Intermodel CMIP5 relationships in the baseline Southern Ocean climate system and with future projections

Jules B. Kajtar[1], Agus Santoso[2], Matthew Collins[3], Andrea Sardinha Taschetto[2], Matthew H. England[2], and Leela M. Frankcombe[4]

[1]University of Tasmania
[2]University of New South Wales
[3]University of Exeter
[4]Climate Change Research Centre

November 21, 2022

Abstract

Climate models exhibit a broad range in the simulated properties of the global climate. In the early historical period, the absolute global mean surface air temperature of models contributing to the fifth phase of the Coupled Model Intercomparison Project (CMIP5) spans a range of ~12-15 °C. Other climate parameters are linked to the global mean temperature, such as sea ice area, atmospheric circulation patterns, and by extension cloudiness, precipitation and albedo. Accurate representation of the baseline climate state is crucial for meaningful future climate projections, since the baseline conditions may dictate the capacity for change. For example, a model with initially smaller sea ice area has less potential to lose sea ice as the planet warms. Amongst the CMIP5 models, it is found that in the baseline climate state there are coherences between Southern Ocean temperature, outgoing shortwave radiation, cloudiness, the position of the mid-latitude eddy-driven jet, and Antarctic sea ice area. The baseline temperature relationship extends to projected future changes in the same set of variables. The tendency for models with initially cooler Southern Ocean surface temperature to exhibit more global warming, and vice versa for initially warmer models, can therefore be linked to baseline Southern Ocean climate system biases. A first look at emerging data from CMIP6 reveals a shift of the tendency towards the Antarctic region, potentially linked to a reduction in biases over the Southern Ocean, which prompts an examination of biases in the Antarctic region as more CMIP6 model data becomes available.

Intermodel CMIP5 relationships in the baseline Southern Ocean climate system and with future projections

Jules B. Kajtar (orcid.org/0000-0003-0114-6610)[1,2,3], **Agus Santoso** (orcid.org/0000-0001-7749-8124)[3,4,5], **Matthew Collins** (orcid.org/0000-0003-3785-6008)[1], **Andréa S. Taschetto** (orcid.org/0000-0001-6020-1603)[3,4], **Matthew H. England** (orcid.org/0000-0001-9696-2930)[3,4], **Leela M. Frankcombe** (orcid.org/0000-0001-6578-3200)[3,4]

[1] College of Engineering, Mathematics, and Physical Sciences, University of Exeter, Exeter, UK
[2] Institute for Marine and Antarctic Studies, University of Tasmania, Tasmania, Australia
[3] Australian Research Council Centre of Excellence for Climate Extremes, Australia
[4] Climate Change Research Centre, University of New South Wales, NSW, Australia
[5] Centre for Southern Hemisphere Oceans Research (CSHOR), CSIRO Oceans and Atmosphere, Hobart, Tasmania, Australia
Corresponding author: Jules B. Kajtar; jules.kajtar@utas.edu.au.

Keywords: Southern Ocean, CMIP5, baseline climate, climate sensitivity

Key points:

- There are robust intermodel correlations across elements of the Southern Ocean climate system in historical CMIP5 simulations.
- The baseline Southern Ocean temperature relationship extends to projected changes in radiation, cloudiness, the jet latitude and sea ice.
- Models with initially cooler Southern Ocean tend to warm more globally, due to an apparent greater capacity for change.

Abstract

Climate models exhibit a broad range in the simulated properties of the global climate. In the early historical period, the absolute global mean surface air temperature of models contributing to the fifth phase of the Coupled Model Intercomparison Project (CMIP5) spans a range of ~12-15 °C. Other climate parameters are linked to the global mean temperature, such as sea ice area, atmospheric circulation patterns, and by extension cloudiness, precipitation and albedo. Accurate representation of the baseline climate state is crucial for meaningful future climate projections, since the baseline conditions may dictate the capacity for change. For example, a model with initially smaller sea ice area has less potential to lose sea ice as the planet warms. Amongst the CMIP5 models, it is found that in the baseline climate state there are coherences between Southern Ocean temperature, outgoing shortwave radiation, cloudiness, the position of the mid-latitude eddy-driven jet, and Antarctic sea ice area. The baseline temperature relationship extends to projected future changes in the same set of variables. The tendency for models with initially cooler Southern Ocean surface temperature to exhibit more global warming, and vice versa for initially warmer models, can therefore be linked to baseline Southern Ocean climate system biases. A first look at emerging data from CMIP6 reveals a shift of the tendency towards the Antarctic region, potentially linked to a reduction in biases over the Southern Ocean, which prompts an examination of biases in the Antarctic region as more CMIP6 model data becomes available.

Plain Language Summary

Modern simulations of the Earth's climate system differ in some of their large-scale features. For example, in models reported on by the Intergovernmental Panel on Climate Change (IPCC) in the Fifth Assessment Report (AR5), the globally averaged baseline surface temperature ranges between 12 and 15 °C. Global mean temperature is known to be linked to other features, such as wind, clouds, and rainfall. Correctly simulating the present-day climate is important, so that we can have more confidence in the possible futures they simulate under different levels of anthropogenic greenhouse gas emissions. In this study, strong relationships are found between modelled Southern Ocean temperature and the amount of sea ice and clouds they simulate. In addition, it is found that the initial Southern Ocean temperature is also related to changes in sea ice and cloud simulated in the future. A model that is cooler initially, for example, tends to have more sea ice and cloud, but also loses more sea ice and cloud in the future, and simulates more global warming.

1 Introduction

The sensitivity of the Earth's climate to greenhouse gas forcing is arguably the key quantity that drives efforts to mitigate the risks of human-induced climate change. But the level of climate sensitivity remains highly uncertain. The most typical measure, equilibrium climate sensitivity (ECS), is defined as the global temperature change in response to a doubling to atmospheric CO_2. The Intergovernmental Panel on Climate Change (IPCC) estimated a likely range in ECS of 1.5-4.5°C in the Fifth Assessment Report (AR5; Stocker et al., 2013). A more recent review, using multiple lines of evidence, narrows the range to 2.6-3.9°C (Sherwood et al., 2020). Climate models of all levels of sophistication have been used to estimate climate sensitivity, but modern efforts focus on the use of general circulation models (GCMs) and Earth system models (ESMs) which include biogeochemical processes. No two climate models are identical, with some exhibiting low sensitivity and others high (e.g. Flato et al., 2013; Forster et al., 2013; Zhai et al., 2015). Furthermore, no model is perfect, and all exhibit some level of bias when compared with observational data. One approach to reducing the level of uncertainty in climate sensitivity is that of 'emergent constraints' (Hall et al., 2019). Emergent constraints aim to find links between the bias of particular variables in the baseline climate, and their evolution under radiative forcing. If a robust relationship emerges, across a wide range of different climate models, then it might be reasonable to expect more 'realistic' models in the baseline would provide more realistic projections. The emergent constraints approach crucially depends upon drawing from a large number of unique climate models. Increasing availability of such model data, as most notably facilitated by the Coupled Model Intercomparison Project (CMIP), allows for deeper studies into the impact of model biases on future projections.

Baseline global mean surface temperature (GMST) has been explored as just one of many possible constraints on climate sensitivity. CMIP, phase 5, (CMIP5) models exhibit a wide range in long-term averaged absolute GMST over the historical period (~12-15 °C; Flato et al., 2013). However, no statistically significant relationship has been found between baseline temperature and ECS in CMIP5 (Flato et al., 2013), nor with future temperature change in the RCP4.5 scenario (Hawkins & Sutton, 2016), though it has been noted that there is an absence of models with overly warm baseline temperature and strong global warming (Hawkins & Sutton, 2016). Simulated absolute temperature is generally considered unimportant, since it is more crucial to initialise a model with near-zero net TOA energy balance (Hawkins & Sutton, 2016). In addition, projected changes are usually measured with respect to the baseline, or unforced, climate, and therefore represented as anomalies. In this study, the potential influence of baseline absolute temperature on climate projections is revisited. The particular focus is on Southern Ocean processes, which play an important role in regulating the global climate.

Past studies have noted possible relationships between elements of the Southern Hemisphere climate system and climate sensitivity, in both CMIP3 and CMIP5. CMIP3 models exhibited a strong relationship between Southern Hemisphere net top-of-atmosphere (TOA) radiation and climate sensitivity (Trenberth & Fasullo, 2010). Whilst showing that the intermodel correlation is strong, Trenberth & Fasullo (2010) also acknowledge that the relationship is likely due to large model biases in the Southern Hemisphere climate system. Larger errors related to the negative bias in cloud amount in CMIP3, they argued, may lead to smaller sensitivity. Grise et al. (2015) find a weaker intermodel correlation between climate sensitivity and Southern Hemisphere net TOA radiation amongst CMIP5 models. They show that the relationship only exists amongst a subset of CMIP5 models with unrealistically bright clouds in the Southern Hemisphere sub-tropics – a characteristic that is typical amongst CMIP3 models. Thus, the apparent link between net TOA radiation and climate sensitivity is not supported by a real physical mechanism, and manifests merely as a result of model biases. Southern Ocean cloudiness and net radiation were therefore deemed inappropriate for constraining equilibrium climate sensitivity (Grise et al., 2015).

In this study we focus on relationships between baseline parameters in the Southern Ocean climate system, and their projected changes. The analysis is primarily of intermodel correlations: motivated by the emergent constraints approach, and armed with a complete suite of CMIP5 simulations, the relationships between a range of variables is explored. We start by revisiting the intermodel correlation between the absolute baseline Southern Ocean surface temperature and GMST change in CMIP5 models, for which a compelling relationship is found. Our aim is to illuminate the role of baseline biases in the Southern Ocean system in CMIP5, and their impact on projected changes, locally and globally. We take a first look at CMIP6, but because the emerging story appears to be one of Antarctic biases, further exploration is left for a future study. After outlining data and methods (Section 2), the relationships between baseline absolute temperature and a range of other climate parameters for both baseline and future changes are examined (Section 3). Finally, the conclusions of this study are summarised in Section 4.

2 Data and Methods

This study focusses on CMIP5 models, but some preliminary analysis is conducted on available CMIP6 output. Equilibrium climate sensitivity (ECS) in CMIP models is generally estimated from the 150 year *abrupt4xCO2* experiment, in which atmospheric CO_2 is instantaneously quadrupled initially (Andrews et al., 2012; Gregory et al., 2004). This method for computing ECS avoids having to slowly evolve CO_2 forcing, and then reach equilibrium, which can take several hundreds, if not thousands, of simulated years (Grose et al., 2018). Another measure for climate sensitivity is the transient climate response (TCR), which is a measure of the global temperature change after 70 years of simulating an annual 1% increase in CO_2. A dedicated CMIP experiment, namely the *1pctCO2* scenario, is run by some modelling groups to compute the TCR. Although TCR might be considered more useful for climate projections over the next few decades (Knutti et al., 2017), ECS unexpectedly correlates more strongly with projected changes under the representative concentration pathway (RCP) scenarios (Grose et al., 2018). On the other hand, although the range in ECS across CMIP6 models is substantially larger than in CMIP5, and is in fact the largest range of any generation dating back to the 1990s, TCR is only slightly larger in CMIP6 as compared to CMIP3 and CMIP5 (Meehl et al., 2020). Since fewer models have data available from the *abrupt4xCO2* or *1pctCO2* experiments, in this study global mean surface temperature (GMST) change from the baseline in the *historical* experiments through to the end of the 21st century under the RCP8.5 emissions scenario experiments (*rcp85*), is taken as a proxy for climate sensitivity. The *rcp85* scenario was chosen since it has the strongest forcing, and therefore the largest projected changes, which helps to draw out possible correlations.

Baseline temperature is the equilibrium temperature that models achieve after a 'spin-up' period. However, the baseline may not be perfectly equilibrated due to model 'drift'– spurious long-term changes unrelated to external forcing nor internal variability, which may be a result of insufficient spin-up integration (Sen Gupta et al., 2013). Here the baseline temperature is evaluated in the early part of the *historical* simulations, corresponding with the late 19th Century. Greenhouse gas forcing may cause some temperature change in this early period, but the *historical* simulations are analysed in preference over the pre-industrial control (*piControl*) simulations, since *piControl* experiment data are available from fewer models than from *historical*. In addition, the aforementioned issue of model drift afflicts both sets of experiments (Sen Gupta et al., 2013). The baseline period is taken as 1861-1900, early in the historical simulations and soon after the pre-industrial state. For projected changes, a difference is taken over the future period 2061-2100 and the baseline. Long reference periods of 40 years were chosen to reduce the influence of internal decadal variability as much as possible.

The primary climate variables analysed in this study are surface temperature (CMIP variable name: *tas*), top-of-atmosphere (TOA) outgoing shortwave radiation (*rsut*), total cloud fraction (*clt*), surface zonal wind stress (*tauu*), and sea ice concentration (*sic*). All available CMIP5 monthly data for each of the five variables were gathered from both *historical* and *rcp85* experiment sets. Annual averages were computed, and then the data were regridded to a common 1×1° global grid. For models with multiple ensemble members, a single model ensemble average was taken. Utilising only models for which all five variables were available over the period 1861-2100 (after appending *rcp85* to *historical*), resulted in a set of 40 CMIP5 models (Table 1). Net TOA radiation is also analysed for the same 40 models, but only in the *historical* experiments. It is computed as TOA incident shortwave radiation (*rsdt*) minus TOA outgoing shortwave (*rsut*) minus TOA outgoing longwave (*rlut*). Some relationships with ECS are computed, which is only available for 30 CMIP5 models (Table 1). Surface air temperature is analysed later in a group of CMIP6 models, using *historical* data together with *ssp585* (Table 2).

TOA outgoing shortwave radiation is analysed here in preference over net radiation, since it gives a better sense of albedo effects, which is of greater interest. The surface zonal wind stress is used to estimate the mean latitude of the eddy-driven jet. After regridding zonal wind stress to the common 1×1° global grid, annual and zonal means are taken. The jet latitude is then computed by fitting a quadratic polynomial to the latitude and two neighbouring grid points where zonal wind stress is maximal in the Southern Hemisphere. This method is similar to that of Kidston & Gerber (2010), but they use 10m zonal wind, for which the computed eddy-driven jet latitude is similar.

Regression coefficients are calculated using ordinary least squares, and quoted correlation values are Pearson's correlation coefficients. 'Intermodel' correlations or regressions refer to the relationship between two variables across the models (i.e. 40 CMIP5 models in most instances). The symbol r denotes intermodel correlation. The 95% and 99% statistical significance levels of correlations are quoted in various cases, and tested using a Student's t-distribution. For a sample size of 40, correlations with magnitude greater than ~0.31 are significant at the 95% level (p = 0.05), and ~0.40 at the 99% level (p = 0.01).

In comparisons with reanalysis, the NOAA-CIRES-DOE Twentieth Century Reanalysis, version 3 (20CRv3) product is utilised (Compo et al., 2011), in which the five primary variables are available.

3 Results

3.1 Baseline temperature and climate sensitivity

Across CMIP5 models, the global mean surface temperature (GMST) in the baseline period spans a range of 2.8 °C (12.1-14.9 °C; Fig. 1a). Following the evolution of global temperature through the historical simulations, and extending with the RCP8.5 emissions scenario, the projected range in absolute temperature is 3.9 °C (15.5-19.4 °C in the future period 2061-2100; Fig. 1a). By considering GMST anomalies with respect to the baseline period in each model, the projected range across all models is 2.5 °C (2.9-5.3 °C). The range of simulated absolute baseline temperatures therefore represents a considerable source of uncertainty in future projections.

There is no significant relationship between baseline GMST and GMST change across models under the RCP8.5 scenario (Fig. 1b). Consistently, no significant relationship was found between baseline GMST and climate sensitivity (Flato et al., 2013), nor with GMST change in the RCP4.5 scenario (Hawkins & Sutton, 2016). Though it has been noted that there is an absence of models with warm baseline temperature simulating strong global warming (Hawkins & Sutton, 2016). However, there is a striking feature in the spatial pattern of intermodel correlation between grid-point baseline surface temperature and GMST change (Fig. 1c). Most of the Southern Ocean baseline temperature is correlated with GMST change (with a grid-point maximum of r = -0.64). The intermodel correlation of the baseline temperature averaged over 35-55 °S and GMST change is -0.53 (Fig. 1d). Hence, models with initially cooler Southern Ocean surface temperature tend to warm more globally, vice versa for models with initially warmer Southern Ocean. Another apparent feature in the spatial pattern is the north-south hemisphere contrast (Fig. 1c), which appears somewhat analogous with the projection of a faster warming of the Northern Hemisphere than the Southern Hemisphere (e.g. Xie et al., 2010). This could indicate the role of the Southern Ocean on the interhemispheric warming pattern. There is a statistically significant intermodel correlation between north-south hemisphere contrast in the baseline and GMST change (r= 0.55). As the significant inter-model correlations occur most prominently in the Southern Ocean, we focus on the Southern Ocean baseline temperature for further analysis.

The Southern Ocean region of statistically significant intermodel correlation is globally the most spatially vast and coherent. Due to its unique geographical configuration, marked by a circumpolar circulation under the influence of the prevalent westerly winds, the Southern Ocean is known to play an important role in the global thermohaline circulation and the uptake of heat and carbon (Manabe et al., 1991; Marshall & Speer, 2012; Mikaloff Fletcher et al., 2006; Toggweiler & Samuels, 1995). Heat and carbon uptake by the Southern Ocean is also simulated as being strong in CMIP5 models (Frölicher et al., 2015). We explore whether a physical mechanism explains the statistically significant negative intermodel correlation between Southern Ocean surface temperature and future GMST change. Another question of interest is around what processes set the absolute temperature of the baseline Southern Ocean in climate models.

In contrast to Grise et al. (2015), here the intermodel relationship between baseline surface air temperature (as opposed to net TOA radiation) and global mean temperature change (as opposed to ECS; see their Fig. 2a) is shown. By analysing surface air temperature change, rather than ECS, the model set is greatly expanded (output from 40 models here, c.f. 20 models in Grise et al., 2015). Even though global mean temperature change is strongly related to ECS (Fig. 2a) and baseline surface air temperature is strongly related to net TOA radiation (Fig. 2b), the patterns shown in Fig. 1c and by Grise et al. (2015; their Fig. 2a for CMIP5) are substantially different. Some exploration reveals that selected baseline years (1861-1900 as opposed to 1990-1999, when anthropogenic forcings are stronger), the length of the baseline period (40 years as opposed to 10 years, which can be influenced by decadal variability), and the set of sampled models, all modify the pattern to some extent. However, exchanging only net TOA radiation with surface air temperature considerably strengthens the intermodel correlations over the Southern Ocean (c.f. Fig. 2c and 2d). Our focus is on the Southern Ocean baseline surface temperature, which directly contributes to the global mean temperature, and its interaction with key processes in the region, such as sea ice, cloud cover, and westerly jet.

3.2 Links between surface temperature and baseline climate

There are statistically significant intermodel regressions and correlations between the Southern Ocean baseline temperature and a range of other baseline climate variables in the domain such as shortwave radiation, cloud cover, the meridional position of the westerly eddy-driven jet, and sea ice area (Fig. 3). Here the spatial patterns are shown as intermodel regressions (Fig. 3a,c,e,g), rather than as intermodel correlations (Fig. 1,2). The baseline surface temperature area-averaged over 35-55°S is strongly negatively correlated with TOA outgoing shortwave radiation in that region (Fig. 3a,b). TOA outgoing shortwave is due mainly to albedo effects, and unsurprisingly the Southern Ocean baseline temperature is also negatively correlated with cloud fraction (Fig. 3c,d). Taken together, these results (Fig. 3a-d) show that models with warmer Southern Ocean surface temperatures tend to have less cloud cover and therefore less outgoing shortwave radiation, and vice versa for models with cooler Southern Ocean surface temperature. The direct relationship between baseline Southern Ocean cloud cover and TOA outgoing shortwave radiation is particularly strong (r = 0.75; Fig. 4).

The intermodel relationship showing that models with warmer Southern Ocean surface temperature have less cloud may seem counterintuitive, but the tendency for higher temperatures leading to increased cloudiness is more typical in the tropics. Higher temperatures throughout the tropical atmospheric column typically leads to greater cloud water content, due to an increased moist adiabatic lapse rate (Betts & Harshvardhan, 1987; Frey et al., 2018). On the other hand, cooler surface temperatures in the midlatitudes promote subsidence and the formation of reflective low-level clouds (Grise & Medeiros, 2016; Klein & Hartmann, 1993), and low clouds over the Southern Ocean exhibit the strongest sensitivity to surface temperature (Wall et al., 2017). The striking negative intermodel regression over the Southern Ocean band (Fig. 3c) is likely due to cooler Southern Ocean surface promoting more subsidence, producing more cloud, increasing albedo, and in a feedback loop, driving lower surface temperature.

The latitude of the Southern Hemisphere eddy-driven jet also appears to be related to Southern Ocean surface temperature across models (Fig. 3e,f). It has previously been shown that the jet latitude is biased equatorward in all CMIP3 models (Kidston & Gerber, 2010), and there is little improvement in CMIP5 (Barnes & Polvani, 2013). In models with cooler Southern Ocean surface temperature, the jet is more equatorward, and vice versa for models with warmer Southern Ocean (r = -0.55; Fig. 3f). Kidston et al. (2011) found a seasonal link between jet latitude and sea ice area, but only during the cold season. Here, the direct relationship between baseline sea ice area and jet latitude is found to be weak (r = -0.18; Fig. 4). Since the storm tracks are embedded in the eddy-driven jet, it is not surprisingly to find a relationship with cloudiness (r = 0.38; Fig. 4) and therefore also with TOA outgoing shortwave radiation, due to albedo effects (r = 0.67; Fig. 4).

There is a statistically significant relationship between baseline Southern Ocean temperature and Antarctic sea ice area (r = -0.36; Fig. 3h), with stronger regression relationships around the edge of the sea ice region (Fig. 3g), and correlation coefficients as high as -0.93. This is due to the strong link between local surface temperature and the presence of sea ice: surface temperature is substantially lower when sea ice is present, as opposed to when it is warmed by the open ocean below. But in a positive feedback, cooler temperature also permits sea ice expansion. Conversely, higher temperature inhibits sea ice formation, and less sea ice exposes more water to solar radiation. Furthermore, the relatively cooler Antarctic waters are transported northward via Ekman advection. The feedback is illustrated to some extent in composite patterns of the 10 models with warmest and coolest baseline Southern Ocean surface temperature (Fig. 5). There are strong temperature anomalies with respect to the model mean over the Antarctic sea ice region, in both baseline (Fig. 5a,b) and projected temperature changes (Fig. 5c,d). Thus, the intermodel relationship is physically consistent in that warmer models have less sea ice and vice versa. Sea ice area correlates poorly with other variables across CMIP5, although there is a weak but statistically significant relationship with Southern Ocean cloud cover (r = -0.35; Fig. 4), which may be a result of sea ice suppressing evaporation (Bromwich et al., 2012).

3.3 Baseline temperature and future projections

Thus far it has been shown that there are a range of physically consistent intermodel relationships between the baseline Southern Ocean surface temperature and a range of other climate variables. But how is all of this relevant to the climate sensitivity? The relationships between baseline temperature and future changes in variables will now be examined. Changes in variables are computed under the RCP8.5 scenario, over the same baseline (1861-1900) and future (2061-2100) periods indicated in Fig. 1a. A strong Southern Ocean signature also emerges in all of the spatial patterns of intermodel regressions (Fig. 6), but here the intermodel correlations are positive.

The change in TOA outgoing shortwave radiation over the Southern Ocean is not consistent across models. In most models, outgoing radiation is reduced, but in a small number it is enhanced (Fig. 6b). Nevertheless, there is a statistically significant relationship between Southern Ocean temperature and the change in TOA outgoing shortwave radiation, such that there is a greater reduction in radiation (i.e., increased heat flux into the ocean) for initially cooler models (r = 0.49;

Fig. 6b). Similarly, initially cooler models tend to lose more cloud cover under global warming (r = 0.50; Fig. 6d). The similarities of the patterns in Fig. 6a and Fig. 6c again reflect the strong link between outgoing shortwave radiation and cloud cover. The relationships with baseline temperature (Fig. 6b,d) emerge despite the fact that there is no statistically significant relationship between the baseline and change in radiation (r = -0.29; Fig. 4), nor between the baseline and change in cloud cover (r = -0.25; Fig. 4). In other words, the baseline Southern Ocean temperature is a stronger predictor of changes in outgoing radiation and cloud cover than the baseline in each of these variables.

The eddy-driven jet is projected to shift poleward under all scenarios of climate change (Arblaster & Meehl, 2006; Miller et al., 2006; Simpson & Polvani, 2016). Furthermore, the future change in jet latitude appears to be closely connected to its baseline latitude, as was seen in CMIP3 models (Kidston & Gerber, 2010), and previously reported for CMIP5 (Simpson & Polvani, 2016). The more equatorward the jet is situated initially, the more poleward it shifts under global warming (r = -0.62; Fig. 4). Since this correlation is between a variable and its change, the change contains a component of the baseline (i.e. A vs. B-A), and it is therefore necessary to verify if the intermodel correlation is significant between the baseline latitude and the future latitude (i.e. A vs. B). In this case, the relationship is robust (r = 0.92; figure not shown).

The Southern Ocean baseline temperature is also a robust predictor of future jet migration, with initially cooler models exhibiting a larger shift in jet latitude (r = 0.57; Fig. 6f). Unlike the Southern Ocean baseline temperature (Fig. 1d), the baseline jet latitude position is not found to be a predictor for global mean surface temperature change (r = 0.28; Fig. 4). Bracegirdle et al., (2018) found that sea ice is more closely related to changes in jet strength, where CMIP5 models with greater historical sea ice area exhibit less jet strengthening in the future. They likewise find that links between sea ice and jet shift are weak, albeit with some apparent seasonal relationships.

The spatial pattern of the intermodel regression between Southern Ocean baseline temperature and sea ice area change (Fig. 6g) is similar to the pattern with baseline sea ice area (Fig. 3g), also showing that local surface temperature nearer to the sea ice region is more important. But it is clear that models with initially more sea ice, which correspond with models having cooler baseline Southern Ocean, also lose more sea ice under global warming (r = -0.82; Fig. 4). As with the test for jet latitude, the intermodel correlation between baseline and future sea ice is likewise robust (r = 0.87; figure not shown).

The relationships between baseline Southern Ocean temperature and other baseline variables might be viewed as negative feedbacks (Fig. 3), whereas the relationships between baseline temperature and projected changes in other variables as positive feedbacks (Fig. 6). When viewed as feedbacks, the baseline and future relationships may appear to be counterintuitive. Apart from a small number of exceptions, the future changes in all variables across most models are negative (Fig. 6b,d,f,h); the biggest exception being TOA outgoing shortwave radiation, for which five models exhibit increases under global warming. Therefore, the relationships between baseline temperature and future changes in other variables can be summarised as less change in initially warmer models, and greater change in initially cooler models. And since it has been found that models with initially cooler Southern Ocean warm more globally, the relationships between Southern Ocean temperature change and future changes in other variables reinforce these relationships (Fig. 7). Hence, models with greater Southern Ocean temperature change exhibit greater change in other variables (i.e. a negative feedback).

Some of the relationships in future changes may emerge simply due to a larger capacity for change. For example, a model with more sea ice initially has more capacity to lose sea ice as the planet warms. Similarly, models with more equatorward eddy-driven jet initially, have more capacity to shift poleward. Poleward shift of the jet under global warming is one of the most robust projections across models. Thus, this perspective on capacity for change may provide a useful clue as to how the GMST change is related to baseline Southern Ocean temperature amongst CMIP5 models.

3.4 Baseline Southern Ocean temperature: an integrating factor

The preceding analysis reveals that baseline surface temperature of the Southern Ocean in CMIP5 is a crucial variable in setting not only the baseline state of the Southern Ocean climate system, but also its future evolution and that of the global mean surface temperature (Fig. 1). It is not necessarily possible to conclude that it is the single key variable, since many variables are linked with one another to some extent. However, it is nevertheless illuminating that the baseline temperature is the only variable that exhibits statistically significant correlations at the 95% level with all other baseline variables and their future changes (Fig. 4).

Model developers have tended to approach the problem of Southern Ocean biases by using the clouds as the controlling variable. Models generally do not simulate enough cloud cover over the Southern Ocean (Fig. 8d), leading to too much incoming shortwave radiation at the surface, warm sea surface temperature biases, reduced sea ice, and the shift in the eddy-driven jet (Hyder et al., 2018; Williams et al., 2017). Since cloud schemes involve the fastest dynamical processes in the chain of causality, they are generally the aspect that model developers have found easiest to manipulate and modify. An apparent consequence of modified cloud schemes has been an increase in climate sensitivities in many state-of-the-art models (Bodas-Salcedo et al., 2019; Zhu & Poulsen, 2020). The findings of this study suggest that an investigation of the processes that set the baseline temperature in models may be more fruitful.

In an attempt to test the influence of baseline biases on GMST projections, the 40 CMIP5 models were subsampled according to whether they are biased above or below the reanalysis, for each of the baseline variables (i.e. models to the left and right of reanalysis in Fig. 8). In the first test, models were split into two groups according to whether their GMST is less than or greater than in the NOAA 20th century reanalysis, with both models and reanalysis averaged over 1961-2000. The period 1961-2000 was chosen (as in Fig. 8), since observations, and therefore the reanalysis, are more uncertain in the earlier baseline period. The projected GMST change in the future period (2061-2100) minus the baseline period (1861-1900) was then examined in the two sets. The mean GMST change in the warmer model set is less than the cooler model set, but there are only 7 models in the warmer set (Fig. 8a). A two sample Student's *t*-test for different means, but assuming unequal variance, reveals that the model-means of future warming in the two sets are not significantly different ($p = 0.33$). However, unsurprisingly, if the models are separated based on Southern Ocean surface temperature, then the two sets are different at the 95% significance level. This is consistent with the intermodel correlation between baseline Southern Ocean surface temperature and GMST change (Fig. 1d). The test on GMST changes was then repeated after subsampling models based on 1961-2000 mean values of each of the other variables shown in Fig. 8c-f. None of the differences in sets were statistically significant, indicating that only the baseline Southern Ocean temperature bias is a robust determinant of GMST change.

Another question that arises from the findings of this study is whether the GMST projections can be constrained by observations, in essence by following the emergent constraints approach. To this end, we took the subset of models that are closest to the reanalysis, in a given variable, and tested whether the GMST projections in that subsample are different to the remaining, more-biased, models. The process of subsampling was conducted for each of the baseline variables examined in this study. The purpose of this exercise, in other words, is to test whether reducing the bias in any baseline variable may significantly alter warming projections or climate sensitivity. As a first test, the 13 models (approximately one third) with GMST closest to that in the NOAA 20th century reanalysis over 1961-2000 were subsampled. The global temperature change in the future period (2061-2100) minus the baseline period (1861-1900) in the model subset was then compared to that in the remaining 27 models. The 13 least biased models warm slightly less than the remaining 27 models. However, a two sample Student's *t*-test for different means, but assuming unequal variance, reveals that the model-means of future warming in the two sets are not significantly different (p = 0.21). Similarly, a two sample Kolomogorov-Smirnov test for the sets coming from different continuous distributions, or a two sample *F*-test for different variances, do not suggest that there are statistically significant differences in future projections between the two sets. This test was repeated after subsampling models based on 1961-2000 mean values of the other variables shown in Fig, 8b-f. For Southern Ocean TOA outgoing shortwave radiation, cloud cover, and Antarctic sea ice area, the 13 least biased models exhibit greater GMST change than in the remaining models. But for eddy-driven jet latitude and baseline Southern Ocean temperature, the 13 least biased models warm less. Subsampling based on eddy-driven jet latitude exhibits the largest differences between pairs of subsets. However, none of the pairs of subsets, for any variable, are significantly different under any of the aforementioned statistical tests. Altering the number of models in the subsample set made little difference. Based on these tests, efforts to constrain GMST projections or climate sensitivity by subsampling less biased models does not seem plausible for CMIP5 models.

3.5 A first look at CMIP6

The following is only a preliminary investigation of the surface temperature relationship in CMIP6. Although the correlations are mostly negative, the statistically significant intermodel correlation over the Southern Ocean seen in CMIP5 (Fig. 1) is not present across the 33 models analysed thus far in CMIP6 (Fig. 9). However, the region of statistically significant intermodel correlation is shifted to the south: the baseline surface temperature over most of the Antarctic sea ice region is negatively correlated with global mean surface temperature change. Note that CMIP6 uses updated historical forcings, and the Shared Socioeconomic Pathway 5-8.5 (SSP5-8.5) is not identical to RCP8.5 in CMIP5 (O'Neill et al., 2016), but the differences are not expected to have appreciable impact on the analysis here.

The changing nature of intermodel relationships across model generations should not be too surprising. As key biases are tackled, and reduced or altered, different intermodel features may arise. For instance, as noted earlier, the strong CMIP3 intermodel relationship between Southern Hemisphere net TOA radiation and climate sensitivity (Trenberth & Fasullo, 2010) was substantially weaker across CMIP5 models (Grise et al., 2015). Similarly, the CMIP5 Southern Ocean temperature relationship with GMST change is weaker in CMIP6, though the cause of this has not yet been revealed, and will be explored in a future study.

The analysis of CMIP6 is not taken further in this study for two reasons. Firstly, at the time of writing, the variables analysed in CMIP5 in this study were only sparsely available in CMIP6 across both historical and scenario runs. Over 130 models have registered their source identifiers for CMIP6 with the World Climate Research Programme (WCRP)[1], so many more simulations are expected to be available over the coming months and years. Secondly, the altered pattern in CMIP6 (c.f. Fig. 1c and Fig. 9) preliminarily indicates that different processes or biases are at play. It is likely that CMIP6 analyses will reveal a different story altogether: about the Antarctic region, rather than Southern Ocean dynamics. A new future study will focus on unravelling the processes underpinning this higher latitude link between baseline Southern Ocean surface temperature biases and future warming.

Despite the current relatively small sample of CMIP6 models from the eventual number expected, some findings relevant to this study have emerged in the literature. It has been found, for example, that 10 out of 27 CMIP6 models analysed simulate higher equilibrium climate sensitivity than any of those in CMIP5 (Zelinka et al., 2020). Although the shift in ECS range is statistically

[1] https://wcrp-cmip.github.io/CMIP6_CVs/

insignificant, the higher sensitivity is due to a stronger reduction of lower level cloud cover under global warming, particularly in the Southern Hemisphere extratropics (Zelinka et al., 2020). Efforts to understand the plausibility of models with higher sensitivity is underway, with the recognition that substantially more CMIP6 simulations are expected. In terms of the global energy budget, CMIP6 is in better agreement with reference estimates than earlier model generations, and particularly for shortwave clear-sky budgets (Wild, 2020).

CMIP6 appears to show a stronger intermodel relationship between the global temperature trends of the recent past (i.e. 1981-2014) with both equilibrium climate sensitivity and transient climate response, as compared with CMIP5 (Tokarska et al., 2020). This opens the potential for future warming estimates to be constrained by observations, as more CMIP6 models become available.

With regards to the other variables examined in this study, CMIP6 exhibits mixed results to date. Despite a larger under-representation in boreal summer Antarctic sea ice area in CMIP6 (Roach et al., 2020), there are nevertheless some positive signs of improvement. For example, there is a reduction in the intermodel spread of seasonal sea ice variations, and the regional distribution is improved, compared to CMIP5 (Roach et al., 2020). The Southern Hemisphere jet stream and storm tracks are also less biased in CMIP6, exhibiting higher mean jet latitude (Bracegirdle et al., 2020; Curtis et al., 2020; Goyal et al., 2020; Priestley et al., 2020), and therefore reduced jet shift under future warming (Curtis et al., 2020). The reduced biases in the simulation of the jet stream is likely due to increased horizontal atmospheric resolution (Curtis et al., 2020). Along with improvements to the representation of surface wind stress forcing, the simulated strength of the Antarctic Circumpolar Current and associated density gradients have improved in CMIP6 (Beadling et al., 2020). The simulated mean sea level has also improved in the Southern Ocean (Lyu et al., 2020).

4 Conclusions

A summary of the relationships revealed in this study is shown in Fig. 10. The schematic highlights the series of CMIP5 model tendencies for those with warmer or cooler Southern Ocean baseline temperature. The schematic is an attempt to illustrate only overall model tendencies: not every model with an initial cool Southern Ocean, for instance, will have all of the features shown in Fig. 10a. The relationships illustrated for the baseline state are physically consistent, i.e. with warmer Southern Ocean there is a tendency for less cloud, and therefore less TOA outgoing shortwave radiation, less sea ice and a more poleward eddy-driven jet (Fig. 10a). Under global warming, in models with initially warmer Southern Ocean, there are lower reductions in sea ice, clouds, and TOA outgoing shortwave radiation, and smaller latitudinal shifts in the eddy-driven jet. Conversely, in models with initially cooler Southern Ocean, there is a tendency for initially larger cloud and sea ice area, higher TOA outgoing shortwave and an eddy-driven jet that is positioned more equatorward (Fig. 10b). Under global warming, initially cooler models tend to simulate a greater poleward jet shift, and a greater reduction in outgoing shortwave, clouds, and sea ice cover. The apparent influence of these relationships in the Southern Ocean climate system is that models with initially warmer Southern Ocean exhibit less global warming, and initially cooler models exhibit more global warming. As noted earlier, this relationship with the amount of global warming may be a result of potential capacity for change, e.g. models with more sea ice initially have greater potential to lose sea ice.

The baseline temperature appears to be a crucial variable in the Southern Ocean in CMIP5, since each of the other variables inspected exhibits a strong intermodel correlation with it, but not necessarily amongst themselves. But we do not necessarily suggest that models be tuned for baseline temperature, even if that is plausible. All variables are linked to one another: tuning a model for one particular baseline parameter, would invariably alter the baseline states of other parameters, but not necessarily favourably. For instance, the intermodel correlations imply that an attempt to cool the Southern Ocean surface in a model which is too warm might shift its jet latitude equatorward (Fig. 3f), but it was shown that most models already have an equatorward bias in the jet latitude (Fig. 8e). There are similar inconsistencies when comparing with observations in other variables: the model mean of Southern Ocean temperature is close to reanalysis (Fig. 8b), and while cloud cover (Fig. 8d) and sea ice (Fig. 8f) is under-represented, there is a tendency for too much TOA outgoing shortwave radiation (Fig. 8c). In a set of simple tests, we also found that it is not necessarily possible to constrain the spread in CMIP5 GMST projections by subsampling models that align more closely with reanalysis.

The primary finding of this study is that there are strong intermodel relationships in the baseline Southern Ocean climate system in CMIP5, but notably, the relationships are physically consistent. The relationships likely emerge due to biases, i.e. broad ranges in values of the simulated baseline variables. Furthermore, the baseline Southern Ocean biases consistently influence simulated changes under global warming. For example, initially cooler models tend to warm more into the future, partly because they have more sea ice initially, and therefore more capacity to lose sea ice.

It was shown that in the first available CMIP6 models, the baseline temperature relationship with global mean temperature change is less pronounced over the Southern Ocean, which may be a result of model improvements. The position of the Southern Hemisphere mid-latitude jet, for instance, appears to be less biased (more poleward) across CMIP6 models. Instead, stronger intermodel correlations emerge in the Antarctic sea ice region, suggesting that biases in polar region dynamics in CMIP6, rather than Southern Ocean dynamics in CMIP5, have a greater influence on global changes. As more CMIP6 model output becomes available, an examination of biases in the Antarctic region and their impact will be conducted.

Acknowledgements

This work was supported by the Natural Environment Research Council (SMURPHS project NE/N005783/1 and NE/N006348/1), and by the Australian Research Council Centre of Excellence for Climate Extremes (CE170100023). A.S. is supported by the Centre for Southern Hemisphere Oceans Research, a joint research centre between QNLM and CSIRO, and by the Earth Systems and Climate Change Hub of the Australian Government's National Environmental Science Program. A.S.T. is supported by the Australian Research Council (FT160100495). L.M.F. is supported by the Australian Research Council (DE170100367).

We acknowledge the World Climate Research Programme (WCRP) Working Group on Coupled Modeling, which is responsible for the Coupled Model Intercomparison Project (CMIP), and we thank the climate modeling groups for producing and making their model output available (https://esgf-node.llnl.gov). CMIP5 and CMIP6 model outputs were made available with the assistance of resources from the National Computational Infrastructure (NCI), which is supported by the Australian Government. We thank the NOAA Physical Sciences Laboratory for making the Twentieth Century Reanalysis data available. Support for the Twentieth Century Reanalysis Project version 3 dataset is provided by the U.S. Department of Energy, Office of Science Biological and Environmental Research (BER), by the National Oceanic and Atmospheric Administration Climate Program Office, and by the NOAA Physical Sciences Laboratory.

References

Andrews, T., Gregory, J. M., Webb, M. J., & Taylor, K. E. (2012). Forcing, feedbacks and climate sensitivity in CMIP5 coupled atmosphere-ocean climate models. *Geophysical Research Letters, 39*(9), L09712. https://doi.org/10.1029/2012GL051607

Arblaster, J. M., & Meehl, G. A. (2006). Contributions of external forcings to southern annular mode trends. *Journal of Climate*, *19*(12), 2896–2905. https://doi.org/10.1175/JCLI3774.1

Barnes, E. A., & Polvani, L. M. (2013). Response of the midlatitude jets, and of their variability, to increased greenhouse gases in the CMIP5 models. *Journal of Climate*, *26*(18), 7117–7135. https://doi.org/10.1175/JCLI-D-12-00536.1

Beadling, R. L., Russell, J. L., Stouffer, R. J., Mazloff, M., Talley, L. D., Goodman, P. J., et al. (2020). Representation of Southern Ocean properties across Coupled Model Intercomparison Project generations: CMIP3 to CMIP6. *Journal of Climate*, *33*, 6555–6581. https://doi.org/10.1175/JCLI-D-19-0970.1.

Betts, A. K., & Harshvardhan. (1987). Thermodynamic Constraint on the Cloud Liquid Water Feedback in Climate Models. *Journal of Geophysical Research*, *92*(D7), 8483–8485.

Bodas-Salcedo, A., Mulcahy, J. P., Andrews, T., Williams, K. D., Ringer, M. A., Field, P. R., & Elsaesser, G. S. (2019). Strong Dependence of Atmospheric Feedbacks on Mixed-Phase Microphysics and Aerosol-Cloud Interactions in HadGEM3. *Journal of Advances in Modeling Earth Systems*, *11*, 1735–1758. https://doi.org/10.1029/2019MS001688

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*, 195–211. https://doi.org/10.1175/JCLI-D-17-0320.1

Bracegirdle, T. J., Holmes, C. R., Hosking, J. S., Marshall, G. J., Osman, M., Patterson, M., & Rackow, T. (2020). Improvements in Circumpolar Southern Hemisphere Extratropical Atmospheric Circulation in CMIP6 Compared to CMIP5. *Earth and Space Science*, *7*, e2019EA001065. https://doi.org/10.1029/2019EA001065

Bromwich, D. H., Nicolas, J. P., Hines, K. M., Kay, J. E., Key, E. L., Lazzara, M. A., et al. (2012). Tropospheric clouds in Antarctica. *Reviews of Geophysics*, *50*(1), RG1004. https://doi.org/10.1029/2011RG000363

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., et al. (2011). The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, *137*(654), 1–28. https://doi.org/10.1002/qj.776

Curtis, P. E., Ceppi, P., & Zappa, G. (2020). Role of the mean state for the Southern Hemispheric jet stream response to CO2 forcing in CMIP6 models. *Environmental Research Letters*, *15*, 064011.

Flato, G. M., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W. J., & Cox, P. M. (2013). Evaluation of Climate Models. In *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. 741–866).

Forster, P. M., Andrews, T., Good, P., Gregory, J. M., Jackson, L. S., & Zelinka, M. D. (2013). Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models. *Journal of Geophysical Research Atmospheres*, *118*(3), 1139–1150. https://doi.org/10.1002/jgrd.50174

Frey, W. R., Morrison, A. L., Kay, J. E., Guzman, R., & Chepfer, H. (2018). The Combined Influence of Observed Southern Ocean Clouds and Sea Ice on Top-of-Atmosphere Albedo. *Journal of Geophysical Research: Atmospheres*, *123*(9), 4461–4475. https://doi.org/10.1029/2018JD028505

Frölicher, T. L., Sarmiento, J. L., Paynter, D. J., Dunne, J. P., Krasting, J. P., & Winton, M. (2015). Dominance of the Southern Ocean in anthropogenic carbon and heat uptake in CMIP5 models. *Journal of Climate*, *28*(2), 862–886. https://doi.org/10.1175/JCLI-D-14-00117.1

Goyal, R., Sen Gupta, A., Jucker, M., & England, M. H. (2020). Historical and projected changes in the Southern Hemisphere surface westerlies. *Submitted to Geophysical Research Letters*, (under review).

Gregory, J. M., Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., et al. (2004). A new method for diagnosing radiative forcing and climate sensitivity. *Geophysical Research Letters*, *31*, L03205. https://doi.org/10.1029/2003GL018747

Grise, K. M., & Medeiros, B. (2016). Understanding the varied influence of midlatitude jet position on clouds and cloud radiative effects in observations and global climate models. *Journal of Climate*, *29*(24), 9005–9025. https://doi.org/10.1175/JCLI-D-16-0295.1

Grise, K. M., Polvani, L. M., & Fasullo, J. T. (2015). Reexamining the relationship between climate sensitivity and the Southern Hemisphere radiation budget in CMIP models. *Journal of Climate*, *28*(23), 9298–9312. https://doi.org/10.1175/JCLI-D-15-0031.1

Grose, M. R., Gregory, J., Colman, R., & Andrews, T. (2018). What Climate Sensitivity Index Is Most Useful for Projections? *Geophysical Research Letters*, *45*(3), 1559–1566. https://doi.org/10.1002/2017GL075742

Sen Gupta, A., Jourdain, N. C., Brown, J. N., & Monselesan, D. P. (2013). Climate drift in the CMIP5 models. *Journal of Climate*, *26*(21), 8597–8615. https://doi.org/10.1175/JCLI_D_12_00521.1

Hall, A., Cox, P. M., Huntingford, C., & Klein, S. (2019). Progressing emergent constraints on future climate change. *Nature Climate Change*, *9*(4), 269–278. https://doi.org/10.1038/s41558-019-0436-6

Hawkins, E., & Sutton, R. T. (2016). Connecting climate model projections of global temperature change with the real world. *Bulletin of the American Meteorological Society*, *97*(6), 963–980. https://doi.org/10.1175/BAMS-D-14-00154.2

Hyder, P., Edwards, J. M., Allan, R. P., Hewitt, H. T., Bracegirdle, T. J., Gregory, J. M., et al. (2018). Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nature Communications*, *9*(1), 3625. https://doi.org/10.1038/s41467-018-05634-2

Kidston, J., & Gerber, E. P. (2010). Intermodel variability of the poleward shift of the austral jet stream in the CMIP3 integrations linked to biases in 20th century climatology. *Geophysical Research Letters*, *37*(9), L09708. https://doi.org/10.1029/2010GL042873

Kidston, J., Taschetto, A. S., Thompson, D. W. J., & England, M. H. (2011). The influence of Southern Hemisphere sea-ice extent on the latitude of the mid-latitude jet stream. *Geophysical Research Letters*, *38*(15), L15804. https://doi.org/10.1029/2011GL048056

Klein, S. A., & Hartmann, D. L. (1993). The seasonal cycle of low stratiform clouds. *Journal of Climate*, *6*, 1587–1606.

Knutti, R., Rugenstein, M. A. A., & Hegerl, G. C. (2017). Beyond equilibrium climate sensitivity. *Nature Geoscience*, *10*(10), 727–736. https://doi.org/10.1038/NGEO3017

Lyu, K., Zhang, X., & Church, J. A. (2020). Regional Dynamic Sea Level Simulated in the CMIP5 and CMIP6 Models: Mean Biases, Future Projections, and Their Linkages. *Journal of Climate*, *33*(15), 6377–6398. https://doi.org/10.1175/jcli-d-19-1029.1

Manabe, S., Stouffer, R. J., Spelman, M. J., & Bryan, K. (1991). Transient responses of a coupled ocean--atmosphere model to gradual changes of atmospheric CO2. Part I. Annual mean response. *Journal of Climate*, *4*(8), 785–818.

Marshall, J., & Speer, K. (2012). Closure of the meridional overturning circulation through Southern Ocean upwelling. *Nature Geoscience*, *5*(3), 171–180.

Meehl, G. A., Senior, C. A., Eyring, V., Flato, G., Lamarque, J.-F., Stouffer, R. J., et al. (2020). Context for interpreting equilibrium climate sensitivity and transient climate response from the CMIP6 Earth system models. *Science Advances*, *6*(26), eaba1981. https://doi.org/10.1126/sciadv.aba1981

Mikaloff Fletcher, S. E., Gruber, N., Jacobson, A. R., Doney, S. C., Dutkiewicz, S., Gerber, M., et al. (2006). Inverse estimates of anthropogenic CO2 uptake, transport, and storage by the ocean. *Global Biogeochemical Cycles*, *20*(2). https://doi.org/10.1029/2005GB002530

Miller, R. L., Schmidt, G. A., & Shindell, D. T. (2006). Forced annular variations in the 20th century Intergovernmental Panel on Climate Change Fourth Assessment Report models. *Journal of Geophysical Research: Atmospheres*, *111*(18), D18101. https://doi.org/10.1029/2005JD006323

O'Neill, B. C., Tebaldi, C., Van Vuuren, D. P., Eyring, V., Friedlingstein, P., Hurtt, G., et al. (2016). The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, *9*(9), 3461–3482. https://doi.org/10.5194/gmd-9-3461-2016

Priestley, M. D. K., Ackerley, D., Catto, J. L., Hodges, K. I., McDonald, R. E., & Lee, R. W. (2020). An overview of the extratropical storm tracks in CMIP6. *Journal of Climate*, *33*, 6315–6343. https://doi.org/10.1175/JCLI-D-19-0928.1.

Roach, L. A., Dörr, J., Holmes, C. R., Massonnet, F., Blockley, E. W., Notz, D., et al. (2020). Antarctic Sea Ice Area in CMIP6. *Geophysical Research Letters*, *47*, e2019GL086729. https://doi.org/10.1029/2019GL086729

Sherwood, S., Webb, M. J., Annan, J. D., Armour, K. C., Forster, P. M., Hargreaves, J. C., et al. (2020). An assessment of Earth's climate sensitivity using multiple lines of evidence. *Reviews of Geophysics*. https://doi.org/10.1029/2019RG000678

Simpson, I. R., & Polvani, L. M. (2016). Revisiting the relationship between jet position, forced response, and annular mode variability in the southern midlatitudes. *Geophysical Research Letters*, *43*(6), 2896–2903. https://doi.org/10.1002/2016GL067989

Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., et al. (2013). *Climate change 2013: The physical science basis*. *Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change*. Cambridge university press Cambridge, United Kingdom and New York, NY, USA.

Toggweiler, J. R., & Samuels, B. (1995). Effect of Drake Passage on the global thermohaline circulation. *Deep Sea Research Part I: Oceanographic Research Papers*, *42*(4), 477–500.

Tokarska, K. B., Stolpe, M. B., Sippel, S., Fischer, E. M., Smith, C. J., Lehner, F., & Knutti, R. (2020). Past warming trend constrains future warming in CMIP6 models. *Science Advances*, *6*(March), eaaz9549.

Trenberth, K. E., & Fasullo, J. T. (2010). Simulation of present-day and twenty-first-century energy budgets of the southern oceans. *Journal of Climate*, *23*(2), 440–454. https://doi.org/10.1175/2009JCLI3152.1

Wall, C. J., Hartmann, D. L., & Ma, P.-L. (2017). Instantaneous linkages between clouds and large-scale meteorology over the Southern Ocean in observations and a climate model. *Journal of Climate*, *30*(23), 9455–9474. https://doi.org/10.1175/JCLI-D-17-0156.1

Wild, M. (2020). The global energy balance as represented in CMIP6 climate models. *Climate Dynamics*, *55*, 553–577. https://doi.org/10.1007/s00382-020-05282-7

Williams, K. D., Copsey, D., Blockley, E. W., Bodas-Salcedo, A., Calvert, D., Comer, R., et al. (2017). The Met Office Global Coupled Model 3.0 and 3.1 (GC3.0 and GC3.1) Configurations. *Journal of Advances in Modeling Earth Systems*, *10*(2), 357–380. https://doi.org/10.1002/2017MS001115

Xie, S.-P., Deser, C., Vecchi, G. A., Ma, J., Teng, H., & Wittenberg, A. T. (2010). Global warming pattern formation: Sea surface temperature and rainfall. *Journal of Climate*, *23*(4), 966–986. https://doi.org/10.1175/2009JCLI3329.1

Zelinka, M. D., Myers, T. A., McCoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., et al. (2020). Causes of higher climate sensitivity in CMIP6 models. *Geophysical Research Letters*, *47*, e2019GL085782. https://doi.org/10.1029/2019GL085782

Zhai, C., Jiang, J. H., & Su, H. (2015). Long-term cloud change imprinted in seasonal cloud variation: More evidence of high climate sensitivity. *Geophysical Research Letters*, *42*(20), 8729–8737. https://doi.org/10.1002/2015GL065911

Zhu, J., & Poulsen, C. J. (2020). On the increase of climate sensitivity and cloud feedback with warming in the Community Atmosphere Models. *Geophysical Research Letters*, *47*, e2020GL089143. https://doi.org/doi:10.1029/2020GL089143

	Model name	Ensemble members used	Exceptions	ECS (°C)
1	ACCESS1-0	r1i1p1		3.83
2	ACCESS1-3	r1i1p1		3.54
3	bcc-csm1-1	r1i1p1		2.83
4	bcc-csm1-1-m	r1i1p1		2.91
5	BNU-ESM	r1i1p1		4.04
6	CanESM2	r1i1p1, r2i1p1, r3i1p1, r4i1p1, r5i1p1		3.71
7	CCSM4	r1i1p1, r2i1p1, r3i1p1, r4i1p1, r5i1p1, r6i1p1	Missing for *clt*: r6i1p1	2.95
8	CESM1-BGC	r1i1p1		2.89[a]
9	CESM1-CAM5	r1i1p1, r2i1p1, r3i1p1	Missing for *clt*: r1i1p1, r2i1p1	4.10[b]
10	CMCC-CESM	r1i1p1		
11	CMCC-CM	r1i1p1		
12	CMCC-CMS	r1i1p1		
13	CNRM-CM5	r1i1p1, r2i1p1, r4i1p1, r6i1p1, r10i1p1		3.25
14	CSIRO-Mk3-6-0	r1i1p1, r2i1p1, r3i1p1, r4i1p1, r5i1p1, r6i1p1, r7i1p1, r8i1p1, r9i1p1, r10i1p1		4.06
15	FGOALS-g2	r1i1p1		3.35
16	FGOALS-s2	r1i1p1, r2i1p1, r3i1p1	Missing for *rsut* and *clt*: r1i1p1	4.19
17	FIO-ESM	r1i1p1, r2i1p1, r3i1p1		
18	GFDL-CM3	r1i1p1		4.00
19	GFDL-ESM2G	r1i1p1		2.43
20	GFDL-ESM2M	r1i1p1		2.45
21	GISS-E2-H	r1i1p1, r1i1p2, r1i1p3, r2i1p1, r2i1p3	Missing for *tauu*: r2i1p1, r2i1p3	2.30
22	GISS-E2-H-CC	r1i1p1		
23	GISS-E2-R	r1i1p1, r1i1p2, r1i1p3, r2i1p1, r2i1p3	Missing for *tas*: r1i1p3 Missing for *tauu*: r2i1p1, r2i1p3	2.11
24	GISS-E2-R-CC	r1i1p1		
25	HadGEM2-AO	r1i1p1, r2i1p1, r3i1p1	Missing for *clt*: r2i1p1, r3i1p1	
26	HadGEM2-CC	r1i1p1		
27	HadGEM2-ES	r1i1p1, r2i1p1, r3i1p1, r4i1p1		4.58
28	inmcm4	r1i1p1		2.08
29	IPSL-CM5A-LR	r1i1p1, r2i1p1, r3i1p1, r4i1p1		4.13
30	IPSL-CM5A-MR	r1i1p1		4.14
31	IPSL-CM5B-LR	r1i1p1		2.60
32	MIROC-ESM-CHEM	r1i1p1		
33	MIROC-ESM	r1i1p1		4.66
34	MIROC5	r1i1p1, r2i1p1, r3i1p1		2.71
35	MPI-ESM-LR	r1i1p1, r2i1p1, r3i1p1		3.63
36	MPI-ESM-MR	r1i1p1		3.45
37	MRI-CGCM3	r1i1p1		2.61
38	MRI-ESM1	r1i1p1		
39	NorESM1-M	r1i1p1		2.82
40	NorESM1-ME	r1i1p1		2.99[c]

Table 1. List of CMIP5 models and ensemble members analysed in this study. Ensemble members from the *historical* experiments were matched with ensemble members with the same identifiers from the *rcp85* experiments. The five primary variables analysed in this study (*tas*, *rsut*, *clt*, *tauu*, and *sic*) were available from all models and ensemble members, unless noted under 'Exceptions'. The equilibrium climate sensitivity (ECS) is recorded for models where available, and taken from Caldwell et al., (2016; their Table 1 and Equation 2), with three exceptions: [a]Nohara et al., (2015), [b]Meehl et al., (2013), [c]Seland et al., (2020).

	Model name	Ensemble member used
1	ACCESS-CM2	r1i1p1f1
2	ACCESS-ESM1-5	r1i1p1f1
3	AWI-CM-1-1-MR	r1i1p1f1
4	BCC-CSM2-MR	r1i1p1f1
5	CAMS-CSM1-0	r1i1p1f1
6	CanESM5-CanOE	r1i1p2f1
7	CanESM5	r1i1p1f1
8	CESM2	r1i1p1f1
9	CESM2-WACCM	r1i1p1f1
10	CNRM-CM6-1	r1i1p1f2
11	CNRM-CM6-1-HR	r1i1p1f2
12	CNRM-ESM2-1	r1i1p1f2
13	EC-Earth3	r1i1p1f1
14	EC-Earth3-Veg	r1i1p1f1
15	FGOALS-f3-L	r1i1p1f1
16	FGOALS-g3	r1i1p1f1
17	FIO-ESM-2-0	r1i1p1f1
18	GFDL-CM4	r1i1p1f1
19	GFDL-ESM4	r1i1p1f1
20	HadGEM3-GC31-LL	r1i1p1f3
21	INM-CM4-8	r1i1p1f1
22	INM-CM5-0	r1i1p1f1
23	IPSL-CM6A-LR	r1i1p1f1
24	KACE-1-0-G	r1i1p1f1
25	MCM-UA-1-0	r1i1p1f2
26	MIROC6	r1i1p1f1
27	MIROC-ES2L	r1i1p1f2
28	MPI-ESM1-2-HR	r1i1p1f1
29	MPI-ESM1-2-LR	r1i1p1f1
30	MRI-ESM2-0	r1i1p1f1
31	NESM3	r1i1p1f1
32	NorESM2-LM	r1i1p1f1
33	UKESM1-0-LL	r1i1p1f2

Table 2. List of CMIP6 models and ensemble members analysed in this study. Ensemble members from the *historical* experiments were matched with ensemble members with the same identifiers from the *ssp585* experiments. Only the *tas* variable was analysed in CMIP6.

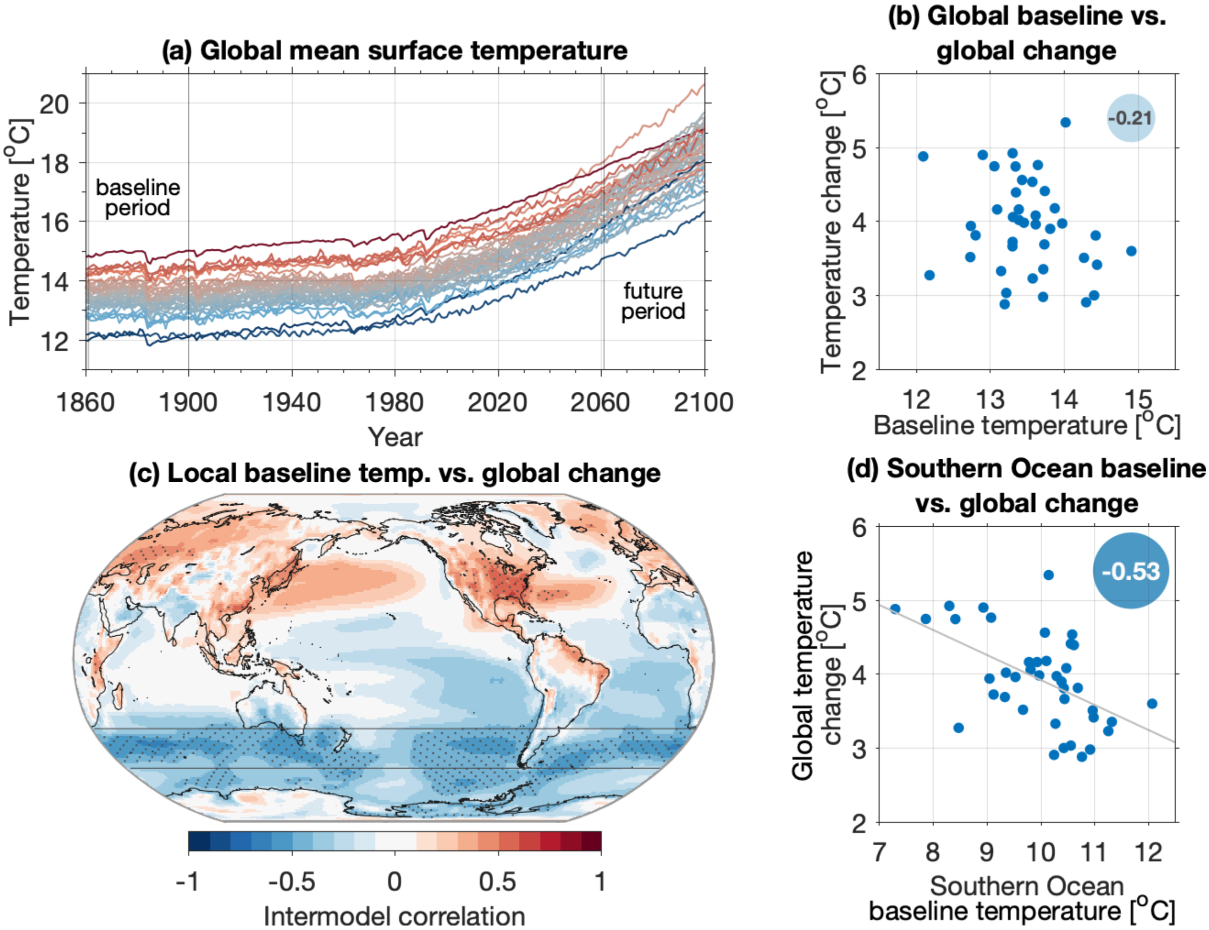

Figure 1. Surface air temperature relationships across CMIP5 models. **a.** Absolute annual global mean surface temperature (GMST) in CMIP5 *historical* simulations with *rcp85* extension. The baseline (1861-1900) and future (2061-2100) periods are indicated. The timeseries' are qualitatively shaded by baseline GMST: initially cooler models in blue and warmer models in red. **b.** GMST averaged over the baseline period, versus the GMST change (i.e. average over future period minus average over baseline). The intermodel correlation (r = -0.21) is quoted, but $p > 0.05$. **c.** Intermodel correlation between grid-point (local) baseline surface air temperature and GMST change. Stippling indicates where correlations are statistically significant at the 99% level. The Southern Ocean region (55-35 °S) used throughout this study is indicated. **d.** Baseline surface air temperature averaged over the Southern Ocean, versus the GMST change. The intermodel correlation (r = -0.53) is statistically significant at $p < 0.01$.

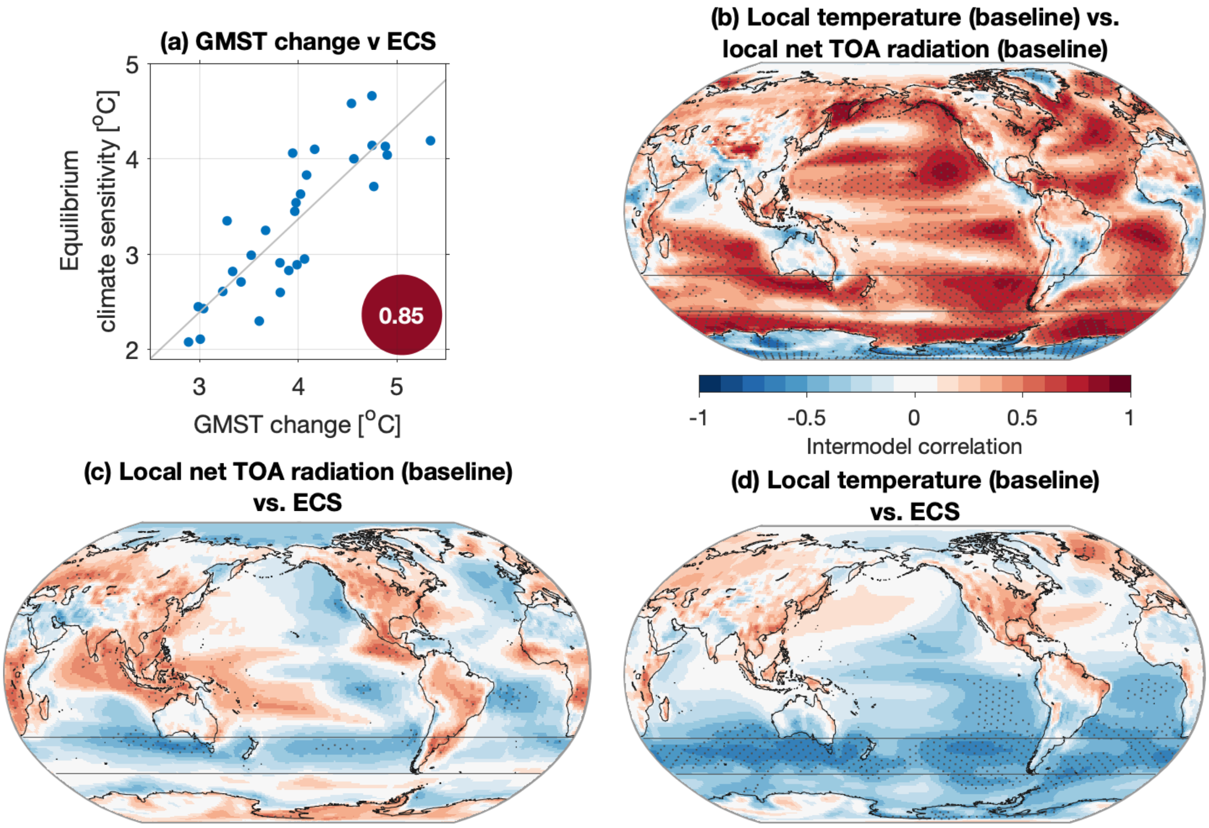

Figure 2. Relationships between equilibrium climate sensitivity (ECS) and other parameters. **a.** Global mean surface air temperature (GMST) change versus ECS, for the 30 models for which the ECS value is available (Table 1). The intermodel correlation (r = 0.85) is statistically significant at p < 0.01. **b.** Intermodel correlation between grid-point (local) baseline surface air temperature and grid-point baseline net top-of-atmosphere (TOA) radiation, across all 40 models. **c.** Intermodel correlation between grid-point (local) net TOA radiation and ECS (30 models). **d.** Intermodel correlation between grid-point (local) baseline surface air temperature and ECS (30 models). Stippling denotes statistically significant correlations at p < 0.01.

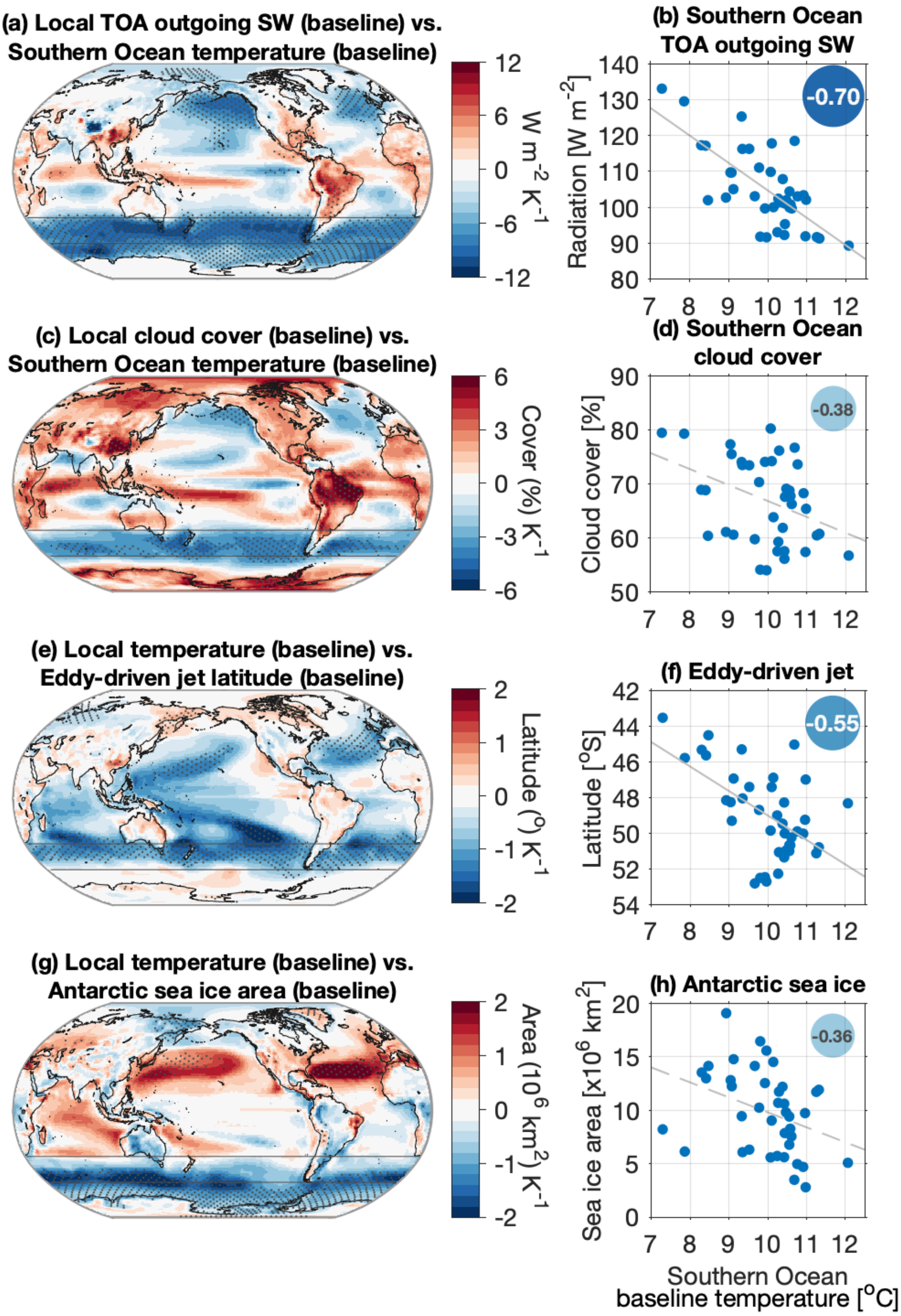

Figure 3. CMIP5 intermodel relationships between baseline temperature and other baseline variables. ‘Local’ denotes the variable inspected at each grid-point across the globe. The left panels show intermodel regressions, expressed as relationships per unit Kelvin, between **a.** TOA outgoing shortwave radiation and Southern Ocean average temperature; **c.** cloud fraction and Southern Ocean average temperature; **e.** temperature and eddy-driven jet latitude; and **g.** temperature and total Antarctic sea ice area. Stippling denotes statistically significant regressions at $p < 0.01$. The right panels show Southern Ocean average baseline surface temperature (abscissa) versus baseline **b.** Southern Ocean average TOA outgoing shortwave radiation; **d.** Southern Ocean average cloud fraction; **f.** eddy-driven jet latitude; and **h.** total Antarctic sea ice area. Intermodel correlations are quoted in the top right. Solid lines of best-fit denote $p < 0.01$, and dashed lines denote $0.01 < p < 0.05$.

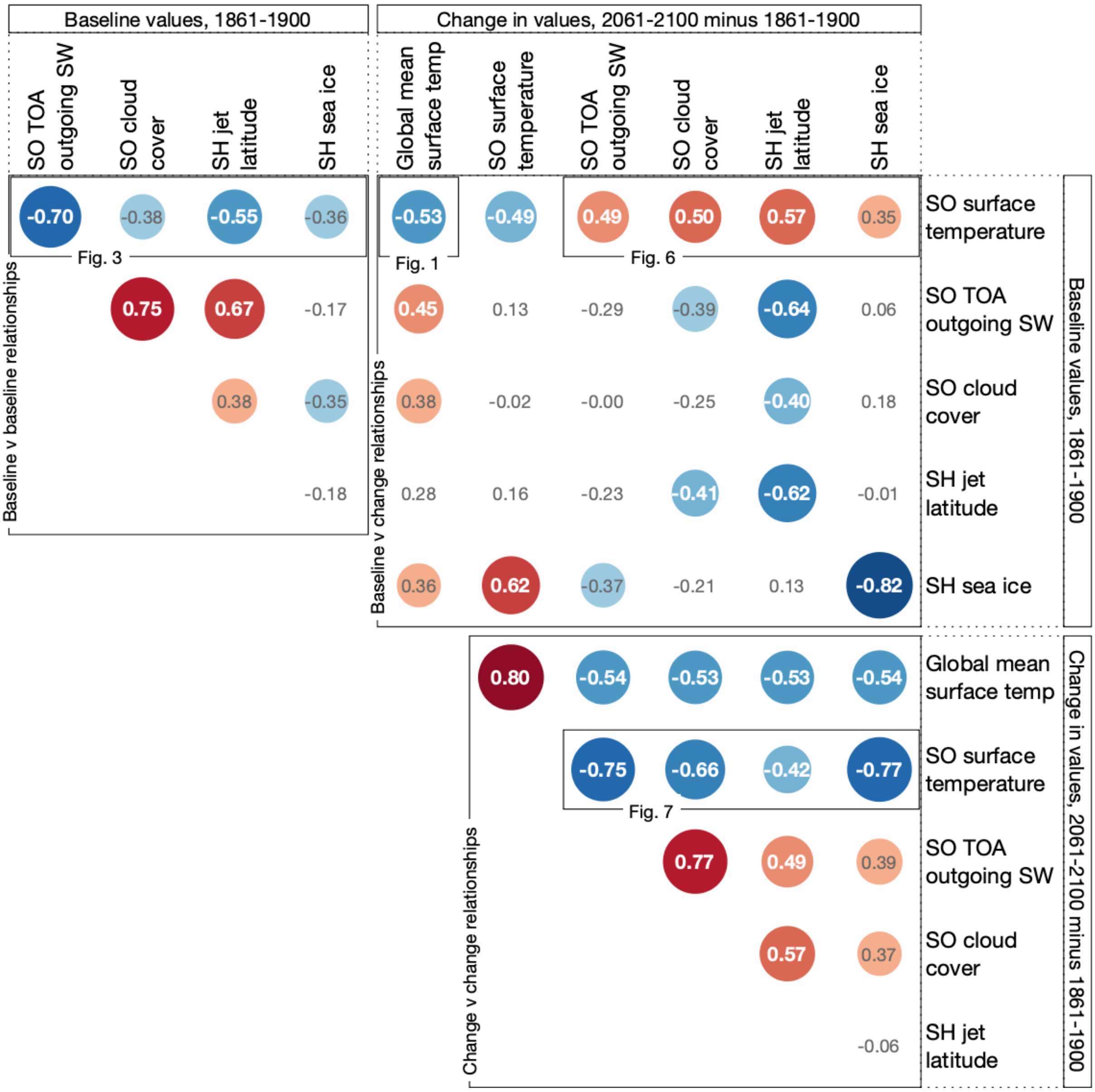

Figure 4. Intermodel correlations between baseline values and future change values of all variables analysed in this study. Red or blue shaded circles denotes positive or negative correlations, respectively, where darker shades are larger circles denote stronger correlations. Correlations that are statistically significant the 99% level are quoted in white text, and shaded circles are shown only where the correlations are statistically significant at the 95% level. The correlations shown in Figs. 1, 3, 6, and 7 are indicated.

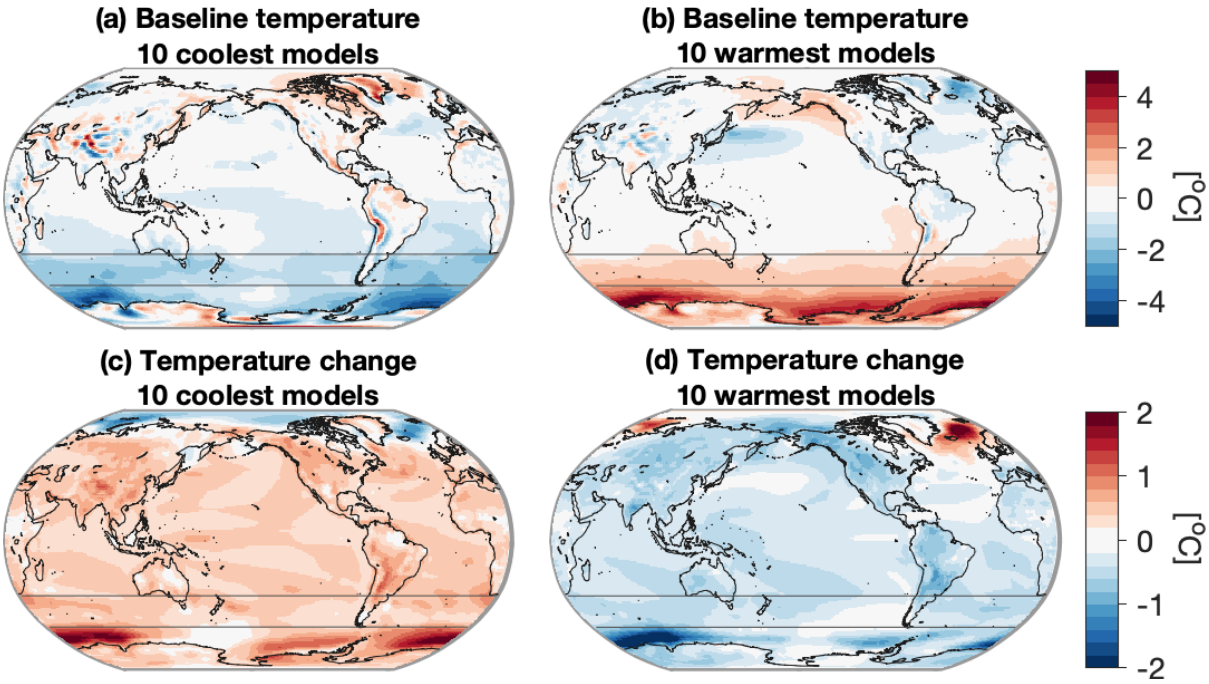

Figure 5. Composites of surface temperature in CMIP5 models by those with coolest and warmest baseline Southern Ocean surface temperature. **a.** Mean baseline surface temperature of 10 models with coolest Southern Ocean baseline temperature, shown as anomalies with respect to the model mean. **b**. As in (a), but for the 10 warmest models. **c**. Mean temperature change of 10 models with coolest Southern Ocean baseline temperature, shown as anomalies with respect to the model mean. **d**. As in (c), but for the 10 warmest models.

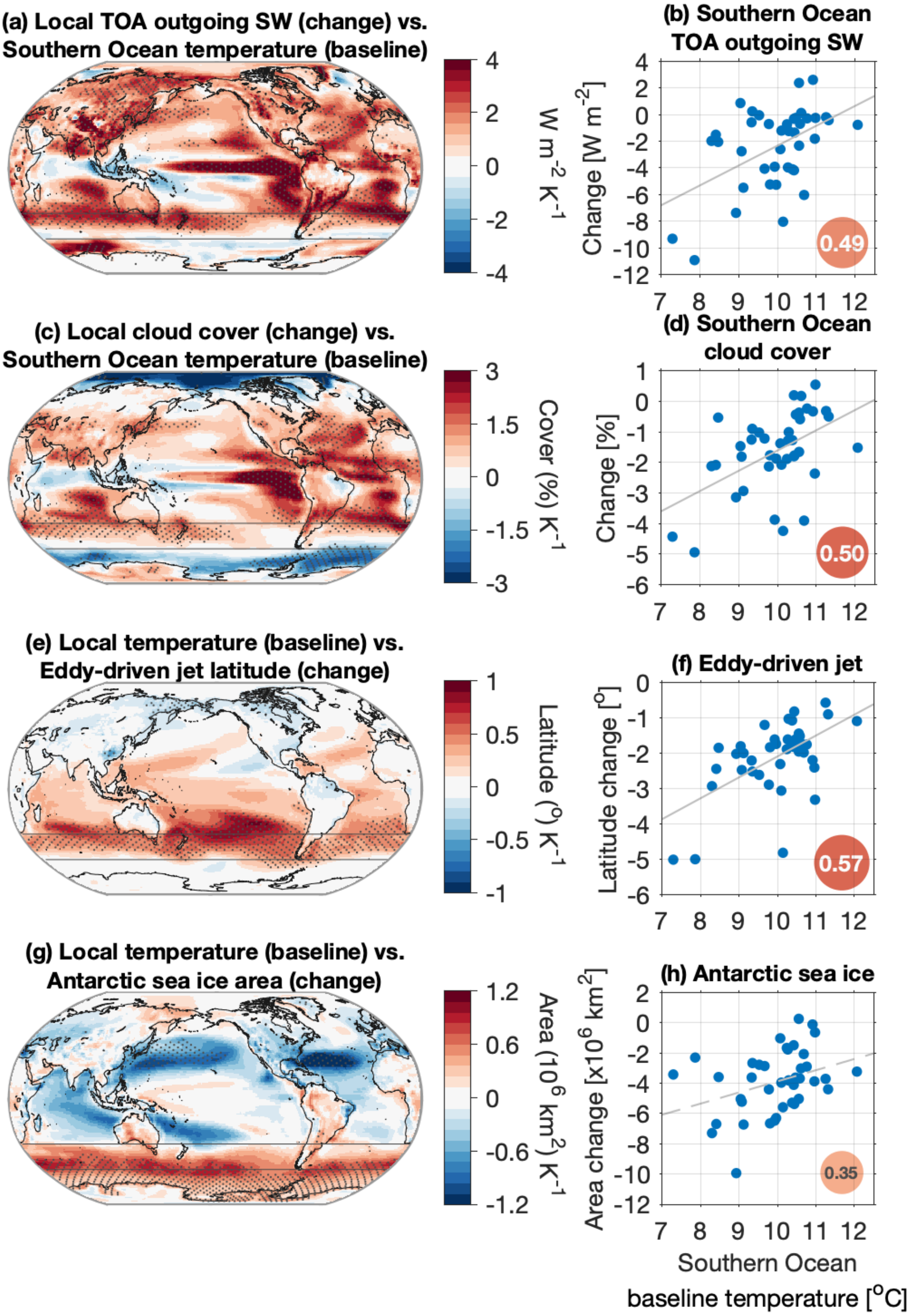

Figure 6. CMIP5 intermodel relationships between baseline temperature and future changes in other variables. 'Local' denotes the variable inspected at each grid-point across the globe. The left panels show intermodel regressions, expressed as relationships per unit Kelvin, between **a.** TOA outgoing shortwave radiation and Southern Ocean average temperature; **c.** cloud fraction and Southern Ocean average temperature; **e.** temperature and eddy-driven jet latitude; and **g.** temperature and total Antarctic sea ice area. Stippling denotes statistically significant regressions at $p < 0.01$. The right panels show Southern Ocean average baseline surface temperature (abscissa) versus future changes in **b.** Southern Ocean average TOA outgoing shortwave radiation; **d.** Southern Ocean average cloud fraction; **f.** eddy-driven jet latitude; and **h.** total Antarctic sea ice area. Intermodel correlations are quoted in the top right. Solid lines of best-fit denote $p < 0.01$, and dashed lines denote $0.01 < p < 0.05$.

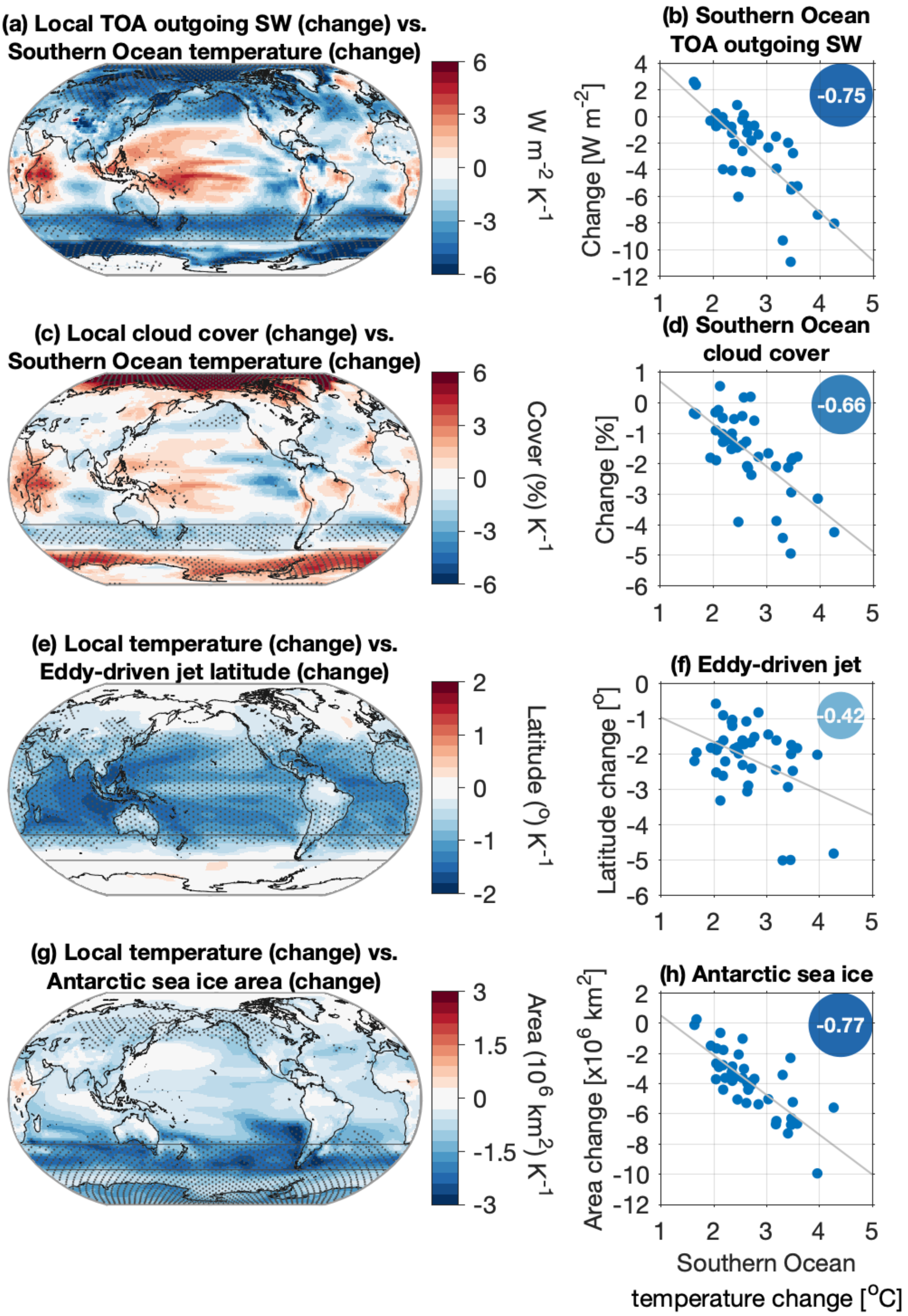

Figure 7. CMIP5 intermodel relationships between future changes in temperature and future changes in other variables. 'Local' denotes the variable inspected at each grid-point across the globe. The left panels show intermodel regressions, expressed as relationships per unit Kelvin change, between **a.** TOA outgoing shortwave radiation and Southern Ocean average temperature; **c.** cloud fraction and Southern Ocean average temperature; **e.** temperature and eddy-driven jet latitude; and **g.** temperature and total Antarctic sea ice area. Stippling denotes statistically significant regressions at $p < 0.01$. The right panels show Southern Ocean average baseline surface temperature (abscissa) versus future changes in **b.** Southern Ocean average TOA outgoing shortwave radiation; **d.** Southern Ocean average cloud fraction; **f.** eddy-driven jet latitude; and **h.** total Antarctic sea ice area. Intermodel correlations are quoted in the top right. Solid lines of best-fit denote $p < 0.01$.

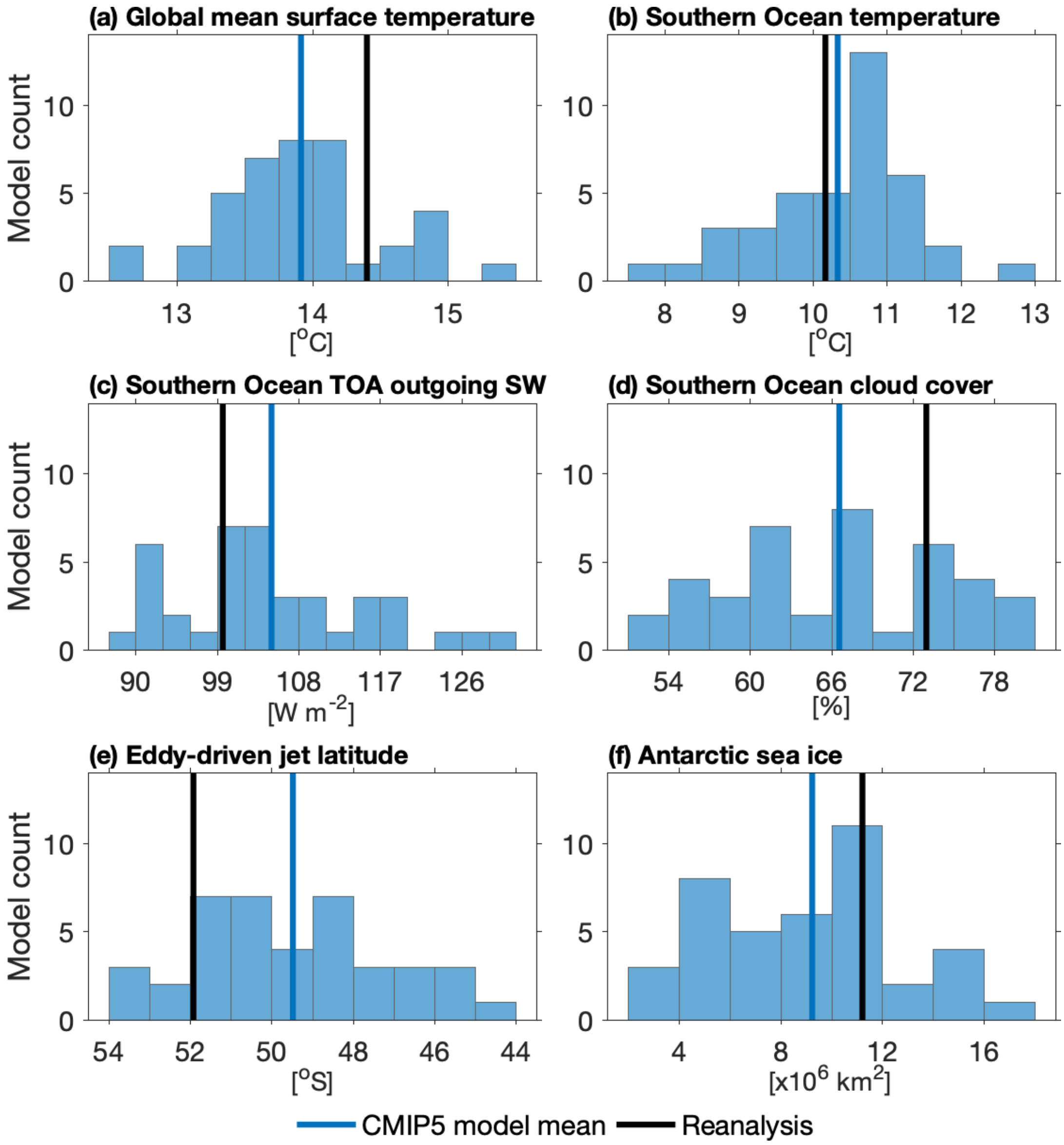

Figure 8. Histograms of parameters in CMIP5 models, averaged over the period 1961-2000. The vertical blue line denotes the model mean, and the black line denotes the NOAA-CIRES-DOE Twentieth Century Reanalysis. **a.** Global mean surface air temperature. **b.** Surface air temperature averaged over the Southern Ocean. **c.** Top-of-atmosphere outgoing shortwave radiation averaged over the Southern Ocean. **d.** Cloud cover averaged over the Southern Ocean. **e.** Eddy-driven jet latitude. **f.** Antarctic sea ice area.

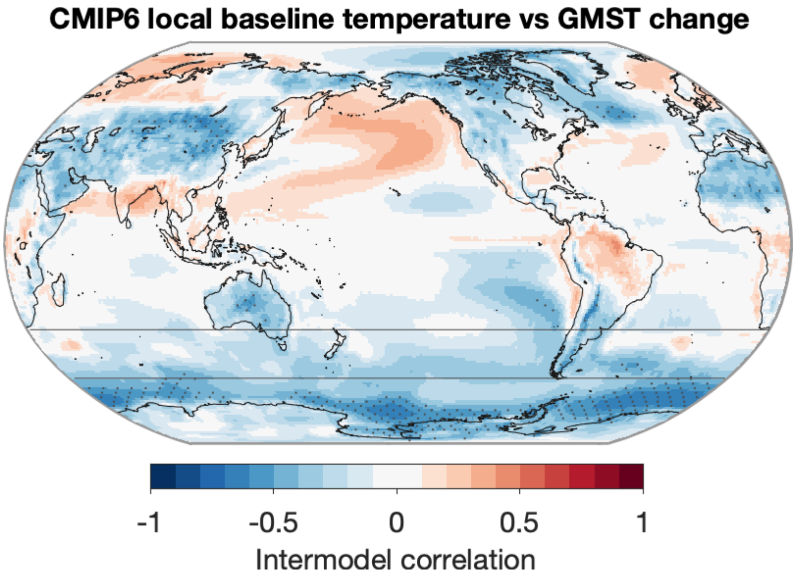

Figure 9. Intermodel correlation in 33 CMIP6 models (Table 2) between grid-point (local) baseline surface air temperature and global mean surface air temperature change. Stippling indicates where correlations are statistically significant at the 99% level. CMIP6 surface air temperature is analysed in the *historical* simulations with *ssp585* extension.

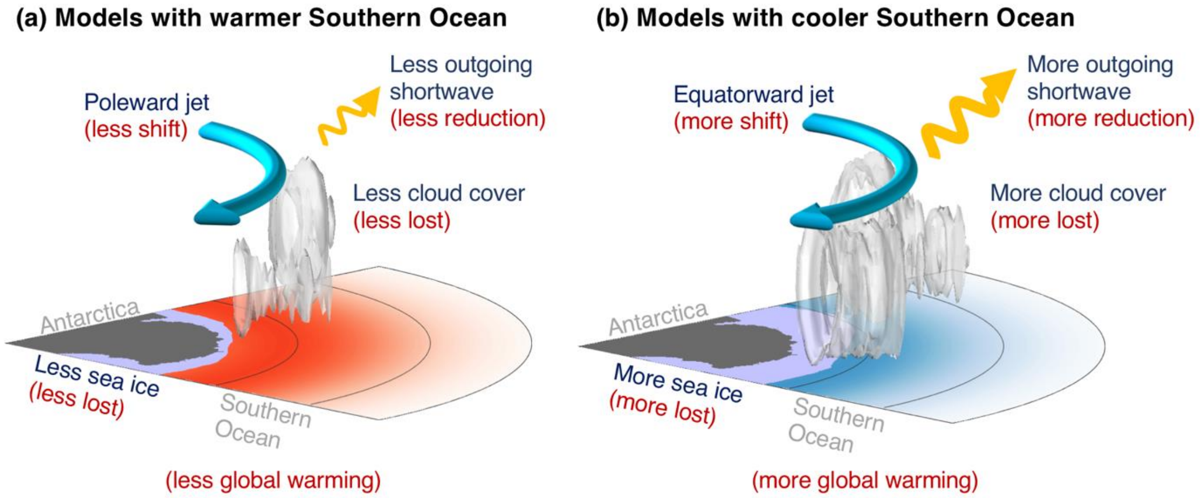

Figure 10. Schematic summary of model tendencies in CMIP5. Red text in parentheses indicate changes under global warming. **a.** Models with warmer baseline Southern Ocean surface air temperature. **b.** Models with cooler baseline Southern Ocean surface air temperature.