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

C. Agosta, X. Fettweis, and R. Datta

Main changes in the revised version:

- Added JRA-55 reanalyses.

- Added realizations r2i1p1 and r3i1p1 when available.

- Used rmse (=CPI) as a systematic measure instead of bias/crmse.

- Used the average of rank per variable for ranking the models

Interactive comment

P. Uotila, petteri.uotila@fmi.fi

Received and published: 12 June 2015

A nice and interesting analysis. I think you should add a proper CMIP5 acknowledgement following <http://cmip-pcmdi.llnl.gov/cmip5/citation.html> to recognise the hard work carried out by the participating climate modelling teams. Thanks.

C. Agosta et al., cecile.agosta@gmail.com

Received and published: 12 June 2015

Dear Dr. Uotila,

We are very sorry to have neglected to acknowledge the CMIP5 modelling teams and the CMIP5 project. Thank you for reminding us. We will add the modelling groups in our Table 1 and add the following sentence in the revised version of the paper: "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 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."

We have also registered the TCD article at CMIP5 (http://cmip.llnl.gov/cmip5/publications/).

Best regards, Cécile Agosta

▶ Correction done.

Responses to Referees

Anonymous Referee #1

Received and published: 28 June 2015

This paper is one of the many papers that try to evaluate the potential of CMIP simulations for downscaling future climate change projections with RCMs. It focuses specifically on the Antarctic, and more so on projections of the Antarctic surface mass balance. The limits of RCMs for the Antarctic are usually somewhere over the Southern Ocean for very good reasons. This paper therefore evaluates and ranks the Southern Ocean climate simulated by CMIP5 models for their (putative) usefulness as driving Antarctic RCMs. Although I am sure that this paper is not not a groundbreaking one, and does not give a definite answer to the question asked (How can we really be sure a given GCM is useful for this purpose?), it is nevertheless a useful contribution to the general discussion of how to chose driving climate models for RCM-based downscaling, over the Antarctic and elsewhere. It is generally well written and clearly structured, although I have the impression that the English could be improved at some instances (but I'm not a native speaker).

▸ Thank you for the comments and suggestions below that have helped to greatly improve our manuscript. Indeed, here we give some initial clues for evaluating CMIP5 models, and this work will be the basis for further investigations of the relationships between SMB components and large-scale forcing fields with the regional climate model MAR.

As a general point, I would have liked to see an evaluation of the Antarctic climate (not necessarily the surface mass balance) simulated at least by the coupled models identified here as the "best" ones. If one can show, a posteriori, that the climate models that correctly simulate the Southern Ocean climate also do a good job over the Antarctic, at least in the mid troposphere and further up (where RCMs arguably do not add much value to driving climate models as the added value is often limited to near-surface fields), then confidence in the pertinence of the selected criteria (and consequently, the proposed "ranking" of the climate models for the specific purpose) could be increased.

▸ Following your advice, we investigated GCM performance in reproducing ERA-Interim circulation pattern in the mid and upper troposphere by considering the geopotential heights at 500 hPa and 250hPa (zg500/zg250), as flowlines approximately follow isolines of zg500 and zg250. We did not mask zg500 and zg250 over the ice-sheet and we found a very strong correlation between the crmse of zg500 and zg250 and the psl crme (R^2=0.94/0.96 respectively). This means that the circulation in the mid and upper troposphere for the entire Antarctic area is well-summarized by the psl over the ocean (Figure S8).

▸ We also looked at the temperature biases at 500hPa and 250hPa and observed that all CMIP5 GCMs have large biases in the upper troposphere, with low correlation with other criteria. RCMs are sometimes forced from the top in the upper troposphere or in the stratosphere with wind fields and temperature. This is the case for the MAR model, and our experience for Greenland (Fettweis et al. 2013) is that the temperature forcing in the upper troposphere have little impact on the climate of the mid and lower troposphere simulated by the RCM. Thus we decided not to consider this field for GCM evaluation.

▸ Fettweis X., Franco B., Tedesco M., van Angelen J.H., Lenaerts J.T., van den Broeke M.R., & Gallée H. (2013) 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.

Without this, it somewhat troubles me that some of the models identified as apt for the Antarctic were recently discarded for driving RCMs in other regions (e.g. McSweeney et al., Clim Dyn 2015; Jury et al., J. Climate).

▶ It is indeed an interesting result, a first attempt at an explanation for this would be that the Antarctic climate in coupled models is dependent on the correct modelling of the ocean and the sea-ice, which are a weakness in many of the CMIP5 models.

Specific comments.

- Abstract, L19-22 : "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." I think this sentence cannot be understood by anyone who has not read the paper. The abstract should be able to stand alone. The word "initial" is misleading: it's the present-day simulated coupled model sea-ice extent, not the one a model is initialized with.

▶ We changed the sentence to the following : « Finally, climate change over the Southern Ocean 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. »

- P. 3115, L. 1-2: "Antarctic mass budget is 10 times lower in magnitude than the individual input/output components." Is the same true for projected changes ? Please justify.

▶ Projected change could be of the same magnitude as the input and output components, but this is a very uncertain issue. Conford et al. 2015 focus on the West Antarctic ice-sheet, where most of the dynamical changes are expected, and they find that a change in SMB could counteract ice dynamics acceleration depending on the warming scenario (see their Figures 5 and 6). When looking for other estimates, the last estimates of changes in ice-dynamics for the most active area in Antarctica, the Amundsen Sea Embayment (ASE), are given by Mouginot et al. 2014: « Between 2003 and 2010, the [calving acceleration] rate is 9.5 Gt/yr2, mainly due to the acceleration of Pine Island Glacier. The record flux year is 2007 with a total ASE ice discharge increasing by 20 Gt/yr in 1 year. After 2010, the total ice flux increases at 2.3 Gt/yr2, similar to the rate prior to 1996 (Figure4). ». Those rates are to compare for example to a change in Antarctic SMB estimated to ~3 Gt/yr2 for A1B (2080-2099 - 1980-1999) in Agosta et al 2013. We did not include this discussion in the text because it would require a lengthy paragraph, since there are a lack of studies on this topic, and it is not the core of our study.

▶ Cornford S.L., Martin D.F., Payne A.J., Ng E.G., Brocq A.M.L., Gladstone R.M., Edwards T.L., Shannon S.R., Agosta C., van den Broeke M.R., Hellmer H.H., Krinner G., Ligtenberg S.R.M., Timmermann R., & Vaughan D.G. (2015) Century-scale simulations of the response of the West Antarctic Ice Sheet to a warming climate. The Cryosphere, 9, 1579–1600.

▶ Mouginot J., Rignot E., & Scheuchl B. (2014) Sustained increase in ice discharge from the Amundsen Sea Embayment, West Antarctica, from 1973 to 2013. Geophysical Research Letters, 41, 1576–1584.

- Maybe you could modify "uncertainties of input" to "uncertainties of change of input" in the following sentence: "Consequently, when using the input-output method, uncertainty in mass change equals the sum of the uncertainties of input and output estimates."

▶ In fact we wanted to say: "Consequently, when using the input-output method, uncertainty in the total mass budget equals the sum of the uncertainties of input and output estimates » (corrected in the text).

- P.3116, L7. "while GCMs results": "GCM results" is better English I think

▶ Corrected

- P.3116, L8. "GCMs results might be biased there because surface schemes are not properly adapted." Why should they be better over the ocean?

▶ Indeed GCMs have no reason to be better over the ocean, so we changed the sentence for the following: « We did not include 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. »

- P. 3316, L19. "we considered the first realization only (r1i1p1)" Did you check whether r2 would change the results?

▶ We initially chose to focus on the r1i1p1 realization because 6-hourly outputs needed for RCM forcing are only available for this realization. However, following your advise, we checked whether r2i1p1 and r3i1p1 could change our results, and it appeared that they seem quite robust (see revised Fig. 2(b)). Averaged over 30yrs, the identified biaises are independent of the realization.

- P3117, L.20. define crmse, not rmse. You use crmse afterwards without introducing the acronym

▶ Corrected (we removed crmse from the study).

- P. 3118, l. 14. Indexes -> indices

▶ Corrected (we now consistently refer to climate prediction indexes instead of indices).

- Shouldn't section 3.1 be part of the "Methods" section ? At least the justification of the chosen variables seems to belong to the Methods in my sense.

▶ You are right. We have moved section 3.1 into section 2.3, in the « method » section.

- P. 3121: "The 5 models with the highest skill scores are MIROC-ESM/MIROC-ESM-CHEM (but show incorrect circulation patterns). . . ". Is that really correct English?

▶ We corrected the English.

- The order of figures in supplementary information is confusing. Text first mentions S8 and S9, then S2 to S7. S1 is only mentioned in the annex, p.3125...

▶ Yes, we re-ordered figures in supplementary information so that they appear in the order of mention in the text.

- p.3122, line 21-25: "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." How much does this statement depend on the red circle (is this BNU?) that seems to be somewhat of an outlier? Would the relationship between sea-ice change and present sea-ice extent still be significant without this one model?

▸ Indeed there is a missing information in the legend, dots with black filling were not considered for computing the regression as they were outliers.

▸ You will find slightly modified values for msie and tos in the revised version of the manuscript because we noticed that there were aberrant value around the ice-sheet because the land mask was too narrow for some models. Consequently, to compute statistics, we extended our land mask of two pixels around Antarctica so that there was no more aberrant values for any model.

- Basically, the negative correlation shows large sea-ice changes for models that have large initial sea ice extent. Is that really surprising? This relationship is necessarily stabilized by the fact that for a given temperature change, a high-sea-ice-biased climate model will have a large delta SIE because the area of the marginal sea-ice region (say, the outermost 500 km that disappear because of the warming) scales almost linearly with its colatitude (roughly we are talking about circles around the pole). In other words, I wonder whether there would be a (significant) relationship if the sea-ice change were measured not in terms of sea-ice extent, but in terms of northward retreat of the sea-ice edge?

▸ Following your suggestion we did the same analysis with the meridional sea-ice position (sea-ice concentration average per longitude, without area-weight), in addition to msie, and we found the same relationship. Consequently it seems that it is the northward retreat of sea-ice that is responsible for the sea-ice decrease and not the concentric shape of the sea-ice around the pole.

- Concerning the circulation criteria you chose: Not clear to me whether the criteria you have chosen are only postulated here to be the ones that influence SMB modeling (with some good arguments) or whether there are any independent proofs to this? Section 3.1 gives good arguments but are there any references, previous model simulations or anything else, that really show that these criteria are necessary and sufficient?

▸ We added a sentence in the discussion to address this issue: «The variable selection is primarily based on our experience of forcing evaluation for regional climate modelling of the Greenland ice-sheet SMB \citep{Fettweis:2013gx}, 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. »

▸ There is, to our knowledge, no article discussing SMB components biases computed with a RCM in respect to the skills of GCM forcing fields for Antarctica. We wish to do this study later with MAR, on the basis of GCM biases highlighted here.

- Are there more ensemble members of the identified "good" models? If yes, please check whether the results are robust.

▸ Following your suggestion we included all available r2 and r3 realizations and we showed that the results are robust (revised Fig. 2(b))

- p.3124, line 26: "We observe that 850hPa 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,. . ." Is that really surprising? Almost any climate variable scales with temperature under climate change...

▸ Indeed this is not surprising, but this sentence aims to highlight that change in climate variables over the Antarctic region is poorly related to change in global temperature.

J. Lenaerts (Referee #2)

j.lenaerts@uu.nl

This manuscript evaluates the ability of the CMIP5 models to represent Antarctic climate, to ultimately present a ranking of the best models to use for RCM simulations. This work will be of interest for RCM users and ice sheet modellers, and fits well within the scope of The Cryosphere. The paper is well written, methods and results well explained, figures are of good quality, and content is original. I recommend publication in TC after the authors could respond to one general and several technical comments outlined below.

▶ Thank you for your constructive and helpful comments and suggestions that helped to greatly improve our manuscript.

General comment

P 3124, L 12-16: While the authors use sea surface temperature bias (in summer) as one of their metrics, they suggest here that it is not important after all, since it does not affect SMB simulations (at least not for Greenland). Moreover, sea-ice extent (in winter) seems to be much more important, also in the perspective of climate change. Based on that, models with a strong 'tos' bias but better 'msie' (e.g. CESM1-CAM5, ACCESS1.0) could ultimately produce much more realistic results than other models, although this is not accounted for when weighing the metrics equally. The authors should discuss the contradiction in the manuscript, and/or considering removing 'tos' as a metric, and/or apply uneven weighing.

▶ We first wanted to weight the tos metric by a factor 0.5, but it raised the problem of the weight choice. We did not want to remove tos from the analyses, because even if we expect it to have minor impact on RCM's results (according to Noël et al. (2014)), it is still a direct forcing of the RCM and this hypothesis has not yet been tested for Antarctica. However we noticed that given the strong biases in tos for all CMIP5 models (CPIs values are typically 2 times larger for tos that for atmospheric variables), too much weight was assigned to the tos variable in the total score as computed in the initial version of the manuscript. This assessment was the origin of the modified ranking of the models proposed in the revised manuscript, which is now based on the average of ranks. The main benefit of this method is that it gives an equal weight to each variable for the global ranking. In addition, we showed that this method is robust and we could give estimates of the ranking uncertainties related to the multi-decadal variability.

Technical comments

P 3114, L 24: Rewrite: ' Mass change of the Antarctic ice sheet (AIS)

▶ Corrected: « The mass balance of the Antarctic ice-sheet is a major source of uncertainty in estimates of projected sea-level rise. »

P 3119, L23: observationS

▶ Corrected

P 3122: this page should be checked, it seems to be forgotten by the authors, or added

later.

▶ Indeed we re-wrote large part of section 3.2 (Climate change)

There are many typing errors on this page:

L 6-7: put the '2' after 'R' in superscript

▶ Corrected

L 8: these two variables

▶ Corrected

L16: extent are strongly

▶ Corrected

L20: plays a major

▶ Sentence modified

P 3123, L9-10: This is remarkable result: many CMIP5 models are actually more similar than ERA-Interim than NCEP or NCEP-2! This could be highlighted more, since this proves once again how unreliable NCEP is on the Southern Hemisphere.

▶ Yes it is true, we highlighted it in the second paragraph of section 3.1 and in the second paragraph of the discussion (section 4).

P 3124, L 30: these two variables

▶ Sentence modified

P 3125: These types of simulations reduce.

▶ Sentence modified

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

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

Abstract. The ~~Antarctic~~ 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 ~~surface~~

5   ~~mass balance~~ 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

10   fields which may have the greatest impact on SMB components simulated by RCMs. Our inter-comparison includes ~~5~~ 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 ~~seems to be~~ is the most pertinent

15   choice for ~~regional climate modeling~~ forcing RCMs over Antarctica, followed by ~~CMCC-CM,~~ ~~MIROC-ESM/MIROC-ESM-CHEM and~~ ACCESS1-0, CESM1-BGC, CESM1-CAM5, NorESM1-M, CCSM4 and EC-EARTH. Finally, climate change over the Southern Ocean ~~is much more~~ ~~dependent on the initial state of winter~~ in CMIP5 is less sensitive to the global warming signal than it is to the present-day simulated sea-ice extent and ~~on the local feedback between~~ to the feedback

20   between sea-ice decrease and air temperature increase ~~and winter sea-ice extent decrease than on the~~ ~~global warming signal~~ around Antarctica.

1 Introduction

~~Mass change in Antarctica is a major component of~~ The mass balance of the Antarctic ice-sheet is a major source of uncertainty in estimates of projected sea-level ~~change~~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. ~~Antarctic mass budget~~ 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 in ~~mass change~~ 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 ~~change itself. That is why efforts are made~~ budget itself. This drives efforts to better estimate and reduce uncertainty on each of these two components.

The ~~Antarctic SMB~~ 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 cannot be reliably deduced from reanalyses or global climate models (GCMs) because their horizontal resolution ($\sim 100\,\mathrm{km}$) 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~~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 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~~ (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 ~~: temperature,~~ (air temperature, air humidity, surface pressure~~and oceanic conditions~~, sea-ice concentration and sea surface temperature).

After describing models~~and skill scores, we explain the selection of metrics, perform multi-metric~~, 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.

2 Data and methods

2.1 CMIP5 climate models and reanalyses

Monthly means fields from 41 CMIP5 models and ~~5~~ 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 (~~0.5° × 1.5°~~ $1.5° \times 1.5°$) with a spatial domain extending south of 40° S over the ocean. We did not include land and ice-covered areas ~~, because~~ because (i) RCM lateral boundaries are ~~usually set far from these areas while GCMs results might be biased there because surface schemes are not properly adapted~~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~~s~~ 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~~ 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 $1370\,\mathrm{ppm}$ 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~~ 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 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 climate (1980–2010). However, comparisons with ~~four other reanalyses are~~ 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 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).

We will later define ~~metrics~~ 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 ~~metrics~~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 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~~Measures

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

$$b = \langle \mu_t^{\mathrm{m}} - \mu_t^{\mathrm{o}} \rangle_{xy},$$

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

~~where m and o exponents are for model outputs and observations respectively, μ_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:~~

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

~~where σ_t^{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. (??) scaled by $\langle \sigma_t^{\mathrm{o}} \rangle_{xy}$, so that $\mathrm{CPI}^2 = \mathrm{bi}^2 + \mathrm{ci}^2$.~~

~~The CPI index is~~ 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 ~~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~~It is defined as follows:

$$\mathrm{CPI_s} = \sqrt{\left\langle \left(\mu_{\mathrm{s}}^{\mathrm{m}} - \mu_{\mathrm{s}}^{\mathrm{o}} \right)^2 \right\rangle_{xy} / \langle \sigma_{\mathrm{s}}^{\mathrm{o}} \rangle_{xy}^2} = \mathrm{rmse_s} / \left\langle \sigma_{\mathrm{s}}^{\mathrm{o}} \right\rangle_{xy}, \tag{1}$$

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, $\langle . \rangle_{xy}$ is the area-weighted spatial average, and rmse_s is the spatial root mean square error for the season s.

When aggregating several seasons, we compute the ~~combined index~~ CPI 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~~ the seasonal indexes:

$$CPI = \sqrt{\sum_s \mathrm{CPI}_s^2} \tag{2}$$

~~3 Results~~

2.3 ~~Metric~~ Variable 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~~ observations so that reanalyses could confidently be considered an "observation". Consequently, we focus on the variables detailed below.

2.3.1 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~~ CPIs are similar (see Fig. ~~S9~~S1), suggesting that the most relevant metric for psl is the combination of the four seasons' ~~crmse~~ CPI values, denoted by psl[ann]~~c (for psl annual crmse index)~~.

2.3.2 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~~ CPIs are both relevant and

155 ~~closely related~~ similar (see Fig. ~~S9~~S1), the combined ~~index~~ CPI 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 (~~CPI, denoted by

160 ta850[s/w]~~b)~~, and the summer ta850 ~~bias index (denoted by~~ CPI, denoted by ta850[sum]~~b)~~.

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 ~~precipitation amount~~ amount of precipitation in RCMs. It is affected by the same strong seasonal cycle as temperature since the maximum water vapor

165 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~~ Consequently we chose to focus on the summer/winter prw ~~bias index metric~~CPI, which we denote by prw[s/w]~~b~~.

2.3.4 Surface oceanic conditions

170 Since most RCMs are not coupled with an oceanic model, sea surface temperature (tos) and sea-ice concentration from ~~forcing~~ the forcing GCM 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~~ regarding normality issues). Sea-ice and open water extents are complementary and show very

175 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 (~~CPI (msie[win]~~b~~) and ~~the summer tos bias index (~~summer tos CPI (tos[sum]~~b~~).

3 Results

180 ### 3.1 ~~Multi-metric~~ multi-variable analysis

~~From the six selected metrics detailed above (psl[ann]c, 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.~~ 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

185 the variables, we first ranked the models by CPI values for each variable and then we computed the 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

6

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 the modified rank when using CPIs plus/minus multi-decadal variabilities while not changing CPIs for other models. Ranks and their associated ranges ~~Skill scores associated with the total index and its individual components~~ are detailed in Table 1 and ~~shown~~ the impact on the average of ranks is displayed in Fig. ~~2. In this figure, models are ranked by their total skill score (black line)~~ 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. ~~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~~ 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 sea-ice, with ice concentrations never exceeding 55 % far from the coast (Compo et al., 2011), which explains its ~~very low skill score~~ low CPI for winter meridional sea-ice extent. ~~For the remaining metrics (tos[sum]b, psl[ann]c, ta850[s/w]b and ta850[sum]b), the four reanalyses are not significantly different~~ With regards to the other variables, the five reanalyses do not differ significantly from ERA-Interim over 1980–2010.

~~Among~~ Each of the CMIP5 models ~~, none show a null bias for all six metrics~~ shows at least one variable ranked under the median value except ACCESS1-3. The 5 models with the highest ~~skill scores are MIROC-ESM/MIROC-ESM-CHEM (but show incorrect circulation patterns),~~ average ranks are ACCESS1-3 ~~(but shows a strong warm bias for summer surface ocean temperature)~~ and ACCESS1-0, although they show a significant warm bias for summer sea surface temperature, ~~, CMCC-CM (but shows a moderate surface ocean temperature and a wet bias for precipitable water),~~ CESM1-BGC, although it shows incorrect circulation pattern, CESM1-CAM5 ~~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~~ , although they show a moderate cold bias for summer air temperature and a wet bias for ~~winter~~ precipitable water. ~~Four~~ Two other models have only one strong bias compared to ERA-Interim: ~~ACCESS1-0~~ CCSM4, showing a significant overestimation of winter meridional sea-ice extent, and EC-EARTH~~(,~~ showing a strong warm bias for summer sea surface temperature (precipitable water was unavailable). ~~surface ocean temperature) and CCSM4~~

7

~~and CESM1-BGC (showing strong overestimations of winter meridional sea-ice extent.~~ Detailed maps of spatial anomalies ~~relatively~~ <u>relative</u> 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 ~~relate~~ <u>relates</u> 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], tos[sum], prw[s/w] and ta850[s/w]~~ <u>changes</u> (2079–2100 mean minus 1980–2010 mean) ~~to our four bias metrics~~<u>in winter sea-ice extent, summer sea surface temperature, precipitable water and 850hPa air temperature to biases for the 1980–2010 period</u>. We found that variable evolutions are significantly correlated to the ~~initial state of~~ <u>bias in</u> winter sea-ice ~~msie[win]b~~ ~~(~~<u>extent</u> ($p < 0.01$, Fig. 3, 1st column), but are poorly correlated to ~~other metrics (not shown)~~<u>biases of other variables</u>.

~~We see that changes in prw[sum/win] and tos[sum]~~ <u>Changes in precipitable water and in summer sea surface temperature</u> are very strongly correlated with changes in ~~ta850[s/w] (R2~~<u>850</u>hPa <u>air temperature ($R^2 > 0.8$)</u>. ~~Evolution of~~ <u>Changes in winter</u> sea-ice ~~extent is~~ <u>are</u> also strongly correlated with changes in ~~ta850[s/w] (R2~~<u>850</u>hPa <u>air temperature ($R^2 = 0.67$</u>0.68), but ~~is~~ <u>are</u> just as well correlated with ~~msie[win]b (R2~~<u>the winter sea-ice bias ($R^2 = 0.55$</u>0.62), such that ~~theses~~ <u>these</u> two variables together explain ~~78~~<u>more than 80</u> % of the variance of the ~~change in~~ <u>changes in winter</u> sea-ice~~extent~~. This suggests that studying the ~~change~~ <u>changes</u> in air temperature and <u>in</u> sea-ice ~~extent~~ is sufficient for understanding <u>the</u> changes in the four ~~studied variables~~ <u>variables studied</u>.

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~~31 % of the variance of ~~Δta850[s/w]~~ <u>850</u>hPa <u>air temperature</u> is explained by ~~msie[win]b~~ <u>the winter sea-ice bias</u> and almost the same amount of variance (36 %) is explained by ~~Δtas40S40N[ann]~~ <u>global warming</u> (Fig. 3, 1st row), ~~while msie[win]b and Δtas40S40N[ann] are not correlated~~ <u>despite winter sea-ice bias and global warming signals being uncorrelated</u> with each other. ~~On the other hand~~<u>Additionally</u>, changes in sea-ice extent ~~is strongly correlated with msie[win]b, but is~~ <u>are</u> 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~~ <u>simulated state under present-day climate and</u> that (ii) ~~this feedback also play a major role in the air temperature increase over the Antarctic region~~<u>both decreasing sea-ice extent and increasing air temperature are influenced heavily by the local feedback between these two variables</u>. 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~~ <u>present-day</u> 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 ice-sheet surface mass balance. This study does not give an absolute ranking of CMIP5 climate models over Antarctica as it is deliberately ~~biased~~ driven by the choice of forcing fields for regional models. The three main ~~bias factors~~ factors impacting on the ranking are the choice of reference fields, the ~~score computation and the variables selection~~ variables selection and the measure computation.

We chose ERA-Interim as the reference field because it ~~was~~ has been shown to be the most reliable contemporary global reanalysis over Antarctica (Bromwich et al., 2011; Bracegirdle and Marshall, 2012) and included ~~four~~ 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, ~~surface ocean~~ sea surface temperature, sea level pressure and sea-ice concentration, except for ~~20CR~~NOAA-20CR-v2, for which sea-ice was misspecified (Compo et al., 2011). For precipitable water, however, we found that ~~NCEP1, NCEP2 and 20CR~~ 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 ~~NCEP2~~NCEP-DOE-v2. 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 and S6), 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~~ 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 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 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.

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 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 ~~surface ocean~~ sea surface 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 ~~.~~ (Fig. 3). 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)~~ CESM1-CAM5 and NorESM1-M, which projects future high (low) 850 hPa air temperature increase and winter sea-ice ~~extent~~ decrease, respectively. However ~~the general circulation in MIROC-ESM is incorrect and NorESM1-M is~~ both models are too cold in summer, ~~directly impacting~~ 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 with the 1980–2010 winter sea-ice ~~extent~~ bias explain more than 80 % of the variance of the change in ~~surface ocean temperature, precipitable water~~ and precipitable water, summer sea surface temperature and winter sea-ice extent, while ~~theses~~ these last 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 winter sea-ice ~~extent~~ decrease.

335   Finally, ~~one mean of reducing the uncertainty of climate change in Antarctica would be to focus on amip-type~~ Krinner et al. (2014) <u>suggested that uncertainties of climate projections over Antarctica could be better quantified by using AMIP-kind</u> projections, for which sea surface conditions are computed as anomalies of the observed state~~as in . This kind of run reduces biases for present-day simulations~~. <u>We believe that if sea surface conditions do not improve in the next CMIP experiment,</u>

340 <u>this method would be valuable, since AMIP experiments show reduce biases compared to historical experiments</u> (see Fig. S10)~~and eliminates uncertainties related to the initial state of the~~ <u>, but a correction should be applied on anomalies to take into account the present-day</u> sea-ice ~~extent for simulations of the future. However these are not currently available in CMIP5.~~ <u>bias of the forcing simulation.</u>

Appendix A: Mean climate and multi-decadal variability

345

We computed the six selected metrics prw[s/w]~~b~~, psl[ann]~~e~~, 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~~S9).

350 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~~ NOAA-20CR-v2 presents spurious trends during the

355 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

360 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 ($\sim -1.7\,°C$), which may induce grid points with strongly skewed distributions. However

365 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.

References

Kobayashi, S., Ota, Y., Harada, Y., Ebita, 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.

390   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.

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.

395   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.

Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Matsui, N., Allan, R. J., Yin, X., Gleason, B. E., Vose, R. S.,

400   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.

Connolley, W. M. and Bracegirdle, T. J.: An Antarctic assessment of IPCC AR4 coupled models, Geophys. Res.

405   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, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge Sanz, B. M., Morcrette, J. J., Park, B. K.,

410   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. (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

415   atmospheric climate model MAR, The Cryosphere, 7, 469–489, doi:10.5194/tc-7-469-2013, 2013.

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.,

420   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.

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

425   multiple climate models, J. Climate, 23, 2739–2758, 2010.

13

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

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.

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.

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., 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.: 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 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., 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, 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.
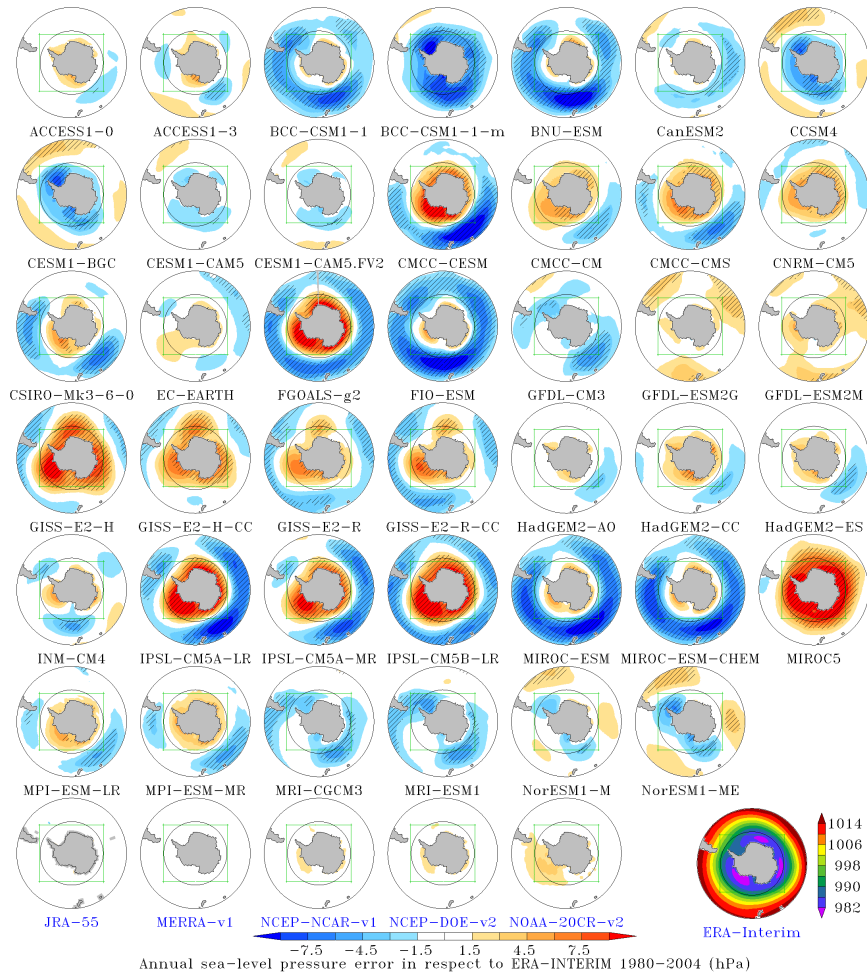
Figure 1. Mean differences of sea-level pressure between models and ERA-Interim over the period 1980–2010 (in hPa). CMIP5 model names are in black and reanalysis names are in blue. Hashes are for areas where the difference is higher that two time ERA-Interim annual sea-level pressure standard deviation 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~~ 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 ~~bias (msie[win]b~~msie[win], blue diamonds), summer sea s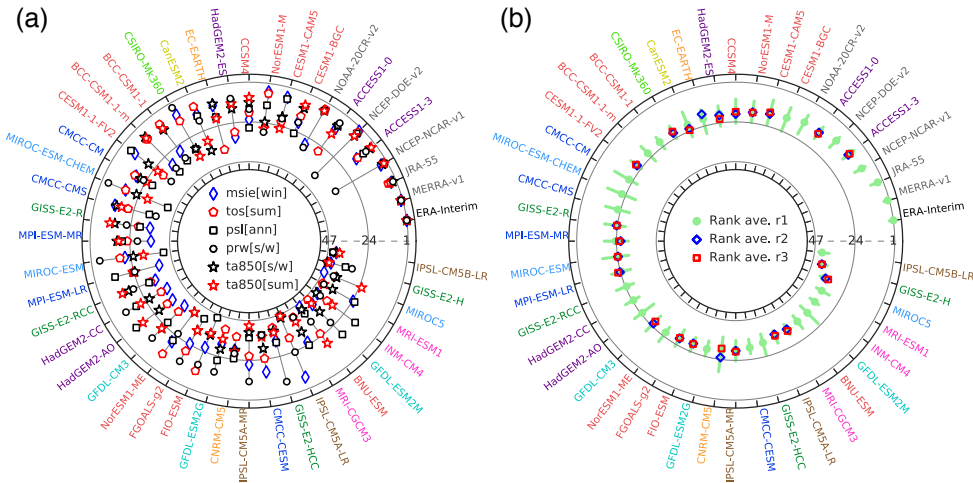urface temperature ~~bias (tos[sum]b~~tos[sum], red pentagons), annual sea-level pressure ~~ermse (psl[ann]e~~psl[ann], black squares), summer/winter precipitable water ~~bias~~ (prw[s/~~w]b~~w], black circles), summer/winter 850 hPa air temperature ~~bias~~ (ta850[s/~~w]b~~w], black stars), and summer 850 hPa air temperature ~~bias (ta850[sum]b~~ta850[sum], red stars). Models are ordered by 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 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. ~~The black line is for the total score computed from the combination of components scores. Blue lines are upper and lower bounds for 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.~~

16
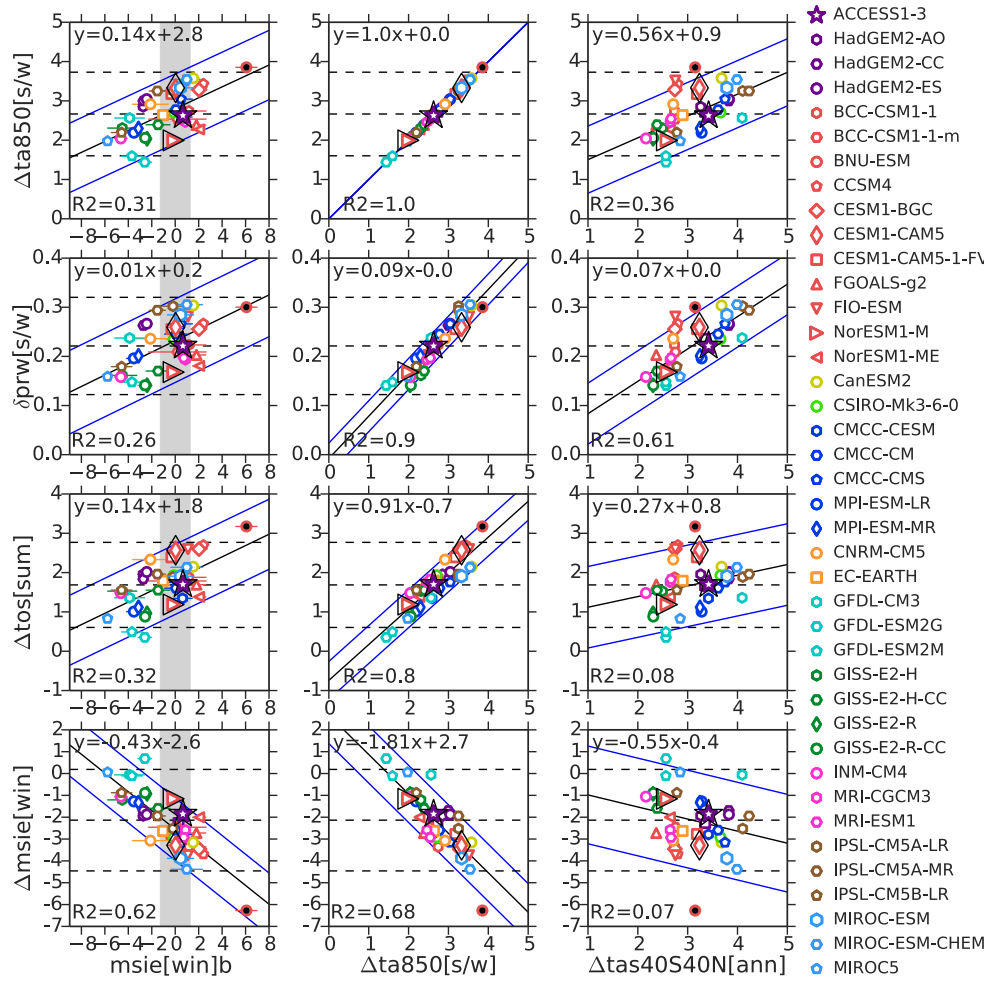
Figure 3. *Y* axes: evolution in time (2070–2100 minus 1980–2010) of summer/winter 850 hPa air temperature (Δta850[s/w]), summer/winter precipitable water (δprw[s/w]), summer ~~sea~~ surface ~~ocean~~ temperature (Δtos[sum]) and winter meridional sea-ice extent scaled by ERA-Interim standard deviation 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 ~~, 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), ~~MIROC-ESM~~ CESM1-CAM5 (~~hexagon~~thin diamond), 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).

17

Table 1. Reanalyses (first 6 rows) and CMIP5 models ~~and reanalyses~~ details. ~~Bias and crmse~~ Climate prediction indexes (CPI) are given plus/minus estimate of ~~their~~ the multi-decadal variability~~(reference: ERA-Interim)~~. ~~Indexes names in bold~~ Ranks are ~~those selected~~ given with in parenthesis the modified rank when using CPI plus/minus multi-decadal variability for the ~~analysis~~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. ~~When~~, and when combining several seasons ~~,~~we give the mean standard deviation plus/minus (maximum − minimum) / 2.

Name	Modelling groups	Lat. grid spacing	msie[win] CPI	Rank	prw[s/w] CPI	Rank	psl[ann] CPI	Rank	ta850[s/w] CPI	Rank	ta850[sum] CPI	Rank	tos[sum] CPI	Rank
ERA-Interim	ECMWF	0.7°	–		$0.75\pm0.1\,\mathrm{kg\,m^{-2}}$		$3.2\pm0.5\,\mathrm{hPa}$		$0.95\pm0.06\,\mathrm{K}$		$0.89\,\mathrm{K}$		$0.56\,\mathrm{K}$	
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±0.3	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±0.2	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] CPI	Rank	prw[s/w] CPI	Rank	psl[ann] CPI	Rank	ta850[s/w] CPI	Rank	ta850[sum] CPI	Rank	tos[sum] CPI	Rank
									CPI and ranks					
GFDL-ESM2G	NOAA GFDL	2.0°	4.0 ± 0.9	32 (28-40)	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 ± 0.2	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 ± 0.2	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 ± 0.5	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 ± 0.2	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 ± 0.4	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)