

Geophysical Research Letters



RESEARCH LETTER

10.1029/2018GL081796

Key Points:

- Some CMIP5 models capture Antarctic sea ice area, but quantifying their dynamic and thermodynamic processes reveals serious deficiencies
- One model is right for the right reasons; realistic Antarctic sea ice can be simulated using existing model physics and resolution
- Realistic representation of sea ice processes in this model is only possible because it has realistic sea ice drift

Supporting Information:

- Supporting Information S1
- Table S1

Correspondence to:

C. R. Holmes,
calmes@bas.ac.uk

Citation:

Holmes, C. R., Holland, P. R., & Bracegirdle, T. J. (2019). Compensating biases and a noteworthy success in the CMIP5 representation of Antarctic sea ice processes. *Geophysical Research Letters*, 46. <https://doi.org/10.1029/2018GL081796>

Received 21 DEC 2018

Accepted 22 MAR 2019

Accepted article online 28 MAR 2019

Compensating Biases and a Noteworthy Success in the CMIP5 Representation of Antarctic Sea Ice Processes

C.R. Holmes¹ , P.R. Holland¹, and T.J. Bracegirdle¹

¹British Antarctic Survey, Cambridge, UK

Abstract Coupled Model Intercomparison Project phase 5 (CMIP5) climate models simulate a wide range of historical sea ice areas. Even models with areas close to observed values may contain compensating errors, affecting reliability of their projections. This study focuses on the seasonal cycle of sea ice, including analysis of model concentration budgets. Many models have insufficient autumn ice growth, leading to large winter biases. A subset of models accurately represent sea ice evolution year-round. However, comparing their winter ice concentration budget to observations reveals a range of behaviors. At least one model has an accurate ice budget, which is only possible due to realistic ice drifts. The CMIP5 generation of model physics and resolution is therefore structurally capable of accurately representing processes in Antarctic sea ice. This implies that substantially improved projections of Antarctic dense ocean water formation and ice sheet melting are possible with appropriate subsetting of existing climate models.

Plain Language Summary Antarctic sea ice simulated by numerical climate models generally exhibits large differences compared to satellite-based estimates. Therefore, attempts to use these models to understand how Antarctic sea ice—and as a result Antarctic climate, sea level rise, and ocean circulation—will change in the future are severely compromised. This study looks beyond standard measures of sea ice such as the area of ocean covered by ice. Instead, we calculate how much of the change in sea ice over the course of a year is caused by local melting and freezing on the one hand and ice transport (for example by winds and ocean currents) on the other. Of an already small number of models which do show a realistic amount of sea ice, we identify that some do so for the wrong reasons. However, there is one model using established model components which shows realistic behavior. This model is able to match observations of sea ice drift. This demonstrates that the existing climate models are capable of representing Antarctic sea ice without further increases in sophistication. Focusing on such models for climate projections may improve the reliability of projections of the future.

1. Introduction

Antarctic sea ice constitutes a critical part of the global climate system. Over the seasonal cycle, it changes in area by a factor of 6 through highly coupled ice, atmosphere, and ocean processes. Variability and trends in ice extent are of major importance to regional climate (e.g., Turner, Bracegirdle, et al., 2013), ecosystems (e.g., Cavanagh et al., 2017; Jenouvrier et al., 2014), and near-surface atmospheric conditions across the continent (e.g., Krinner et al., 2014) which have implications for the mass balance of the ice sheet. Sea ice formation, transport, and melt play an important role in setting the salinity structure of the Southern Ocean, affecting global ocean circulation (e.g., Abernathy et al., 2016; Haumann et al., 2016).

Projecting future changes in Antarctic sea ice is therefore of paramount importance. However, the models of the Coupled Model Intercomparison Project phase 5 (CMIP5) (Taylor et al., 2012) vary widely in their simulation of Antarctic sea ice (e.g., Turner, Maksym, et al., 2013; Zunz et al., 2013). Known discrepancies from satellite observations also include recent trends (although this may be a result of internal variability; Swart & Fyfe, 2013; Zunz et al., 2013) and the frequency distribution of ice concentration (Roach et al., 2018).

It has long been established (Bracegirdle et al., 2015; Rääsänen, 2007) that there is a relationship between a model's historical sea ice area and how much the Antarctic region warms under greenhouse gas forcing. This implies that large sea ice biases are prohibitive in making reliable projections of future change. Indeed, the Intergovernmental Panel on Climate Change Fifth Assessment Report states that confidence is low in projections of Antarctic sea ice due to biases in representation of historical climate (Collins et al., 2013). However, a key question is whether models with small sea ice area biases provide more trustworthy

©2019. The Authors.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

projections? To address this, one should establish whether the few models that simulate realistic sea ice area do so for the right reasons. Assessing processes in models with realistic ice cover is therefore necessary.

Holland and Kwok (2012) introduced a methodology to decompose the daily evolution of sea ice concentration (SIC) into contributions from thermodynamic and dynamic processes, using satellite-derived measurements of ice drift and concentration. They used this to quantify large-scale processes in autumn and winter sea ice, such as divergence and freezing in the pack, advection to the ice edge, and melting at the boundary. Uotila et al. (2014) applied this technique to assess the CMIP5 models ACCESS1-0 and ACCESS1-3 (supporting information Table S1) and the ACCESS-OM ocean-sea ice model forced with prescribed atmospheric conditions. They found that the models had too strong ice edge advection as a result of ice drift biases, but this was balanced by excessive melting to produce realistic ice evolution. They noted that observational uncertainties are large near the ice edge. Lecomte et al. (2016) examined the ice concentration budget in IPSL-CM5A-MR and CCSM4, highlighting advective biases at the sea ice edge and underestimations of ice velocity divergence in the pack. Schroeter et al. (2018) analyzed the ice volume budget of the CMIP5 models, which provides a more comprehensive picture of sea ice processes, but model evaluation was not possible in this case due to the lack of ice volume observations.

Holland and Kimura (2016) used new satellite data to produce new observed SIC budgets, providing an opportunity to extend the analyses of Uotila et al. (2014) and Lecomte et al. (2016) to the full seasonal cycle and to higher spatial resolution. Here, we address how representative the biases found in these earlier papers are of the full CMIP5 multimodel ensemble of coupled climate models, both in terms of seasonal cycles and in the underlying budget terms. Our key questions are (1) Do models with a realistic annual cycle of ice area have compensating biases in ice processes, and are any models sufficiently realistic to suggest reliability in projections? (2) Where process errors exist, can they be tied to dynamic or thermodynamic processes, and is there a clear link to the sea ice model used, to model resolution, or to forcing from coupled components?

2. Data and Methods

2.1. Model and Observational Data

For observed budgets, we use daily data from satellite observations as in Holland and Kimura (2016, henceforth HK16). Drift fields were derived from Advanced Microwave Scanning Radiometer for the Earth Observing System (AMSR-E) brightness temperatures on a 60-km polar stereographic grid. SIC were derived from AMSR-E brightness temperatures on a 12.5-km grid using the Enhanced NASA Team algorithm (Cavalieri et al., 2014) and binned onto the same 60-km grid. Analysis is conducted during the period 2003–2010, for which AMSR-E data are available. HK16 assess the sensitivity of the derived observational budgets to this choice of input data sets, arguing that their chosen data are the best available.

For model budgets, we use CMIP5 data from the simulations with historical forcing prior to 2005 and RCP4.5 (Meinshausen et al., 2011) forcing from 2006 onward. Models are analyzed if they have daily data for both simulations for SIC and ice drift (CMIP5 variables “sic,” “usi,” and “vsi”), with the exception of CanESM and MIROC-ESM, due to their low resolution. This leaves 23 models from 13 modeling centers (full information, including the ensemble members used and a summary of the model grids, are in supporting information Text S1 and Table S1). Budgets are calculated on the polar stereographic grid of the observations, so linear interpolation is first performed using the `scipy` `griddata` routine. Velocities are rotated into the polar stereographic coordinate directions before regridding.

To aid in interpretation of the budgets, we compare European Centre for Medium-Range Weather Forecasts Re-Analysis-Interim reanalysis (Dee et al., 2011) monthly mean sea level pressures and 10-m wind components with equivalent CMIP5 diagnostics.

2.2. Budget Calculations

HK16 note that the tracking procedure for ice drifts yields fewer data at the ice edge. To give the fairest comparison between observations and models, our analysis masks all data where SIC < 15%, in both models and observations. Fifteen percent is a standard threshold in sea ice analyses, and its use here is further supported by the limited number of drift observations at these low concentrations (supporting information Figure S1).

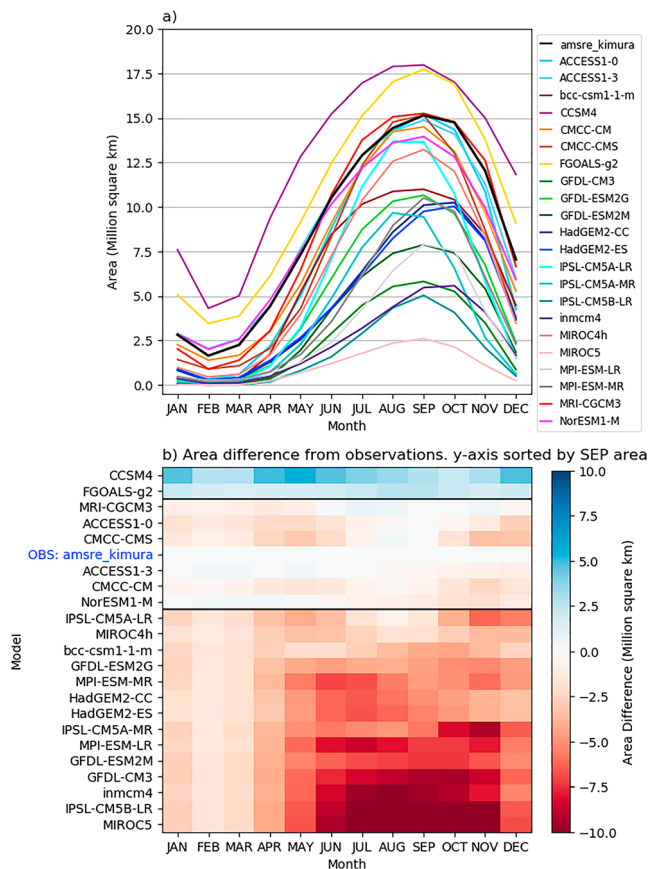


Figure 1. Seasonal evolution of ice area in CMIP5 models. (a) Seasonal cycle of ice area in the 22 CMIP5 models with data available for budget evaluation. (b) Monthly difference from observations for each model; models are ranked by September sea ice area. Black horizontal lines demark the subset of six models which are considered in subsequent analysis. For both plots, areas are calculated on the polar stereographic grid using a common land mask. CMIP5 = Coupled Model Intercomparison Project phase 5.

We calculate the terms in the SIC budget equation:

$$\frac{\partial C}{\partial t} = -\mathbf{u} \cdot \nabla C - C \nabla \cdot \mathbf{u} + \text{residual},$$

where C is SIC and \mathbf{u} is the vector drift field. The budget expresses the change in concentration (left-hand side) as a sum of components from advection $\mathbf{u} \cdot \nabla C$, divergence $C \nabla \cdot \mathbf{u}$, both directly calculated, and the residual. The residual encapsulates both thermodynamics (melting and freezing) and redistribution processes, such as ridging, whereby ice thickness is gained at the expense of concentration. HK16 demonstrated that, as also argued by Uotila et al. (2014), the residual is overwhelmingly thermodynamic; exceptions are convergent regions of high concentration near the coast.

To calculate budgets, we first apply a 7×7 cell square-window filter to both drift and SIC in both the models and observations, in order to smooth grid-scale noise in the ice drift observations, and in both SIC and drift in the models. The land mask and low SIC mask are re-applied to the smoothed fields. The budget terms are then calculated daily as in HK16, applying a 3-day time mean to advection and divergence terms. Finally, monthly averages are calculated. The output fields are masked using the divergence field both before and after the time average. Divergence has the smallest coverage because it uses spatial derivatives of velocity which, in observations, are not available everywhere there is SIC data.

3. Results

3.1. Ice Area

Figure 1 shows modeled and satellite-derived ice areas for the budget period 2003–2010; there are negative sea ice area biases in most models analyzed here, which are large in many cases. Previous studies (e.g., Turner, Bracegirdle, et al., 2013; Zunz et al., 2013) suggested that the multimodel mean ice extent is only slightly lower than observations. The set of models assessed here (supporting information Table S1) differs from these earlier studies, due to incomplete model availability in the earlier studies

and the restriction to models with daily data in the present study. In addition, earlier studies used the period 1979–2005; differing trends in models and observations (Zunz et al., 2013) mean that the later 2003–2010 period generally results in lower modeled ice coverage (supporting information Figure S2) and higher observed ice coverage.

Absolute biases are largest in winter and spring (June to November, Figure 1b). The biases in summer are also large, and greatest in relative terms, with some models having virtually no summer sea ice (Figure 1a). Since the ice concentration budget examines the rate of change of SIC, dC/dt (HK16), it is helpful to frame biases in terms of the seasonal evolution of ice area. Negative biases increase in magnitude between March and July in several models, implying insufficient ice expansion, and the consequent low bias in winter maximum ice area subsequently implies insufficient melt between December and February (Figure 1b).

As a starting point in our analysis of whether models get the right ice area for the right reasons, we identify a subset of models with close-to-observed ice area. These models are then examined in more detail using budgets and other diagnostics to determine whether they also represent key processes correctly. Monthly anomalies from observations (Figure 1b) highlight two possible classes of models. One group, encompassing ACCESS and CMCC models, MRI-CGCM3, and NorESM1-M, simulates ice area reasonably well year-round. The other group includes a number of models which have sizeable biases in all months, negative in all models except FGOALS-g2 and CCSM4. Only small differences were found between the members of single-model ensembles (Figures S2 and S3), which implies that the differing biases are the result of

model structural differences, rather than internal variability. The class separation is supported by a quantitative analysis; a model is defined as “good” if the 95% uncertainty estimate of the observed ice area (calculated from the mean and standard deviation of the eight years of data) encompasses the model mean for any month. This is also the subset with lowest root-mean-square error as calculated from the 12 monthly climatologies.

We next seek to understand whether total ice area is a useful indicator of whether a model is correctly simulating ice processes, by examining budget processes in the good models. We use the HK16 data to perform the first comparison of the full seasonal cycle, but we also focus particularly on September since (i) the amplitude of the annual cycle is dominated by the September maximum, (ii) the winter budget is a test of both dynamics and thermodynamics, unlike the thermodynamics-dominated budgets in other months, and (iii) most data are available for winter months.

3.2. Ice Concentration and Drift

Figure 2 shows maps of September SIC, ice drift, and budget terms, for observations and the good model subset. For ACCESS and CMCC we only show the model configuration which performs best year-round, namely, ACCESS1-3 and CMCC-CM, as the pairs are similar in winter for both ice area and budget terms (supporting information Figures S4–S7). As expected from the ice areas (Figure 1), all models have ice cover which is broadly consistent with the satellite observations (Figure 2a1–2a5).

Ice drifts are largest near the ice edge in both satellite observations and models (Figure 2b1–2b5), with three of the models displaying generally faster drifts than observed. The observed and modeled drifts show known features such as coastal easterlies, export in the western Ross and Weddell seas, generally northward drift, and links to the atmospheric circulation (sea level pressure contours in Figures 2d–2f). However, MRI-CGCM3 simulates near-zero drift (Figure 2b4) and lacks these features; as pointed out by Marzocchi and Jansen (2017), this is due to the erroneous zero wind drag over ice in this model, such that the ice drift is largely decoupled from the wind field.

3.3. Ice Concentration Budgets

Focusing initially on observations (first column of Figure 2; see also HK16), net September dC/dt is generally small (Figure 2c1). There are largely compensating regions of positive dC/dt (blue) and negative dC/dt (red), consistent with small total area changes in September, the month of maximum area. Advection contributes positively to the budget at the ice edge (a SIC source; Figure 2d1, blue). Sea level pressure contours, and vector ice drifts, demonstrate that this arises from wind-driven advection of ice down the concentration gradient. In the pack, divergent motion acts as a local sink of ice concentration (Figure 2e1, red), but with a compensating positive residual (thermodynamic) term (Figure 2f1, blue) that implies an ice concentration source through freezing. A negative residual at the ice edge in the Ross and Weddell seas, where the advection term is positive, implies that a sink of ice due to melting balances the source due to advection and thus constrains the ice edge. Echoing HK16, these results demonstrate that dynamics and thermodynamics are tightly coupled in the Antarctic in winter.

Most of the models show compensating regions of ice gain and loss (Figures 2c2–2c5), although MRI-CGCM3 has very small values everywhere, consistent with the lack of ice drift. Similarly, all models broadly capture a positive advection term at the ice edge, which is linked to northeastward advection (drift vectors). The CMCC models display a weak localized advective ice sink (red) in the western Ross and Weddell seas, which the sea level pressure links to winds advecting ice up-gradient. In general, there is divergent ice loss in the pack (red, Figures 2e2–2e5). This varies greatly between models in both magnitude and smoothness. Convergent features near the Antarctic Peninsula relate to ice drift toward the coast (HK16). Again, the divergence in MRI-CGCM3 is clearly deficient.

To elucidate the contribution of dynamics (divergence and advection) and thermodynamics (the residual) to the total ice budget, we convert the grid point terms to units of square kilometers per month and spatially integrate the time average of each term. A common land mask is applied before integrating terms, so that identified discrepancies are not directly a result of differences in ocean area due to model resolution.

Figure 3 shows these results. Again, we focus first on observations (left column). The black cross shows that ice expansion (integrated dC/dt) is weakly positive in September, with ice area slightly increasing during the

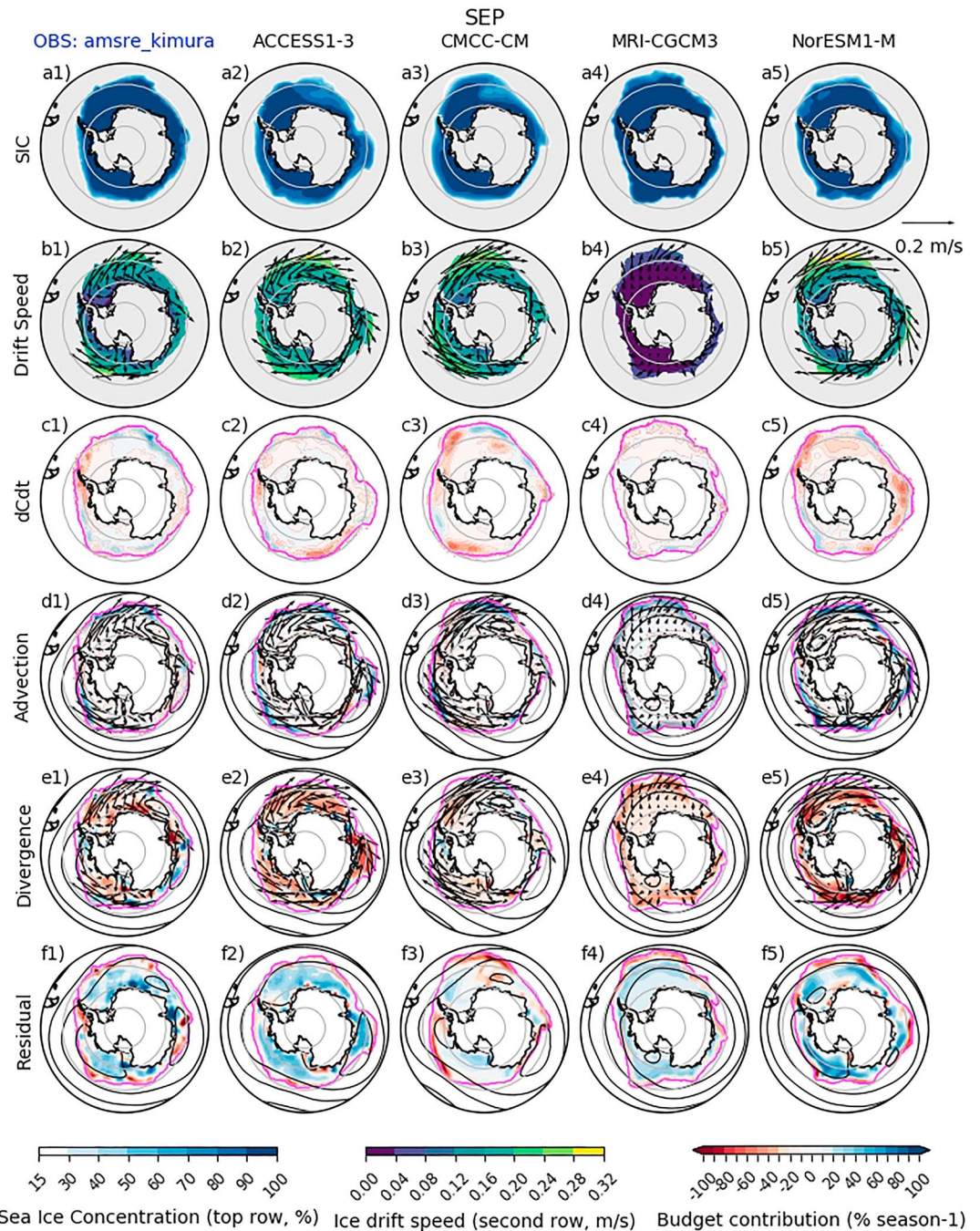


Figure 2. State variables and concentration budget terms for September in “good” model subset. For sea ice concentration (SIC) budget terms, blue indicates a SIC source and red indicates a sink. (a1–a5) SIC (shading). (b1–b5) Drift speed (shading) and drift vectors. (c1–c5) dC/dt (shading) and SIC 15% contour (magenta). (d1–d5) Tendency due to advection (shading), sea level pressure (black contours, interval 10 hPa), ice drift vectors, and SIC 15% contour. (e1–e5) Same as d1–d5, but divergence. (f1–f5) Same as d1–d5, but residual term, without ice drift vectors.

month. There are large positive contributions (an ice source) from advective transports (blue triangles) consistent with the region of positive advection at the ice edge (Figure 2d1) and negligible negative contributions (red triangle). In contrast, there are large negative contributions (an ice sink) from divergent transports (red star) consistent with the divergent interior ice pack (Figure 2e1) and smaller positive contributions from local convergence (blue star). The magnitude of the negative residual (red

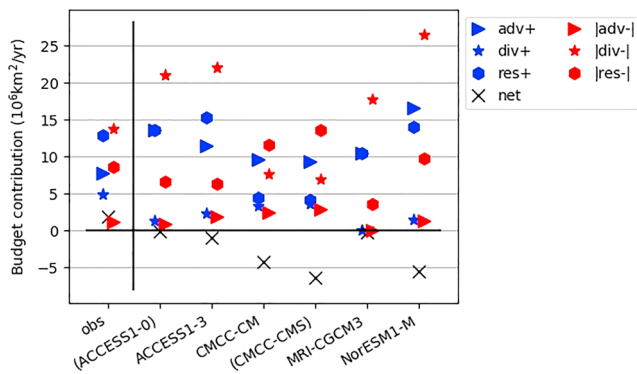


Figure 3. Spatial integrals of September budget components for observations and the “good” model subset. Total expansion (the integral of dC/dt) is marked with a black cross. Other terms are split into a sum over only regions where their time mean is positive (sources, blue) or negative (sinks, red, negated). These terms are denoted by symbols: advection (triangle), divergence (star), and residual (hexagon).

hexagon) is consistent with the dominant balance being that of melting balancing the advection (blue triangle). The magnitude of the positive residual (blue hexagon) is consistent with freezing balancing the divergence (red star).

In the models, the advective ice expansion (Figure 3, blue triangles), which is largely at the ice edge, is stronger than observed by up to a factor of 2. The negative residual (red hexagons) has a variety of biases between the models. In ACCESS1-3, ACCESS1-0, MRI-CGCM3, and NorESM1-M, the integrated ice sink due to divergence (red stars) is too strong and much larger than the residual ice source (blue hexagons). This is consistent with refreezing being insufficient to balance the divergent ice sink in the models, in contrast to observations, and so with near-zero or negative expansion (black crosses) in the models and thus a slightly early start to the melt season. In contrast, in CMCC-CM and CMCC-CMS the divergent ice sink and residual ice source are both much smaller than observed. This is consistent with a partially compensating bias, whereby too weak dynamics are accompanied by too weak thermodynamics.

The divergence and residual terms differ most between the models (Figures 2 and 3): What, then, sets these differences? Intermodel differences in divergence are not just a result of different drift speeds (Lecomte et al., 2016) since CMCC-CM and ACCESS1-3 have similar speed (Figure 2b) but different divergence (Figure 2e). We focus therefore on the spatial patterns of ice drift for these two models (Figure 2e and supporting information Figure S4).

In the regions of ice export in the western Ross and Weddell seas, CMCC models display northward drifts which are constrained to very southerly latitudes, while further north the drift is too sharply westerly. In contrast, the ACCESS models capture the ice drift well. In turn, this is linked to the simulation of sea level pressure; both ACCESS models clearly better represent the location and longitudinal extent of the climatological low-pressure areas in this month.

In conclusion, we can discriminate between the models with most realistic sea ice cover on the basis of the model budgets. While spatial integrals provide a helpful synthesis, they hide information; MRI-CGCM3 has realistic integrated values (Figure 3), but the budget terms distribution is erroneous. NorESM1-M has excessive divergence and advection, linked to too-strong drifts. CMCC-CM has insufficient divergence and residual and spatial deficiencies in both terms. While ACCESS1-3 divergence is a little stronger and smoother than the observations, it is possible the observations are artificially noisy due to uncertainty in the observed ice drift (HK16). Furthermore, our results point to appropriate representation of both the atmosphere-ice coupling and the zonally asymmetric atmospheric circulation as being important to this accurate representation of sea ice processes. As for sea ice area (section 3.1), we find that internal variability is too small to affect these conclusions, implying that our budget analysis reveals model structural differences (supporting information Text S3 and Figures S8 and S9).

We find that ACCESS1-3 provides a realistic representation of the processes acting in Antarctic sea ice. This assessment is based on three factors: (i) It successfully reproduces the seasonal cycle of ice area (Figure 1); (ii) it successfully reproduces the spatial patterns of the winter maximum ice concentration budget (Figure 2), for the right reasons; and (iii) it successfully reproduces the seasonal cycle of the ice concentration budget (Figures S7 and S8). Argument (i) can only be made of the six good models. Of those models, argument (ii) can only be made for the ACCESS models. MRI-CGCM3 and NorESM1-M have no ice drift and too rapid ice drift, respectively, while the CMCC models contain too little divergence and freezing in the inner pack, as their climatological wind patterns are wrong. This strengthens the conclusions of Agosta et al. (2015), who concluded that the ACCESS models have most faithful representation of near-surface Antarctic and Southern Ocean climate of all CMIP5 models.

3.4. Role of Sampling Errors in Observations

Uotila et al. (2014) found a greater discrepancy in advection between observations and the ACCESS models than the present study. We found that changing the threshold for calculation of the budgets from $SIC > 15\%$

to SIC > 0% brought our results more in line with theirs (supporting information Text S2 and Table S2). We propose that the removal of low SIC data masks the models in the regions where the observations are often missing data and therefore provides a fairer comparison between observations and models. This therefore supports the hypothesis of Uotila et al. (2014) that observational uncertainty in part explained the large discrepancies found by their analysis.

3.5. Budgets in Other CMIP5 Models and Other Months

Supporting information Figures S5 and S6 show the September dynamics and residual budget terms in all models, ranked by September area (the good model subset analyzed above are panels o to u). For models with strong negative biases in ice area (panels a–i), it is not particularly meaningful to compare budget terms, since ice processes cannot occur where there is no ice. However, most have a clear residual (thermodynamic) ice source in the interior as expected (Figure S6).

Of more interest are the year-round budget terms (Figure S7), because by definition biases in winter area result from biases in evolution in earlier months. In the ice growth season, many models have too weak (March–May) or too late (June onward) ice expansion (Figure S7b). In March–May, this can be linked to ice dynamics and the residual source term both being too weak (Figure S7c–S7e and Table S5). These integrals are affected by the ice area, which varies strongly between the models, but the results clearly indicate a role of deficient ice dynamics in the autumn in the failure of the models to simulate sufficient winter sea ice (Table S5). Many models also have too rapid or early winter retreat (Figure S7b), although all terms vary greatly between models in this season. Based on the autumn results, improving the simulation of sea ice dynamics during the season of freeze-up is crucial to improving simulations of winter maximum ice area.

Finally, the simulations within the good model subset vary in their simulation of summer ice area (Figure 1). We focus on the pairs [ACCESS1-0, ACCESS1-3] and [CMCC-CM, CMCC-CMS] (Figure S4). There is too little summer ice in ACCESS1-0, while ACCESS1-3 is closer to observations, as previously documented in Uotila et al. (2014). Second, CMCC-CMS has a greater ice retreat in summer and expansion in winter (Figure 1) than CMCC-CM. The underlying reasons for these differences are briefly discussed in supporting information Text S4.

4. Conclusions

It is generally viewed that CMIP5 representation of Antarctic sea ice is poor; the Intergovernmental Panel on Climate Change Fifth Assessment Report states low confidence in projected changes due to the “inability of almost all models” to accurately simulate mean and variability in historical sea ice (Collins et al., 2013). In contrast, the present analysis suggests that a subset of good models are able to reproduce the observed seasonal sea ice cycle. However, ice concentration budget terms vary greatly within the good models, suggesting a range of skill in simulating ice processes. In winter, when clear interactions between dynamic and thermodynamic processes are evident in observed budgets, the models vary in their fidelity of simulating these different contributions. ACCESS climate models appear to have excellent representation of the budget terms, with this model's previously documented overcompensation between dynamics and thermodynamics (Schroeter et al., 2018; Uotila et al., 2014) found here to be largely an artifact of incomplete observational sampling.

Our evidence therefore suggests that ACCESS1-3 (and ACCESS1-0) has a realistic winter ice cover for the right reasons. Individually, none of the model components in these two models is structurally unique among the CMIP5 models, and the model resolution is not exceptional either (supporting information Table S1). Therefore, the inability of other models to match observations of ice concentration and its governing budget is not, in general, fundamentally caused by inadequacy in model structure. While CMIP6 will bring increased spatial resolution in many ocean models, and updated sea ice code, these results suggest that this is unlikely to produce a step change in Antarctic sea ice representation. Although the main focus of our paper is to analyze sea ice budgets, we show evidence that points to accurate representation of the circumpolar pressure trough and its climatological lows as key to a faithful representation of Antarctic sea ice, particularly during the winter maximum. This motivates more detailed future research, building upon the zonally averaged evidence for this linkage presented in Bracegirdle et al. (2018), and strengthens the evidence that modeling groups should focus on these atmospheric processes.

There is a strong relationship between the sea ice area in a coupled model's "present" and the projected changes in sea ice area, temperature, and precipitation in that model (Bracegirdle et al., 2015). However, this study shows cases where a climate model's apparently realistic representation of current sea ice area results from incorrect coupled ocean-ice-atmosphere processes. These dynamic and thermodynamic processes, as represented here by the concentration budget, must be correct in order to correctly force the ocean, due to the role of sea ice in ocean water mass modification (Abernathy et al., 2016; Haumann et al., 2016). Therefore, sea ice that is "right for the wrong reasons" could lead to misplaced confidence in projections of change in all aspects of Antarctic climate. For example, MRI-CGCM3 has a realistic sea ice extent both in present-day and paleoclimate simulations, but the ocean state in paleoclimate simulations is a clear outlier compared to other models (Marzocchi & Jansen, 2017). Therefore, the lack of wind stress over ice in this model, and its consequent dynamic biases (Figure 2), may have significant implications for 21st-century projections of oceanic variables. Appropriate choice of existing models, as well as work to improve the atmosphere-sea ice dynamic coupling in future models, would therefore be expected to provide more reliable projections of future change in Antarctic sea ice. Future work will investigate the relationship between sea ice budget accuracy and projected changes in Antarctic climate.

Acknowledgments

The World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and the climate modeling groups (supporting information Table S1) are thanked for producing and making available their model output. For CMIP, the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. The original CMIP5 data can be accessed through the ESGF data portals (see <http://pcmdi-cmip.llnl.gov/cmip5/availability.html>). ECMWF is thanked for providing the ERA-Interim data set. The budgets based on satellite observations are available at the website (<https://data.bas.ac.uk/>). The authors were funded by the Natural Environment Research Council as part of the British Antarctic Survey Polar Science for Planet Earth Programme and grant NE/N01829X/1.

References

- Abernathy, R. P., Cerovecki, I., Holland, P. R., Newsom, E., Mazloff, M., & Talley, L. D. (2016). Water-mass transformation by sea ice in the upper branch of the Southern Ocean overturning. *Nature Geoscience*, 9(8), 596–601. <https://doi.org/10.1038/ngeo2749>
- Agosta, C., Fettweis, X., & Datta, R. (2015). Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance. *The Cryosphere*, 9(6), 2311–2321. <https://doi.org/10.5194/tc-9-2311-2015>
- Bracegirdle, T. J., Hyder, P., & Holmes, C. R. (2018). CMIP5 diversity in southern westerly jet projections related to historical sea ice area: Strong link to strengthening and weak link to shift. *Journal of Climate*, 31(1), 195–211. <https://doi.org/10.1175/JCLI-D-17-0320.1>
- Bracegirdle, T. J., Stephenson, D. B., Turner, J., & Phillips, T. (2015). The importance of sea ice area biases in 21st century multimodel projections of Antarctic temperature and precipitation. *Geophysical Research Letters*, 42, 10–832. <https://doi.org/10.1002/2015GL067055>
- Cavaleri, D. J., T. Markus, & J. C. Comiso (2014). AMSR-E/Aqua daily L3 12.5 km brightness temperature, sea ice concentration, & snow depth polar grids, version 3. NASA National Snow and Ice Data Center Distributed Active Archive Center, accessed 20 February 2015. https://doi.org/10.5067/AMSR-E/AE_SI12.003
- Cavanagh, R. D., Murphy, E. J., Bracegirdle, T. J., Turner, J., Knowland, C. A., Corney, S. P., et al. (2017). A synergistic approach for evaluating climate model output for ecological applications. *Frontiers in Marine Science*, 4, 308. <https://doi.org/10.3389/fmars.2017.00308>
- Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., et al. (2013). Long-term climate change: Projections, commitments and irreversibility. In T. F. Stocker, et al. (Eds.), *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 1029–1136). Cambridge: Cambridge University Press.
- Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., et al. (2011). The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656), 553–597. <https://doi.org/10.1002/qj.828>
- Haumann, F. A., Gruber, N., Münnich, M., Frenger, I., & Kern, S. (2016). Sea-ice transport driving Southern Ocean salinity and its recent trends. *Nature*, 537(7618), 89–92. <https://doi.org/10.1038/nature19101>
- Holland, P. R., & Kimura, N. (2016). Observed concentration budgets of Arctic and Antarctic sea ice. *Journal of Climate*, 29(14), 5241–5249. <https://doi.org/10.1175/JCLI-D-16-0121.1>
- Holland, P. R., & Kwok, R. (2012). Wind-driven trends in Antarctic sea-ice drift. *Nature Geoscience*, 5(12), 872–875. <https://doi.org/10.1038/ngeo1627>
- Jenouvrier, S., Holland, M., Stroeve, J., Serreze, M., Barbraud, C., Weimerskirch, H., & Caswell, H. (2014). Projected continent-wide declines of the emperor penguin under climate change. *Nature Climate Change*, 4(8), 715–718. <https://doi.org/10.1038/nclimate2280>
- Krinner, G., Llargeron, C., Ménégou, M., Agosta, C., & Brutel-Vuilmet, C. (2014). Oceanic forcing of Antarctic climate change: A study using a stretched-grid atmospheric general circulation model. *Journal of Climate*, 27(15), 5786–5800. <https://doi.org/10.1175/JCLI-D-13-00367.1>
- Lecomte, O., Goosse, H., Fichefet, T., Holland, P. R., Uotila, P., Zunz, V., & Kimura, N. (2016). Impact of surface wind biases on the Antarctic sea ice concentration budget in climate models. *Ocean Modelling*, 105, 60–70. <https://doi.org/10.1016/j.ocemod.2016.08.001>
- Marzocchi, A., & Jansen, M. F. (2017). Connecting Antarctic sea ice to deep-ocean circulation in modern and glacial climate simulations. *Geophysical Research Letters*, 44, 6286–6295. <https://doi.org/10.1002/2017GL073936>
- Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J. F., et al. (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climate Change*, 109(1–2), 213–241. <https://doi.org/10.1007/s10584-011-0156-z>
- Raaiäsaänen, J. (2007). How reliable are climate models? *Tellus A: Dynamic Meteorology and Oceanography*, 59(1), 2–29. <https://doi.org/10.1111/j.1600-0870.2006.00211.x>
- Roach, L. A., Dean, S. M., & Renwick, J. A. (2018). Consistent biases in Antarctic sea ice concentration simulated by climate models. *The Cryosphere*, 12(1), 365–383. <https://doi.org/10.5194/tc-12-365-2018>
- Schroeter, S., Hobbs, W., Bindoff, N. L., Massom, R., & Matear, R. (2018). Drivers of Antarctic sea ice volume change in CMIP5 models. *Journal of Geophysical Research: Oceans*, 123, 7914–7938. <https://doi.org/10.1029/2018JC014177>
- Swart, N. C., & Fyfe, J. C. (2013). The influence of recent Antarctic ice sheet retreat on simulated sea ice area trends. *Geophysical Research Letters*, 40, 4328–4332. <https://doi.org/10.1002/grl.50820>
- Taylor, K. E., Stouffer, R. J., & Meehl, G. A. (2012). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93, 485–498. <https://doi.org/10.1175/BAMS-D-11-00094.1>

- Turner, J., Bracegirdle, T. J., Phillips, T., Marshall, G. J., & Hosking, J. S. (2013). An initial assessment of Antarctic sea ice extent in the CMIP5 models. *Journal of Climate*, *26*(5), 1473–1484. <https://doi.org/10.1175/JCLI-D-12-00068.1>
- Turner, J., Maksym, T., Phillips, T., Marshall, G. J., & Meredith, M. P. (2013). The impact of changes in sea ice advance on the large winter warming on the western Antarctic Peninsula. *International Journal of Climatology*, *33*(4), 852–861. <https://doi.org/10.1002/joc.3474>
- Uotila, P., Holland, P. R., Vihma, T., Marsland, S. J., & Kimura, N. (2014). Is realistic Antarctic sea-ice extent in climate models the result of excessive ice drift. *Ocean Modelling*, *79*, 33–42. <https://doi.org/10.1016/j.ocemod.2014.04.004>
- Zunz, V., Goosse, H., & Massonnet, F. (2013). How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent? *The Cryosphere*, *7*(2), 451–468. <https://doi.org/10.5194/tc-7-451-2013>