The Cryosphere Discuss., 9, 3113–3136, 2015
www.the-cryosphere-discuss.net/9/3113/2015/
doi:10.5194/tcd-9-3113-2015


The Cryosphere Discussions
Open Access







# Evaluation of the CMIP5 models in the aim of regional modelling of the Antarctic surface mass balance

**C. Agosta[1], X. Fettweis[1], and R. Datta[2]**

[1]Department of Geography, Climatology Laboratory, Université de Liège, Liège, Belgium
[2]City College of New York, City University of New York, New York, NY, 10031, USA

Received: 22 April 2015 – Accepted: 11 May 2015 – Published: 11 June 2015

Correspondence to: C. Agosta (cecile.agosta@gmail.com)

Published by Copernicus Publications on behalf of the European Geosciences Union.

## Abstract

The Antarctic surface mass balance (SMB) cannot be reliably deduced from global climate models (GCMs), both because their spatial resolution is insufficient and because their physics are not adapted for cold and snow-covered regions. By contrast, regional climate models (RCMs) adapted for polar regions can physically and dynamically downscale surface mass balance components over the ice-sheet using large scale forcing at their boundaries. Polar-oriented RCMs require appropriate GCM fields for forcing because the response of the cryosphere to a warming climate is dependent on its initial state and is not linear with respect to temperature increase. In this context, we evaluate current climate in 41 climate models from the Coupled Model Intercomparison Project Phase 5 (CMIP5) dataset over Antarctica by focusing on forcing fields which may have the greatest impact on SMB components simulated by RCMs. Our inter-comparison includes 5 reanalyses, among which ERA-Interim reanalysis is chosen as a reference over 1979–2014. Model efficiency is assessed taking into account the multi-decadal variability of the fields over the 1850–1980 period. We show that less than 10 CMIP5 models show reasonable biases compared to ERA-Interim, among which ACCESS1-3 seems to be the most pertinent choice for regional climate modeling over Antarctica, followed by CMCC-CM, MIROC-ESM/MIROC-ESM-CHEM and NorESM1-M. Finally, climate change over the Southern Ocean is much more dependent on the initial state of winter sea-ice extent and on the local feedback between air temperature increase and winter sea-ice extent decrease than on the global warming signal.

## 1 Introduction

Mass change in Antarctica is a major component of sea-level change. Projections of Antarctic mass changes are based on the input-output method, in which ice-sheet surface mass balance (SMB, input) and ice-sheet dynamics (output), are modeled separately. Antarctic mass budget is 10 times lower in magnitude than the individual input/output components. Consequently, when using the input-output method, uncertainty in mass change equals the sum of the uncertainties of input and output estimates, which are of the same order of magnitude as the mass change itself. That is why efforts are made to better estimate and reduce uncertainty on each of these two components.

The Antarctic SMB is driven by snowfall at the ice-sheet margins, although sublimation, melt, refreezing, and drifting snow can be of importance locally. These components cannot be reliably deduced from reanalyses or global climate models (GCMs) because their horizontal resolution ($\sim 100$ km) is insufficient and because their physics are not adapted for cold and snow-covered regions. Polar-oriented regional climate models (RCMs) are able to fill this gap because their physics have been specifically developed/calibrated for these areas. Forced with reanalyses, their results can be evaluated directly against meteorological, remote-sensing and SMB observations available in these high latitude regions. With regard to climate change, the response of the cryosphere will depend both on its initial state and on the climate change signal. Consequently, RCM results will rely on the ability of GCMs to adequately simulate the current climate as well as on GCM estimates of future changes.

Unlike previously published evaluations of the CMIP5 models over Antarctica which focus on specific fields such as westerly winds (Bracegirdle et al., 2014) or sea-ice (Turner et al., 2013; Mahlstein et al., 2013; Shu et al., 2015), in this paper we aim to evaluate the CMIP5 fields that will be used as input for RCMs and that may have the greatest impact on RCM-based SMB components: temperature, humidity, surface pressure and oceanic conditions.

After describing models and skill scores, we explain the selection of metrics, perform multi-metric analysis and establish relationships between climate change in GCMs and their representation of current climate. We conclude by discussing potential sources of bias in our method and by summarizing our main outcomes.

## 2 Data and methods

### 2.1 CMIP5 climate models and reanalyses

Monthly means fields from 41 CMIP5 models and 5 reanalyses, listed in Table 1, are compared in this work. All data were bi-linearly interpolated onto a common regular longitude-latitude horizontal grid ($0.5^{\circ} \times 1.5^{\circ}$ with a spatial domain extending south of 40° S over ocean. We did not include land and ice-covered areas, because RCM lateral boundaries are usually set far from these areas while GCMs results might be biased there because surface schemes are not properly adapted. Seasonal values are defined by 3 months means, with winter consisting of June–July–August for atmospheric variables and July–August–September for oceanic variables. All other seasons are defined with a similar one-month lag for oceanic variables.

CMIP5 data were retrieved from the Historical (1850–2005 period) and representative concentration pathway 8.5 "RCP85" (2006–2100 period) coupled ocean-atmosphere experiments. The RCP85 scenario is an upper range of plausible future emission for which greenhouse gas radiative forcing continues to rise throughout the 21st century until the 1370 ppm $CO_2$ equivalent (Moss et al., 2010). In this scenario, stratospheric ozone recovery is represented across the CMIP5 models, with recovery over Antarctica to near pre-ozone hole amounts by 2100. For each CMIP5 model, we considered the first realization only (r1i1p1) and merged Historical and RCP85 to form continuous time series from 1850 to 2100. Given the high number of models investigated, we highlighted models which contained obvious similarities in code or were produced by the same institution (colors in Figs. 2 and 3), following the work of Knutti et al. (2013, colors in their Fig. 1).

Recent reanalysis inter-comparisons have shown the European Centre for Medium-Range Weather Forecasts "Interim" re-analysis (ERA-Interim, Dee et al., 2011) to be the most reliable contemporary global reanalysis over Antarctica (Bromwich et al., 2011; Bracegirdle and Marshall, 2012), prompting our choice of ERA-Interim as a reference for representing the current climate (1980–2010). However, comparisons with four other reanalyses are also performed in our study: the National Aeronautics and Space Administration Modern Era Retrospective-Analysis for Research and Applications (MERRA, Rienecker et al., 2011); the National Centers for Environmental Prediction (NCEP)/National Center for Atmospheric Research Global Reanalysis 1 (NCEP1, Kalnay et al., 1996); the NCEP/Department of Energy Atmospheric Model Intercomparison Project 2 reanalysis (NCEP2, Kanamitsu et al., 2002); and the National Oceanic and Atmospheric Administration (NOAA) Twentieth Century Reanalysis v2 (20CR, Compo et al., 2011).

We will later define metrics to compare CMIP5 GCMs outputs with ERA-Interim over the period 1980–2010 (31 years). In order to reduce the sensitivity of our comparisons to the choice of this reference period, we computed the multi-decadal intrinsic variability of those metrics. Over the Antarctic region considered, CMIP5 GCM metrics show no significant trends until the 1980's, but evolve significantly afterwards. Consequently, we estimated the multi-decadal climate variability of each metric for every CMIP5 GCM by considering the variability of the 31 year running metric during the stable period 1850–1980. We present this estimate in details in Appendix A. The multi-decadal variability estimate gives an error bar around the reference period value, which depends on each metric and each model (Table 1).

## 2.2 Indexes and scores

Spatial bias $b$ and centered root mean square error (rmse) $c$ are measures which are easy to interpret and are defined formally as follows:

$$\begin{cases} b = \langle \mu_t^{\mathrm{m}} - \mu_t^{\mathrm{o}} \rangle_{xy}, \\ c = \sqrt{\left\langle \left(\mu_t^{\mathrm{m}} - \mu_t^{\mathrm{o}} - b\right)^2 \right\rangle_{xy}}, \end{cases} \quad (1)$$

where m and o exponents are for model outputs and observations respectively, $\mu_t$ is the time average of annual or seasonal values for each grid point and $\langle . \rangle_{xy}$ is the area-weighted spatial average.

The climate prediction index (CPI) introduced by Murphy et al. (2004) is widely used in climatology studies for model evaluation and weighted projections (for example Connolley and Bracegirdle, 2007; Franco et al., 2011). It is directly related to the bias and the centred rmse (crmse) by the following relationship:

$$\mathrm{CPI} = \sqrt{\left\langle \left(\mu_t^{\mathrm{m}} - \mu_t^{\mathrm{o}}\right)^2 \right\rangle_{xy} / \left\langle \sigma_t^{\mathrm{o}} \right\rangle_{xy}^2} = \sqrt{b^2 + c^2} / \left\langle \sigma_t^{\mathrm{o}} \right\rangle_{xy}, \tag{2}$$

where $\sigma_t^{\mathrm{o}}$ is the temporal standard deviation of annual or seasonal observation values for each grid point. We therefore define the bias index bi and the crmse index ci as the bias $b$ and crmse $c$ of Eq. (1) scaled by $\langle \sigma_t^{\mathrm{o}} \rangle_{xy}$, so that $\mathrm{CPI}^2 = \mathrm{bi}^2 + \mathrm{ci}^2$.

The CPI index is based on statistical theory for normally-distributed variables, which gives that the probability that a realisation $r$ belongs to a population of mean $\mu$ and a standard deviation $\sigma$ which is proportional to $\exp(-(|r-\mu|/\sigma)^2/2)$. We therefore define the skill score associated with the index ind as $\exp(-\mathrm{ind}^2/2)$, as in Murphy et al. (2004). When considering a combination of several indexes, we compute the combined index as the root mean square of its components' indexes. Indexes vary between 0 and +infinity (close to 0 if model compares very well with ERA-Interim), whereas skill scores vary between 0 and 1 (close to 1 is close to ERA-Interim). It is worth noting that with this definition, CPI is the combination of bi and ci.

## 3 Results

### 3.1 Metric selection

A metric is the association between an index/score and a variable. Our variable selection is based on three criteria: (i) the variable should be a forcing field for RCMs, (ii) the variable should have an impact on RCM-modeled SMB, and (iii) the variable should be constrained with sufficient observation so that reanalyses could confidently be considered an "observation". Consequently, we focus on the variables detailed below.

### 3.1.1 Sea level pressure

Sea level pressure (psl) is a proxy for the large-scale circulation patterns which significantly impact the precipitation patterns simulated by RCMs. The psl spatial anomalies compared to ERA-Interim for the period 1980–2010 are shown in Fig. 1. The circulation patterns are mainly described by the spatial variability of psl, evaluated by the crmse index. The crmse index variability also drives the CPI variability (see Fig. S8 in the Supplement), which is why we will focus on crmse only. We observe that the four seasonal psl crmse indexes are similar (see Fig. S9), suggesting that the most relevant metric for psl is the combination of the four seasons' crmse values, denoted by psl[ann]c (for psl annual crmse index).

### 3.1.2 Air temperature at 850 hPa

The air temperature in the free atmosphere (here at 850 hPa; ta850) has an impact on phase changes in RCMs (refreeze/melt of snowpack, snow/rain fall). It also controls the maximal water vapor content of the atmosphere. Because of its pronounced seasonal cycle, ta850 presents large temporal variability in autumn and spring, such that seasonal means are not reliable for these seasons, though it is more stable in summer and winter. As summer and winter indexes are both relevant and closely related (see Fig. S9), the combined index of these two seasons form a robust metric. However, special attention should be given to summer ta850 (denoted by ta850[sum]), since it has the highest impact on the melt/refreezing amounts and on the hydrometeors' phase changes. Finally, CPI variability is controlled by the bias variability for ta850 (see Fig. S8). In conclusion, the most relevant metrics for our study are the summer/winter ta850 bias index (denoted by ta850[s/w]b) and the summer ta850 bias index (denoted by ta850[sum]b).

### 3.1.3 Precipitable water

Column-integrated atmospheric water vapor, or precipitable water (prw), is a proxy for the humidity content of the atmosphere which impacts the precipitation amount in RCMs. It is affected by the same strong seasonal cycle as temperature since the maximum water vapor content of an air parcel is related to the temperature through the Clausius-Clapeyron relationship. Consequently, as with ta850, seasonal prw is relevant when its value reaches its minima and maxima, i.e. in winter and summer. The CPI is also controlled by the bias for prw, so we chose to focus on the summer/winter prw bias index metric, which we denote by prw[s/w]b.

### 3.1.4 Surface oceanic conditions

Since most RCMs are not coupled with an oceanic model, sea surface temperature (tos) and sea-ice concentration from forcing are used to simulate oceanic conditions in the RCM's integration domain. Instead of sea-ice concentration, we considered the meridional sea-ice extent (msie), defined as sea-ice concentration times cell area summed for each longitude (see Appendix B on normality issues). Sea-ice and open water extents are complementary and show very strong seasonal cycles. Consequently, seasonal analyses for these oceanic variables should refer to winter msie and summer tos. As with ta850 and prw, their CPI variability is controlled by the bias variability. Accordingly, the most relevant metrics for oceanic variables are the winter msie bias index (msie[win]b) and the summer tos bias index (tos[sum]b).

## 3.2 Multi-metric analysis

From the six selected metrics detailed above (psl[ann], ta850[s/w]b, ta850[sum]b, prw[s/w]b, tos[sum]b, and msie[win]b), we computed the total index and its associated multi-decadal variability. Skill scores associated with the total index and its individual components are detailed in Table 1 and shown in Fig. 2. In this figure, models are





TCD
9, 3113–3136, 2015

ranked by their total skill score (black line). However, the pertinent criteria to evaluate model performance is the distance between the total skill score plus multi-decadal variability (external blue line) and ERA-Interim, taking into account ERA-Interim multi-decadal variability (grey crown).

With regard to reanalyses, only MERRA shows similarity to ERA-Interim for the six metrics. NCEP1, NCEP2 and 20CR share a significant positive bias in precipitable water. 20CR presents a misspecification of sea-ice, with ice concentrations never exceeding 55 % far from the coast (Compo et al., 2011), which explains its very low skill score for winter meridional sea-ice extent. For the remaining metrics (tos[sum]b, psl[ann], ta850[s/w] and ta850[sum]), the four reanalyses are not significantly different from ERA-Interim over 1980–2010.

Among CMIP5 models, none show a null bias for all six metrics. The 5 models with the highest skill scores are MIROC-ESM/MIROC-ESM-CHEM (but show incorrect circulation patterns), ACCESS1-3 (but shows a strong warm bias for summer surface ocean temperature), CMCC-CM (but shows a moderate warm bias for summer surface ocean temperature and a wet bias for precipitable water), BCC-CSM1-1-m (but shows a strong wet bias for precipitable water and an incorrect sea-ice spatial distribution) and NorESM1-M (but shows a moderate cold bias for summer air temperature and a wet bias for winter precipitable water). Four other models have only one strong bias compared to ERA-Interim: ACCESS1-0 and EC-EARTH (showing a strong warm bias for summer surface ocean temperature) and CCSM4 and CESM1-BGC (showing strong overestimations of winter meridional sea-ice extent). Detailed maps of spatial anomalies relatively to ERA-Interim similar to Fig. 1 can be found in Fig. S2 to S7.

### 3.3 Climate change

Knutti et al. (2010) showed that model skills in simulating present-day climate conditions relate only weakly to the magnitude of predicted change for surface temperature, except for sea-ice covered regions in winter. We looked for emergent constraints for our region by correlating projected future changes in msie[win],





TCD
9, 3113–3136, 2015

tos[sum], prw[s/w] and ta850[s/w] (2079–2100 mean minus 1980–2010 mean) to our four bias metrics. We found that variable evolutions are significantly correlated to the initial state of winter sea-ice msie[win]b (Fig. 3, 1st column), but are poorly correlated to other metrics (not shown).

We see that changes in prw[sum] and tos[sum] are very strongly correlated with changes in ta850[s/w] ($R2 > 0.8$). Evolution of sea-ice extent is also strongly correlated with changes in ta850[s/w] ($R2 = 0.67$), but is just as well correlated with msie[win]b ($R2 = 0.55$), such that theses two variables together explain 78 % of the variance of the change in sea-ice extent. This suggests that studying the change in air temperature and sea-ice extent is sufficient for understanding changes in the four studied variables.

We introduce mid-latitude (40° S to 40° N) annual surface air temperature change, denoted by Δtas40S40N[ann], as a proxy for the global warming signal. We see that 32 % of the variance of Δta850[s/w] is explained by msie[win]b and almost the same amount of variance (36 %) is explained by Δtas40S40N[ann] (Fig. 3, 1st row), while msie[win]b and Δtas40S40N[ann] are not correlated with each other. On the other hand, changes in sea-ice extent is strongly correlated with msie[win]b, but is not significantly correlated with the global warming signal (Fig. 3, 4th row). This means that (i) the decrease in sea-ice extent is mainly driven by its initial state over current climate and the local feedbacks between air temperature increase and sea-ice decrease and that (ii) this feedback also play a major role in the air temperature increase over the Antarctic region. This section highlights the importance of simulating current climate conditions correctly, as future projected anomalies in climate over Antarctica will be significantly dependent of the conditions of winter sea ice cover over the Historical period.

## 4 Discussion and conclusions

The main goal of this work was to provide a fair overview of the strengths and weaknesses of model outputs from the last multi-model ensemble CMIP5 as a first and essential step toward regional modeling of the Antarctic surface mass balance. This study does not give an absolute ranking of CMIP5 climate models over Antarctica as it is deliberately biased by the choice of forcing fields for regional models. The three main bias factors are the choice of reference fields, the score computation and the variables selection.

We chose ERA-Interim as the reference field because it was shown to be the most reliable contemporary global reanalysis over Antarctica (Bromwich et al., 2011; Bracegirdle and Marshall, 2012) and included four other reanalyses into our study to assess our knowledge of the current state of the Antarctic climate. Our results show that these reanalyses are not significantly different from ERA-Interim for 850 hPa air temperature, surface ocean temperature, sea level pressure and sea-ice concentration, except for 20CR, for which sea-ice was misspecified (Compo et al., 2011). For precipitable water, however, we found that NCEP1, NCEP2 and 20CR reanalyses from NOAA share a significant positive bias when compared to ERA-Interim. This bias was already noted by Nicolas and Bromwich (2011) for NCEP2. The same paper shows that ERA-Interim has a constant bias of $-0.6\,\mathrm{kg\,m^{-2}}$ compared to the SSM/I satellite data for the 60–50° S area. We compared ERA-Interim with the most recent version of Satellite Microwave Radiometer brightness temperatures converted to precipitable water using the RSS Version-7 algorithm over the 1988–2014 period (RemoteSensingSystems, 2013). We see a bias of only $-0.25\,\mathrm{kg\,m^{-2}}$ for the 60–50° S area and of $-0.21\,\mathrm{kg\,m^{-2}}$ for the 60–40° S area, for all seasons. This bias is much lower than those encountered between ERA-Interim and models (see Figs. S4 and S5), leading us to believe that ERA-Interim can be confidently used as a reference for precipitable water in this region.

With regard to score computation, we focused on widely used and easy-to-interpret measures: the bias and the centred rmse scaled by the observed inter-annual variability of each variable. These measures are based on statistical theory for normally-distributed variables, which we verified as applicable to our dataset.

In selecting variables, we sought to focus on a limited number of variables and to avoid redundancy. We considered psl rather than 500 hPa geopotential height because the latter can be strongly impacted by air temperature biases at low atmospheric levels. Another variable that could be of importance for modeling surface mass balance is the meridional moisture flux (mmf), calculated by integrating specific humidity times meridional wind from the surface to the top of the atmosphere. This depends on available precipitable water as well as large-scale circulation, driving moisture advection into the Antarctic domain. However mmf is dominated by time-varying synoptic-scale motions, also called transient eddies (Tsukernik and Lynch, 2013), which are captured at the sub-daily time step. This means that a study of meridional moisture flux requires 6H outputs for all models, which we were not able to obtain. It would be of interest to put the vertical integral of northward and eastward water vapour flux as a standard output in the next CMIP.

In the context of these choices, ACCESS1-3 is the CMIP5 model showing the best performance for modeling surface mass balance. It has a significant warm bias for summer surface ocean temperature, but shows no significant biases for the 5 other selected metrics. As shown by Noël et al. (2014) over Greenland, biases in sea surface temperatures only marginally impact the SMB simulated by RCMs. In addition, ACCESS1-3 variable evolutions are close to the multi-model ensemble mean evolutions. Two other models with high skill scores could also be of particular interest because they cover the range of plausible variable evolutions: MIROC-ESM (or identically MIROC-ESM-CHEM) and NorESM1-M, which projects future high (low) 850 hPa air temperature increase and sea-ice extent decrease, respectively. However the general circulation in MIROC-ESM is incorrect and NorESM1-M is too cold in summer, directly impacting the melt increase projected by RCMs.

With regard to climate change estimates from CMIP5, we see no significant change in sea-level pressure patterns for RCP85 during the 21st century, whereas the other variables evolve significantly from the 1980's to 2100. We observe that 850 hPa air temperature change combined with the 1980–2010 sea-ice extent bias explain more than 80 % of the variance of the change in surface ocean temperature, precipitable water and sea-ice extent, while theses two variables have, respectively, moderate and null correlation with the global warming signal. This demonstrates the importance of a robust evaluation over the current climate, as the future projected climate anomalies over Antarctica could be significantly dependent on a model's ability to properly simulate present-day sea-ice extent. In addition, we believe that a better understanding of climate change over the Antarctic region would be achieved with a better quantification of the feedback between free atmosphere warming and sea-ice extent decrease.

Finally, one mean of reducing the uncertainty of climate change in Antarctica would be to focus on amip-type projections, for which sea surface conditions are computed as anomalies of the observed state as in Krinner et al. (2014). This kind of run reduces biases for present-day simulations (see Fig. S10) and eliminates uncertainties related to the initial state of the sea-ice extent for simulations of the future. However these are not currently available in CMIP5.

## Appendix A: Mean climate and multi-decadal variability

We computed the six selected metrics prw[s/w]b, psl[ann]c, ta850[s/w]b, ta850[sum]b, tos[sum]b, and msie[win]b for the 41 CMIP5 GCMs on 31 years moving average between 1850 and 2100 in respect to ERA-Interim over the period 1980–2010. We observed that all metrics showed no significant trends from 1850 to 1980 whereas they evolved significantly afterwards (see Fig. S1). We estimated the multi-decadal climate variability of each CMIP5 GCM and each metric by computing the range of this metric (maximum minus minimum) during this stable 1850–1980 period. Subsequently, we focused on the period 1980–2010 covered by ERA-Interim and we considered the 1980–2010 metrics values plus/minus the multi-decadal variability estimate computed over 1850–1980. With regards to the reanalyses, 20CR presents spurious trends during the 1971–1980 period and the others do not cover a substantial portion of the stable period. Consequently we approximate their multi-decadal variability by the 90th percentile of CMIP5 multi-decadal variabilities.

## Appendix B: Normality issues

Indexes defined in Sect. 2.2 should be applied on normally-distributed variables to be valid. We checked that seasonal atmospheric variables follow normal distributions against time for all grid points. However, sea-ice concentration have bounded distributions, hence we apply the scores on msie instead.

Furthermore, msie has a lower bound of 0 and tos has a lower bound of the freezing point of sea water ($\sim -1.7\,^{\circ}$C), which may induce grid points with strongly skewed distributions. However our work focuses on seasons of maximal extent of sea-ice (winter) and free ocean (summer), so the impact of grid points with a skewed distribution is negligible.

**The Supplement related to this article is available online at doi:10.5194/tcd-9-3113-2015-supplement.**

*Acknowledgements.* We acknowledge ETH Zurich for facilitating access to the CMIP archive and we particularly thank Urs Beyerle for his precious help. We thank Hubert Gallée for fruitful discussions and helpful advises.

## References

Bracegirdle, T. J. and Marshall, G. J.: The reliability of Antarctic tropospheric pressure and temperature in the latest global reanalyses, J. Climate, 25, 7138–7146, 2012. 3116, 3123

Bracegirdle, T. J., Turner, J., Hosking, J. S., and Phillips, T.: Sources of uncertainty in projections of twenty-first century westerly wind changes over the Amundsen Sea, West Antarctica, in CMIP5 climate models, Clim. Dynam., 43, 2093–2104, 2014. 3115

Bromwich, D. H., Nicolas, J. P., and Monaghan, A. J.: An assessment of precipitation changes over Antarctica and the Southern Ocean since 1989 in contemporary global reanalyses, J. Climate, 24, 4189–4209, doi:10.1175/2011JCLI4074.1, 2011. 3116, 3123

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S., Rutledge, G., Bessemoulin, P., Brönnimann, S., Brunet, M., Crouthamel, R. I., Grant, A. N., Groisman, P. Y., Jones, P. D., Kruk, M. C., Kruger, A. C., Marshall, G. J., Maugeri, M., Mok, H. Y., Nordli, Ø., Ross, T. F., Trigo, R. M., Wang, X. L., Woodruff, S. D., and Worley, S. J.: The twentieth century reanalysis project, Q. J. Roy. Meteor. Soc., 137, 1–28, 2011. 3117, 3121, 3123

Connolley, W. M. and Bracegirdle, T. J.: An Antarctic assessment of IPCC AR4 coupled models, Geophys. Res. Lett., 34, L22505, doi:10.1029/2007GL031648 2007. 3118

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge Sanz, B. M., Morcrette, J. J., Park, B. K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J. N., and Vitart, F.: The ERA-interim reanalysis: configuration and performance of the data assimilation system, Q. J. Roy. Meteor. Soc., 137, 553–597, 2011. 3116

Franco, B., Fettweis, X., Erpicum, M., and Nicolay, S.: Present and future climates of the Greenland ice sheet according to the IPCC AR4 models, Clim. Dynam., 36, 1897–1918, 2011. 3118

Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha, S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W., Higgins, W., Janowiak, J., Mo, K. C., Ropelewski, C., Wang, J., Jenne, R., and Joseph, D.: The NCEP/NCAR 40-year reanalysis project, B. Am. Meteorol. Soc., 77, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2, 1996. 3117

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L.: NCEP–DOE AMIP-II reanalysis (R-2), B. Am. Meteorol. Soc., 83, 1631–1643, 2002. 3117

Knutti, R., Furrer, R., Tebaldi, C., Cermak, J., and Meehl, G. A.: Challenges in combining projections from multiple climate models, J. Climate, 23, 2739–2758, 2010. 3121

Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: generation CMIP5 and how we got there, Geophys. Res. Lett., 40, 1194–1199, 2013. 3116, 3134, 3136

Krinner, G., Largeron, C., Menegoz, M., Agosta, C., and Brutel-Vuilmet, C.: Oceanic forcing of antarctic climate change: a study using a stretched-grid atmospheric general circulation model, J. Climate, 27, 1–47, 2014. 3125

Ligtenberg, S. R. M., van de Berg, W. J., van den Broeke, M. R., Rae, J. G. L., and van Meijgaard, E.: Future surface mass balance of the Antarctic ice sheet and its influence on

sea level change, simulated by a regional atmospheric climate model, Clim. Dynam., 41, 867–884, 2013. 3133
Mahlstein, I., Gent, P. R., and Solomon, S.: Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations, J. Geophys. Res.-Atmos., 118, 5105–5110, 2013. 3115
Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., Emori, S., Kainuma, M., Kram, T., Meehl, G. A., Mitchell, J. F. B., Nakicenovic, N., Riahi, K., Smith, S. J., Stouffer, R. J., Thomson, A. M., Weyant, J. P., and Wilbanks, T. J.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747–756, 2010. 3116
Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: Quantification of modelling uncertainties in a large ensemble of climate change simulations, Nature, 430, 768–772, 2004. 3118
Nicolas, J. P. and Bromwich, D. H.: Precipitation changes in high southern latitudes from global reanalyses: a cautionary tale, Surv. Geophys., 32, 475–494, 2011. 3123
Noël, B., Fettweis, X., van de Berg, W. J., van den Broeke, M. R., and Erpicum, M.: Sensitivity of Greenland Ice Sheet surface mass balance to perturbations in sea surface temperature and sea ice cover: a study with the regional climate model MAR, The Cryosphere, 8, 1871–1883, doi:10.5194/tc-8-1871-2014, 2014. 3124
RemoteSensingSystems: The Monthly Mean Total Precipitable Water Data Set on a 1 degree grid made from Remote Sensing Systems Version-7 Microwave Radiometer Data, updated 04/2015, available at: www.remss.com, last access: 21 April 2015, 2013. 3123
Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Julio Bacmeister, Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., Bloom, S., Chen, J., Collins, D., Conaty, A., da Silva, A., Gu, W., Joiner, J., Koster, R. D., Lucchesi, R., Molod, A., Owens, T., Pawson, S., Pegion, P., Redder, C. R., Reichle, R., Robertson, F. R., Ruddick, A. G., Sienkiewicz, M., and Woollen, J.: MERRA: NASA's modern-era retrospective analysis for research and applications, J. Climate, 24, 3624–3648, 2011. 3117
Shu, Q., Song, Z., and Qiao, F.: Assessment of sea ice simulations in the CMIP5 models, The Cryosphere, 9, 399–409, doi:10.5194/tc-9-399-2015, 2015. 3115
Tsukernik, M. and Lynch, A. H.: Atmospheric meridional moisture flux over the Southern Ocean: a story of the Amundsen Sea, J. Climate, 26, 8055–8064, 2013. 3124





TCD
9, 3113–3136, 2015

Turner, J., Bracegirdle, T. J., Phillips, T., Marshall, G. J., and Hosking, J. S.: An initial assessment of Antarctic sea ice extent in the CMIP5 models, J. Climate, 26, 1473–1484, 2013. 3115





TCD
9, 3113–3136, 2015

TCD
9, 3113–3136, 2015



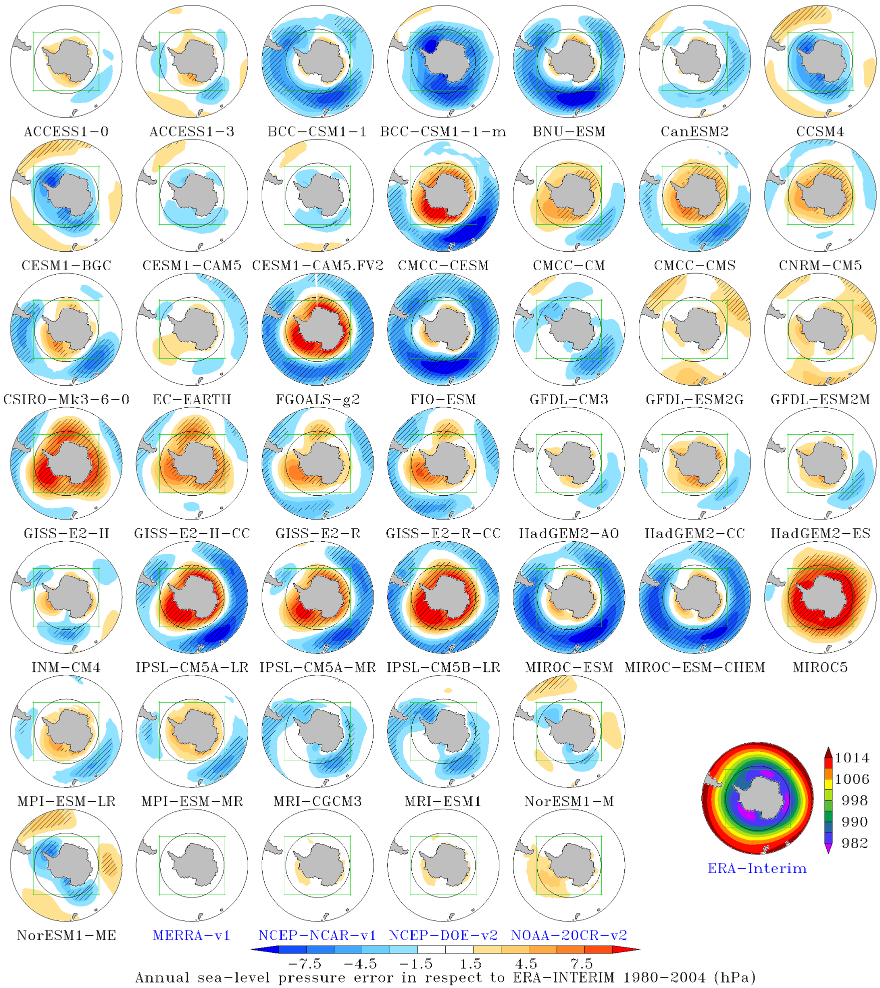


**Table 1.** CMIP5 models and reanalyses details. Bias and crmse indexes are given plus/minus estimate of their multi-decadal variability (reference: ERA-Interim). Indexes names in bold are those selected for the analysis. On the ERA-Interim line, we give the ERA-Interim standard deviation of spatially-averaged annual values, which are the scaling factors for the indexes. When combining several seasons, we give the mean standard deviation plus/minus (maximum – minimum) / 2.

| Name | Short name | Lat. grid spacing | Bias and crmse indexes | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | msie[win] **bias** | crmse | prw[s/w] **bias** | crmse | psl[ann] bias | **crmse** | ta850[s/w] **bias** | crmse | ta850[sum] **bias** | crmse | tos[sum] **bias** | crmse |
| ERA-Interim | ERA-Int | 0.7° | – | | 0.75 ± 0.1 kg m$^{-2}$ | | 3.2 ± 0.5 hPa | | 0.95 ± 0.06 K | | 0.89 K | | 0.55 K | |
| MERRA-v1 | MERRA | 0.5° | −0.0 ± 1.3 | 0.4 ± 0.9 | 0.5 ± 0.7 | 0.2 ± 0.2 | 0.0 ± 0.2 | 0.1 ± 0.4 | 0.3 ± 0.7 | 0.2 ± 0.2 | 0.2 ± 0.6 | 0.2 ± 0.2 | 0.1 ± 0.9 | 0.3 ± 0.4 |
| NCEP-DOE-v2 | NCEP2 | 2.5° | 1.1 ± 1.3 | 1.4 ± 0.9 | 2.3 ± 0.7 | 1.1 ± 0.2 | 0.1 ± 0.2 | 0.3 ± 0.4 | 0.5 ± 0.7 | 0.8 ± 0.2 | 0.5 ± 0.6 | 0.7 ± 0.2 | −0.3 ± 0.9 | 0.4 ± 0.4 |
| NCEP-NCAR-v1 | NCEP1 | 2.5° | 1.3 ± 1.3 | 1.3 ± 0.9 | 1.7 ± 0.7 | 1.0 ± 0.2 | 0.1 ± 0.2 | 0.2 ± 0.4 | 0.4 ± 0.7 | 0.7 ± 0.2 | 0.5 ± 0.6 | 0.5 ± 0.2 | −0.1 ± 0.9 | 0.4 ± 0.4 |
| NOAA-20CR-v2 | 20CR | 2.0° | −3.8 ± 1.3 | 1.9 ± 0.9 | 1.8 ± 0.7 | 0.6 ± 0.2 | 0.2 ± 0.2 | 0.3 ± 0.4 | 0.3 ± 0.7 | 0.9 ± 0.2 | −0.5 ± 0.6 | 0.7 ± 0.2 | 0.3 ± 0.9 | 0.6 ± 0.4 |
| ACCESS1-0 | ACCE-0 | 1.25° | 0.2 ± 0.3 | 1.7 ± 0.4 | 0.7 ± 0.5 | 0.7 ± 0.2 | 0.1 ± 0.2 | 0.6 ± 0.2 | 0.4 ± 0.4 | 1.0 ± 0.1 | 0.3 ± 0.4 | 1.3 ± 0.1 | 3.0 ± 0.6 | 2.0 ± 0.2 |
| ACCESS1-3 | ACCE-3 | 1.25° | 0.6 ± 0.3 | 1.9 ± 0.1 | 0.6 ± 0.3 | 0.9 ± 0.2 | 0.2 ± 0.1 | 0.7 ± 0.2 | 0.3 ± 0.3 | 0.8 ± 0.1 | −0.0 ± 0.3 | 0.8 ± 0.1 | 1.7 ± 0.5 | 2.0 ± 0.1 |
| BCC-CSM1-1 | BCC-1 | 2.8° | 1.8 ± 1.4 | 2.8 ± 1.6 | 1.4 ± 0.4 | 1.2 ± 0.2 | 0.9 ± 0.2 | 0.9 ± 0.2 | 0.4 ± 0.3 | 1.2 ± 0.2 | −0.5 ± 0.4 | 1.0 ± 0.2 | 0.9 ± 0.5 | 1.9 ± 0.4 |
| BCC-CSM1-1-m | BCC-1-m | 1.0° | 0.8 ± 2.4 | 3.9 ± 2.8 | 1.6 ± 0.6 | 1.1 ± 0.1 | 1.0 ± 0.1 | 1.0 ± 0.2 | 0.4 ± 0.7 | 1.0 ± 0.2 | 0.0 ± 0.6 | 1.0 ± 0.3 | 0.8 ± 1.1 | 2.1 ± 0.5 |
| BNU-ESM | BNU-ESM | 2.8° | 6.0 ± 0.8 | 2.8 ± 0.5 | 1.3 ± 0.6 | 1.5 ± 0.2 | 1.0 ± 0.2 | 1.5 ± 0.3 | 1.4 ± 0.6 | 1.8 ± 0.2 | −0.0 ± 0.6 | 1.5 ± 0.2 | 1.9 ± 0.8 | 2.6 ± 0.3 |
| CanESM2 | CanESM | 2.8° | 1.4 ± 0.5 | 1.6 ± 0.3 | 1.2 ± 0.4 | 0.7 ± 0.1 | 0.3 ± 0.1 | 0.6 ± 0.2 | 1.6 ± 0.4 | 1.0 ± 0.1 | −1.5 ± 0.5 | 0.9 ± 0.1 | −0.2 ± 0.5 | 2.2 ± 0.2 |
| CCSM4 | CCSM4 | 1.25° | 2.6 ± 0.4 | 1.3 ± 0.1 | 0.6 ± 0.5 | 1.2 ± 0.2 | 0.4 ± 0.1 | 0.9 ± 0.2 | 0.5 ± 0.5 | 1.1 ± 0.2 | −0.5 ± 0.5 | 1.0 ± 0.1 | 0.5 ± 0.6 | 2.8 ± 0.3 |
| CESM1-BGC | CES-BGC | 1.25° | 2.2 ± 0.6 | 1.4 ± 0.2 | 0.7 ± 0.6 | 1.2 ± 0.1 | 0.3 ± 0.1 | 0.9 ± 0.2 | 0.4 ± 0.6 | 1.0 ± 0.2 | −0.3 ± 0.5 | 0.9 ± 0.2 | 0.7 ± 0.7 | 2.6 ± 0.2 |
| CESM1-CAM5 | CES-C5 | 1.25° | 0.4 ± 0.7 | 1.6 ± 0.2 | 1.0 ± 0.5 | 0.9 ± 0.1 | 0.2 ± 0.2 | 0.6 ± 0.2 | 1.0 ± 0.5 | 0.9 ± 0.1 | −1.2 ± 0.5 | 1.1 ± 0.1 | 1.9 ± 0.8 | 2.2 ± 0.1 |
| CESM1-CAM5-1-FV2 | CES-C5-FV2 | 1.25° | 0.0 ± 0.4 | 1.7 ± 0.1 | 1.7 ± 0.3 | 1.2 ± 0.1 | 0.1 ± 0.1 | 0.5 ± 0.1 | 0.9 ± 0.3 | 1.0 ± 0.1 | −1.1 ± 0.3 | 1.2 ± 0.1 | 3.0 ± 0.4 | 2.3 ± 0.1 |
| CMCC-CESM | CM-CESM | 3.75° | 0.7 ± 1.0 | 2.1 ± 0.5 | 2.0 ± 0.3 | 1.3 ± 0.2 | 0.3 ± 0.2 | 1.7 ± 0.5 | 0.2 ± 0.3 | 1.7 ± 0.2 | −0.1 ± 0.3 | 2.2 ± 0.2 | 1.3 ± 0.4 | 3.2 ± 0.3 |
| CMCC-CM | CM-CM | 0.75° | 0.3 ± 0.9 | 2.3 ± 0.7 | 1.3 ± 0.4 | 0.8 ± 0.1 | 0.2 ± 0.2 | 1.0 ± 0.3 | 0.3 ± 0.4 | 1.3 ± 0.1 | 0.0 ± 0.4 | 1.6 ± 0.1 | 1.2 ± 0.6 | 2.5 ± 0.2 |
| CMCC-CMS | CM-CMS | 1.8° | 0.8 ± 1.0 | 2.0 ± 0.6 | 2.1 ± 0.2 | 1.1 ± 0.2 | 0.1 ± 0.2 | 1.1 ± 0.4 | 0.3 ± 0.3 | 1.2 ± 0.2 | 0.2 ± 0.2 | 1.5 ± 0.2 | 1.5 ± 0.5 | 2.5 ± 0.2 |
| CNRM-CM5 | CNRM | 1.4° | −1.7 ± 1.6 | 3.0 ± 0.9 | 1.4 ± 0.5 | 0.9 ± 0.1 | 0.2 ± 0.2 | 0.9 ± 0.3 | 1.1 ± 0.6 | 1.2 ± 0.2 | 1.1 ± 0.6 | 1.3 ± 0.2 | 4.2 ± 0.8 | 2.6 ± 0.5 |
| CSIRO-Mk3-6-0 | CSIRO | 1.9° | 0.1 ± 0.6 | 2.0 ± 0.2 | 0.4 ± 0.5 | 0.7 ± 0.1 | 0.3 ± 0.2 | 1.0 ± 0.3 | 1.3 ± 0.5 | 1.2 ± 0.1 | −1.6 ± 0.4 | 1.4 ± 0.1 | −1.1 ± 0.6 | 3.2 ± 0.2 |
| EC-EARTH | ECEARTH | 1.125° | −0.6 ± 0.4 | 1.8 ± 0.4 | – | – | 0.2 ± 0.2 | 0.7 ± 0.3 | 0.2 ± 0.3 | 1.2 ± 0.2 | 0.3 ± 0.3 | 1.5 ± 0.1 | 4.1 ± 0.5 | 2.6 ± 0.2 |
| FGOALS-g2 | FGOALS | 2.8° | 2.1 ± 0.7 | 2.4 ± 0.3 | 0.8 ± 0.4 | 0.9 ± 0.1 | 0.3 ± 0.2 | 1.9 ± 0.4 | 1.0 ± 0.4 | 1.5 ± 0.2 | −1.0 ± 0.4 | 1.8 ± 0.2 | −0.4 ± 0.5 | 3.0 ± 0.3 |
| FIO-ESM | FIO-ESM | 2.875° | 1.7 ± 0.2 | 2.9 ± 0.3 | 0.7 ± 0.4 | 1.1 ± 0.1 | 1.3 ± 0.2 | 1.4 ± 0.3 | 1.2 ± 0.3 | 1.4 ± 0.2 | −1.6 ± 0.3 | 1.4 ± 0.2 | −0.7 ± 0.3 | 2.5 ± 0.2 |
| GFDL-CM3 | GFDL-CM3 | 1.8° | −3.6 ± 1.3 | 3.3 ± 0.8 | 1.0 ± 0.4 | 0.8 ± 0.1 | 0.5 ± 0.1 | 0.8 ± 0.2 | 0.1 ± 0.4 | 1.3 ± 0.2 | −0.1 ± 0.4 | 1.6 ± 0.2 | 3.6 ± 0.6 | 2.9 ± 0.2 |

**Table 1.** Continued.

| Name | Short name | Lat. grid spacing | Bias and crmse indexes | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | msie[win] **bias** | crmse | prw[s/w] **bias** | crmse | psl[ann] bias | **crmse** | ta850[s/w] **bias** | crmse | ta850[sum] **bias** | crmse | tos[sum] **bias** | crmse |
| GFDL-ESM2G | GFDL-2G | 2.0° | −1.9 ± 0.6 | 2.9 ± 0.6 | 0.9 ± 0.3 | 0.9 ± 0.2 | 0.5 ± 0.2 | 0.7 ± 0.2 | 0.8 ± 0.3 | 1.4 ± 0.2 | 1.1 ± 0.3 | 1.7 ± 0.2 | 5.1 ± 0.4 | 2.7 ± 0.5 |
| GFDL-ESM2M | GFDL-2M | 2.0° | −3.3 ± 1.0 | 3.7 ± 1.0 | 1.3 ± 0.5 | 0.9 ± 0.2 | 0.4 ± 0.2 | 0.7 ± 0.4 | 1.2 ± 0.5 | 1.5 ± 0.2 | 1.6 ± 0.5 | 1.9 ± 0.2 | 6.2 ± 0.9 | 3.3 ± 0.4 |
| GISS-E2-H | GISS-H | 2.5° | −5.2 ± 1.3 | 4.3 ± 0.6 | 1.5 ± 0.8 | 1.2 ± 0.1 | 0.7 ± 0.2 | 1.3 ± 0.3 | 2.7 ± 0.9 | 1.6 ± 0.1 | 3.1 ± 0.8 | 1.8 ± 0.1 | 8.6 ± 1.3 | 3.0 ± 0.2 |
| GISS-E2-H-CC | GISS-HCC | 2.5° | −1.5 ± 1.3 | 4.1 ± 1.0 | 0.4 ± 0.7 | 1.0 ± 0.1 | 0.5 ± 0.2 | 1.0 ± 0.3 | 1.5 ± 0.8 | 1.3 ± 0.2 | 1.8 ± 0.8 | 1.4 ± 0.1 | 5.8 ± 1.0 | 2.6 ± 0.2 |
| GISS-E2-R | GISS-R | 2.5° | −3.3 ± 0.5 | 3.7 ± 0.3 | 1.1 ± 0.4 | 1.0 ± 0.1 | 0.1 ± 0.2 | 1.0 ± 0.3 | 0.4 ± 0.5 | 1.2 ± 0.2 | 0.4 ± 0.5 | 1.2 ± 0.1 | 3.2 ± 0.6 | 2.4 ± 0.2 |
| GISS-E2-R-CC | GISS-RCC | 2.5° | −3.4 ± 0.3 | 3.5 ± 0.2 | 1.0 ± 0.3 | 1.0 ± 0.1 | 0.1 ± 0.2 | 1.0 ± 0.3 | 0.5 ± 0.4 | 1.2 ± 0.2 | 0.5 ± 0.4 | 1.2 ± 0.2 | 3.4 ± 0.5 | 2.5 ± 0.2 |
| HadGEM2-AO | Had-AO | 1.25° | −2.3 ± 0.5 | 3.7 ± 0.7 | – | – | 0.2 ± 0.2 | 0.7 ± 0.2 | 1.2 ± 0.4 | 1.0 ± 0.1 | 1.0 ± 0.4 | 1.1 ± 0.1 | 3.9 ± 0.7 | 2.2 ± 0.1 |
| HadGEM2-CC | Had-CC | 1.25° | −2.2 ± 0.3 | 4.0 ± 0.3 | 0.8 ± 0.2 | 0.8 ± 0.1 | 0.2 ± 0.2 | 0.8 ± 0.2 | 0.8 ± 0.2 | 1.2 ± 0.1 | 0.5 ± 0.2 | 1.5 ± 0.1 | 3.5 ± 0.2 | 2.8 ± 0.2 |
| HadGEM2-ES | Had-ES | 1.25° | −1.9 ± 0.5 | 3.3 ± 0.6 | 0.8 ± 0.4 | 0.7 ± 0.1 | 0.2 ± 0.2 | 0.7 ± 0.3 | 0.8 ± 0.3 | 1.0 ± 0.1 | 0.4 ± 0.4 | 1.2 ± 0.1 | 3.1 ± 0.5 | 2.3 ± 0.2 |
| INM-CM4 | INM-CM4 | 1.5° | −5.1 ± 0.6 | 3.2 ± 0.2 | 2.6 ± 0.5 | 1.2 ± 0.1 | 0.2 ± 0.1 | 0.8 ± 0.2 | 1.6 ± 0.4 | 1.8 ± 0.1 | 0.9 ± 0.4 | 1.8 ± 0.2 | 3.6 ± 0.5 | 2.9 ± 0.2 |
| IPSL-CM5A-LR | IPSLA-LR | 1.9° | −0.2 ± 1.0 | 1.8 ± 0.4 | 1.1 ± 0.5 | 1.0 ± 0.2 | 0.5 ± 0.3 | 2.2 ± 0.4 | 2.3 ± 0.4 | 1.6 ± 0.2 | −3.1 ± 0.4 | 1.9 ± 0.2 | −1.2 ± 0.4 | 4.3 ± 0.2 |
| IPSL-CM5A-MR | IPSLA-MR | 1.3° | −1.4 ± 0.9 | 2.0 ± 0.6 | 0.8 ± 0.5 | 0.9 ± 0.2 | 0.3 ± 0.2 | 1.6 ± 0.4 | 1.5 ± 0.4 | 1.4 ± 0.2 | −2.1 ± 0.4 | 1.5 ± 0.2 | −0.9 ± 0.5 | 3.6 ± 0.3 |
| IPSL-CM5B-LR | IPSLB-LR | 1.3° | −5.8 ± 1.3 | 3.4 ± 0.4 | 3.5 ± 1.1 | 1.3 ± 0.2 | 0.2 ± 0.2 | 2.0 ± 0.3 | 1.5 ± 0.9 | 1.7 ± 0.1 | 1.2 ± 1.0 | 2.0 ± 0.1 | 6.2 ± 1.4 | 3.7 ± 0.3 |
| MIROC-ESM | MIR-E | 2.8° | 0.7 ± 1.0 | 2.8 ± 0.4 | 0.5 ± 0.6 | 0.9 ± 0.1 | 1.2 ± 0.1 | 1.3 ± 0.3 | 0.4 ± 0.6 | 1.4 ± 0.1 | −0.6 ± 0.6 | 1.7 ± 0.1 | −0.1 ± 0.8 | 2.9 ± 0.2 |
| MIROC-ESM-CHEM | MIR-E-C | 2.8° | 1.2 ± 1.2 | 2.4 ± 0.4 | 0.5 ± 0.7 | 0.8 ± 0.1 | 1.2 ± 0.2 | 1.3 ± 0.4 | 0.5 ± 0.7 | 1.3 ± 0.1 | −0.7 ± 0.6 | 1.6 ± 0.1 | −0.3 ± 0.9 | 2.8 ± 0.2 |
| MIROC5 | MIROC5 | 1.4° | −6.4 ± 0.4 | 4.4 ± 0.2 | 2.4 ± 0.4 | 1.1 ± 0.1 | 1.3 ± 0.2 | 1.4 ± 0.3 | 1.1 ± 0.4 | 1.6 ± 0.1 | 0.4 ± 0.4 | 1.5 ± 0.1 | 4.2 ± 0.5 | 3.3 ± 0.2 |
| MPI-ESM-LR | MPI-LR | 1.9° | −3.1 ± 0.6 | 3.3 ± 0.6 | 1.3 ± 0.4 | 0.8 ± 0.1 | 0.1 ± 0.2 | 0.7 ± 0.3 | 0.5 ± 0.3 | 1.3 ± 0.2 | −0.6 ± 0.3 | 1.5 ± 0.2 | 1.1 ± 0.3 | 3.2 ± 0.2 |
| MPI-ESM-MR | MPI-MR | 1.8° | −2.9 ± 0.4 | 3.1 ± 0.3 | 1.5 ± 0.4 | 0.7 ± 0.2 | 0.1 ± 0.2 | 0.8 ± 0.4 | 0.4 ± 0.3 | 1.2 ± 0.2 | −0.5 ± 0.3 | 1.4 ± 0.2 | 1.3 ± 0.2 | 3.0 ± 0.2 |
| MRI-CGCM3 | MRI-C3 | 1.1° | 1.8 ± 0.5 | 2.8 ± 0.4 | 3.0 ± 0.2 | 1.0 ± 0.1 | 0.5 ± 0.2 | 0.7 ± 0.3 | 1.4 ± 0.2 | 1.1 ± 0.2 | 1.8 ± 0.2 | 1.4 ± 0.2 | 5.8 ± 0.3 | 3.5 ± 0.2 |
| MRI-ESM1 | MRI-ESM1 | 1.1° | 1.6 ± 0.4 | 2.7 ± 0.2 | 3.3 ± 0.4 | 1.1 ± 0.1 | 0.5 ± 0.2 | 0.7 ± 0.3 | 1.7 ± 0.3 | 1.1 ± 0.2 | 2.1 ± 0.3 | 1.3 ± 0.2 | 6.3 ± 0.4 | 3.5 ± 0.2 |
| NorESM1-M | Nor-M | 1.9° | 0.3 ± 0.5 | 1.5 ± 0.4 | 1.3 ± 0.3 | 1.1 ± 0.1 | 0.2 ± 0.1 | 0.7 ± 0.3 | 1.1 ± 0.3 | 0.9 ± 0.2 | −1.5 ± 0.4 | 0.7 ± 0.2 | 0.0 ± 0.4 | 1.8 ± 0.1 |
| NorESM1-ME | Nor-ME | 1.9° | 2.3 ± 0.5 | 1.3 ± 0.2 | 0.9 ± 0.3 | 1.3 ± 0.2 | 0.2 ± 0.1 | 0.8 ± 0.2 | 1.7 ± 0.2 | 1.1 ± 0.2 | −2.3 ± 0.3 | 0.7 ± 0.1 | −1.4 ± 0.2 | 2.0 ± 0.1 |



ACCESS1−0
ACCESS1−3
BCC−CSM1−1
BCC−CSM1−1−m
BNU−ESM
CanESM2
CCSM4
CESM1−BGC
CESM1−CAM5
CESM1−CAM5.FV2
CMCC−CESM
CMCC−CM
CMCC−CMS
CNRM−CM5
CSIRO−Mk3−6−0
EC−EARTH
FGOALS−g2
FIO−ESM
GFDL−CM3
GFDL−ESM2G
GFDL−ESM2M
GISS−E2−H
GISS−E2−H−CC
GISS−E2−R
GISS−E2−R−CC
HadGEM2−AO
HadGEM2−CC
HadGEM2−ES
INM−CM4
IPSL−CM5A−LR
IPSL−CM5A−MR
IPSL−CM5B−LR
MIROC−ESM
MIROC−ESM−CHEM
MIROC5
MPI−ESM−LR
MPI−ESM−MR
MRI−CGCM3
MRI−ESM1
NorESM1−M
NorESM1−ME
MERRA−v1
NCEP−NCAR−v1
NCEP−DOE−v2
NOAA−20CR−v2
ERA−Interim
1014
1006
998
990
982
−7.5 −4.5 −1.5 1.5 4.5 7.5
Annual sea−level pressure error in respect to ERA−INTERIM 1980−2004 (hPa)

**Figure 1.** Mean differences of sea-level pressure between models and ERA-Interim over the period 1980–2010 (in hPa). CMIP5 model names are in black and reanalysis names are in blue. Hashes are for areas where the difference is higher that two time ERA-Interim annual sea-level pressure standard deviation over the same period. External circle is 40° S and intermediate black circle is 60° S. Green rectangle is a typical domain boundary for regional climate models over Antarctica (e.g. Ligtenberg et al., 2013). ERA-Interim sea-level pressure over the period 1980–2010 is displayed in the low-right panel (in hPa).

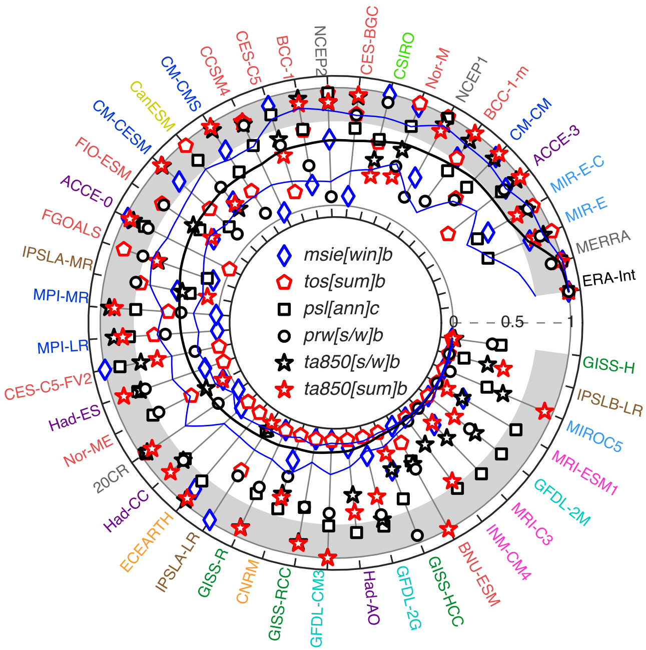



**Figure 2.** Scores for winter meridional sea-ice extent bias (msie[win]b, blue diamonds), summer sea surface temperature bias (tos[sum]b, red pentagons), annual sea-level pressure crmse (psl[ann]c, black squares), summer/winter precipitable water bias (prw[s/w]b, black circles), summer/winter 850 hPa air temperature bias (ta850[s/w]b, black stars), and summer 850 hPa air temperature bias (ta850[sum]b, red stars). The black line is for the total score computed from the combination of components scores. Blue lines are upper and lower bounds for the total score taking into account multi-decadal variabilities of components. The grey crown width is the combination of 90th percentiles of CMIP5 GCMs multi-decal variabilities. Scores range from 0 (worst, internal circle) to 1 (best, external circle). Models with obvious similarities in code or produced by the same institution are marked with the same color (clusters), following Knutti et al. (2013).

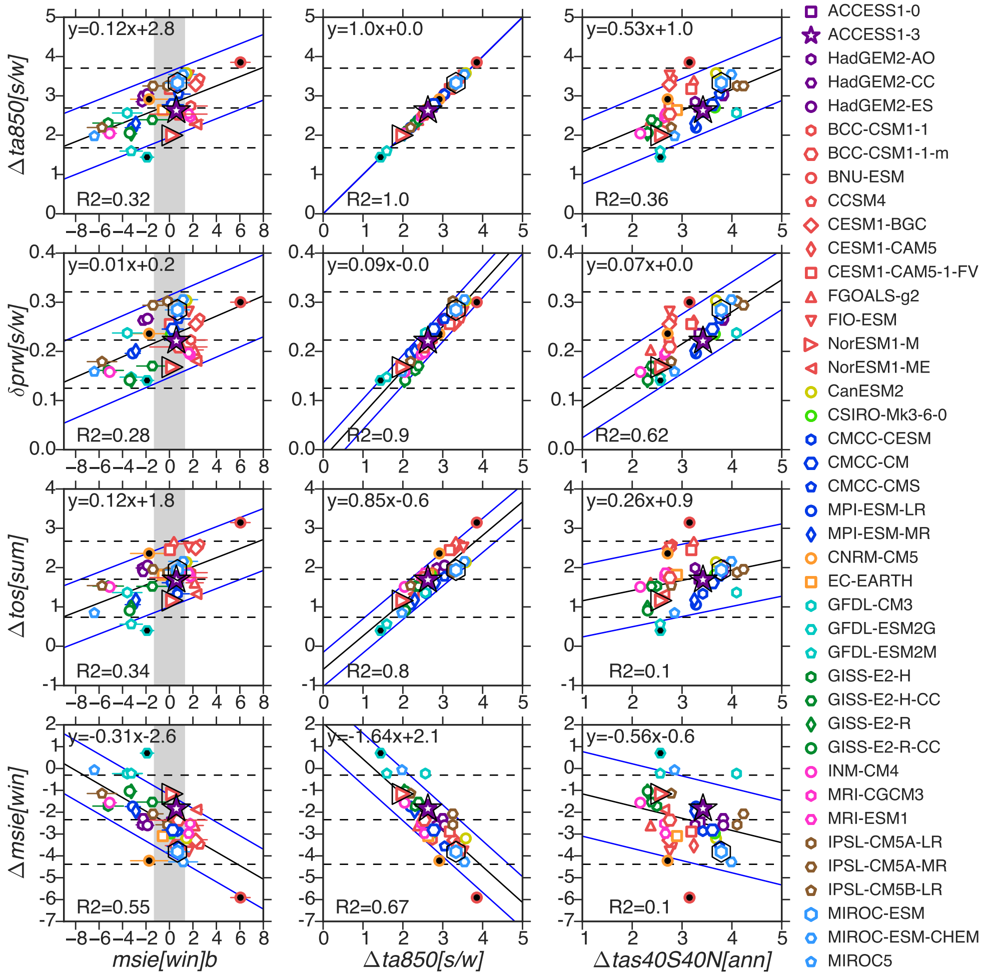

y=0.12x+2.8
R2=0.32
y=1.0x+0.0
R2=1.0
y=0.53x+1.0
R2=0.36
y=0.01x+0.2
R2=0.28
y=0.09x-0.0
R2=0.9
y=0.07x+0.0
R2=0.62
y=0.12x+1.8
R2=0.34
y=0.85x-0.6
R2=0.8
y=0.26x+0.9
R2=0.1
y=-0.31x-2.6
R2=0.55
y=-1.64x+2.1
R2=0.67
y=-0.56x-0.6
R2=0.1
Δ ta850[s/w]
δprw[s/w]
Δ tos[sum]
Δ msie[win]
msie[win]b
Δ ta850[s/w]
Δ tas40S40N[ann]
ACCESS1-0
ACCESS1-3
HadGEM2-AO
HadGEM2-CC
HadGEM2-ES
BCC-CSM1-1
BCC-CSM1-1-m
BNU-ESM
CCSM4
CESM1-BGC
CESM1-CAM5
CESM1-CAM5-1-FV
FGOALS-g2
FIO-ESM
NorESM1-M
NorESM1-ME
CanESM2
CSIRO-Mk3-6-0
CMCC-CESM
CMCC-CM
CMCC-CMS
MPI-ESM-LR
MPI-ESM-MR
CNRM-CM5
EC-EARTH
GFDL-CM3
GFDL-ESM2G
GFDL-ESM2M
GISS-E2-H
GISS-E2-H-CC
GISS-E2-R
GISS-E2-R-CC
INM-CM4
MRI-CGCM3
MRI-ESM1
IPSL-CM5A-LR
IPSL-CM5A-MR
IPSL-CM5B-LR
MIROC-ESM
MIROC-ESM-CHEM
MIROC5

**Figure 3.** *Y* axes: evolution in time (2070–2100 minus 1980–2010) of summer/winter 850 hPa air temperature (Δta850[s/w]), summer/winter precipitable water (*δ*prw[s/w]), summer surface ocean temperature (Δtos[sum]) and winter meridional sea-ice extent scaled by ERA-Interim standard deviation of annual values (Δmsie[win]). The Δ symbol is for absolute differences and the *δ* symbol for absolute differences divided by 1980–2010 mean value. *X* axes: winter msie bias (msie[win]b), Δta850[s/w] and evolution in time of annual surface air temperature between 40° S and 40° N (Δtas40S40N[ann]). Horizontal coloured lines in the first column are two time the multi-decadal variability of msie[win]b. The grey band width is two times the 90th percentile of msie[win]b multi-decadal variabilities. Three of the five highest-scores models are highlighted with black contours: ACCESS1-3 (star), MIROC-ESM (hexagon), and NorESM1-M (triangle). Models with obvious similarities in code or produced by the same institution are marked with the same color (clusters), following Knutti et al. (2013).





TCD
9, 3113–3136, 2015