

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

C. Agosta¹, X. Fettweis¹, and R. Datta²

¹Department of Geography, Climatology Laboratory, Université de Liège, Liège, Belgium

²City College of New York, City University of New York, New York, NY, 10031, USA

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

Abstract. The surface mass balance (SMB) of the Antarctic ice-sheet 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 SMB 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 6 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 is the most pertinent choice for forcing RCMs over Antarctica, followed by ACCESS1-0, CESM1-BGC, CESM1-CAM5, NorESM1-M, CCSM4 and EC-EARTH. Finally, climate change over the Southern Ocean in CMIP5 is less sensitive to the global warming signal than it is to the present-day simulated sea-ice extent and to the feedback between sea-ice decrease and air temperature increase around Antarctica.

1 Introduction

The mass balance of the Antarctic ice-sheet is a major source of uncertainty in estimates of projected sea-level rise. 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. The mass budget of the Antarctic ice-sheet is 10 times lower in magnitude than the individual input/output components. Consequently, when using the input-output method, uncertainty
25 in the total mass budget equals the sum of the uncertainties of input and output estimates, which are of the same order of magnitude as the mass budget itself. This drives efforts to better estimate and reduce uncertainty on each of these two components.

The SMB of the Antarctic ice-sheet is driven by snowfall at the ice-sheet margins, although sublimation, melt, refreezing, and drifting snow can be of importance locally. These components
30 cannot be reliably deduced from reanalyses or global climate models (GCMs) because their horizontal resolution (~ 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
35 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. Accordingly, 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
40 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 (atmospheric fields at lateral boundaries and surface oceanic conditions into the integration domain) and those that may have the greatest impact on RCM-based SMB components (air temperature, air humidity, surface pressure, sea-ice concentration and sea surface
45 temperature).

After describing models, measures and variable selection in Section 2, we perform multi-variable analysis and establish relationships between climate change in GCMs and their representation of current climate in Section 3. We conclude by discussing potential sources of bias in our method and by summarizing our main outcomes.

50 **2 Data and methods**

2.1 CMIP5 climate models and reanalyses

Monthly means fields from 41 CMIP5 models and 6 reanalyses, listed in Table 1, are compared in this work. All data were bi-linearly interpolated onto a common regular longitude-latitude horizontal grid ($1.5^\circ \times 1.5^\circ$) with a spatial domain extending south of 40° S over the ocean. We did not include
55 land and ice-covered areas because (i) RCM lateral boundaries are set over the ocean when possible and (ii) RCMs are never forced by GCM outputs over the land surface, except for the initialization.

Seasonal values are defined by 3-month 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.

60 CMIP5 data was 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 CO₂ equivalent (Moss et al., 2010). In this scenario, stratospheric ozone recovery is represented across the CMIP5
65 models, with recovery over Antarctica to near pre-ozone hole amounts by 2100. We merged Historical and RCP85 to form continuous time series from 1850 to 2100. We focused on the first realization (r1i1p1), but also considered r2i1p1 and r3i1p1 realizations, when available, to check the robustness of our results. 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
70 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, 1979–present, 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
75 climate (1980–2010). However, comparisons with five other reanalyses were also performed in our study: the Japanese 55-year Reanalysis from the Japan Meteorological Agency (JRA-55, 1958–present, Kobayashi et al., 2015), the National Aeronautics and Space Administration Modern Era Retrospective-Analysis for Research and Applications (MERRA, 1979–present, Rienecker et al., 2011); the National Centers for Environmental Prediction (NCEP)/National Center for
80 Atmospheric Research Global Reanalysis 1 (NCEP-NCAR-v1, 1948–present, Kalnay et al., 1996); the NCEP/Department of Energy Atmospheric Model Intercomparison Project 2 reanalysis (NCEP-DOE-v2, 1979–present, Kanamitsu et al., 2002); and the National Oceanic and Atmospheric Administration (NOAA) Twentieth Century Reanalysis v2 (NOAA-20CR-v2, 1870–2012, Compo et al., 2011).

85 We will later define measures 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 measures. 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
90 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 Measures

95 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 based on statistical theory for normally-distributed variables, which gives that the probability that a realisation r belongs to a population of mean μ and a standard deviation σ is proportional to $\exp(-(|r - \mu|/\sigma)^2/2)$. It is defined as follows:

$$100 \text{ CPI}_s = \sqrt{\langle (\mu_s^m - \mu_s^o)^2 \rangle_{xy} / \langle \sigma_s^o \rangle_{xy}^2} = \text{rmse}_s / \langle \sigma_s^o \rangle_{xy}, \quad (1)$$

where the index s denotes the season, s and o exponents are for model outputs and observations respectively, μ_s is the time average of seasonal values for each grid point, σ_s^o is the temporal standard deviation of seasonal observation values for each grid point, $\langle \cdot \rangle_{xy}$ is the area-weighted spatial average, and rmse_s is the spatial root mean square error for the season s .

105 When aggregating several seasons, we compute the CPI as the root mean square of the seasonal indexes:

$$\text{CPI} = \sqrt{\sum_s \text{CPI}_s^2} \quad (2)$$

2.3 Variable selection

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 observations so that reanalyses could confidently be considered an “observation”. Consequently, we focus on the variables detailed below.

2.3.1 Sea level pressure

115 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. We observe that the four seasonal psl CPIs are similar (see Fig. S1), suggesting that the most relevant metric for psl is the combination of the four seasons’ CPI values, denoted by psl[ann].

2.3.2 Air temperature at 850 hPa

120 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 CPIs are both relevant and similar (see

125 Fig. S1), the combined CPI of these two seasons form a robust metric. However, special attention should be given to summer ta850, since it has the highest impact on the melt/refreezing amounts and on the hydrometeors' phase changes. In conclusion, the most relevant metrics for our study are the summer/winter ta850 CPI, denoted by ta850[s/w], and the summer ta850 CPI, denoted by ta850[sum].

130 2.3.3 Precipitable water

Column-integrated atmospheric water vapor, or precipitable water (prw), is a proxy for the humidity content of the atmosphere, which impacts the amount of precipitation 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, 135 seasonal prw is relevant when its value reaches its minima and maxima, i.e. in winter and summer. Consequently we chose to focus on the summer/winter prw CPI, which we denote by prw[s/w].

2.3.4 Surface oceanic conditions

Since most RCMs are not coupled with an oceanic model, sea surface temperature (tos) and sea-ice concentration from the forcing GCM are used to simulate oceanic conditions in the RCM's 140 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 regarding 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 CPI (msie[win]) and summer tos CPI (tos[sum]).

145 3 Results

3.1 multi-variable analysis

The CPI values range from 0 to ~ 7 for msie[win] and tos[sum] and from 0 to ~ 3 for the other variables (Table 1). In order to obtain a global metric which gives an equal weight to each of the variables, we first ranked the models by CPI values for each variable and then we computed the 150 average of ranks. More oriented comparisons can be carried out by assigning different weights to the variables of greatest interest. A variable-by-variable comparison remains the most objective when an only skill score is used to evaluate a model. In Fig. 2(a) we show for each model the ranks of its variables, with models ordered according to the average of ranks. We evaluate the effect of multi-decadal variability of the variables on the ranking by computing for each model and each variable 155 the modified rank when using CPIs plus/minus multi-decadal variabilities while not changing CPIs for other models. Ranks and their associated ranges are detailed in Table 1 and the impact on the average of ranks is displayed in Fig. 2(b) (green lines). In addition, the average of ranks for the

first realization (r1i1p1) is similar to that of the 2nd and 3rd realizations when available (Fig. 2(b), markers), which is a good indicator of the robustness of the method.

160 As expected, the five reanalyses march to the head of the podium, although the ACCESS models perform surprisingly, with ACCESS1-3 overtaking NCEP-DOE-v2 as well as NOAA-20CR-v2 and with ACCESS1-0 overtaking NOAA-20CR-v2. These results are explained by the significant positive bias in precipitable water shared by NCEP-NCAR-v1, NCEP-DOE-v2 and NOAA-20CR-v2 compared to the other reanalyses. In addition, NOAA-20CR-v2 presents a misspecification of
165 sea-ice, with ice concentrations never exceeding 55 % far from the coast (Compo et al., 2011), which explains its low CPI for winter meridional sea-ice extent. With regards to the other variables, the five reanalyses do not differ significantly from ERA-Interim over 1980–2010.

Each of the CMIP5 models shows at least one variable ranked under the median value except ACCESS1-3. The 5 models with the highest average ranks are ACCESS1-3 and ACCESS1-0,
170 although they show a significant warm bias for summer sea surface temperature, CESM1-BGC, although it shows incorrect circulation pattern, CESM1-CAM5 and NorESM1-M, although they show a moderate cold bias for summer air temperature and a wet bias for precipitable water. Two other models have only one strong bias compared to ERA-Interim: CCSM4, showing a significant overestimation of winter meridional sea-ice extent, and EC-EARTH, showing a strong warm bias
175 for summer sea surface temperature (precipitable water was unavailable). Detailed maps of spatial anomalies relative to ERA-Interim similar to Fig. 1 can be found in Fig. S2 to S7.

3.2 Climate change

Knutti et al. (2010) showed that model skills in simulating present-day climate conditions relates only weakly to the magnitude of predicted change for surface temperature, except for sea-ice
180 covered regions in winter. We looked for emergent constraints for our region by correlating projected changes (2079–2100 mean minus 1980–2010 mean) in winter sea-ice extent, summer sea surface temperature, precipitable water and 850hPa air temperature to biases for the 1980–2010 period. We found that variable evolutions are significantly correlated to the bias in winter sea-ice extent ($p < 0.01$, Fig. 3, 1st column), but are poorly correlated to biases of other variables.

185 Changes in precipitable water and in summer sea surface temperature are very strongly correlated with changes in 850hPa air temperature ($R^2 > 0.8$). Changes in winter sea-ice are also strongly correlated with changes in 850hPa air temperature ($R^2 = 0.68$), but are just as well correlated with the winter sea-ice bias ($R^2 = 0.62$), such that these two variables together explain more than 80 % of the variance of the changes in winter sea-ice. This suggests that studying the changes in air
190 temperature and in sea-ice is sufficient for understanding the changes in the four variables studied.

We introduce mid-latitude (40° S to 40° N) annual surface air temperature change as a proxy for the global warming signal. We see that 31 % of the variance of 850hPa air temperature is explained by the winter sea-ice bias and almost the same amount of variance (36 %) is explained by global

warming (Fig. 3, 1st row), despite winter sea-ice bias and global warming signals being uncorrelated
195 with each other. Additionally, changes in sea-ice extent are 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 simulated state under present-day climate and that (ii) both decreasing sea-ice extent and
increasing air temperature are influenced heavily by the local feedback between these two variables.
This section highlights the importance of simulating current climate conditions correctly, as future
200 projected anomalies in climate over Antarctica will be significantly dependent of the conditions of
winter sea ice cover over the present-day 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
205 modeling of the Antarctic ice-sheet surface mass balance. This study does not give an absolute
ranking of CMIP5 climate models over Antarctica as it is deliberately driven by the choice of
forcing fields for regional models. The three main factors impacting on the ranking are the choice of
reference fields, the variables selection and the measure computation.

We chose ERA-Interim as the reference field because it has been shown to be the most reliable
210 contemporary global reanalysis over Antarctica (Bromwich et al., 2011; Bracegirdle and Marshall,
2012) and included five 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, sea surface temperature, sea level pressure and
sea-ice concentration, except for NOAA-20CR-v2, for which sea-ice was misspecified (Compo
215 et al., 2011). For precipitable water, however, we found that NCEP-NCAR-v1, NCEP-DOE-v2 and
NOAA-20CR-v2 reanalyses from NOAA share a significant positive bias when compared to ERA-
Interim. This bias was already noted by Nicolas and Bromwich (2011) for NCEP-DOE-v2. The
same paper shows that ERA-Interim has a constant bias of -0.6 kg m^{-2} compared to the SSM/I
satellite data for the $60\text{--}50^\circ \text{ S}$ area. We compared ERA-Interim with the most recent version of
220 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 kg m^{-2} for the $60\text{--}50^\circ \text{ S}$ area and of -0.21 kg m^{-2} for the $60\text{--}40^\circ \text{ S}$ area, for all
seasons. This bias is much lower than those encountered between ERA-Interim and models (see
Figs. S5 and S6), leading us to believe that ERA-Interim can be confidently used as a reference for
225 precipitable water in this region.

The variable selection is primarily based on our experience of forcing evaluation for regional
climate modelling of the Greenland ice-sheet SMB (Fettweis et al., 2013), with adaptations specific
to the Antarctic ice-sheet, for which precipitation is the major component of SMB and where melt

amounts are expected to increase significantly during the century. We sought to focus on a limited
230 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, while the centered patterns of the two variables are strongly correlated (see Fig. S8). 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
235 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
240 of northward and eastward water vapour flux as a standard output in the next CMIP.

With regard to measure computation, we focused on the widely used climate prediction index, a
measure based on statistical theory for normally-distributed variables which we verified as applicable
to our dataset. In order to give the same weight to the six selected variables, we chose to first rank
CMIP5 models by variable according to their CPI and then use the average of ranks. The use of the
245 first 3 realizations showed the robustness of the ranking, after which we also evaluated the impact of
multi-decadal variability on the ranks.

In the context of these choices, ACCESS1-3 is the CMIP5 model showing the best performance
for modeling surface mass balance with a RCM. It has a significant warm bias for summer sea surface
temperature, but shows no significant biases for the 5 other metrics. As shown by Noël et al. (2014)
250 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 (Fig. 3). Two other models with high skill scores could also be of particular interest
because they cover the range of plausible variable evolutions: CESM1-CAM5 and NorESM1-M,
which projects future high (low) 850 hPa air temperature increase and winter sea-ice decrease,
255 respectively. However both models are too cold in summer, which may impact 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 (see Fig. S9), whereas the other variables evolve
significantly from the 1980's to 2100. We observe that 850 hPa air temperature change combined
260 with the 1980–2010 winter sea-ice bias explain more than 80 % of the variance of the change in
precipitable water, summer sea surface temperature and winter sea-ice extent, while these last two
variables have 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.
265 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 winter sea-ice decrease.

Finally, Krinner et al. (2014) suggested that uncertainties of climate projections over Antarctica could be better quantified by using AMIP-kind projections, for which sea surface conditions are computed as anomalies of the observed state. We believe that if sea surface conditions do not improve in the next CMIP experiment, this method would be valuable, since AMIP experiments show reduce biases compared to historical experiments (see Fig. S10), but a correction should be applied on anomalies to take into account the present-day sea-ice bias of the forcing simulation.

Appendix A: Mean climate and multi-decadal variability

We computed the six selected metrics $prw[s/w]$, $psl[ann]$, $ta850[s/w]$, $ta850[sum]$, $tos[sum]$, and $msie[win]$ 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. S9). 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, NOAA-20CR-v2 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 ($\sim -1.7^\circ\text{C}$), 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/tc-0-1-2015-supplement](https://doi.org/10.5194/tc-0-1-2015-supplement).

Acknowledgements. We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy’s
300 Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. ERA-Interim data is provided by the European Centre for Medium-Range Weather Forecasts, from their Web site at <http://www.ecmwf.int/en/research/climate-reanalysis/era-interim>. JRA-55 data is provided by the CISL Research Data Archive, managed by NCAR’s Data Support Section, from their Web site at <http://rda.ucar.edu>.
305 MERRA-v1 data is provided by the Goddard Earth Sciences Data and Information Services Center, from their Web site at <http://disc.sci.gsfc.nasa.gov/mdisc/data-holdings>. NCEP-NCAR-v1, NCEP-DOE-v2 and NOAA-20CR-v2 data are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at <http://www.esrl.noaa.gov/psd/>. Support for the Twentieth Century Reanalysis Project dataset is provided by the U.S. Department of Energy, Office of Science Innovative and Novel Computational Impact on Theory and
310 Experiment (DOE INCITE) program, and Office of Biological and Environmental Research (BER), and by the National Oceanic and Atmospheric Administration Climate Program Office. 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

- 315 Kobayashi, S., Ota, Y., Harada, Y., Ebata, A., Moriya, M., Onoda, H., Onogi, K., Kamahori, H., Kobayashi, K., Endo, H., Miyaoka, K., and Takahashi, K.: The JRA-55 Reanalysis: General Specifications and Basic Characteristics, *J. Meteorol. Soc. Japan*, 93, 5–48, doi:10.2151/jmsj.2015-001, 2015.
- 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, doi:10.1007/s00382-013-2032-1, 2014.
- 320 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, doi:10.1175/JCLI-D-11-00685.1, 2012.
- 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.
- 325 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.
- 330 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.
- 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, I., 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.
- 335 (Fettweis et al., 2013)
- Fettweis, X., Franco, B., Tedesco, M., van Angelen, J. H., Lenaerts, J. T., van den Broeke, M. R., and Gallée, H.: Estimating Greenland ice sheet surface mass balance contribution to future sea level rise using the regional atmospheric climate model MAR, *The Cryosphere*, 7, 469–489, doi:10.5194/tc-7-469-2013, 2013.
- 340 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.
- 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.
- 345 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.
- 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.

- Knutti, R., Masson, D., and Gettelman, A.: Climate model genealogy: generation CMIP5 and how we got there, 355 *Geophys. Res. Lett.*, 40, 1194–1199, 2013.
- Krinner, G., Langeron, 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.
- Ligtenberg, S. R. M., van de Berg, W. J., van den Broeke, M. R., Rae, J. G. L., and van Meijgaard, E.: Future 360 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.
- 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.
- Moss, R. H., Edmonds, J. A., Hibbard, K. A., Manning, M. R., Rose, S. K., van Vuuren, D. P., Carter, T. R., 365 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.
- Murphy, J. M., Sexton, D. M. H., Barnett, D. N., Jones, G. S., Webb, M. J., Collins, M., and Stainforth, D. A.: 370 Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, 430, 768–772, 2004.
- 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.
- Noël, B., Fettweis, X., van de Berg, W. J., van den Broeke, M. R., and Erpicum, M.: Sensitivity of Greenland 375 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.
- 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.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Julio Bacmeister, Liu, E., Bosilovich, M. G., 380 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.
- Shu, Q., Song, Z., and Qiao, F.: Assessment of sea ice simulations in the CMIP5 models, *The Cryosphere*, 9, 385 399–409, doi:10.5194/tc-9-399-2015, 2015.
- 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.
- 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.

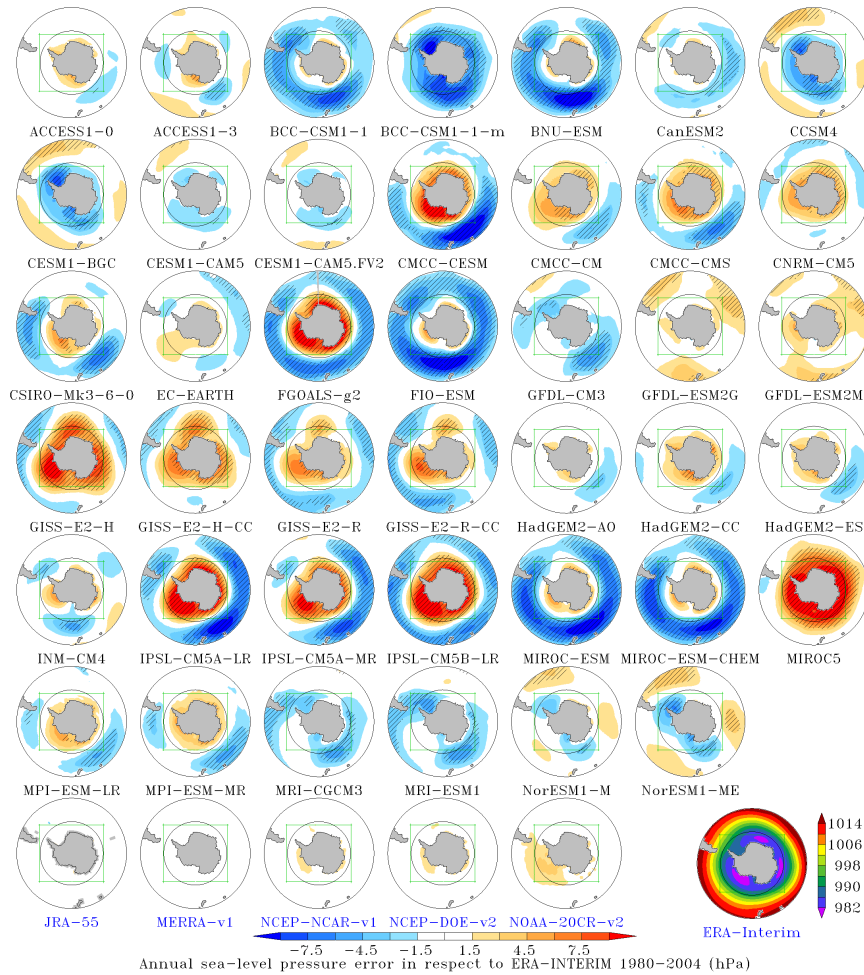


Figure 1. Mean differences of sea-level pressure between models and ERA-Interim over the period 1980–2010 (in hPa). CMIP5 model names are in black and reanalysis names are in blue. Hashes are for areas where the difference is higher than two times ERA-Interim annual sea-level pressure standard deviation 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).

Table 1. Reanalyses (first 6 rows) and CMIP5 models details. Climate prediction indexes (CPI) are given plus/minus estimate of the multi-decadal variability. Ranks are given with in parenthesis the modified rank when using CPI plus/minus multi-decadal variability for the considered model while not changing CPI of other models. 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, and when combining several seasons we give the mean standard deviation plus/minus (maximum – minimum) / 2.

Name	Modelling groups	Lat. grid spacing	msie[win]		prw[s/w]		psl[ann]		CPI and ranks					
			CPI	Rank	CPI	Rank	CPI	Rank	ta850[s/w]	Rank	ta850[sum]	Rank	tos[sum]	Rank
ERA-Interim	ECMWF	0.7°	–	–	0.75 ± 0.1 kg m ⁻²		3.2 ± 0.5 hPa		0.95 ± 0.06 K					
JRA-55	JMA	1.25°	0.5 ± 1.0	4 (2-5)	0.6 ± 0.5	3 (2-10)	0.2 ± 0.4	4 (2-9)	0.7 ± 0.4	3 (3-10)	0.8 ± 0.4	5 (3-11)	0.9 ± 0.9	6 (2-7)
MERRA-v1	NASA	0.5°	0.1 ± 1.0	2 (2-5)	0.5 ± 0.5	2 (2-5)	0.1 ± 0.4	2 (2-6)	0.3 ± 0.4	2 (2-2)	0.3 ± 0.4	2 (2-2)	0.2 ± 0.9	2 (2-6)
NCEP-DOE-v2	NCEP-DOE	2.5°	0.4 ± 1.0	3 (2-5)	2.5 ± 0.5	40 (37-42)	0.3 ± 0.4	5 (2-10)	1.0 ± 0.4	7 (3-23)	0.9 ± 0.4	7 (3-14)	0.4 ± 0.9	4 (2-6)
NCEP-NCAR-v1	NCEP-NCAR	2.5°	0.5 ± 1.0	5 (2-6)	2.0 ± 0.5	36 (28-39)	0.2 ± 0.4	3 (2-7)	0.8 ± 0.4	4 (3-14)	0.7 ± 0.4	3 (3-11)	0.3 ± 0.9	3 (2-6)
NOAA-20CR-v2	NOAA	2.0°	3.6 ± 1.0	29 (23-38)	1.9 ± 0.5	31 (21-37)	0.3 ± 0.4	6 (2-14)	1.0 ± 0.4	6 (3-23)	0.9 ± 0.4	6 (3-13)	0.6 ± 0.9	5 (2-6)
ACCESS1-0	CSIRO-BOM	1.25°	1.9 ± 0.4	11 (6-17)	1.0 ± 0.3	7 (4-16)	0.6 ± 0.2	9 (7-21)	1.1 ± 0.1	9 (6-11)	1.3 ± 0.1	15 (12-15)	3.7 ± 0.4	28 (25-32)
ACCESS1-3	CSIRO-BOM	1.25°	2.1 ± 0.2	15 (12-18)	1.1 ± 0.2	8 (5-15)	0.7 ± 0.2	10 (7-22)	0.9 ± 0.2	5 (3-8)	0.8 ± 0.2	4 (3-7)	2.7 ± 0.3	14 (11-22)
BCC-CSM1-1	BCC	2.8°	3.1 ± 0.5	28 (23-29)	1.9 ± 0.3	33 (28-37)	1.3 ± 0.2	35 (35-37)	1.2 ± 0.3	12 (6-27)	1.1 ± 0.3	11 (6-15)	2.1 ± 0.4	8 (7-13)
BCC-CSM1-1-m	BCC	1.0°	4.0 ± 1.5	31 (20-42)	1.9 ± 0.4	34 (26-37)	1.4 ± 0.1	37 (35-37)	1.1 ± 0.3	8 (4-22)	1.0 ± 0.3	9 (3-15)	2.2 ± 0.5	10 (7-15)
BNU-ESM	GCESS	2.8°	6.7 ± 0.9	46 (45-47)	2.0 ± 0.4	35 (28-39)	1.8 ± 0.3	41 (38-47)	2.3 ± 0.4	44 (38-45)	1.5 ± 0.3	19 (12-31)	3.3 ± 0.4	26 (18-28)
CanESM2	CCCma	2.8°	2.1 ± 0.5	14 (8-22)	1.3 ± 0.4	18 (6-30)	0.7 ± 0.2	15 (8-26)	1.9 ± 0.4	37 (28-44)	1.8 ± 0.4	31 (16-38)	2.2 ± 0.3	9 (8-10)
CCSM4	NSF-DOE-NCAR	1.25°	2.7 ± 0.5	23 (16-28)	1.3 ± 0.1	17 (12-20)	1.0 ± 0.2	28 (14-34)	1.2 ± 0.4	13 (5-29)	1.1 ± 0.4	10 (3-19)	2.9 ± 0.2	19 (16-22)
CESM1-BGC	NSF-DOE-NCAR	1.25°	2.4 ± 0.7	19 (11-27)	1.4 ± 0.2	19 (12-27)	0.9 ± 0.2	26 (14-34)	1.1 ± 0.5	10 (3-27)	1.0 ± 0.5	8 (3-16)	2.7 ± 0.1	15 (14-17)
CESM1-CAM5	NSF-DOE-NCAR	1.25°	1.6 ± 0.3	7 (6-11)	1.4 ± 0.3	20 (9-29)	0.6 ± 0.2	8 (7-15)	1.3 ± 0.4	19 (6-30)	1.6 ± 0.4	26 (12-33)	3.0 ± 0.5	22 (12-26)
CESM1-1-FV2	NSF-DOE-NCAR	1.25°	1.7 ± 0.1	10 (7-10)	2.1 ± 0.2	37 (32-37)	0.6 ± 0.1	7 (7-10)	1.3 ± 0.2	20 (11-27)	1.6 ± 0.2	27 (16-32)	3.9 ± 0.3	31 (28-32)
CMCC-CESM	CMCC	3.75°	2.3 ± 0.7	17 (7-26)	2.4 ± 0.3	39 (38-41)	1.7 ± 0.5	39 (35-47)	1.8 ± 0.2	31 (29-37)	2.2 ± 0.2	38 (35-41)	3.3 ± 0.3	25 (23-27)
CMCC-CM	CMCC	0.75°	2.3 ± 0.6	18 (10-25)	1.5 ± 0.3	23 (13-30)	1.0 ± 0.4	29 (8-35)	1.3 ± 0.2	21 (11-27)	1.6 ± 0.1	25 (18-28)	2.8 ± 0.2	17 (14-22)
CMCC-CMS	CMCC	1.8°	2.0 ± 0.6	13 (6-22)	2.4 ± 0.3	38 (38-41)	1.1 ± 0.4	34 (12-37)	1.2 ± 0.2	14 (8-27)	1.5 ± 0.2	17 (14-29)	3.0 ± 0.3	21 (14-24)
CNRM-CM5	CNRM-CERFACS	1.4°	3.8 ± 1.5	30 (19-41)	1.7 ± 0.4	28 (14-36)	0.9 ± 0.3	25 (8-34)	1.6 ± 0.4	30 (17-40)	1.7 ± 0.4	29 (16-35)	4.7 ± 0.9	38 (29-41)
CSIRO-Mk3-6-0	CSIRO-QCCCE	1.9°	1.6 ± 0.2	9 (6-10)	0.8 ± 0.2	4 (3-7)	1.0 ± 0.3	32 (20-35)	1.8 ± 0.3	32 (27-42)	2.1 ± 0.4	37 (32-43)	2.5 ± 0.1	13 (11-13)
EC-EARTH	EC-EARTH	1.125°	2.0 ± 0.4	12 (7-18)	–	–	0.8 ± 0.3	19 (7-33)	1.2 ± 0.3	11 (6-27)	1.5 ± 0.1	20 (16-28)	4.9 ± 0.4	39 (37-40)
FGOALS-g2	LASG-IAP	2.8°	2.9 ± 0.4	25 (23-28)	1.2 ± 0.3	13 (5-27)	1.8 ± 0.4	42 (36-47)	1.8 ± 0.3	34 (28-42)	2.0 ± 0.3	34 (30-40)	3.0 ± 0.2	20 (17-23)
FIO-ESM	FIO	2.875°	3.1 ± 0.3	27 (24-28)	1.3 ± 0.2	16 (11-25)	1.9 ± 0.2	46 (40-47)	1.9 ± 0.3	35 (28-42)	2.1 ± 0.3	36 (32-42)	2.5 ± 0.3	12 (11-16)
GFDL-CM3	NOAA GFDL	1.8°	5.2 ± 1.0	41 (35-45)	1.2 ± 0.2	14 (8-20)	1.0 ± 0.2	27 (18-34)	1.3 ± 0.2	22 (11-27)	1.6 ± 0.1	22 (16-29)	4.4 ± 0.6	36 (31-39)

Table 1. Continued.

Name	Modelling groups	Lat. grid spacing	msie[win]		prw[s/w]		psl[ann]		CPI and ranks			tos[sum]		
			CPI	Rank	CPI	Rank	CPI	Rank	CPI	Rank	Rank	CPI	Rank	
GFDL-ESM2G	NOAA GFDL	2.0°	4.0 ± 0.9	32 (32-40)	1.2 ± 0.1	15 (12-18)	0.9 ± 0.2	22 (9-34)	1.6 ± 0.2	29 (26-33)	2.0 ± 0.2	35 (33-38)	5.6 ± 0.5	41 (40-41)
GFDL-ESM2M	NOAA GFDL	2.0°	5.4 ± 1.4	42 (32-46)	1.5 ± 0.4	27 (12-33)	0.8 ± 0.4	17 (7-34)	1.9 ± 0.3	36 (28-42)	2.4 ± 0.4	43 (35-45)	7.1 ± 0.9	46 (42-46)
GISS-E2-H	NOAA GFDL	2.5°	6.0 ± 1.3	45 (39-46)	1.9 ± 0.4	32 (21-37)	1.4 ± 0.3	36 (35-39)	3.2 ± 0.7	47 (46-47)	3.6 ± 0.7	46 (46-47)	9.2 ± 1.2	47 (47-47)
GISS-E2-H-CC	NOAA GFDL	2.5°	4.1 ± 0.7	34 (29-39)	1.1 ± 0.4	9 (4-20)	1.0 ± 0.3	33 (13-35)	2.0 ± 0.4	40 (31-44)	2.3 ± 0.5	39 (32-45)	6.5 ± 0.9	42 (41-46)
GISS-E2-R	NOAA GFDL	2.5°	4.2 ± 0.3	36 (31-37)	1.5 ± 0.3	22 (12-30)	1.0 ± 0.3	30 (15-34)	1.2 ± 0.2	15 (8-27)	1.3 ± 0.1	12 (10-15)	3.8 ± 0.5	29 (25-33)
GISS-E2-R-CC	NOAA GFDL	2.5°	4.2 ± 0.1	35 (34-36)	1.4 ± 0.3	21 (12-29)	1.0 ± 0.3	31 (16-35)	1.3 ± 0.2	18 (11-27)	1.3 ± 0.2	14 (11-15)	4.1 ± 0.4	32 (28-36)
HadGEM2-AO	MOHC	1.25°	4.6 ± 0.8	38 (30-42)	–	–	0.7 ± 0.2	11 (7-26)	1.6 ± 0.3	28 (18-35)	1.5 ± 0.2	16 (14-28)	4.4 ± 0.6	34 (29-39)
HadGEM2-CC	MOHC	1.25°	4.7 ± 0.3	39 (37-40)	1.1 ± 0.1	11 (6-13)	0.8 ± 0.2	18 (8-29)	1.4 ± 0.1	27 (19-27)	1.5 ± 0.1	21 (16-28)	4.4 ± 0.3	35 (33-38)
HadGEM2-ES	MOHC	1.25°	4.1 ± 0.7	33 (29-39)	1.1 ± 0.2	10 (5-15)	0.7 ± 0.3	12 (7-30)	1.2 ± 0.2	16 (8-27)	1.3 ± 0.2	13 (10-15)	3.8 ± 0.5	30 (27-33)
INM-CM4	INM	1.5°	5.8 ± 0.6	44 (42-45)	2.8 ± 0.4	42 (40-43)	0.8 ± 0.2	16 (8-29)	2.4 ± 0.2	45 (43-45)	2.0 ± 0.1	33 (33-37)	4.6 ± 0.4	37 (33-39)
IPSL-CM5A-LR	IPSL	1.9°	1.6 ± 0.6	8 (6-15)	1.5 ± 0.4	24 (9-34)	2.0 ± 0.4	47 (39-47)	2.8 ± 0.4	46 (46-47)	3.6 ± 0.4	47 (46-47)	4.3 ± 0.2	33 (32-36)
IPSL-CM5A-MR	IPSL	1.3°	2.5 ± 0.6	22 (12-26)	1.2 ± 0.4	12 (5-27)	1.6 ± 0.4	38 (35-46)	2.0 ± 0.3	41 (31-44)	2.5 ± 0.4	45 (38-45)	3.5 ± 0.4	27 (24-30)
IPSL-CM5B-LR	IPSL	1.3°	5.8 ± 0.7	43 (41-45)	3.8 ± 1.0	45 (42-45)	1.8 ± 0.3	45 (38-47)	2.2 ± 0.4	43 (36-45)	2.3 ± 0.2	41 (37-45)	6.8 ± 1.0	44 (42-46)
MIROC-ESM	MIROC	2.8°	2.5 ± 0.5	21 (13-25)	1.0 ± 0.4	6 (4-20)	1.8 ± 0.2	44 (39-47)	1.4 ± 0.4	26 (8-34)	1.8 ± 0.4	32 (16-38)	2.9 ± 0.3	18 (14-24)
MIROC-ESM-CHEM	MIROC	2.8°	2.3 ± 0.8	16 (6-26)	0.9 ± 0.4	5 (2-19)	1.8 ± 0.3	43 (38-47)	1.4 ± 0.5	23 (5-34)	1.8 ± 0.4	30 (16-37)	2.8 ± 0.4	16 (11-23)
MIROC5	MIROC	1.4°	7.3 ± 0.4	47 (47-47)	2.6 ± 0.3	41 (38-42)	1.7 ± 0.3	40 (38-47)	2.0 ± 0.2	39 (35-42)	1.6 ± 0.1	23 (16-29)	5.2 ± 0.4	40 (39-40)
MPI-ESM-LR	MPI-M	1.9°	4.8 ± 0.6	40 (37-41)	1.5 ± 0.3	25 (16-30)	0.7 ± 0.3	13 (7-29)	1.4 ± 0.2	25 (15-27)	1.6 ± 0.2	24 (16-31)	3.2 ± 0.2	24 (23-26)
MPI-ESM-MR	MPI-M	1.8°	4.5 ± 0.3	37 (35-40)	1.7 ± 0.3	29 (20-34)	0.8 ± 0.4	20 (7-34)	1.3 ± 0.3	17 (8-27)	1.5 ± 0.3	18 (12-31)	3.1 ± 0.1	23 (22-24)
MRI-CGCM3	MRI	1.1°	3.0 ± 0.3	26 (23-28)	3.2 ± 0.2	43 (43-43)	0.9 ± 0.3	24 (9-34)	1.8 ± 0.1	33 (31-37)	2.3 ± 0.2	40 (38-43)	6.7 ± 0.2	43 (42-44)
MRI-ESM1	MRI	1.1°	2.8 ± 0.4	24 (19-28)	3.5 ± 0.4	44 (43-45)	0.9 ± 0.2	23 (10-34)	2.0 ± 0.2	38 (31-42)	2.5 ± 0.2	44 (40-45)	7.1 ± 0.3	45 (44-46)
NorESM1-M	NCC	1.9°	1.5 ± 0.4	6 (6-11)	1.7 ± 0.2	30 (26-34)	0.7 ± 0.3	14 (7-30)	1.4 ± 0.3	24 (11-30)	1.6 ± 0.4	28 (12-33)	1.9 ± 0.1	7 (7-7)
NorESM1-ME	NCC	1.9°	2.4 ± 0.5	20 (12-25)	1.5 ± 0.2	26 (17-30)	0.8 ± 0.2	21 (10-29)	2.0 ± 0.2	42 (34-43)	2.4 ± 0.3	42 (37-45)	2.5 ± 0.1	11 (11-13)

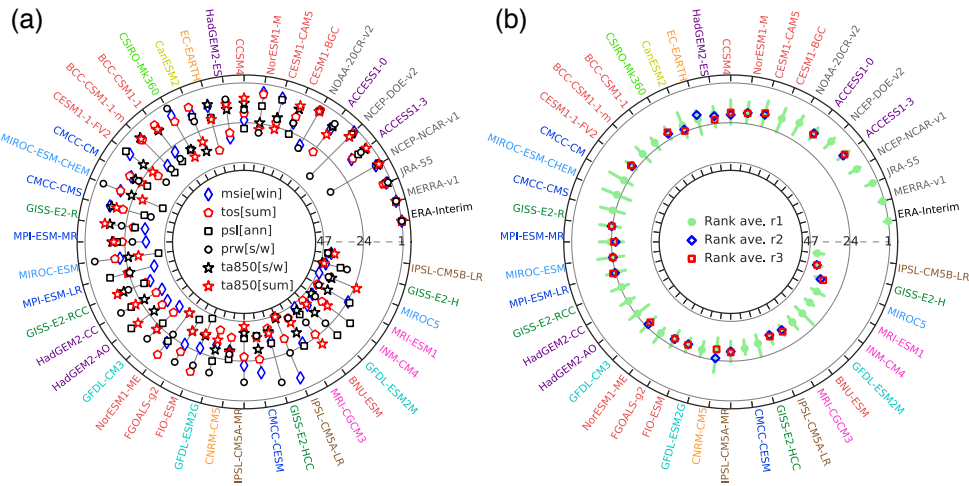


Figure 2. Model ranking according to CPI values: external circle is for rank 1 (ERA-Interim) while internal circle is for rank 47 (largest CPI). Models with obvious similarities in code or produced by the same institution are marked with the same color (clusters), following Knutti et al. (2013). (a) Model rank for winter meridional sea-ice extent (msie[win], blue diamonds), summer sea surface temperature (tos[sum], red pentagons), annual sea-level pressure (psl[ann], black squares), summer/winter precipitable water (prw[s/w], black circles), summer/winter 850 hPa air temperature (ta850[s/w], black stars), and summer 850 hPa air temperature (ta850[sum], red stars). Models are ordered by the average of ranks. (b) Average of ranks for r1i1p1 (green dots), r2i1p1 (blue diamonds), and r3i1p1 (red squares) model realizations. When a field was not available for the 2nd or the 3rd realizations we used the CPI value of the 1st realization for computing ranks. Green lines show variations of the average of ranks when using CPIs plus/minus multi-decadal variabilities for the considered model while not changing CPIs for other models.

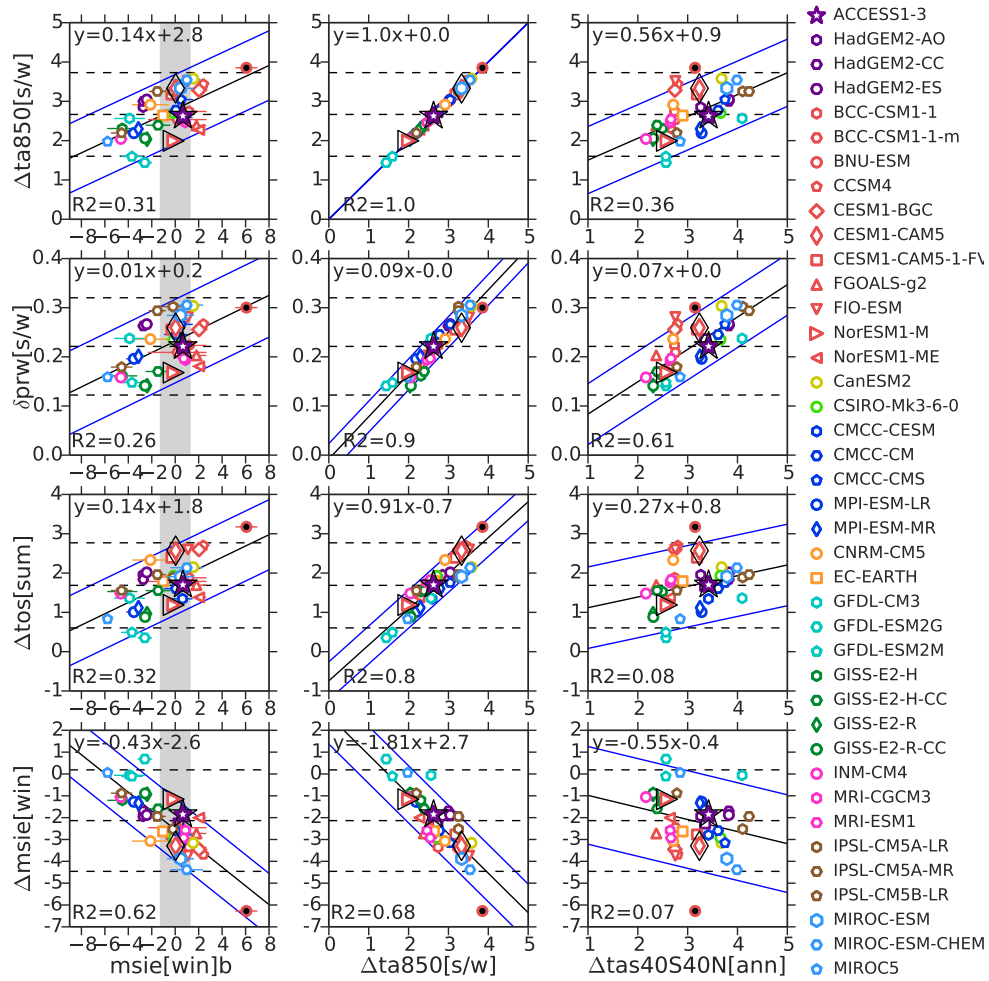


Figure 3. Y axes: evolution in time (2070–2100 minus 1980–2010) of summer/winter 850 hPa air temperature ($\Delta ta_{850}[s/w]$), summer/winter precipitable water ($\delta prw[s/w]$), summer sea surface temperature ($\Delta tos[sum]$) and winter meridional sea-ice extent scaled by ERA-Interim standard deviation of annual values ($\Delta msie[win]$). 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 ($msie[win]b$), $\Delta ta_{850}[s/w]$ and evolution in time of annual surface air temperature between $40^\circ S$ and $40^\circ N$ ($\Delta tas_{40S40N}[ann]$). Horizontal coloured lines in the first column are two time the multi-decadal variability of $msie[win]b$, and the grey band width is two times the 90th percentile of $msie[win]b$ multi-decadal variabilities. Solid black lines are regression lines computed without considering the outlier BNU-ESM (red dot with black face color). Blue lines are vertical shift of the regression line by 1.96 standard deviation of residuals. Three of the five highest-scores models are highlighted with black contours: ACCESS1-3 (star), CESM1-CAM5 (thin diamond), and NorESM1-M (triangle). Models with obvious similarities in code or produced by the same institution are marked with the same color, following Knutti et al. (2013).