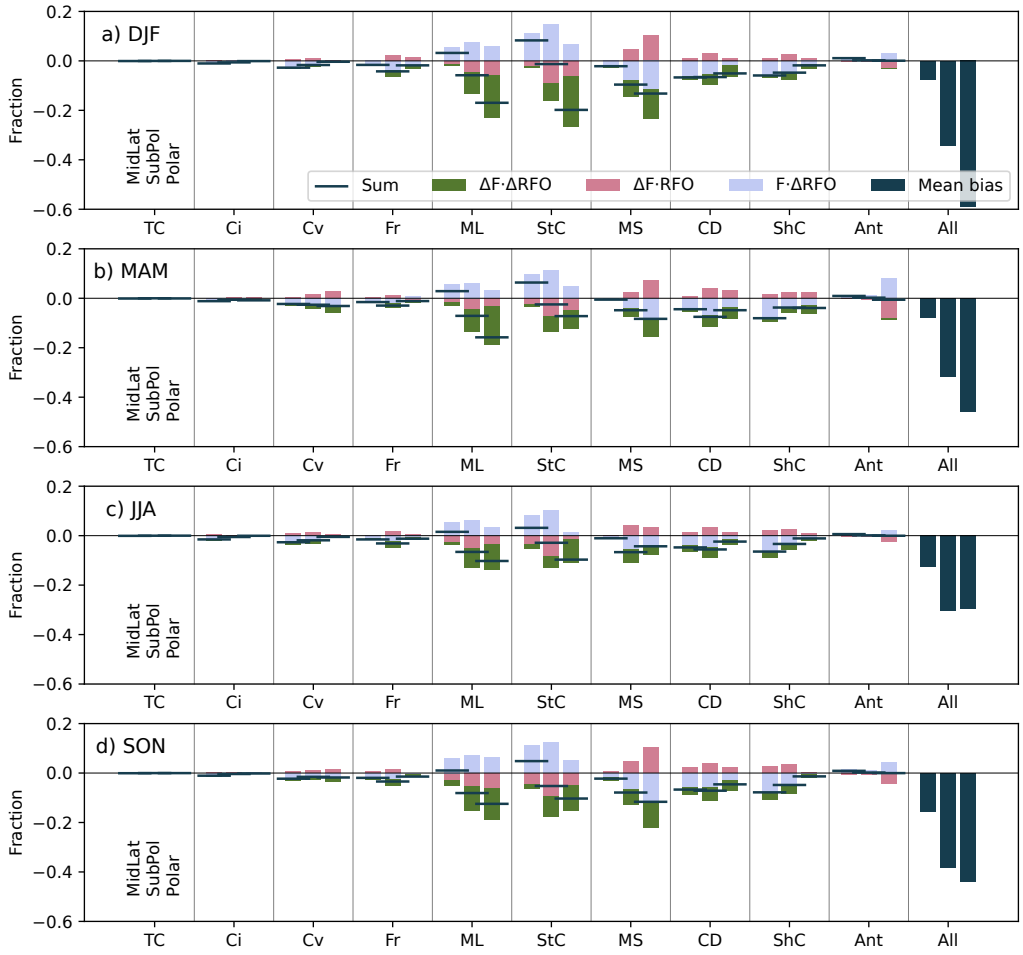

Figure A1. The decomposed mean biases in the CFL (fraction) for each cloud regime (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant) over the three regions from left to right mid-latitudes, sub-polar and polar, for each season (DJF, plot a; MAM, b; JJA, c; SON, d respectively). The sum of the decomposed biases are shown by the horizontal bar, while the the terms of the bias decomposition (see eq. 1) are shown in blue ($F_r^{sat} \cdot \Delta RFO_r$), pink ($\Delta F_r \cdot RFO_r^{sat}$) and green ($\Delta F_r \cdot \Delta RFO_r$). On the far right of each plot is the mean radiative bias for each region/season as a reference

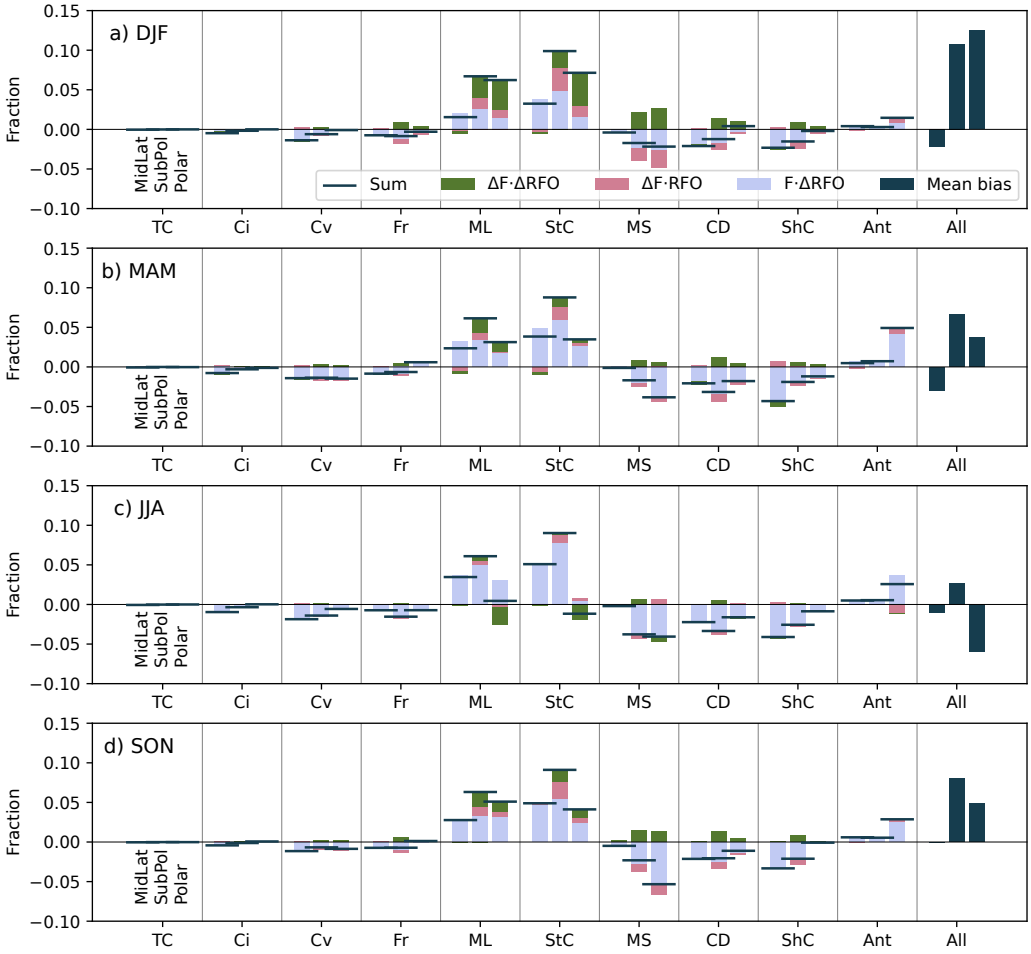

Figure A2. The decomposed mean biases in the CFI (fraction) for each cloud regime (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant) over the three regions from left to right mid-latitudes, sub-polar and polar, for each season (DJF, plot a; MAM, b; JJA, c; SON, d respectively). The sum of the decomposed biases are shown by the horizontal bar, while the the terms of the bias decomposition (see eq. 1) are shown in blue ($F_r^{sat} \cdot \Delta RFO_r$), pink ($\Delta F_r \cdot RFO_r^{sat}$) and green ($\Delta F_r \cdot \Delta RFO_r$). On the far right of each plot is the mean radiative bias for each region/season as a reference

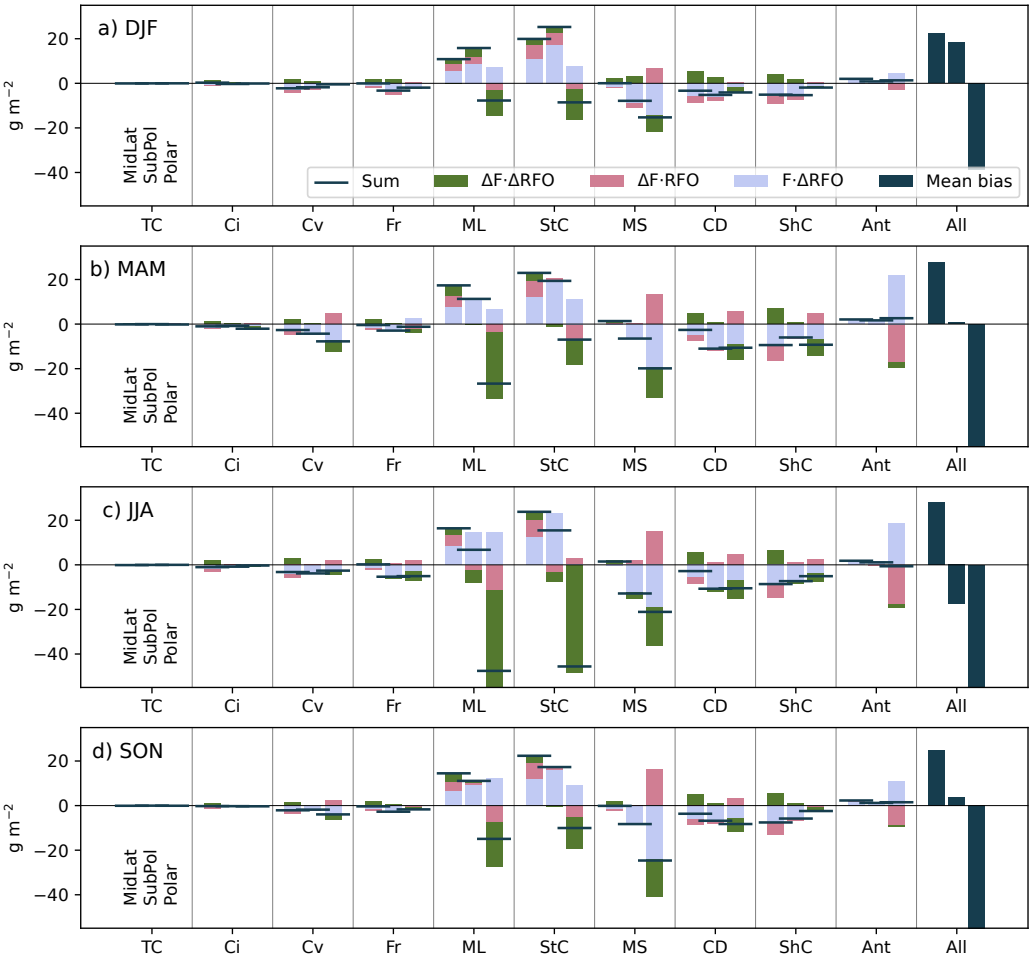

Figure A3. The decomposed mean biases in the LWP (g m^{-2}) for each cloud regime (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant) over the three regions from left to right mid-latitudes, sub-polar and polar, for each season (DJF, plot a; MAM, b; JJA, c; SON, d respectively). The sum of the decomposed biases are shown by the horizontal bar, while the the terms of the bias decomposition (see eq. 1) are shown in blue ($F_r^{sat} \cdot \Delta RFO_r$), pink ($\Delta F_r \cdot RFO_r^{sat}$) and green ($\Delta F_r \cdot \Delta RFO_r$). On the far right of each plot is the mean radiative bias for each region/season as a reference

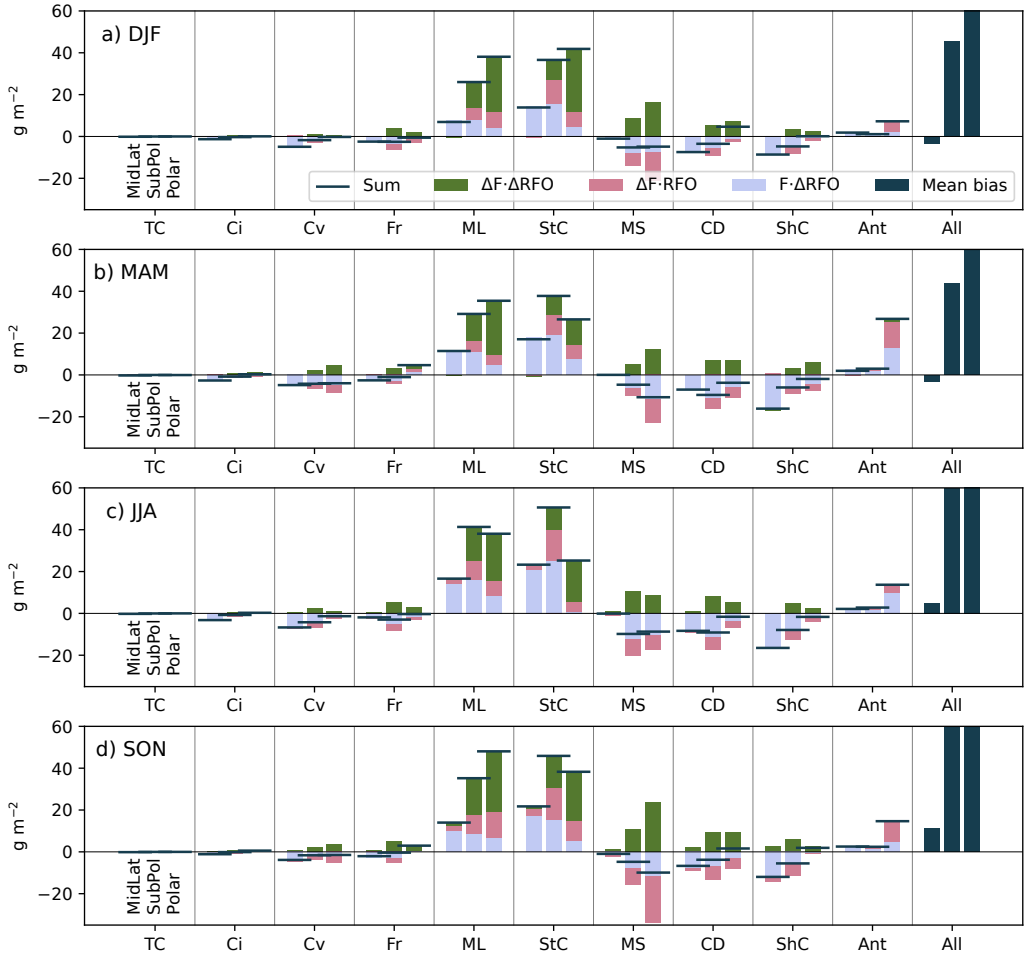

Figure A4. The decomposed mean biases in the IWP (g m^{-2}) for each cloud regime (thin cirrus, TC; cirrus, Ci; convective, Cv; frontal, Fr; mid-level, ML; stratocumulus, StC; marine stratiform, MS; cloud decks, CD; shallow cumulus ShC; Antarctic, Ant) over the three regions from left to right mid-latitudes, sub-polar and polar, for each season (DJF, plot a; MAM, b; JJA, c; SON, d respectively). The sum of the decomposed biases are shown by the horizontal bar, while the the terms of the bias decomposition (see eq. 1) are shown in blue ($F_r^{sat} \cdot \Delta RFO_r$), pink ($\Delta F_r \cdot RFO_r^{sat}$) and green ($\Delta F_r \cdot \Delta RFO_r$). On the far right of each plot is the mean radiative bias for each region/season as a reference

References

Anderberg, M. R.: Cluster Analysis for Applications. A volume in Probability and Mathematical Statistics: A Series of Monographs and Textbooks, vol. 19, Academic Press, New York, https://doi.org/10.1016/C2013-0-06161-0, 1973.

Bender, F. A., Engström, A., Wood, R., and Charlson, R. J.: Evaluation of hemispheric asymmetries in marine cloud radiative properties, Journal of Climate, 30, 4131–4147, https://doi.org/10.1175/JCLI-D-16-0263.1, 2017.

Bi, D., Dix, M., Marsland, S., O'Farrell, S., Sullivan, A., Bodman, R., Law, R., Harman, I., Srbinovsky, J., Rashid, H. A., Dobrohotoff, P., Mackallah, C., Yan, H., Hirst, A., Savita, A., Dias, F. B., Woodhouse, M., Fiedler, R., and Heerdegen, A.: Configuration and spin-up of ACCESS-CM2, the new generation Australian Community Climate and Earth System Simulator Coupled Model, Journal of Southern Hemisphere Earth Systems Science, https://doi.org/10.1071/es19040, 2020.

Bodas-Salcedo, A., Webb, M. J., Bony, S., Chepfer, H., Dufresne, J. L., Klein, S. A., Zhang, Y., Marchand, R., Haynes, J. M., Pincus, R., and John, V. O.: COSP: Satellite simulation software for model assessment, Bulletin of the American Meteorological Society, 92, 1023–1043, https://doi.org/10.1175/2011BAMS2856.1, 2011.

Bodas-Salcedo, A., Williams, K. D., Field, P. R., and Lock, A. P.: The surface downwelling solar radiation surplus over the southern ocean in the met office model: The role of midlatitude cyclone clouds, Journal of Climate, 25, 7467–7486, https://doi.org/10.1175/JCLI-D-11-00702.1, 2012.

Bodas-Salcedo, A., Williams, K. D., Ringer, M. A., Beau, I., Cole, J. N. S., Dufresne, J. L., Koshiro, T., Stevens, B., Wang, Z., and Yokohata, T.: Origins of the solar radiation biases over the Southern Ocean in CFMIP2 models, Journal of Climate, 27, 41–56, https://doi.org/10.1175/JCLI-D-13-00169.1, 2014.

Bodas-Salcedo, A., Andrews, T., Karmalkar, A. V., and Ringer, M. A.: Cloud liquid water path and radiative feedbacks over the Southern Ocean, Geophysical Research Letters, 43, 938–10, https://doi.org/10.1002/2016GL070770, 2016a.

Bodas-Salcedo, A., Hill, P. G., Furtado, K., Williams, K. D., Field, P. R., Manners, J. C., Hyder, P., and Kato, S.: Large contribution of supercooled liquid clouds to the solar radiation budget of the Southern Ocean, Journal of Climate, 29, 4213–4228, https://doi.org/10.1175/JCLI-D-15-0564.1, 2016b.

Bodman, R. W., Karoly, D. J., Dix, M. R., Harman, I. N., Srbinovsky, J., Dobrohotoff, P. B., and Mackallah, C.: Evaluation of CMIP6 AMIP climate simulations with the ACCESS-AM2 model, Journal of Southern Hemisphere Earth Systems Science, https://doi.org/10.1071/ES19033, 2020.

Bony, S. and Dufresne, J. L.: Marine boundary layer clouds at the heart of tropical cloud feedback uncertainties in climate models, Geophysical Research Letters, 32, 1–4, https://doi.org/10.1029/2005GL023851, 2005.

Calinski, T. and Harabasz, J.: A dendrite method for cluster analysis, Communications in Statistics - Theory and Methods, 3, 1–27, https://doi.org/10.1080/03610927408827101, 1974.

Cho, N., Tan, J., and Oreopoulos, L.: Classifying planetary cloudiness with an updated set of modis cloud regimes, Journal of Applied Meteorology and Climatology, 60, 981–997, https://doi.org/10.1175/JAMC-D-20-0247.1, 2021.

Chubb, T. H., Jensen, J. B., Siems, S. T., and Manton, M. J.: In situ observations of supercooled liquid clouds over the Southern Ocean during the HIAPER Pole-to-Pole Observation campaigns, Geophysical Research Letters, 40, 5280–5285, https://doi.org/10.1002/grl.50986, 2013.

Davies, D. L. and Bouldin, D. W.: A Cluster Separation Measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-1, 224–227, https://doi.org/10.1109/TPAMI.1979.4766909, 1979.

Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., Wielicki, B. A., Young, D. F., and Sun, M.: Geostationary enhanced temporal interpolation for ceres flux products, Journal of Atmospheric and Oceanic Technology, 30, 1072–1090, https://doi.org/10.1175/JTECH-D-12-00136.1, 2013.

Doelling, D. R., Sun, M., Nguyen, L. T., Nordeen, M. L., Haney, C. O., Keyes, D. F., and Mlynczak, P. E.: Advances in geostationary-derived longwave fluxes for the CERES synoptic (SYN1deg) product, Journal of Atmospheric and Oceanic Technology, 33, 503–521, https://doi.org/10.1175/JTECH-D-15-0147.1, 2016.

Edwards, J. M. and Slingo, A.: Studies with a flexible new radiation code. I: Choosing a configuration for a large-scale model, Quarterly Journal of the Royal Meteorological Society, 122, 689–719, https://doi.org/10.1002/qj.49712253107, 1996.

Eyring, V., Bony, S., Meehl, G. A., Senior, C. A., Stevens, B., Stouffer, R. J., and Taylor, K. E.: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization, Geoscientific Model Development, 9, 1937–1958, https://doi.org/10.5194/gmd-9-1937-2016, 2016.

Feng, L., Smith, S. J., Braun, C., Crippa, M., Gidden, M. J., Hoesly, R., Klimont, Z., Van Marle, M., Van Den Berg, M., and Van Der Werf, G. R.: The generation of gridded emissions data for CMIP6, Geoscientific Model Development, 13, 461–482, https://doi.org/10.5194/gmd-13-461-2020, 2020.

Fiddes, S. L., Woodhouse, M. T., Nicholls, Z., Lane, T. P., and Schofield, R.: Cloud, precipitation and radiation responses to large perturbations in global dimethyl sulfide, Atmospheric Chemistry and Physics, 18, 10 177–10 198, https://doi.org/10.5194/acp-18-10177-2018, 2018.

Field, P. R. and Wood, R.: Precipitation and cloud structure in midlatitude cyclones, Journal of Climate, 20, 233–254, https://doi.org/10.1175/JCLI3998.1, 2007.

Frey, W. R., Maroon, E. A., Pendergrass, A. G., and Kay, J. E.: Do Southern Ocean Cloud Feedbacks Matter for 21st Century Warming?, Geophysical Research Letters, 44, 447–12, https://doi.org/10.1002/2017GL076339, 2017.

Fricko, O., Havlik, P., Rogelj, J., Klimont, Z., Gusti, M., Johnson, N., Kolp, P., Strubegger, M., Valin, H., Amann, M., Ermolieva, T., Forsell, N., Herrero, M., Heyes, C., Kindermann, G., Krey, V., McCollum, D. L., Obersteiner, M., Pachauri, S., Rao, S., Schmid, E., Schoepp, W., and Riahi, K.: The marker quantification of the Shared Socioeconomic Pathway 2: A middle-of-the-road scenario for the 21st century, Global Environmental Change, 42, 251–267, https://doi.org/10.1016/j.gloenvcha.2016.06.004, 2017.

Gettelman, A. and Sherwood, S. C.: Processes Responsible for Cloud Feedback, https://doi.org/10.1007/s40641-016-0052-8, 2016.

Gettelman, A., Bardeen, C. G., McCluskey, C. S., Järvinen, E., Stith, J., Bretherton, C., McFarquhar, G., Twohy, C., D'Alessandro, J., and Wu, W.: Simulating Observations of Southern Ocean Clouds and Implications for Climate, Journal of Geophysical Research: Atmospheres, 125, https://doi.org/10.1029/2020JD032619, 2020.

Gong, S. L.: A parameterization of sea-salt aerosol source function for sub- and super-micron particles, Global Biogeochemical Cycles, 17, n/a–n/a, https://doi.org/10.1029/2003GB002079, 2003.

Gregory, D. and Rowntree, P. R.: A Mass Flux Convection Scheme with Representation of Cloud Ensemble Characteristics and Stability-Dependent Closure, Monthly Weather Review, 118, 1483–1506, https://doi.org/10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2, 1990.

Hawcroft, M., Haywood, J. M., Collins, M., Jones, A., Jones, A. C., and Stephens, G.: Southern Ocean albedo, inter-hemispheric energy transports and the double ITCZ: global impacts of biases in a coupled model, Climate Dynamics, 48, 2279–2295, https://doi.org/10.1007/s00382-016-3205-5, 2017.

Haynes, J. M., Jakob, C., Rossow, W. B., Tselioudis, G., and Brown, J. B.: Major characteristics of Southern Ocean cloud regimes and their effects on the energy budget, Journal of Climate, 24, 5061–5080, https://doi.org/10.1175/2011JCLI4052.1, 2011.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., Chiara, G., Dahlgren, P., Dee, D., Diamantakis, M., Dragani, R., Flemming, J., Forbes, R., Fuentes, M., Geer, A., Haimberger, L., Healy, S., Hogan, R. J., Hólm, E., Janisková, M., Keeley, S., Laloyaux, P., Lopez, P., Lupu, C., Radnoti, G., Rosnay, P., Rozum, I., Vamborg, F., Villaume, S., and Thépaut, J.: The ERA5 global reanalysis, Quarterly Journal of the Royal Meteorological Society, 146, 1999–2049, https://doi.org/10.1002/qj.3803, 2020.

Hinkelman, L. M. and Marchand, R.: Evaluation of CERES and CloudSat Surface Radiative Fluxes Over Macquarie Island, the Southern Ocean, Earth and Space Science, 7, https://doi.org/10.1029/2020EA001224, 2020.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J. I., Li, M., Liu, L., Lu, Z., Moura, M. C., O'Rourke, P. R., and Zhang, Q.: Historical (1750-2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geoscientific Model Development, 11, 369–408, https://doi.org/10.5194/gmd-11-369-2018, 2018.

Holz, R. E., Platnick, S., Meyer, K., Vaughan, M., Heidinger, A., Yang, P., Wind, G., Dutcher, S., Ackerman, S., Amarasinghe, N., Nagle, F., and Wang, C.: Resolving ice cloud optical thickness biases between CALIOP and MODIS using infrared retrievals, Atmospheric Chemistry and Physics, 16, 5075–5090, https://doi.org/10.5194/acp-16-5075-2016, 2016.

Huang, Y., Siems, S. T., Manton, M. J., Protat, A., and Delanoë, J.: A study on the low-altitude clouds over the Southern Ocean using the DARDAR-MASK, Journal of Geophysical Research Atmospheres, 117, https://doi.org/10.1029/2012JD017800, 2012.

Hubanks, P., Pincus, R., Platnick, S., and Meyer, K.: Level-3 COSP Cloud Properties (MCD06COSP_L3) Combined Terra & Aqua MODIS Global Gridded Product for Climate Studies User Guide, Tech. rep., NASA, https://atmosphere-imager.gsfc.nasa.gov/sites/default/files/ModAtmo/documents/L3_MCD06COSP_UserGuide_v13.pdf, 2020.

Humphries, R. S., Keywood, M. D., Gribben, S., McRobert, I. M., Ward, J. P., Selleck, P., Taylor, S., Harnwell, J., Flynn, C., Kulkarni, G. R., Mace, G. G., Protat, A., Alexander, S. P., and McFarquhar, G.: Southern Ocean latitudinal gradients of cloud condensation nuclei, Atmospheric Chemistry and Physics, 21, 12 757–12 782, https://doi.org/10.5194/acp-21-12757-2021, 2021.

Hurrell, J. W., Hack, J. J., Shea, D., Caron, J. M., and Rosinski, J.: A new sea surface temperature and sea ice boundary dataset for the community atmosphere model, Journal of Climate, 21, 5145–5153, https://doi.org/10.1175/2008JCLI2292.1, 2008.

Hwang, Y. T. and Frierson, D. M.: Link between the double-intertropical convergence zone problem and cloud biases over the southern ocean, Proceedings of the National Academy of Sciences of the United States of America, 110, 4935–4940, https://doi.org/10.1073/pnas.1213302110, 2013.

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., Wood, R. A., Meijers, A. J., Mulcahy, J., Field, P., Furtado, K., Bodas-Salcedo, A., Williams, K. D., Copsey, D., Josey, S. A., Liu, C., Roberts, C. D., Sanchez, C., Ridley, J., Thorpe, L., Hardiman, S. C., Mayer, M., Berry, D. I., and Belcher, S. E.: Critical Southern Ocean climate model biases traced to atmospheric model cloud errors, Nature Communications, 9, https://doi.org/10.1038/s41467-018-05634-2, 2018.

Jakob, C. and Tselioudis, G.: Objective identification of cloud regimes in the Tropical Western Pacific, Geophysical Research Letters, 30, https://doi.org/10.1029/2003GL018367, 2003.

Kay, J. E., Wall, C., Yettella, V., Medeiros, B., Hannay, C., Caldwell, P., and Bitz, C.: Global climate impacts of fixing the Southern Ocean shortwave radiation bias in the Community Earth System Model (CESM), Journal of Climate, 29, 4617–4636, https://doi.org/10.1175/JCLI-D-15-0358.1, 2016.

King, M. D., Menzel, W. P., Kaufman, Y. J., Tanré, D., Gao, B. C., Platnick, S., Ackerman, S. A., Remer, L. A., Pincus, R., and Hubanks, P. A.: Cloud and aerosol properties, precipitable water, and profiles of temperature and water vapor from MODIS, IEEE Transactions on Geoscience and Remote Sensing, 41, 442–456, https://doi.org/10.1109/TGRS.2002.808226, 2003.

Kuma, P., McDonald, A. J., Morgenstern, O., Alexander, S. P., Cassano, J. J., Garrett, S., Halla, J., Hartery, S., Harvey, M. J., Parsons, S., Plank, G., Varma, V., and Williams, J.: Evaluation of Southern Ocean cloud in the HadGEM3 general circulation model and MERRA-2 reanalysis using ship-based observations, Atmospheric Chemistry and Physics, 20, 6607–6630, https://doi.org/10.5194/acp-20-6607-2020, 2020.

Lana, A., Bell, T. G., Simó, R., Vallina, S. M., Ballabrera-Poy, J., Kettle, A. J., Dachs, J., Bopp, L., Saltzman, E. S., Stefels, J., Johnson, J. E., and Liss, P. S.: An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean, Global Biogeochemical Cycles, 25, 1–17, https://doi.org/10.1029/2010GB003850, 2011.

Leinonen, J., Lebsock, M. D., Oreopoulos, L., and Cho, N.: Interregional differences in MODIS-derived cloud regimes, Journal of Geophysical Research, 121, 11 648–11 665, https://doi.org/10.1002/2016JD025193, 2016.

Liss, P. S. and Merlivat, L.: Air-Sea Gas Exchange Rates: Introduction and Synthesis, in: The Role of Air-Sea Exchange in Geochemical Cycling, edited by Buat-Ménard, P., pp. 113–127, Springer Netherlands, Dordrecht, https://doi.org/10.1007/978-94-009-4738-2_5, 1986.

Listowski, C., Delanoë, J., Kirchgaessner, A., Lachlan-Cope, T., and King, J.: Antarctic clouds, supercooled liquid water and mixed phase, investigated with DARDAR: Geographical and seasonal variations, Atmospheric Chemistry and Physics, 19, 6771–6808, https://doi.org/10.5194/acp-19-6771-2019, 2019.

Mace, G. G. and Protat, A.: Clouds over the Southern Ocean as observed from the R/V investigator during CAPRICORN. Part I: Cloud occurrence and phase partitioning, Journal of Applied Meteorology and Climatology, 57, 1783–1803, https://doi.org/10.1175/JAMC-D-17-0194.1, 2018.

Mace, G. G., Protat, A., Humphries, R. S., Alexander, S. P., McRobert, I. M., Ward, J., Selleck, P., Keywood, M., and McFarquhar, G. M.: Southern Ocean Cloud Properties Derived From CAPRICORN and MARCUS Data, Journal of Geophysical Research: Atmospheres, 126, https://doi.org/10.1029/2020JD033368, 2021.

Mann, G. W., Carslaw, K. S., Spracklen, D. V., Ridley, D. A., Manktelow, P. T., Chipperfield, M. P., Pickering, S. J., and Johnson, C. E.: Description and evaluation of GLOMAP-mode: a modal global aerosol microphysics model for the UKCA composition-climate model, Geosci. Model Dev, 3, 519–551, https://doi.org/10.5194/gmd-3-519-2010, 2010.

Mann, G. W., Carslaw, K. S., Ridley, D. A., Spracklen, D. V., Pringle, K. J., Merikanto, J., Korhonen, H., Schwarz, J. P., Lee, L. A., Manktelow, P. T., Woodhouse, M. T., Schmidt, A., Breider, T. J., Emmerson, K. M., Reddington, C. L., Chipperfield, M. P., and Pickering, S. J.: Intercomparison of modal and sectional aerosol microphysics representations within the same 3-D global chemical transport model, Atmospheric Chemistry and Physics, 12, 4449–4476, https://doi.org/10.5194/acp-12-4449-2012, 2012.

Marchant, B., Platnick, S., Meyer, K., Arnold, G. T., and Riedi, J.: MODIS Collection 6 shortwave-derived cloud phase classification algorithm and comparisons with CALIOP, Atmospheric Measurement Techniques, 9, 1587–1599, https://doi.org/10.5194/amt-9-1587-2016, 2016.

Mason, S., Fletcher, J. K., Haynes, J. M., Franklin, C., Protat, A., and Jakob, C.: A hybrid cloud regime methodology used to evaluate Southern Ocean cloud and shortwave radiation errors in ACCESS, Journal of Climate, 28, 6001–6018, https://doi.org/10.1175/JCLI-D-14-00846.1, 2015.

McDonald, A. J., Cassano, J. J., Jolly, B., Parsons, S., and Schuddeboom, A.: An automated satellite cloud classification scheme using self-organizing maps: Alternative ISCCP weather states, Journal of Geophysical Research, 121, 009–13, https://doi.org/10.1002/2016JD025199, 2016.

McFarquhar, G. M., Bretherton, C. S., Marchand, R., Protat, A., DeMott, P. J., Alexander, S. P., Roberts, G. C., Twohy, C. H., Toohey, D., Siems, S., Huang, Y., Wood, R., Rauber, R. M., Lasher-Trapp, S., Jensen, J., Stith, J. L., Mace, J., Um, J., Järvinen, E., Schnaiter, M., Gettelman, A., Sanchez, K. J., McCluskey, C. S., Russell, L. M., McCoy, I. L., Atlas, R. L., Bardeen, C. G., Moore, K. A., Hill, T. C. J., Humphries, R. S., Keywood, M. D., Ristovski, Z., Cravigan, L., Schofield, R., Fairall, C., Mallet, M. D., Kreidenweis, S. M., Rainwater, B., D'Alessandro, J., Wang, Y., Wu, W., Saliba, G., Levin, E. J. T., Ding, S., Lang, F., Truong, S. C. H., Wolff, C., Haggerty, J., Harvey, M. J., Klekociuk, A. R., and McDonald, A.: 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, E894–E928, https://doi.org/10.1175/BAMS-D-20-0132.1, 2021.

Min, Q., Joseph, E., Lin, Y., Min, L., Yin, B., Daum, P. H., Kleinman, L. I., Wang, J., and Lee, Y.-N.: Comparison of MODIS cloud microphysical properties with in-situ measurements over the Southeast Pacific, Atmospheric Chemistry and Physics, 12, 11 261–11 273, https://doi.org/10.5194/acp-12-11261-2012, 2012.

Mulcahy, J. P., Jones, C., Sellar, A., Johnson, B., Boutle, I. A., Jones, A., Andrews, T., Rumbold, S. T., Mollard, J., Bellouin, N., Johnson, C. E., Williams, K. D., Grosvenor, D. P., and McCoy, D. T.: Improved Aerosol Processes and Effective Radiative Forcing in HadGEM3 and UKESM1, Journal of Advances in Modeling Earth Systems, 10, 2786–2805, https://doi.org/10.1029/2018MS001464, 2018.

Mülmenstädt, J., Salzmann, M., Kay, J. E., Zelinka, M. D., Ma, P.-L., Nam, C., Kretzschmar, J., Hörnig, S., and Quaas, J.: An underestimated negative cloud feedback from cloud lifetime changes, Nature Climate Change, 11, 508–513, https://doi.org/10.1038/s41558-021-01038-1, 2021.

Noble, S. R. and Hudson, J. G.: MODIS comparisons with northeastern Pacific in situ stratocumulus microphysics, Journal of Geophysical Research: Atmospheres, 120, 8332–8344, https://doi.org/10.1002/2014JD022785, 2015.

Oreopoulos, L., Cho, N., Lee, D., Kato, S., and Huffman, G. J.: An examination of the nature of global MODIS cloud regimes, Journal of Geophysical Research: Atmospheres, 119, 8362–8383, https://doi.org/10.1002/2013JD021409, 2014.

Oreopoulos, L., Cho, N., Lee, D., and Kato, S.: Radiative effects of global MODIS cloud regimes, Journal of Geophysical Research: Atmospheres, 121, 2299–2317, https://doi.org/10.1002/2015JD024502, 2016.

Painemal, D. and Zuidema, P.: Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements, Journal of Geophysical Research Atmospheres, 116, https://doi.org/10.1029/2011JD016155, 2011.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research 12, 12, 2825–2830, 2011.

Pincus, R., Platnick, S., Ackerman, S. A., Hemler, R. S., and Patrick Hofmann, R. J.: Reconciling simulated and observed views of clouds: MODIS, ISCCP, and the limits of instrument simulators, Journal of Climate, 25, 4699–4720, https://doi.org/10.1175/JCLI-D-11-00267.1, 2012.

Platnick, S., Meyer, K. G., King, M. D., Wind, G., Amarasinghe, N., Marchant, B., Arnold, G. T., Zhang, Z., Hubanks, P. A., Holz, R. E., Yang, P., Ridgway, W. L., and Riedi, J.: The MODIS Cloud Optical and Microphysical Products: Collection 6 Updates and Examples From Terra and Aqua, IEEE Transactions on Geoscience and Remote Sensing, 55, 502–525, https://doi.org/10.1109/TGRS.2016.2610522, 2017.

Protat, A., Schulz, E., Rikus, L., Sun, Z., Xiao, Y., and Keywood, M. D.: Shipborne observations of the radiative effect of Southern Ocean clouds, Journal of Geophysical Research: Atmospheres, 122, 318–328, https://doi.org/10.1002/2016JD026061, 2017.

Saponaro, G., Sporre, M. K., Neubauer, D., Kokkola, H., Kolmonen, P., Sogacheva, L., Arola, A., De Leeuw, G., Karset, I. H., Laaksonen, A., and Lohmann, U.: Evaluation of aerosol and cloud properties in three climate models using MODIS observations and its corresponding COSP simulator, as well as their application in aerosol-cloud interactions, Atmospheric Chemistry and Physics, 20, 1607–1626, https://doi.org/10.5194/acp-20-1607-2020, 2020.

Schuddeboom, A., McDonald, A. J., Morgenstern, O., Harvey, M., and Parsons, S.: Regional Regime-Based Evaluation of Present-Day General Circulation Model Cloud Simulations Using Self-Organizing Maps, Journal of Geophysical Research: Atmospheres, 123, 4259–4272, https://doi.org/10.1002/2017JD028196, 2018.

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

Simmons, J. B., Humphries, R. S., Wilson, S. R., Chambers, S. D., Williams, A. G., Griffiths, A. D., McRobert, I. M., Ward, J. P., Keywood, M. D., and Gribben, S.: Summer aerosol measurements over the East Antarctic seasonal ice zone, Atmospheric Chemistry and Physics, 21, 9497–9513, https://doi.org/10.5194/acp-21-9497-2021, 2021.

Sporre, M. K., O'Connor, E. J., Håkansson, N., Thoss, A., Swietlicki, E., and Petäjä, T.: Comparison of MODIS and VIIRS cloud properties with ARM ground-based observations over Finland, Atmospheric Measurement Techniques, 9, 3193–3203, https://doi.org/10.5194/amt-9-3193-2016, 2016.

Tselioudis, G., Rossow, W., Zhang, Y., and Konsta, D.: Global weather states and their properties from passive and active satellite cloud retrievals, Journal of Climate, 26, 7734–7746, https://doi.org/10.1175/JCLI-D-13-00024.1, 2013.

Tselioudis, G., Rossow, W. B., Jakob, C., Remillard, J., Tropf, D., and Zhang, Y.: Evaluation of Clouds, Radiation, and Precipitation in CMIP6 Models Using Global WeatherStates Derived from ISCCP-H Cloud Property Data, Journal of Climate, pp. 1–42, https://doi.org/10.1175/JCLI-D-21-0076.1, 2021.

Van Marle, M. J., Kloster, S., Magi, B. I., Marlon, J. R., Daniau, A. L., Field, R. D., Arneth, A., Forrest, M., Hantson, S., Kehrwald, N. M., Knorr, W., Lasslop, G., Li, F., Mangeon, S., Yue, C., Kaiser, J. W., and Van Der Werf, G. R.: Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750-2015), https://doi.org/10.5194/gmd-10-3329-2017, 2017.

Walters, D., Baran, A. J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J., Furtado, K., Hill, P., Lock, A., Manners, J., Morcrette, C., Mulcahy, J., Sanchez, C., Smith, C., Stratton, R., Tennant, W., Tomassini, L., Van Weverberg, K., Vosper, S., Willett, M., Browse, J., Bushell, A., Carslaw, K., Dalvi, M., Essery, R., Gedney, N., Hardiman, S., Johnson, B., Johnson, C., Jones, A., Jones, C., Mann, G., Milton, S., Rumbold, H., Sellar, A., Ujiie, M., Whitall, M., Williams, K., and Zerroukat, M.: The Met Office Unified Model Global Atmosphere 7.0/7.1 and JULES Global Land 7.0 configurations, Geoscientific Model Development, 12, 1909–1963, https://doi.org/10.5194/gmd-12-1909-2019, 2019.

Webb, M. J., Andrews, T., Bodas-Salcedo, A., Bony, S., Bretherton, C. S., Chadwick, R., Chepfer, H., Douville, H., Good, P., Kay, J. E., Klein, S. A., Marchand, R., Medeiros, B., Pier Siebesma, A., Skinner, C. B., Stevens, B., Tselioudis, G., Tsushima, Y., and Watanabe, M.:

The Cloud Feedback Model Intercomparison Project (CFMIP) contribution to CMIP6, Geoscientific Model Development, 10, 359–384, https://doi.org/10.5194/gmd-10-359-2017, 2017.

Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Elsevier, 3rd edn., https://doi.org/10.1016/B978-0-12-385022-5.00026-9, 2011.

Williams, K. D. and Tselioudis, G.: GCM intercomparison of global cloud regimes: Present-day evaluation and climate change response, Climate Dynamics, 29, 231–250, https://doi.org/10.1007/s00382-007-0232-2, 2007.

Williams, K. D. and Webb, M. J.: A quantitative performance assessment of cloud regimes in climate models, Climate Dynamics, 33, 141–157, https://doi.org/10.1007/s00382-008-0443-1, 2009.

Wilson, D. R., Bushell, A. C., Kerr-Munslow, A. M., Price, J. D., and Morcrette, C. J.: PC2: A prognostic cloud fraction and condensation scheme. I: Scheme description, Quarterly Journal of the Royal Meteorological Society, 134, 2093–2107, https://doi.org/10.1002/qj.333, 2008.

Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M., and Thuburn, J.: An inherently mass-conserving semi-implicit semi-Lagrangian discretization of the deep-atmosphere global non-hydrostatic equations, Quarterly Journal of the Royal Meteorological Society, 140, 1505–1520, https://doi.org/10.1002/qj.2235, 2014.

Woodward, S.: Modeling the atmospheric life cycle and radiative impact of mineral dust in the Hadley Centre climate model, Journal of Geophysical Research: Atmospheres, 106, 18 155–18 166, https://doi.org/10.1029/2000JD900795, 2001.

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophysical Research Letters, 47, 1–12, https://doi.org/10.1029/2019GL085782, 2020.

Zhang, K., Wan, H., Liu, X., Ghan, S. J., Kooperman, G. J., Ma, P. L., Rasch, P. J., Neubauer, D., and Lohmann, U.: Technical note: On the use of nudging for aerosol-climate model intercomparison studies, Atmospheric Chemistry and Physics, 14, 8631–8645, https://doi.org/10.5194/acp-14-8631-2014, 2014.