Assessing bias-corrections of oceanic surface conditions for atmospheric models

Julien Beaumet[1], Gerhard Krinner[1], Michel Déqué[2], Rein Haarsma[3], and Laurent Li[4]

[1]Univ. Grenoble Alpes, CNRS, IRD, Institut des Géosciences de l'Environnement, F-38000, Grenoble, France
[2]Météo-France, Centre National de Recherche Météorologiques, Toulouse, France
[3]Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands
[4]Laboratoire de Météorologie Dynamique, Université Pierre-Marie Curie, CNRS, Paris, France

Correspondence to: Julien Beaumet (Julien.Beaumet@univ-grenoble-alpes.fr)

Abstract. Future sea-surface temperature and sea-ice concentration from coupled ocean-atmosphere general circulation models such as those from the CMIP5 experiment are often used as boundary ~~forcing~~ forcings for the downscaling of future climate ~~experiment~~experiments. Yet, these models show some considerable biases when compared to the observations over present climate. In this paper, existing methods such as an absolute anomaly method and a quantile-quantile method for sea surface temperature (SST) as well as a look-up table and a relative anomaly method for sea-ice concentration (SIC) are presented. For SIC, we also propose a new analog method. Each method is objectively evaluated with a perfect model test using CMIP5 model ~~experiment~~ experiments and some real-case applications using observations. ~~With~~ We find that with respect to other previously existing methods~~for SIC~~, the analog method is a substantial improvement for the bias correction of future SIC. Consistency between the constructed SST and SIC fields is an important constraint to consider, as is consistency between the prescribed sea-ice ~~concentrations~~concentration and thickness; we show that the latter can be ensured by using a simple parameterization of sea-ice thickness as a function of instantaneous and annual minimum SIC.

1 Introduction-Context

Coupled climate models are the most reliable tools that we have today for large-scale climate projections, such as in the Coupled Model Intercomparison Project ~~,~~ Phase 5 (CMIP5 ~~, project (Taylor et al., 2012)), in which these projections were based on Representative Concentration Pathways (RCPs; Moss et al. (2010))~~(Taylor et al., 2012)). Regional-scale information is obtained by using these global simulations as a basis for downscaling exercises. Dynamical downscaling, as opposed to empirical-statistical downscaling (e.g., Hewitson et al., 2014), is carried out either with Regional Climate Models (RCM) (e.g., Giorgi and Gutowski, 2016) or with high-resolution ~~global atmospheric general~~ atmospheric global circulation models (Haarsma et al., 2016). In both cases, information about the projected changes of sea-surface conditions, such as Sea Surface Temperatures (SST), Sea-Ice Concentration (SIC) and Sea-Ice ~~Thickness~~ thickness (SIT), is required as a lower boundary condition for the higher-resolution models. However, SST and SIC conditions modelled by coupled Atmosphere-Ocean ~~General Circulation Model (AOGCMs pr~~ Global Circulation Models (AOGCMs or CGCMs) show important biases for the present climate ~~(Flato et al., 2013) . It has, for example,~~ (Flato et al., 2013; Li and Xie, 2014; Richter et al., 2014; Levine et al., 2013; Zhang and Zha For example, it has been highlighted that most of the CMIP5 models had difficulties in reliably modelling the seasonal cycle and the trend of sea-ice extent in the Antarctic over the historical period (Turner et al., 2013). Therefore, the validity and reliability of such coupled simulations is questionable for future climate projections (e.g. end of the ~~21st century)~~ 21st century), and so is their use as boundary conditions when performing dynamical downscaling of future climate projections.

Prescribing correct SST is crucial for atmospheric modelling because SST determines heat and moisture exchanges with the atmosphere (Ashfaq et al., 2011; Hernández-Díaz et al., 2017). ~~In high latitude~~ The absence of the Pacific cold tongue bias and the reduction of the double ITCZ problem in AMIP experiments with respect to the CMIP5 model experiments (Li and Xie, 2014) shows the importance of forcing atmospheric model by SST close to the observations. For instance, improvements in the modelling of tropical cyclone activity in the Gulf of Mexico (Holland et al., 2010) and of summer precipitation in Mongolia (Sato et al., 2007) were obtained by bias-correcting SST and other AOGCM outputs before using them as forcing for RCMs. At high latitudes, SIC (Krinner et al., 2008; Screen and Simmonds, 2010; Noël et al., 2014) and, in some cases, SIT (Gerdes, 2006; Krinner et al., 2010) are two additional ~~required and~~ crucial boundary conditions for atmospheric ~~modelling of recent and future climate change~~ models. Krinner et al. (2014) demonstrated that for the Antarctic climate as simulated by an atmospheric model, prescribed SST and sea-ice changes have greater influence than prescribed greenhouse gas concentration changes. ~~Integrated~~ Large-scale average winter sea-ice extent and summer SST have been identified among the key boundary forcings for regional modelling of the Antarctic surface mass balance (Agosta et al., 2013), which is the only potentially significant negative contributor to the global eustatic sea level change ~~in~~ over the course of the ~~21st~~ 21st century (Agosta et al., 2013; Church et al., 2013; Lenaerts et al., 2016). We note that while there is a considerable body of scientific literature on the effect of varying SST and SIC on simulated climate, very few studies focused on the role of varying SIT in atmosphere-only simulations (Gerdes, 2006; Krinner et al., 2010; Semmler et al., 2016) , although air-sea fluxes in the presence of sea ice are strongly influenced by the thickness of the sea ice and the overlying snow cover. Gerdes (2006) and Krinner et al. (2010) have shown that the atmospheric response to changes in Arctic SIT can induce atmospheric signals that are of similar magnitude as those due to changes in sea ice cover. In most atmosphere-only General Circulation Models (AGCMs), SIT will therefore also need to be prescribed along with SST and SIC. When SST and SIC from a coupled climate model are directly used, SIT from that same run should of course be used; however, in case SST and SIC from the coupled model run are bias-corrected, as we strongly suggest here, we argue that SIT should be prescribed in a physically consistent manner in the atmosphere-only simulation.

In this study, we describe, evaluate and discuss different existing and new methods for the construction of bias-corrected future SST~~and SIC~~ , SIC and SIT. These methods generally take into account observed oceanic boundary conditions as well as the climate change signal coming from CMIP5 AOGCM scenarios to build more reliable SST and SIC conditions for future climate, which should reduce the uncertainties when used to force future climate projections. The different methods have been evaluated using a perfect model ~~test~~ approach, and by carrying out real-case applications on observations. Applied changes in mean and variances have been investigated as well as the coherence of SIC and SST after applying bias correction methods. The analysis of the results focuses on methods for sea-ice, as bias correction of SIC is ~~a more complicated~~ more complicated an issue to deal with. For SIT, we propose a diagnostic using SIC following Krinner et al. (1997) . Because there were no reliable observational data sets available until recently (Lindsay and Schweiger, 2015; Kurtz and Markus, 2012, e.g.) , we evaluate here directly diagnosed SIT against new observations. In the following, we present the bias-correction methods, the data and the evaluation methods in section~~2.~~ 2.1. The results of the evaluation are shown in section~~3 and are~~ 3. Because SST and SIC are bias-corrected separately, section 3.3 presents a few considerations about SST and SIC consistency after performing bias corrections. The results are then discussed together with general considerations on bias correction of oceanic surface conditions in section~~4. Finally, our findings are summed up and we~~ 4. Finally, we sum up our findings and draw conclusions in section~~5.~~ 5.

2 Data and methods

2.1 Data

Application and validation of the methods for bias correction have been achieved using ~~observationnal~~ observational SST and SIC data from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) that are generally used as boundary conditions for Atmospheric Model Intercomparion Project (AMIP) experiments (Taylor et al., 2000), called “PCMDI obs.” or “observations” in this paper. The AOGCM’s historical and ~~future simulated~~ projected sea-surface conditions come from CMIP5 simulations (Taylor et al., 2012). Only the first ensemble members of the historical, ~~rcp4.5 and rcp8.5~~ and of the Representative Concentration Pathways (RCPs; Moss et al. (2010)) 4.5 and 8.5 simulations have been considered. Most methods have been tested using CNRM-CM5, IPSL-CM5A-LR and HadGEM-ES coupled GCM. Data from NorESM1-M~~and the~~ , MIROC-ESM, EC-EARTH, CCSM4 models have also been used as analog candidates in the analog method for sea-ice. Prior to any application of the bias correction methods, AOGCMs data have been ~~bilinearly~~ bi-linearly regridded onto a common regular 1°x1° grid. For the evaluation of the diagnosed SIT, we used the Lindsay and Schweiger (2015) data for the Arctic. For the Antarctic, in spite of recent observations with autonomous underwater vehicles by Williams et al. (2015) which tend to suggest occurrence of thicker Antarctic sea-ice than previously acknowledged, we will use the Kurtz and Markus (2012) data because of their large spatial coverage.

2.2 Sea Surface Temperature methods

The bias correction of simulated SST is a ~~fairly~~ relatively easy and a straightforward issue to deal with. ~~Nevertheless, different~~ Different methods have been developed ~~. In this section, we describe an anomaly-based method and a quantile-quantile method. Results from their application are presented in section 3.~~

2.2.1 ~~Anomaly method~~

~~This frequently used method (e.g., Krinner et al., 2008) simply consists~~ and presented in the literature. Here we re-evaluate two different frequently used methods. The first is an absolute anomaly method (Krinner et al., 2008, e.g.,) , which consists of simply adding the SST ~~anomaly coming from the difference between a coupled AOGCM projection and the corresponding historical simulation to the present-day observations. In practice, for each grid point, the difference between the SST~~ difference for a given month ~~in the future from a climate change simulation and the climatological mean SST in the corresponding historical simulation from the same coupled AOGCM is added to the observed climatological mean SST (e.g . PCMDI, 1971-2000):~~

$$SST_{Fut,est} = \overline{SST_{obs}} + \left(SST_{Fut,AOGCM} - \overline{SST_{Hist,AOGCM}}\right)$$

~~In (A1),~~ ~~$SST_{Fut,est}$~~ ~~is the estimated future SST for a given month,~~ ~~$\overline{SST_{obs}}$~~ ~~the observed climatological monthly mean,~~ ~~$SST_{Fut,AOGCM}$~~ ~~the model future SST for a given month in the future AOGCM scenario and~~ ~~$\overline{SST_{Hist,AOGCM}}$~~ ~~the model climatological monthly~~ from an AOGCM scenario to the climatological mean in the ~~AOGCM historical simulation for the same reference period as for the observed climatology. As a result, the reconstructed SST time series has the chronology of the AOGCM projected scenario.~~

2.2.1 ~~Quantile-quantile method~~

~~This method has been proposed and described in Ashfaq et al. (2011) It consists in adding, for each grid point and each calendar month's quantile in the observations, the corresponding quantile change in the GCM data set, i.e. the difference between the maximum SST in the projected scenario and in the historical simulation, between the second highest SSTs in the two simulations, and so on for each ranked SST quantile. However, unlike Ashfaq et al. (2011), we did not create a new SST field for the present by replacing SST from the GCM in the historical period by its corresponding quantile in the observations, but we directly added the quantile change~~ observations. The second is a quantile-quantile method presented in Ashfaq et al. (2011) , where for each quantile and each month, the climate change signal coming from the AOGCM scenario is added to the corresponding quantile ~~of the observational time series (Figure A1). This allows keeping the observations chronology and their inter-annual variability in estimated SSTs for the future. In our results, we noticed a large fine-scale spatial variability of~~ in the ~~constructed bias-corrected SSTs that was due to the large spatial variability of the climate change increments (quantile change) calculated individually for each pixel. To fix this , we applied a slight spatial filtering (3 grid point Hann box filter) of the quantile shifts in order to produce more consistent SST fields~~observations. Presenting these well-known methods in detail is of limited interest for the main part of this paper. However, interested readers can find a more complete description of the methods in Appendix A.

~~Illustration of the quantile-quantile method for min. and max. of SST time series for a grid point in the Central Pacific : GCM historical simulation (blue, left), GCM projected scenario (red, left), observed SST(thin, right), reconstructed future SST (thick, right)~~

2.3 Sea-ice Concentration methods

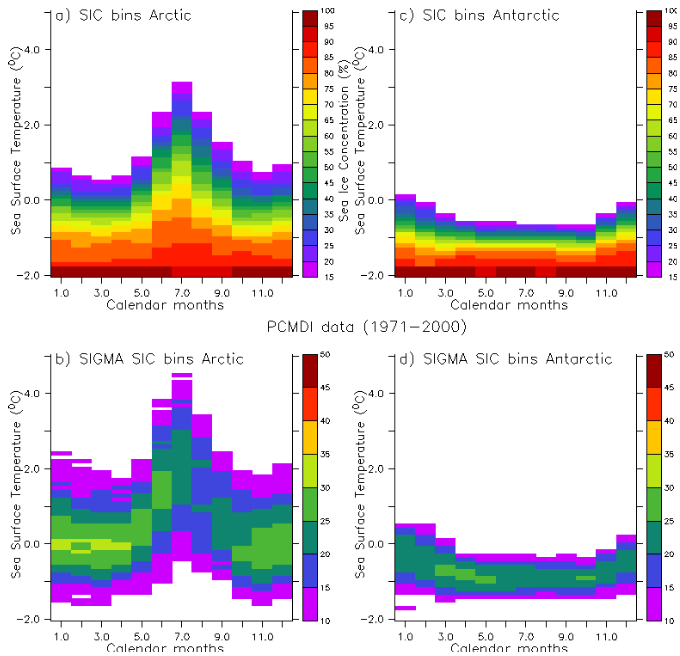


Figure 1. Look Up Tables (top) linking SST and SIC for the Arctic (left) and the Antarctic (right) built using 1971-2000 PCMDI observations and the associated uncertainty (root mean square error) on the computed SIC average (bottom).

~~Sea-ice concentration~~ SIC is more difficult to bias correct because it is a relative quantity that must be strictly bounded between 0 and 100 %. This difficulty led some authors to neglect SIC bias correction altogether in studies with prescribed corrected future SSTs that did not specifically focus on polar regions (e.g., Hernández-Díaz et al., 2017) . In this section, we present three methods: a look-up table, ~~a~~ an iterative relative anomaly and an analog method.

2.3.1 Look-up Table method

This method has been developed at *the Royal Netherlands Meteorological Institute* (KNMI). It is used in Haarsma et al. (2013) and within the framework of the High Resolution Model Intercomparison Project (HighResMIP) (Haarsma et al., 2016). ~~It is based on the assumption~~ A regression of SIC as a function of SST is also used in the HAPPI project (Mitchell et al., 2017) . In this method,the assumption is made that SIC is a function of SST. Therefore, SST are ranked per 0.1 K bins and the corresponding average SIC for each temperature bin between -2 and +5°C is calculated. Relations between SST and SIC have been found to be dependent on seasons and hemispheres. Therefore, using monthly mean values of SST and SIC from historical observations, look-up tables are built, separately for the Arctic and the Antarctic, for each calendar month (Figure 1). Then, with the help of future SSTs, these ~~look-up tables~~ Look-up Tables (LUT) are used to retrieve future SIC.

2.3.2 Iterative relative anomaly method

Here we follow a method described by Krinner et al. (2008). It is based on relative regional sea-ice area (SIA) changes ~~which~~ and is essentially an iterative scheme of mathematical morphology for image erosion and dilation (Haralick et al., 1987). The Arctic and the Antarctic are divided into sectors of equal longitude. In each sector, the average SIA is calculated by spatially integrating SIC. With respect to the method introduced in Krinner et al. (2008), we introduce the use of a quantile-quantile method to determine the targeted SIA in the bias-corrected projection. This targeted SIA is then calculated for each sector and each quantile, with the help of the following equation:

$$SIA_{Fut,est} = SIA_{obs} \cdot \left(\frac{SIA_{Fut,AOGCM}}{SIA_{Hist,AOGCM}} \right) \tag{1}$$

In (2), ~~$SIA_{\mathrm{Fut,est}}$~~ $SIA_{Fut,est}$ is the estimated projected SIA for the current month and sector, ~~SIA_{Obs}~~ SIA_{Obs} the SIA from the observations, and ~~$SIA_{\mathrm{Fut,AOGCM}}$ and $SIA_{\mathrm{Hist,AOGCM}}$~~ $SIA_{Fut,AOGCM}$ and $SIA_{Hist,AOGCM}$ are respectively computed SIA for the corresponding quantile to the observations, using SIC from a future scenario and a historical AOGCM's simulation. Starting from an observed present SIC map and using the computed relative SIA change for a given sector, the decrease (increase) in SIC is then realized using an iterative process: SIC in each grid box is replaced by the minimum (maximum) SIC of all adjacent pixels (Figure 2); the new spatially integrated SIA is calculated and the operation is repeated until the obtained change converges towards the computed targeted SIA retrieved from AOGCM's simulation sea-ice data and observations. Afterwards, the decrease/increase process is repeated on the hemisphere scale in order to ensure that the change in SIC reproduces the total hemispheric SIA change.

2.3.3 Analog method

In this method, we divide the Arctic and the Antarctic into ~~n_{s}~~ n_s geographical sectors that correspond to different seas of the Arctic and the Southern Oceans; we defined ~~n_{s}~~ n_s = 12 sectors for the Arctic and ~~n_{s}~~ n_s = 7 sectors for the Antarctic. For each sector and each month, the quantiles of the sea-ice extent (SIE: total area with SIC above 15%) and the SIA are computed from SIC observations over the AMIP period. Corresponding quantile changes in SIE and SIA are computed using SICs from a CMIP5 AOGCM ~~'s~~ historical simulation and a projected scenario run. Computed quantile changes are then applied to the corresponding quantiles in the observations in order to obtain ~~target~~ targeted future SIA and SIE for each month, quantile and sectors. Then, a library of future SIC fields is built by collecting SIC observations from the AMIP period as well as SIC from CMIP5 projections. The presence of SIC maps from ~~futures~~ AOGCM projections in this library is justified by the need to take into account physically plausible future SIC distributions outside of the current observed range. However, ~~AOGCM that overly~~ AOGCMs that poorly represent sea-ice ~~distribution~~ annual cycle in present-day climate are preferably dismissed from this library. Future SIC is then finally reconstructed by searching the analog for each quantile q, sector s and month m in the library, that is to say the SIC field that minimizes the cost function C expressed by:

$$C_{q,m,s} = \sqrt{\left(\frac{SIA_s - SIA_{T_{(q,m,s)}}}{SIA_{max_{(q,m,s)}}} \right)^2 + \left(\frac{SIE_s - SIE_{T_{(q,m,s)}}}{SIE_{max_{(q,m,s)}}} \right)^2} \tag{2}$$

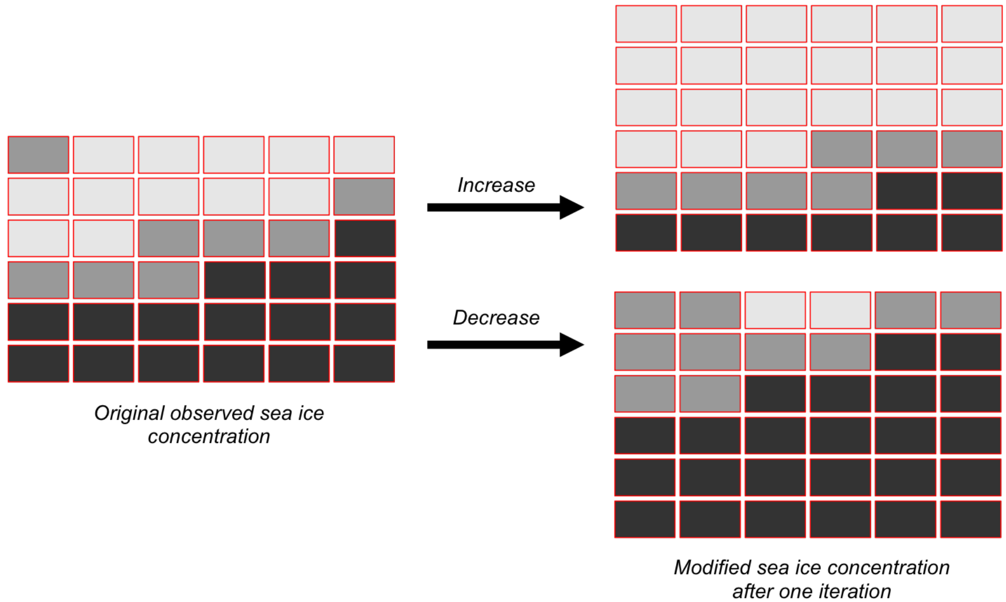


Figure 2. Iteratively constructing a "corrected" future SIC field using the iterative relative anomaly method (see ~~text~~section 2.3.2).

where SIA_s and SIE_s are the SIA and SIE of the processed sectors of the analog candidate from the library, ~~$SIA_{T_{(q,m,s)}}$ and $SIE_{T_{(q,m,s)}}$~~ $SIA_{T_{(q,m,s)}}$ and $SIE_{T_{(q,m,s)}}$ are the targeted ~~future sea-ice area and extent~~ projected SIE and SIA computed using the quantile-quantile method, and ~~$SIA_{max_{(q,m,s)}}$ and $SIE_{max_{(q,m,s)}}$~~ $SIA_{max_{(q,m,s)}}$ and $SIE_{max_{(q,m,s)}}$ are the maximum SIA and SIE of the processed sector. The double criterion on both SIE and SIA was introduced in order to be able to distinguish cases in which the total SIE in a sector is similar but the average SIC is very different (and vice versa). In order to avoid issues introduced by different land masks between AOGCMs and PCMDI data, we filled land grid points with sea-ice using a nearest neighbour method and masked all the grid points with the same land mask built with land fraction from PCMDI data in order to compute SIEs and SIAs for each region with the same reference. Analogs are attributed without taking into account the month of the analog candidate in the library. This allows for instance attributing a summer sea-ice map from present observations for a future winter month reconstructed sea-ice field. For each quantile q, month m and sector s, this procedure yields an hemispheric SIC field ~~$SIC_{opt_{(i,q,m,s)}}$~~ $SIC_{opt_{(i,q,m,s)}}$ that minimizes the cost function for the given sector, month and quantile. For a given month and quantile, there are thus ~~n_s~~ n_s hemispheric SIC fields ~~$SIC_{opt_{(i,q,m,s)}}$~~ $SIC_{opt_{(i,q,m,s)}}$. At each grid point i, the corresponding ~~n_s~~ n_s SIC values are then blended using a weight function ~~$w_{(i,s)}$~~ $w_{(i,s)}$ depending on the distance ~~$d_{(i,s)}$~~ $d_{(i,s)}$ of that grid point to the center of each of the sectors in order to obtain the final reconstructed SIC, ~~$SIC_{(i,q,m)}$~~ $SIC_{(i,q,m)}$, for a given quantile q and month m:

$$SIC_{(i,q,m)} = \sum_{s=1}^{n_s} \left(w_{(i,s)} \times SIC_{opt_{(i,q,m,s)}} \right) \tag{3}$$

with

$$w_{(i,s)} = \left(1 + \left(\frac{d_{(i,s)}}{d_r}\right)^4\right)^{-1} \quad (4)$$

Here, ~~d_r~~ d_r is a reference distance of 500 km, yielding a smooth transition at the boundaries between two adjacent sectors. At the center of a sector, this yields a weight that is very close to 1 for the relevant field that was identified as optimal for that sector and that is close to 0 for the fields identified as optimal for the other sectors; at the boundary between two sectors, the weights are typically 0.5 for the two relevant sectors and close to 0 for the others.

2.4 Sea Ice Thickness method

2.4.1 Diagnosing sea-ice thickness from sea-ice concentration

As described by Krinner et al. (2010) , the parameterization of sea-ice thickness SIT (denoted h_S in the following) as a function of the local instantaneous SIC f and annual-minimum SIC f_{min} is designed such as to yield h_S of the order of 3 meters for multi-year sea ice (deemed to be dominant when the local annual minimum fraction $f_{min} \gg 0$) and h_S below 60cm (with a stronger annual cycle) in regions where sea-ice completely disappears in summer (that is, $f_{min} = 0$), and intermediate values for intermediate cases:

$$h_S = (c_1 + c_2 f_{min}^2) \cdot (1 + c_3(f - f_{min})) \quad (5)$$

with c_1=0.2m, c_2=2.8m and $c3$=2. This corresponds to the observed characteristics of Arctic and Antarctic sea ice, with multi-year sea ice being generally much thicker than first year ice. The parameter c_3 introduces a seasonal ice thickness variation in areas where there is a concomitant seasonal cycle of SIC. A more parsimonious formulation using only two parameters could have been designed to comply with these constraints. However, for the sake of consistency with previous work, we used the equation proposed by Krinner et al. (1997) who designed the parameterization such as to allow for a fairly strong seasonal cycle of SIT also in regions with intermediate values of f_{min}.

2.5 Evaluation

Evaluation of the above methods is mainly achieved with a perfect model approach. ~~In this test~~A perfect model approach usually consists of using model data as a substitute for observations, and trying to predict projected model data from that model; this prediction can then be evaluated against the available model projections (e.g., Hawkins et al., 2011) . In the real world, as observations of future climate are obviously not yet available, an equivalent approach is impossible if one cannot wait long enough for the future to become reality. Another type of perfect model approach are "Big Brother" experiments for evaluating downscaling techniques. In such studies, high-resolution model output is degraded in resolution and downscaling methods are then applied to these low-resolution data. The resulting synthetic high-resolution fields are then compared to the original high-resolution output (e.g., Denis et al., 2002; de Elía et al., 2006) . Here, we consider SST and SIC from the historical simulation of one coupled AOGCM as being the observations. Then, we apply the different bias correction methods using

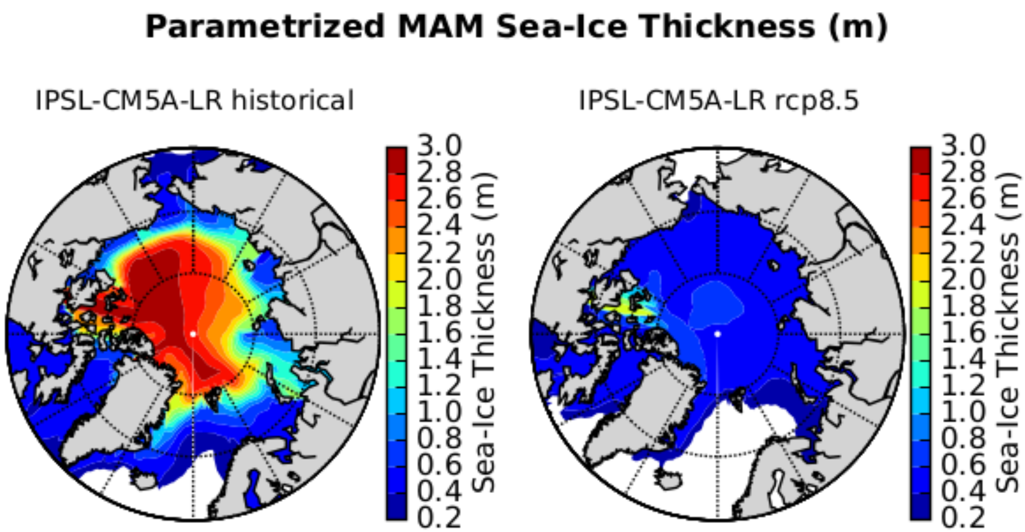


Figure 3. Spring (MAM) estimated mean SIT(m) using parametrization from (Krinner et al., 1997) and IPSL-CM5A-LR SIC data from the historical run (1971-2000, left) and the rcp8.5 scenario (2071-2100, right).

the climate change signal coming from a scenario of the same ~~AOGCM~~model. Obtained projected SST and SIC using this perfect model test are finally compared with original SST and SIC from the AOGCM climate change experiment.

Additionally, we also performed an assessment of real case applications using observations and climate change signals coming from AOGCM projections. Changes in mean and variance in the coupled model projection with respect to the historical simulation are compared to the introduced change in mean and variance in the estimated future SST and SIC using bias correction methods with respect to the observed climatological data. We ~~assume~~ consider here that an ideal bias correction method should reproduce the same change in mean and variance between the observations and the bias-corrected projected SST and SIC as between the used coupled GCM historical simulation and its climate change scenario. For SIT, since the method is a diagnostic using SIC in order to ensure the consistency between these two variables, the evaluation of the method is achieved by comparing estimated SIT with observations that were not available until recently (Lindsay and Schweiger, 2015; Kurtz and Markus, 2012) .

As SST and SIC are bias-corrected separately, section 3.3 presents a few considerations about SST and SIC consistency after performing bias corrections. The effects of the corrections applied *a posteriori* in order to ensure the physical consistency between the two variables are evaluated within the framework of the perfect model test.

3 Results

3.1 Sea Surface Temperatures

3.1.1 Perfect model test

~~In this section, we discuss the application of the perfect model test for both the anomaly and the quantile-quantile method . To apply this test, we used CNRM-CM5 data from the historical simulation over the 1971-2000 period and from the rcp8.5 projection for the 2071-2100 period. Corrected rcp8.5 SST have been compared with the original SST projection. For the anomaly method~~ Absolute anomaly or quantile-quantile methods have been used for SST in previous bias-correction applications cited before in this paper. As a consequence, the utility of a perfect model test here is limited for SSTs, and it was only applied in order to be consistent with the evaluation of the method for SIC. For both methods, the relation between the ~~anomaly-corrected~~ bias-corrected projected SST and the SST directly obtained from the AOGCM projection is trivial when we replace observed SST by ~~SST~~ the one from the AOGCM historical simulation, as for instance in (1). As a result, ~~when comparing corrected rcp SST using the perfect model test and original SST from CNRM-CM5 rcp8.5 scenario, we obtain, by construction, a null bias all over the world (*figure not shown*). For the quantile-quantile method, the bias is also null in most regions. However, since we applied a very slight spatial filtering of the quantile increment, some slight biases (positive or negative) appear in regions of steep SST gradients (i.e. regions with major oceanic currents). Nevertheless, these biases are negligible (a few tenths of degrees Celsius;*figure not shown*)~~the resulting errors were null or close to zero, and the results are therefore not presented or discussed.

3.1.2 Real-case application

Here, we present the application of the anomaly and the quantile-quantile methods in a ~~real case~~ real-case application. For this application, we use SST ~~data~~ from PCMDI observations data set over 1971-2000, from ~~the~~ IPSL-CM5A-LR and CNRM-CM5 historical simulation over the same period, as well as the rcp8.5 scenario over 2071-2100. Histograms of frequency distribution of SST for different regions of the world (Weddell Sea, Central Pacific and North Atlantic) have been plotted in order to compare frequency distributions in the observations, in the GCM historical and future simulations, as well as in the estimated bias-corrected future SST using the quantile-quantile and the anomaly ~~method~~ methods (Figure 4). In this figure, we can appreciate the change in mean and variance between the GCM historical simulation and the GCM future scenario and between the PCMDI observations and ~~in the estimated~~ the bias-corrected SST scenario. ~~This~~ In Figure 4 (bottom), we can see the large cold bias of the AOGCM with respect to the observations in the North Atlantic, as coupled models usually struggle to correctly represent the Atlantic Meridional Overturning Circulation (AMOC). The change in mean and variance due to the climate change signal is more explicitly ~~calculated and~~ presented for the North Atlantic for the application with CNRM-CM5 model in Table 1. Results from the anomaly method and from the quantile-quantile method are very similar, and both methods succeed in applying the same change in mean and variance coming from the AOGCM scenario to the observations when producing bias-corrected SST.

Table 1. Mean and standard deviation ~~difference~~ change between present and future SST data sets for North Atlantic (45°N to 58°N, 105°W to 85°W)

	Mean ~~difference~~ change (°C)	STD ~~difference~~ change (°C)
CNRM-CM5 rcp8.5 - CNRM-CM5 hist	+3.04	+0.59
Anomaly meth. app. - PCMDI obs	+3.06	+0.66
Quantile-quantile meth. app. - PCMDI obs	+3.04	+0.68

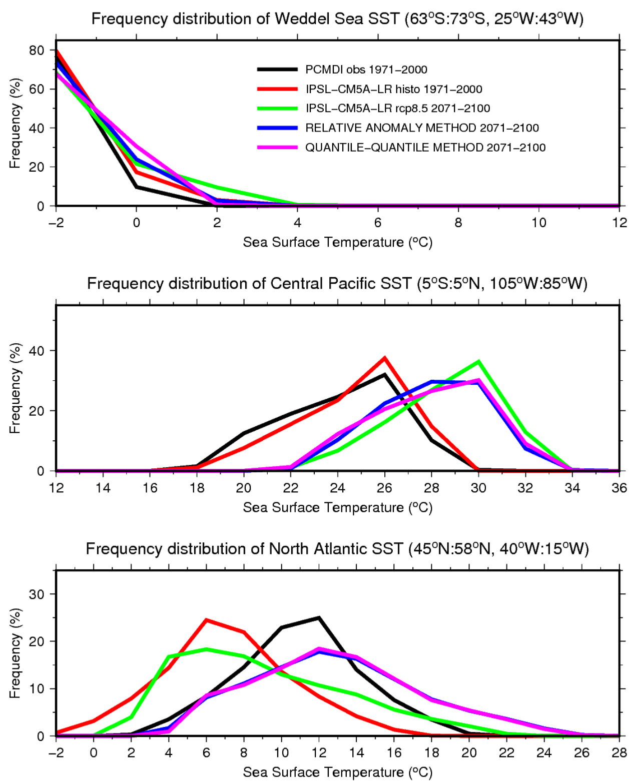


Figure 4. Frequency distribution of SST for PCMDI observations (black), IPSL-CM5A-LR historical (red) over 1971-2000 and rcp8.5 (green), quantile-quantile method (pink) and anomaly method (blue) applications over 2071-2100 for Weddell Sea (top), Central Pacific (center) and North Atlantic (bottom)

3.2 Sea-Ice Concentration

3.2.1 Perfect model test

In this section, we present the results of the application of the perfect model test for the three methods for bias correction of SIC. The term "perfect model test" is not absolutely pertinent for the evaluation of the Look-up Table method, as we first ~~computing look-up tables~~ computed LUTs using SST and SIC from an AOGCM historical simulation. Then, we used the SST of the climate change projection from the same AOGCM and retrieved SIC with the help of the previously computed LUT. An example of computed LUT using data of the historical simulation of CNRM-CM5 can be seen in Figure 5. It is noteworthy that this new ~~look-up table~~ LUT is significantly different from the one using PCMDI observations (Figure 1). Even though, the use of this LUT for the perfect model test instead of LUTs computed using observed SST and SIC over the AMIP period can be discussed, the use of LUT computed using observations would necessarily produce poorer result for the reconstruction of SIC ~~the AOGCM future~~ of the AOGCM's scenario in a perfect model test. Using AOGCM data, inconsistent or missing results were found for most of SST bins at or below the freezing point of sea water (-1.8°C). In order to fill the LUT, we therefore fixed SIC=99% for SST=-2.0°C and linearly interpolated SIC between -1.7°C and -2.0°C.

The perfect model test is more rigorously applied for the evaluation of the relative anomaly and the analog method, as we simply replaced time series of the observed SIC by the one from the AOGCM historical ~~scenario~~ simulation before applying the method without any specific modification or calibration. For the analog method, ~~we mention that~~ the tested AOGCM projection has been excluded from the possible analog candidates before applying the method and the perfect model test.

~~Mean biases~~ Errors (%) after applying the perfect model test are shown for the three methods for the rcp4.5 and rcp8.5 scenarios of the IPSL-CM5A-LR and CNRM-CM5 AOGCM (Figure 6). ~~One can see that the mean bias~~ These errors are generally lower for the LUT method : the mean Root Mean Square Error (RMSE) on the estimation ~~of sea-ice concentration remains reasonable for most of~~ for each scenarios for the Arctic and the Antarctic ~~for the analog method and very small for the look-up table method~~ is 4,8%. The mean error (ME) using this method tends to be positive in the Arctic and negative in the Southern Oceans. Errors using the relative anomaly method exhibits some larger values (mean RMSE = 8%). The errors using the analog method have intermediate values with respect to the first two methods (mean RMSE = 5.9%). Some of the ~~biases~~ errors of the analog method for regions with very complex coastal geography, such as the Canadian Archipelago, are due to the differences in land mask between the tested and the chosen AOGCM as analog candidate, despite the care taken for this issue. ~~Mean bias for the relative anomaly method exhibits some larger values.~~ The pattern of the ~~biases using this method~~ errors using the iterative relative anomaly seems robust between the different AOGCM scenarios. It is also noteworthy that the pattern of the ~~biases~~ errors is also similar between different methods, especially if we consider the results in the Arctic for the scenarios of the CNRM-CM5 model.

With the results of the perfect model test, we also performed a comparison between the frequency distribution of the mean SIC in the AOGCM future scenario (here CNRM-CM5, rcp8.5) and in the corresponding estimation using the bias correction methods (Figure 7). In these plots, we represented the histogram of frequency of ~~sea-ice concentration~~ SIC for four regions: Ross Sea (72°S:77°S; 174°E:163°W), Weddell Sea (63°S:73°S; 45°W:25°W), Arctic Basin (80°N:90°N; 180°W:180°E), and the Canadian Archipelago (66°N:80°N; 130°W:80°W). These regions have been chosen because they are the principal regions where there remains a significant amount of ~~sea-ice~~ sea ice by the end of the 21st century under the rcp8.5 scenario. With the ~~look-up table~~ LUT method (blue lines in Figure 7), the distribution of ~~sea-ice concentration is more or less~~ SIC is quite well reproduced in the Arctic (Figure 7 c and d), whereas in the Antarctic seas the distribution (Figure 7 a and b) exhibits well-marked peaks that we do not find in the GCM data set (black lines). The presence of such peaks is easy to explain by taking into account the structure of the ~~look-up tables~~LUT as i) for a given month, the SIC does not always increase monotonically with decreasing SST, ii) the discrete nature of LUT is not in favour of a continuous SIC frequency distribution. Moreover, using this method, we find a large underestimation of ~~the sea-ice concentrations~~ SIC above 90%, mainly in the Southern Hemisphere, with almost no occurrence of these high ~~sea-ice concentrations~~ SIC values in the estimations using the LUT method for the Ross and Weddell Seas. The frequency distribution of the sea-ice using the relative anomaly method (green lines in Figure 7) ~~seems more reasonable~~is closer to the distribution in the AOGCM, even if there is a slight overestimation of the frequency for concentrations between 70 and 90% and an underestimation for very high ~~sea-ice concentrations~~ SICs (above 90%). Finally, the distribution obtained using the analog method (red lines on Figure 7) is very close to the distribution of the original AOGCM ~~future~~ scenario. The results are robust because differences of sea-ice frequency distribution between ~~future estimation and future AOGCM future scenario~~ bias-corrected projections and AOGCMs scenarios are very similar for ~~rcp4.5 from CNRM-CM5 as well as for both scenarios from IPSL-CM5A-LR~~ other scenarios and coupled models (*figures not shown*).

3.2.2 Real-case application

~~In this section, we~~ We applied the three bias correction methods using PCMDI SIC ~~observations~~ data from the 1971-2000 period, as well as the IPSL-CM5A-LR and CNRM-CM5 historical data over the same period and ~~the~~ data from the rcp4.5 and rcp8.5 ~~future~~ scenarios from 2071-2100 in order to obtain future bias-corrected ~~sea-ice corrections~~SIC. The reliability of the methods is evaluated by comparing the change in mean and variance between the observations ~~over present climate and future estimated sea-ice concentrations to~~ and the bias-corrected projected SICs with the corresponding changes in the ~~climate change simulation~~ original AOGCM scenario with respect to the historical simulation. ~~An~~ We consider here that an ideal method should apply the same statistical changes to observed sea-ice as the one present in the climate change projection used to derive climate change signal.

In Figure 8, the bias-corrected mean ~~sea-ice concentration~~ SIC change is plotted against the corresponding change in mean SIC in the AOGCM ~~future~~ scenario used to determine the climate change signal. All points in the plot are obtained by the same four AOGCM ~~future~~ scenarios as well as the same four "test regions" as in previous section (Ross and Weddell Seas, Arctic Basin, Canadian Archipelago). Similarly, in Figure 9, applied changes in standard deviation for the ~~future estimated~~ bias-corrected projected SIC are plotted against corresponding standard deviation change in the AOGCM climate change experiment.

For the ~~look-up table~~ LUT method (Figure 8a), future SSTs have been bias-corrected using the quantile-quantile method before using computed LUT for the retrieval of future SIC. Using this method, there seem to be no systematic ~~errors~~ error in the applied change in mean SIC. ~~However, the~~ The mean error on the estimation of the change in mean SIC for every regions and scenarios

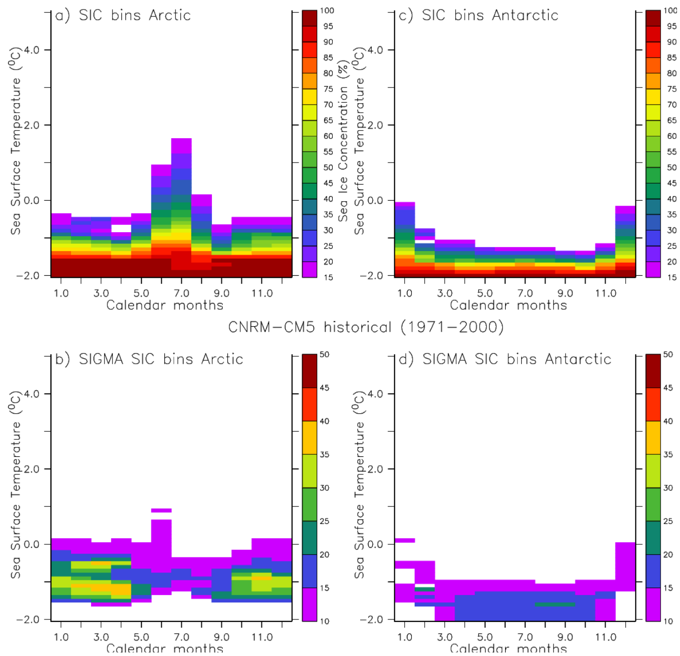


Figure 5. Look-up tables linking SST and SIC for the Arctic (a) and the Antarctic (c) built using 1971-2000 CNRM-CM5 historical simulation data and the associated uncertainty (root mean square error) on the computed SIC average (b,d)

is -2.2% and the RMSE is 42%. The spread of the points seems to increase for stronger decreases ~~of~~ in sea-ice. Main outliers with a high overestimation of the decrease in SIC are ~~constituted by~~ points representing the evolution of sea-ice in the Weddell Sea, mainly for CNRM-CM5 scenarios. If we consider change in SIC variability (Figure 9a), ~~there is a strong systematic bias and the~~ systematic error (-14.9%) and RMSE (69.3%) are strong. The decrease in SIC variability in the ~~future~~ Antarctic seas in the projection is strongly overestimated. Indeed, due to the structure of the ~~look-up table itself~~LUTs themselves, the variability of SIC in ~~future estimations~~ the bias-corrected projections is much lower than in the observations or in the original scenarios. The application of the relative anomaly method shows a more general overestimation (ME = -11.6% ; RMSE = 52.2%) of the decrease in mean SIC (Figure 8b). This overestimation is more pronounced for the Weddell Sea area and for the scenarios of the CNRM-CM5 model. Only the decrease in mean SIC in the Arctic Basin is correctly reproduced with respect to the AOGCMs ~~future~~ scenarios. Concerning the change in SIC variability (Figure 9b), the scores are comparable to the application of the LUT method (ME = -11,6% ; RMSE = 64.7%). The increase in variability in the Arctic Basin and in the Canadian Archipelago is correctly reproduced whereas for the Antarctic seas and particularly the Weddell sector, the decrease in SIC variability is once again ~~massively~~ dramatically overestimated.

Finally, the application of the analog method ~~is able to reproduce a great part~~ gives intermediate scores (ME = -8% ; RMSE = 48.7%) with respect to the two previous methods for the estimation of the change in mean SIC (Figure 8c). ~~Nevertheless,~~ These scores are greatly deteriorated by distinct outliers corresponding to the Weddell Sea sector ~~are once again present~~ for each AOGCM scenario, with ~~a strong~~ an overestimation of the decrease in sea-ice. As for the relative anomaly method, the

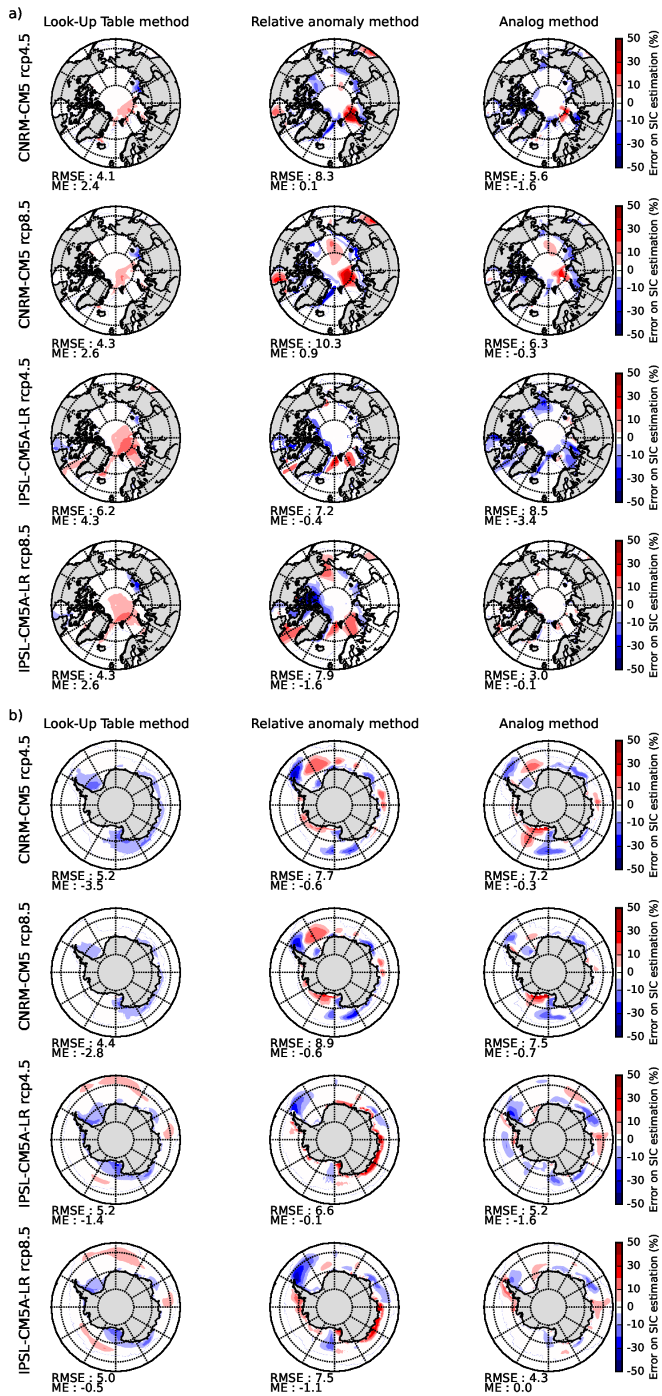


Figure 6. Mean ~~bias~~ error on the estimation of SIC with respect to the original AOGCM future scenario for the LUT, iterative relative anomaly and analog ~~method~~ methods with CNRM-CM5 and IPSL-CM5A-LR rcp4.5 and rcp8.5 scenarios for the Arctic (a) and the Antarctic (b)

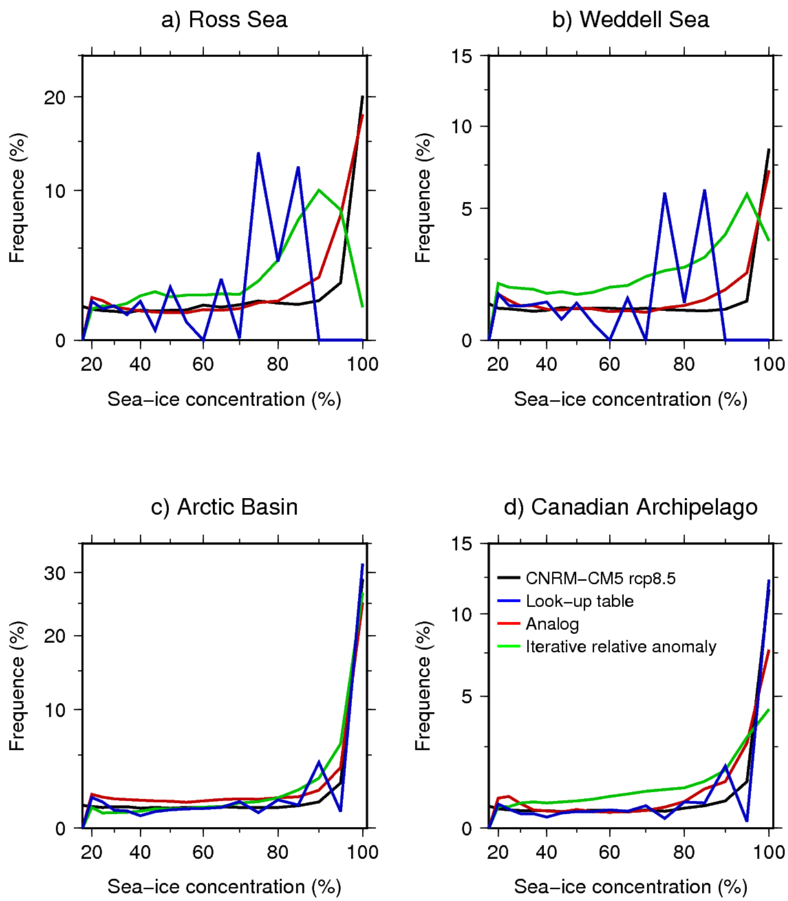


Figure 7. Frequency distribution of SIC in CNRM-CM5 rcp8.5 scenario (black) and in estimation using different methods in a perfect model test: Look-up table (blue), analog (red), and iterative relative anomaly (green). Regions are: a) Ross Sea (72°S:77°S, 174°E:163°W); b) Weddell Sea (63°S:73°S, 43°W:25°W); c) Arctic Basin (80°N:90°N, 180°W:180°E); d) Canadian Archipelago (66°N:88°N, 130°W:80°W)

change in SIC variability (Figure 9c) is correctly reproduced (ME = -9.3% ; RMSE = 60.3%), especially in the Arctic, while there is ~~a strong~~ an overestimation of the decrease in variability around Antarctica, particularly for the Weddell Sea.

3.3 Consistency between Sea Surface Temperature and Sea-ice Concentration~~consistencyAs bias correction~~

As bias corrections of SST and sea-ice are performed separately, the physical consistency between the two variables ~~is assessed~~ needs to be ensured a posteriori. To do so, three different issues are examined:

- There is a considerable amount of sea-ice (>15%) in the corrected scenario where the SST is above fresh water freezing point (273.15K). In this case, we set SST equal to the sea water freezing point (271.35K) for any SIC equal or greater than 50%. If the future calculated SIC is between 15 and 50%, the future SST is obtained by linearly interpolating between the sea water freezing point and the freshwater freezing point.

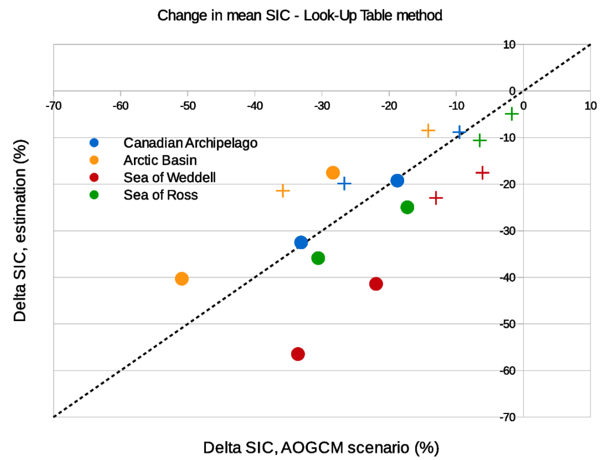
- The future corrected ~~sea-surface temperature~~ SST is below the fresh water freezing point but there is no significant (<15%) SIC in the bias-corrected scenario. In this case, we put the SST of the concerned grid point equal to the fresh water freezing point.

- SST has been used to remove very localized suspicious presence of sea-ice (no-ice) in the Arctic in summer. Any sea-ice for SST above 276.15K has been removed, this temperature being the highest temperature at which significant amount of sea-ice (15%) is found is the Arctic in the computed ~~look-up table~~ LUT using PCMDI data.

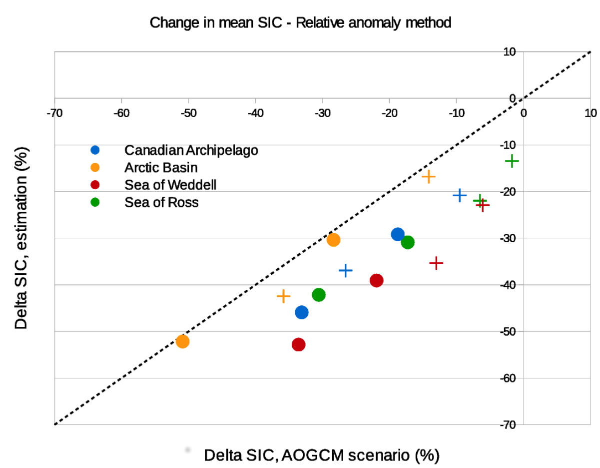
The impact of these modifications has been evaluated using the framework of the perfect model test. After applying the analog method for SIC and the quantile-quantile method for SST in a perfect model approach, we applied the correction for SST and SIC consistency and compared obtained SSTs to the original AOCGM future scenario used to carry out the experiment. The ~~biases~~ error can be seen in Figure 10 for the application of the method with IPSL-CM5A-LR and CNRM-CM5 scenarios. ~~It~~ Error is negligible in most regions. Very locally, it can reach up to 1°C. These regions generally correspond to regions where the analog method has shown some ~~biases~~ errors for the reconstruction of sea-ice especially for CNRM-CM5 scenarios. The occurrences of the three cases mentioned above have been assessed for both the perfect method test and the real-case application. ~~First~~ The first and third cases are very ~~seldom~~ rare and about 1% or less of the global oceanic surfaces experience at least one case during a 30 years experiment. The second case is more frequent, more than 20% of the global oceanic surfaces experience at least one occurrence during a 30 year experiment, while the mean occurrence at each time step is about 1 to 2% of the global oceanic surfaces. This case is responsible for the small (0.25 to 0.5K) but widespread warm bias on SST that can be seen in the Antarctic seas for the reconstruction of IPSL model scenarios in Figure~~10.~~ 10. Nevertheless, this slight decrease in the quality of the reconstruction of SST is worth considering in order to ensure physical consistency between SST and SIC.

4 ~~Discussion~~

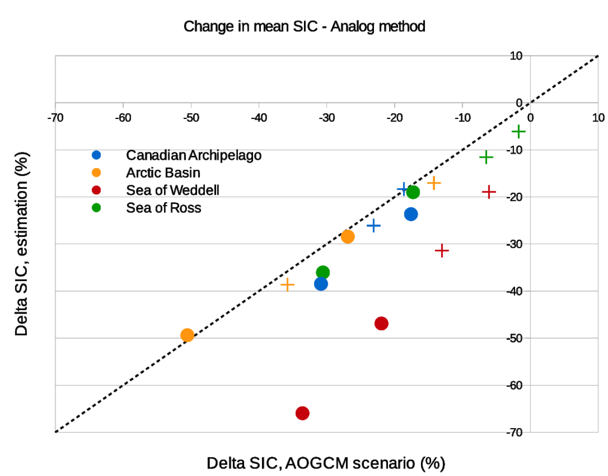
3.1 Sea Ice Thickness



(a)

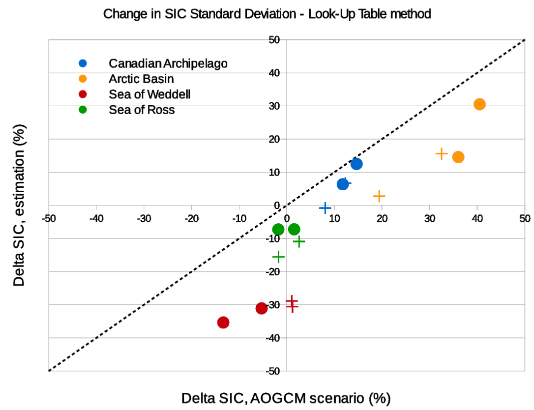


(b)

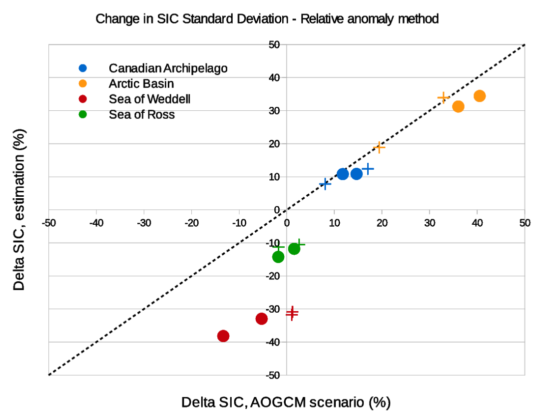


(c)

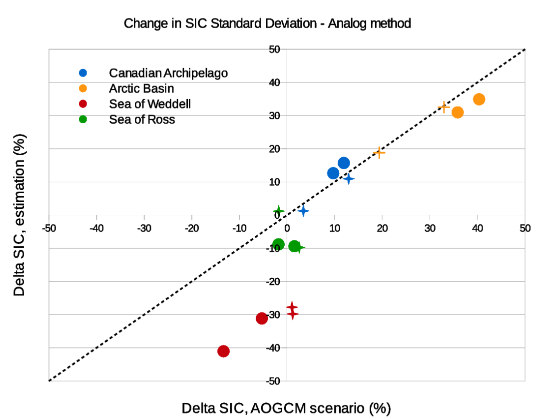
Figure 8. Change in mean ~~estimated future~~ bias-corrected SIC projections using a) ~~look-up table~~Look-Up Table, b) iterative relative anomaly, c) analog ~~method~~ methods against corresponding mean change in the AOGCM ~~future~~ scenario for the four test regions ~~(~~: Canadian Archipelago (blue), Arctic Basin (orange), Weddell Sea (red) and Ross Sea (green). Circles represent scenarios (rcp4.5 and rcp8.5) of



(a)



(b)



(c)

Figure 9. Change in ~~estimated future~~ bias-corrected SIC projections standard deviation using a) ~~look-up table~~Look-Up Table, b) iterative relative anomaly, c) analog ~~method~~ methods against corresponding mean change in the AOGCM ~~future~~ scenario for the four test regions ~~(~~: Canadian Archipelago (blue), Arctic Basin (orange), Weddell Sea (red) and Ross Sea (green). Circles represent scenarios (rcp4.5 and rcp8.5) of CNRM-CM5 and crosses, scenarios of IPSL-CM5A-LR

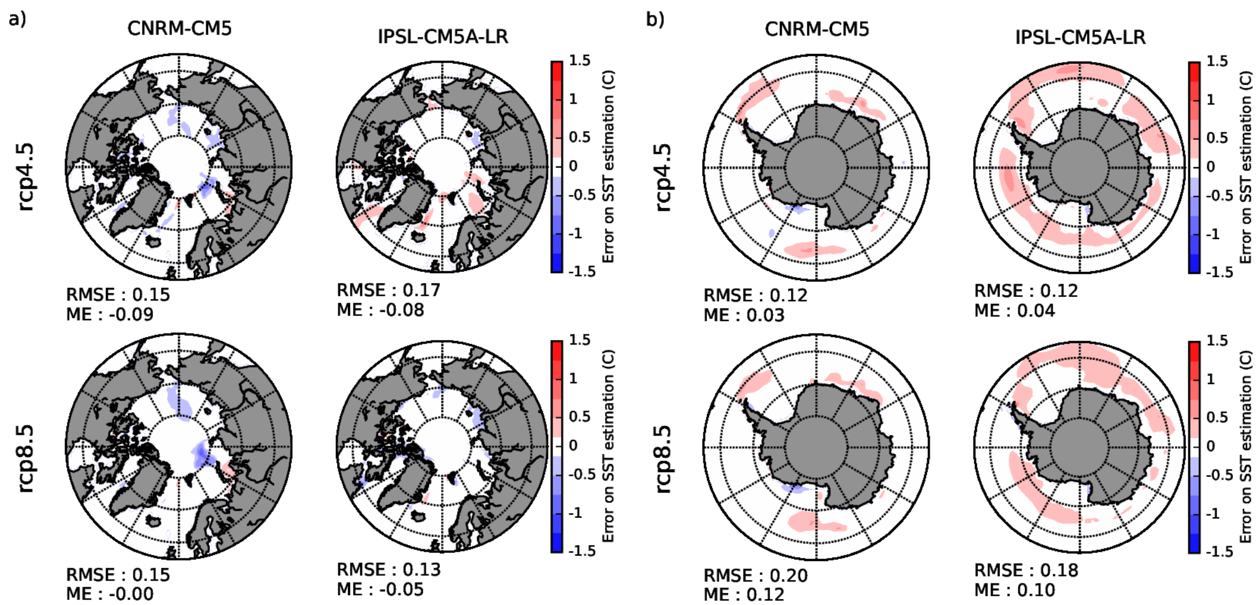


Figure 10. Mean ~~bias~~ error on the estimation of ~~the sea surface temperature~~ SST with respect to the corresponding original AOGCM ~~future~~ scenario after applying the analog method for sea-ice, the quantile-quantile method for SST and the correction for SST and SIC consistency for the Arctic (a) and the Southern Oceans (b)

The original formulation by Krinner et al. (1997) was parameterized for both hemispheres. We will therefore first present results for the original unique parameter set $c_{1,2,3}$ applied to both hemispheres. In a second step, we will present results for separate Arctic and Antarctic parameter sets, yielding a better fit to the observations. The reasoning is that, at the expense of generality of the diagnostic parameterization, one could argue that the strong difference between the Arctic and Antarctic geographic configuration — a closed small ocean favouring ice ridging and thus thicker sea ice in the Arctic, and large open ocean favouring thinner sea ice around Antarctica — justifies choosing different parameter sets for the two hemispheres. As changes of the position of the continents will be irrelevant over the time scales of interest here, climate change experiments will not be adversely affected by this loss of generality.

3.1.1 Option 1: Global parameter set

A comparison between the observed (Lindsay and Schweiger, 2015) and our diagnosed evolution of the Arctic mean SIT is given in Figure 11. The geographical patterns of the observed (in fact, observation-regressed) and parameterized Arctic ice thickness for March and September over the observation period 2000-2013 (Figure 12) do bear some resemblance, but they also show some clear deficiencies of the diagnostic parameterization. The diagnostic parameterization reproduces high SIT north of Greenland and the Canadian Archipelago, linked to persistent strong ice cover, but underestimates maximum ice thickness (due in part to compression caused by the ocean surface current configuration). Thinner sea ice over the seasonally

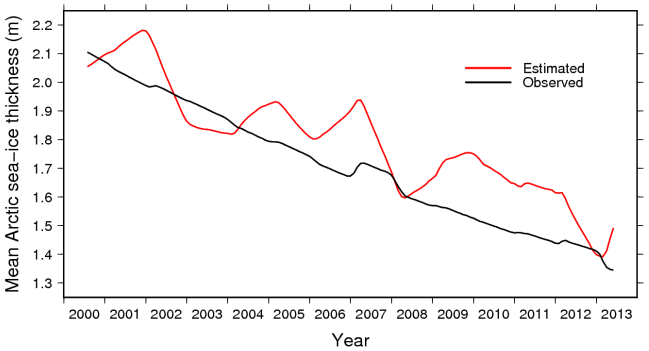


Figure 11. Observed (black, after (Lindsay and Schweiger, 2015)) and diagnosed (red) 12-month moving average mean sea-ice SIT of the Arctic basin (see Figure 12). The global parameter set is used here. Slight differences to Figure 13 of Lindsay and Schweiger (2015) appear because here we mask ice-free (SIC < 15%) areas that have a finite, non-zero ice thickness in the regression proposed by Lindsay and Schweiger (2015) who extend their regression to the entire Arctic Basin at all seasons.

ice-free parts of the basin is reproduced, but it is actually too thin, particularly in winter (for example in the Chukchi Sea). Obvious artifacts appear in September north of about 82°N where the SIC in the ERA-Interim data set clearly bears the signs of limitations due to the absence of satellite data.

Both for spring (Oct-Nov) and fall (May-Jun), our diagnosed SIT (Figure 13) compares generally well with the ICESat data except for an overestimate in the Weddell Sea, at both seasons. The geographical pattern of alternating regions with thin and thick sea ice is remarkably well reproduced.

3.1.2 Option 2: Separate Arctic and Antarctic parameter sets

A slightly better fit for the two poles can be obtained with separate parameters sets. For the Arctic, it seems desirable to increase winter SIT in the Chukchi Sea area (by increasing c_3 slightly) and to decrease the average SIT over the Central Arctic (by decreasing c_2). Figures 14 and 15 show results for the Arctic with c_1=0.2m, c_2=2.4m and c_3=3. The spatial fit is slightly better, but the recent Arctic-mean decadal trend towards decreased average SIT is somewhat less well reproduced. For the Antarctic, the main feature to improve is the maximum ice thickness in the Weddell Sea, which can be decreased by lowering c_2 to 2.0m. The Antarctic parameter set then becomes c_1=0.2m, c_2=2m and c_3=2. The result (Figure 16) is indeed a decreased thickness of the perennial Weddell Sea ice with little impact elsewhere.

In any case, these hemisphere-specific sea-ice parameter sets are not very different from each other and fairly similar to the original formulation.

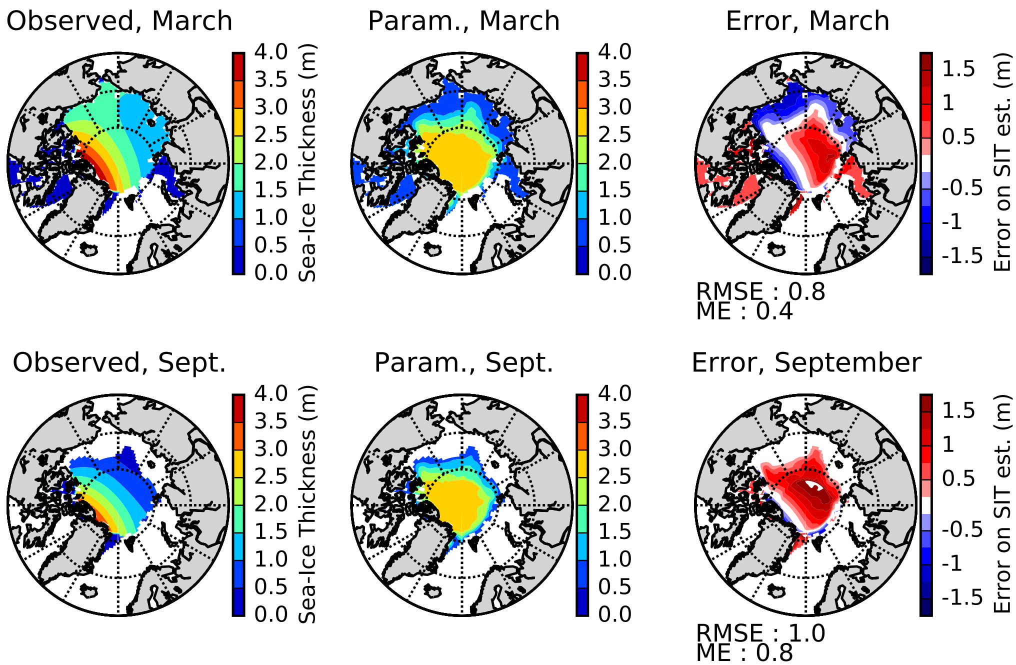


Figure 12. Observed (regressed, Lindsay and Schweiger (2015)) and parameterized Arctic SIT (in m) for March and September, and difference between these (right), with the global parameter set.

4 Discussion

4.1 Sea Surface Temperatures

The bias correction of projected ~~SSTs~~ SST coming from AOGCM scenarios is ~~an issue~~ fairly easy to deal with, and different appropriate solutions have already been proposed in the literature ~~(e.g., Krinner et al., 2008; Ashfaq et al., 2011; Hernández-Díaz et al., 2017)~~ In these papers, it has been demonstrated that the use of bias-corrected SSTs has considerable influences on the modeled climate and its response in projected scenarios for regions and processes as different as precipitation and temperature in the ~~Tropics and~~ tropics,the West African Monsoon ~~as well as for~~ and the climate of Antarctica.

In this paper, we reviewed two existing bias-correction methods and propose a validation that allows objectively evaluating the efficiency of these methods with the use of a perfect model test and a real-case application. Since both methods show no ~~bias~~ biases in the perfect model test and succeed in reproducing the change in mean and variability coming from the AOGCM future scenarios, we can be confident in the use of these methods for bias-correction of future AOGCM scenarios.

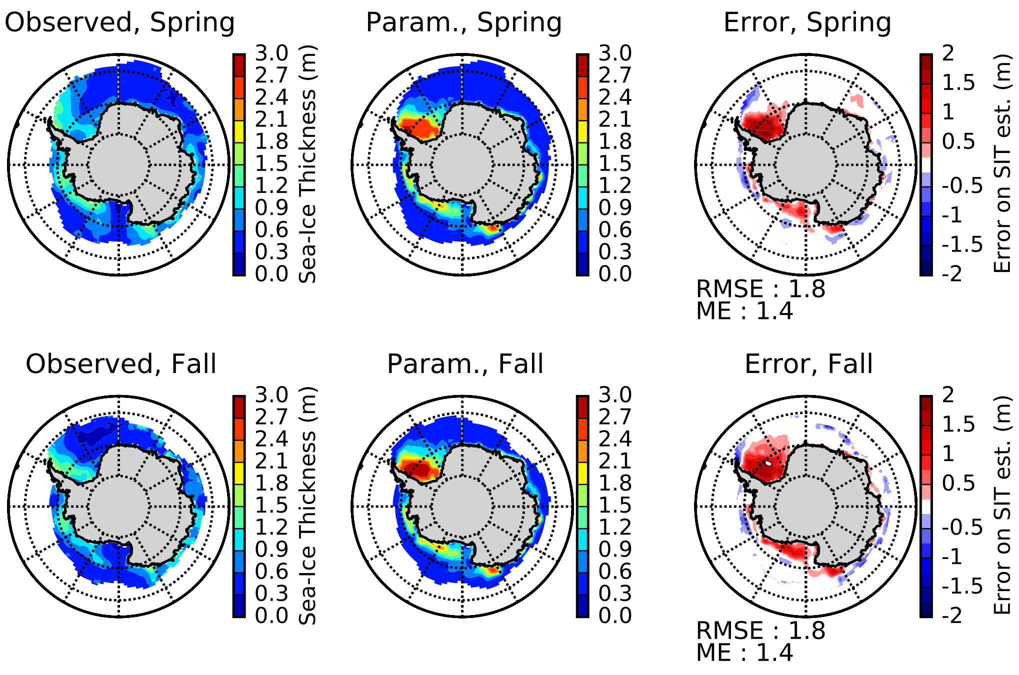


Figure 13. Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for Spring and Fall, and difference between these (right), with the global parameter set.

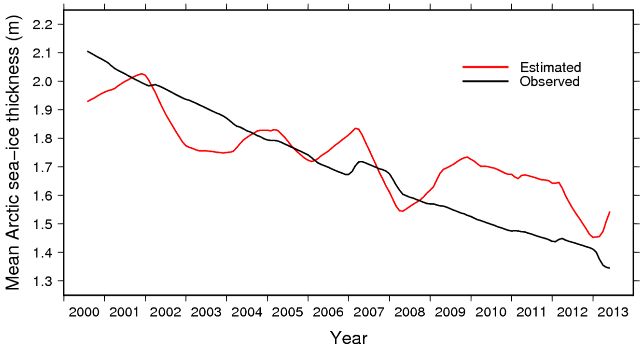


Figure 14. Observed (black, after Lindsay and Schweiger (2015)) and diagnosed (red) 12-month moving average mean SIT of the Arctic basin with the Arctic-specific parameter set.

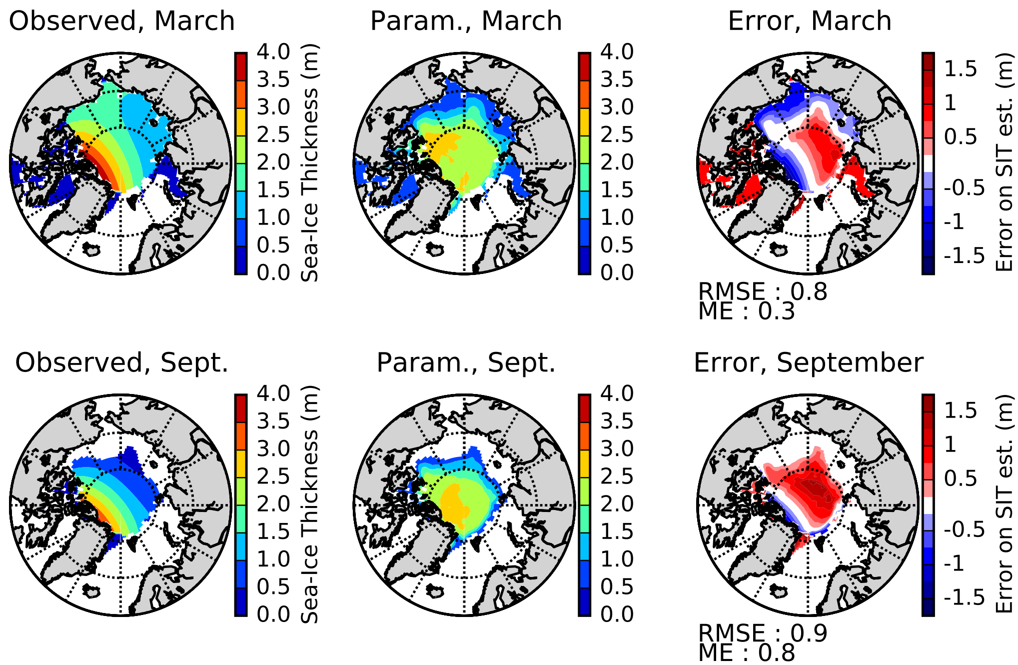


Figure 15. Observed Lindsay and Schweiger (2015) and parameterized Arctic SIT (in m) for March and September, and difference between these (right), with the Arctic-specific parameter set.

4.2 Sea-Ice Concentration

SIC is a quantity that has to remain strictly bounded between 0 and 100%, exhibits some sharp gradients and has to remain physically consistent with SST. Therefore the empirical bias correction of future SIC from coupled models scenarios is a much more complex issue to deal with than the bias correction of SSTs. The absence of satisfying solution proposals for this issue in the literature has led to incorrect bias-correction of future SIC in a recent study (Hernández-Díaz et al., 2017). Yet, the proposal of convenient solutions for the bias correction of sea-ice for ~~future~~ projected scenarios is crucial for the community interested in the downscaling of ~~future climate scenarios~~ climate scenarios experiments for polar regions.

In the perfect model test, we have seen that the ~~look-up table~~ LUT method shows some reduced ~~mean bias~~ errors over most regions (Figure 6). However, we have seen that the frequency distribution of future SIC obtained using this method is ~~different from~~ very different than the original distribution in the AOGCM and unavoidably exhibits some peaks due to the structure of LUT (Figure 7). Moreover, the absence of SIC above 90% in the Antarctic is also a considerable limitation to the method considering the large differences in terms of heat and moisture exchanges in winter between an ocean fully covered by sea-ice and an ocean that exhibits some ice-free channels (Krinner et al., 2010). In addition, the use of SST as a proxy for SIC is physically questionable, as we should expect a large SIC gradient around the freezing point. The fact that both SST and SIC

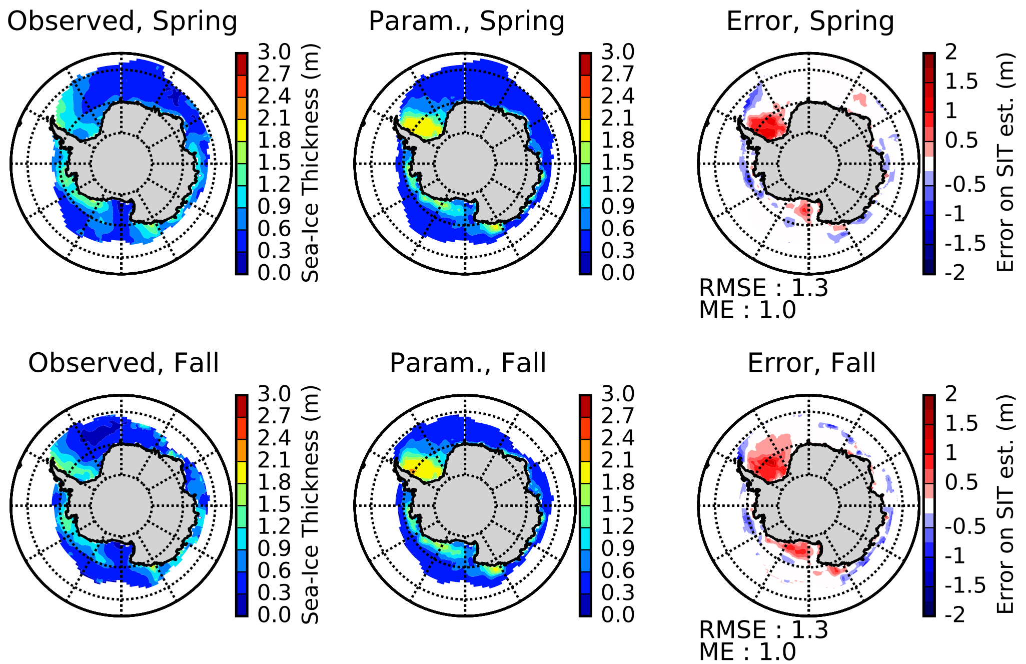


Figure 16. Observed Kurtz and Markus (2012) and parameterized Antarctic SIT (in m) for Spring and Fall, and difference between these (right), with the Antarctic-specific parameter set.

are averaged over a long period (one month) and over a considerable area (1°x1°) is probably the main reason why we find nevertheless a relation between the two variables. The ~~real case~~ real-case application of the method also shows some difficulties for the reconstruction of large decreases in mean SIC (Figure 8a) as well as a poor reconstruction of the change in variability in future SIC (Figure 9a).

The relative anomaly method (Krinner et al., 2008) shows the largest spatial mean ~~biases~~ errors in the perfect model test (Figure 6). The structure of some ~~biases~~ errors seems to be constant across the reconstruction of different climate scenarios used in the perfect model test. The empirical reduction of SIC by an iterative "erosion" from the edges of the sea-ice covered regions has most likely the tendency to overestimate the decrease of sea-ice for some coastal regions, while it probably fails to reproduce some processes involved in the disappearance of sea-ice in the future such as for example the inflow of warmer waters through the Barents Sea or the Bering Strait in the Arctic. The "real-case" application of the relative anomaly method has shown some systematic negative ~~bias~~ errors in the reconstruction of the decrease in mean SIC (Figure 8b) and ~~some important~~ a substantial overestimation of the decrease in variability in the Antarctic seas (Figure 9b).

The evaluation of the analog method with the perfect model test ~~allows to highlight some mean biases locally slightly bigger~~ shows that the mean error can be locally slightly higher than for the ~~look-up table~~ LUT method (Figure 6). However, the frequency distribution of the ~~future estimated~~ bias-corrected SIC perfectly reproduces the frequency distribution of the sea-ice in the original AOGCM ~~future~~ scenario (Figure 7). The real-case application of the method succeeds in reproducing the change in mean and variability of SIC for most of the tested regions and scenarios (Figure 8c). However, the decrease in mean (Figure 8c) and variability (Figure 9c) of the sea-ice in the Antarctic, particularly the Weddell Sea, is also largely overestimated using this method. With respect to the relative anomaly method, the fact that we use observed or ~~AOGCMs modeled~~ AOGCM-simulated sea-ice maps to reconstruct estimated future ~~sea-ice~~ sea ice, and that we use a criterion for both ~~sea-ice area and sea-ice extent~~ SIA and SIE, allows us to better reproduce some critical features of future sea-ice cover, and to obtain a more realistic frequency distribution. It should be noted that in the perfect model test as well as in the real-case application, the original AOGCM is not present among the possible analog candidates. If this is done, the results are even better using this method.

The fact that the analog method and the relative anomaly method share the same ~~bias~~ errors in the real-case application with a strong overestimation of the decrease in mean and variability of the sea-ice in the Weddell Sea in particularly for the scenarios of the CNRM-CM5 model is not a coincidence. For both methods, the targeted future SIE (or SIA) for a given sector is a product of the division of the integrated SIE (SIA) in the AOGCM ~~future~~ scenario by the corresponding quantity in the historical simulation. As a consequence, the targeted projected SIE (SIA) for a given sector and a given month is null when the integrated SIE (SIA) is null in the future AOGCM scenario. Therefore, the bias in the ~~future~~ scenario is not corrected in that case. The fact that both methods overestimate the decrease in sea ice mainly for CNRM-CM5 scenarios is to be linked to the fact that the historical simulation of this AOGCM shows some considerable negative biases for the sea-ice in the Weddell Sea with respect to the observations. Consequently, SIC in the Weddell Sea in CNRM-CM5 ~~future~~ rcp8.5 scenario is low and the number of months with a complete disappearance of sea ice is large. For these months, SIC in these sectors is not bias-corrected with the latter two methods. This means that although the methods described here are in principle applicable to any AOGCM output, it seems to be wise to ~~preferentially select output of reasonably “well-behaving” AOGCMs~~ exclude AOGCMs with large negative bias on sea-ice in their historical simulation as initial material for the bias-correction.

4.3 ~~A note on sea-ice thickness~~Sea-ice Thickness

~~Air-sea fluxes in the presence of sea ice arestrongly influenced by the thickness of the sea ice and the overlying snow cover. Gerdes (2006) and Krinner et al. (2010) have shown that the atmospheric response to changes in Arctic sea-ice thickness is substantial. In most AGCMs, sea-ice thickness will also need to be prescribed along with sea-surface temperature and~~ Given the simplicity of the proposed diagnostic SIT parameterization, the results are, at least in some aspects such as the predicted average Arctic sea-ice ~~concentration. When SST and SICfrom a coupled climate model are directly used, sea-ice thickness from that same run should of course be used; however, in case SST and SIC from~~ thinning, surprisingly good. The Central Arctic SITs results are clearly adversely affected by the input SICs North of 82°N. Arctic winter SIT in the marginal seas appears underestimated. In the Antarctic, the spatial pattern of SIT is very well represented.

We think that in the absence of pan-Arctic and pan-Antarctic satellite-based data before approximately 2000, this parameterization can serve as a surrogate, and that it can, because it seems to have predictive power, also serve for climate change experiments with AGCMs or RCMs. Because of its simplicity, implementing this parameterization should not be too complicated in any case provided the model does explicitly take into account SIT in its computations of heat flow through sea ice. In that case, SIT can either be calculated online (with the need to keep track of annual minimum SIC during the execution of the code) or be input as a daily boundary condition along with the SIC.

Of course, another possibility would be to prescribe SIT anomalies from coupled models. In this case, it would probably be wise to compute the prescribe SIT using its relative thickness changes. For example, in a climate change experiment, this would read $h_{presc}(t) = h_{obs,2003-2008}\, h_{sim}(t)/h_{sim,2003-2008}$. Problems could of course occur in areas where the coupled model ~~run are bias-corrected, as we strongly suggest here, we argue that sea-ice thickness should be prescribed in a physically consistent manner in the atmosphere-only simulation. An in-detail evaluation of sea-ice thickness prescription methods is beyond the scope of the main part of this paper. Therefore, an evaluation and further refinement of a simple parameterization of~~ simulates no sea-ice ~~thickness~~ cover at present. A physically consistent diagnostic parameterization of SIT as a function of ~~instantaneous and annual minimum SIC, initially suggested by Krinner et al. (1997) and used by Krinner et al. (2010) , is presented in the supplementary material of this paper.~~ constructed SIC, as proposed here, would not suffer from such problems.

In any case, it is very probable that Arctic SIT will further decrease as multi-year sea ice will be replaced by a predominantly seasonal sea-ice cover. This should probably be taken into account in future modeling exercises similar to CORDEX or HighResMip, given the non-negligible impact of sea-ice thinning on winter heat fluxes in particular.

4.4 General considerations on bias correction of oceanic forcings

As already mentioned~~before~~, one may doubt whether it is possible to bias-correct a GCM that has overly ~~strong~~ large biases in present-day climate. Indeed, most of the bias-correction methods rely on the hypothesis than the climate change signal coming ~~of~~ from an AOGCM scenario is not dependent on the bias in the historical simulations. This hypothesis can largely be questioned in a non-linear system (formed by SIC and SST). For example, in a model with a ~~strong~~ large negative bias in sea-ice for present-day climate, most of the additional energy due to an enhanced greenhouse effect will be used to heat the ocean, while it would be primarily used to melt sea-ice in a model with a correct initial sea-ice state. For such a model, the reliability of the climate change signal in SST is thus necessarily questionable. The selection of climate models based on their credibility for climate change scenario is a complex issue (Brekke et al., 2008; Baumberger et al., 2017, e.,g.) , dependent on the purposes, the processes and the region of study. Whether the climate change signal should be corrected remains on open question (Ehret et al., 2012) , even though there are good reasons to believe that model biases are time invariant (Maurer et al., 2013) .

Skills of coupled GCMs in reproducing the observed climate and its variability for a region of interest are often evaluated in order to use the GCM output as forcing for downscaling experiments. However, skills of atmospheric ~~GCM~~ GCMs are generally better when forced by observed oceanic boundary conditions ~~(Krinner et al., 2008; Ashfaq et al., 2011; Hernández-Díaz et al., 2017)~~ (Krinn

Similarly, even though bias correction methods have some limitations, for future climate experiments, there are good reasons to believe that simulations produced using bias-corrected oceanic forcings bear reduced uncertainties with respect to simulations realized with “raw” oceanic forcings from coupled model scenarios such as those from the CMIP5 experiments.

Bias-corrected oceanic forcings can be used to force a regional climate model (RCM), but in this case an additional modelling step has to be carried out, as bias-corrected oceanic forcings should be used to force an atmosphere only GCM that will provide atmospheric lateral boundary conditions for the RCM in order to ensure the consistency between oceanic and atmospheric forcings, such as in Hernández-Díaz et al. (2017). In this framework, the use of a variable resolution GCM which allows to directly use bias-corrected oceanic forcings and downscale ~~future climate experiments~~ climate scenarios is an alternative worth considering, as it also allows two-way interactions between the downscaled regions and the general atmospheric circulation.

5 Conclusions

In this paper, we reviewed existing methods for bias correction of SST and SIC and proposed new ones, such as the analog method for sea-ice. We also proposed validation methods that allow objectively evaluating bias-correction methods with the use of a perfect model test and real-case applications.

The bias-correction of SST is an issue that has already been widely addressed in recent papers and its importance for the modeling and downscaling of future climate scenarios has been demonstrated for multiple regions of the world. In our analysis, we were able to demonstrate the reliability and the suitability of absolute anomaly and quantile-quantile methods for the bias correction of future SST scenarios.

The bias correction of SIC is a more difficult issue to address. With the analog method, we propose a method that shows promising results in most cases and that allows reconstructing future SIC with a realistic frequency distribution in the future. However, the fact that the relative anomaly between an AOGCM ~~future scenario and the~~ scenario and its historical simulation is also used in this method in order to determine future targeted sea-ice extent and area, prevent from bias-correcting cases where sea-ice disappears entirely in a given sector or even an hemisphere. Despite the absence of a perfect and definite answer to this issue, we propose a new and improved method as well as a convenient, objective way to evaluate bias correction methods for ~~future~~ climate scenarios. ~~We draw the attention on the bias-correction of sea-ice that~~ The bias correction of sea ice is currently somewhat overlooked by the community. The application of a multivariate bias correction method (Cannon, 2016) is also a perspective that could help with the bias correction of SST and SIC ~~future~~ projected scenarios at the same time. Nevertheless, corrected SIC using the analog method ~~represent~~ represents a substantial improvement with respect to other previously existing bias-correction methods for sea-ice scenarios and will therefore be made available to anyone willing to use them as forcing for bias-corrected downscaling experiments.

Code and data availability. *FORTRAN* code enabling the generation of bias-corrected future SST and SIC using CMIP5 scenarios and PCMDI data as input are publicly available for each method via *https* transfer (https://mycore.core-cloud.net/index.php/s/3Lo3Tlr9wsyUGjk) or *ftp* transfer (ftp://ftp.lthe.fr/pub/beaumet/Sourcecode_SSTSICmethods.tar.gz). Bias-corrected future CMIP5 scenarios (rcp4.5 and 8.5) realized within the frame of this study (IPSL-CM5A-LR and CNRM-CM5) are available as well (https://mycore.core-cloud.net/index.php/s/Q1cIsS71Mo4vC or ftp://ftp.lthe.fr/pub/beaumet/Data_BCSST-SIC.tar.gz).

Appendix A: A ~~simple diagnostic parameterization of sea-ice thickness for AGCM simulations~~Bias correction methods : Sea Surface Temperatures

A1 ~~Introduction - general remarks~~

~~Atmospheric circulation models (AGCMs or regional climate models) require information about the state of~~

A0.1 Anomaly method

This frequently used method (e.g., Krinner et al., 2008) simply consists of adding the SST anomaly coming from the difference between a coupled AOGCM projection and the ~~sea surface as a lower boundary condition. While much attention has been paid to sea-surface temperature (SST) and sea-ice concentration (SIC) in that respect, the issue of prescribing correct (or at least reasonable) sea-ice thickness (SIT) has been somewhat neglected historically. While there is a considerable body of scientific literature on the effect of varying SST and SIC on simulated climate, only very few studies focused on the role of varying SIT in atmosphere-only simulations. The authors are aware of three such studies (Gerdes, 2006; Krinner et al., 2010; Semmler et al., 2016) . Gerdes (2006) concluded that "realistic sea ice thickness changes can induce atmospheric signals that are of similar magnitude as those due to changes in sea ice cover", while Krinner et al. (2010) show that the impact of a variable sea-ice thickness compared to a uniform value is essentially limited to the cold seasons and the lower troposphere, and that sea-ice thickness changes have a significant impact also in the context of climate change simulations. Near-surface temperature changes of the order of a few °C are observed in response to the replacement of a uniform thick Arctic sea-ice cover by variable sea-ice thickness. In this note, a simple diagnostic parameterization initially developed by Krinner et al. (1997) is discussed and evaluated against new Arctic and Antarctic sea-ice thickness data that were not available in the mid-90s.The idea is to propose a simple parameterization of sea-ice thickness that can be used in a variety of climate modelling applications, in particular for AGCM or RCM simulations of climate conditions different than today, from palaeoclimate studies to climate projections.In these applications, this parameterization can be particularly useful in cases where future sea-surface conditions (SST, SIC and SIT)are not directly prescribed from a coupled ESM run, but rather obtained using a bias correction method.~~

A1 ~~Methods~~

A0.1 ~~Diagnosing sea-ice thickness from sea-ice concentration~~

~~As described by Krinner et al. (2010) , the parameterization of sea-ice thickness~~ $\cancel{h_S}$ ~~as a function of the local instantaneous sea-ice fraction f is designed such as to yield~~ $\cancel{h_S}$ ~~of the order of 3 meters for multi-year sea ice (deemed to be dominant when the local annual minimum fraction~~ $\cancel{f_{min} \gg 0}$~~) and~~ $\cancel{h_S}$ ~~below 60cm (with a stronger annual cycle) in regions where sea-ice completely disappears in summer (that is,~~ $\cancel{f_{\mathrm{min}} = 0}$~~), and intermediate values for intermediate cases:~~

$$\cancel{h_S = (c_1 + c_2 f_{min}^2) \cdot (1 + c_3 (f - f_{min}))}$$

~~with c_1=0.2m, c_2=2.8m and c_3=2. This corresponds to the observed characteristics of Arctic and Antarctic sea ice, with multi-year sea ice being generally much thicker than first year ice. The parameter c_3 introduces a seasonal ice thickness variation in areas where there is a concomitant seasonal cycle of sea-ice concentration. A more parsimonious, simply bilinear formulation could have been designed to comply with these constraints. However, for the sake of consistency with previous work, we used the equation proposed by Krinner et al. (1997) who designed the parameterization such as to allow for a fairly strong seasonal cycle of sea-ice thickness also in regions with intermediate values of f_{min}. Figure 3 (from Krinner et al. (2010)) illustrates diagnosed Arctic sea-ice thickness for the present and for the end of the 21st century (2081-2100) using bias-correction applied to sea-ice concentrations from a coupled ESM SRES-A1B simulation (Krinner et al., 2008) .~~

~~Prescribed annual mean Arctic sea-ice thickness (in m) in an AGCM climate change experience with bias-corrected sea-surface conditions, using the proposed diagnostic parameterization (figure from Krinner et al. (2010)). Left: present (1981-2000), right: SRES-A1B for 2081-2100. Bias correction for SST and SIC after Krinner et al. (2008)~~ corresponding historical simulation to the present-day observations. In practice, for each grid point, the difference between the SST for a given month in the future from a climate change simulation and the climatological mean SST in the corresponding historical simulation from the same coupled AOGCM is added to the observed climatological mean SST (e.g . PCMDI, 1971-2000):

$$SST_{Fut,est} = \overline{SST_{obs}} + \left(SST_{Fut,AOGCM} - \overline{SST_{Hist,AOGCM}}\right) \tag{A1}$$

In (A1), $SST_{Fut,est}$ is the estimated future SST for a given month, $\overline{SST_{obs}}$ the observed climatological monthly mean, $SST_{Fut,AOGCM}$ the model future SST for a given month in the future AOGCM scenario and $\overline{SST_{Hist,AOGCM}}$ the model climatological monthly mean in the AOGCM historical simulation for the same reference period as for the observed climatology. As a result, the reconstructed SST time series has the chronology of the AOGCM projected scenario.

A0.1 ~~Data~~

~~In the following, we used sea-ice concentration data extracted from the ERA-Interim output; this is typically the kind of data that would be used in AGCM or RCM simulations.Lindsay and Schweiger (2015) recently proposed a 15-parameter spatial and temporal regression of Arctic sea-ice thickness observations from submarines, aircraft and satellites. We will use these observations here. Kurtz and Markus (2012) have deduced Antarctic SIT from ICESat data for the period 2003-2008. Although observations with autonomous underwater vehicles by Williams et al. (2015) tend to suggest occurrence of thicker Antarctic sea-ice than previously acknowledged, we will use the Kurtz and Markus (2012) data because of their large spatial coverage.~~

A1 ~~Results~~

~~The original formulation by Krinner et al. (1997) was parameterized for both hemispheres. We will therefore first present results for the original unique parameter set $c_{1,2,3}$ applied to both hemispheres. In a second step, we will present results for separate Arctic and Antarctic parameter sets, yielding a better fit to the observations. The reasoning is that, at the expense of generality of the diagnostic parameterization, one could argue that the strong difference between the Arctic and Antarctic geographic configuration — a closed small ocean favouring ice ridging and thus thicker sea ice in the Arctic, and large open~~

~~ocean favouring thinner sea ice around Antarctica — justifies choosing different parameter sets for the two hemispheres. As the position of the continents will not change over the time scales of interest here, climate change experiments will not be adversely affected by this loss of generality.~~

A0.1 ~~Option 1: Global parameter set~~Quantile-quantile method

~~A comparison between the observed (Lindsay and Schweiger, 2015) and our diagnosed evolution of the Arctic mean sea-ice thickness is given in Figure 11. The geographical patterns of the observed (in fact, observation-regressed) and parameterized Arctic ice thickness for March and September over the observation period 2000-2013 (Figure 12) do bear some resemblance, but they also show some clear deficiencies of the diagnostic parameterization. The diagnostic parameterization reproduces high sea-ice thickness north of Greenland and the Canadian Archipelago, linked to persistent strong ice cover, but underestimates maximum ice thickness (due in part to compression caused by the ocean surface current configuration). Thinner sea ice over the seasonally ice-free parts of the basin is reproduced, but it is actually too thin, particularly in winter (for example in the Chukchi Sea). Obvious artifacts appear in September north of about 82°N where the SIC in the ERA-Interim data set clearly bears the signs of limitations due to the absence of satellite data. Both for spring (Oct-Nov) and fall (May-Jun), our diagnosed SIT (Figure 13) compares generally well with the ICESat data except for an overestimate in the Weddell Sea, at both seasons. The geographical pattern of alternating regions with thin and thick sea ice is remarkably well reproduced.~~

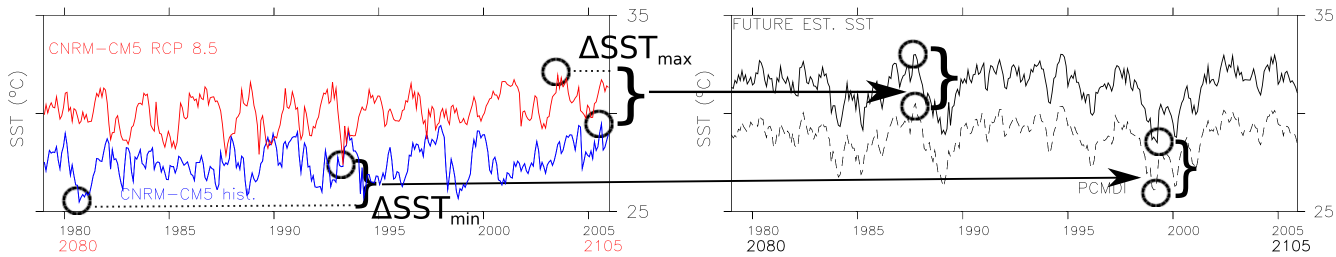
~~Observed (black, after Lindsay and Schweiger, 2015) and diagnosed (red) 12-month moving average mean sea-ice thickness of the Arctic basin (see Figure 12). The global parameter set is used here. Slight differences to Figure 13 of Lindsay and Schweiger (2015) ap because here we mask ice-free (SIC < 15%) areas that have a finite, non-zero ice thickness in the regression proposed by Lindsay and Schweiger (2015) who extend their regression to the entire Arctic Basin at all seasons.~~

~~Observed (regressed, Lindsay and Schweiger (2015)) and parameterized Arctic sea-ice thickness (in m) for March and September, and difference between these (right), with the global parameter set.~~

~~Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for March and September, and difference between these (right), with the global parameter set.~~

A0.2 ~~Option 2: Separate Arctic and Antarctic parameter sets~~

~~A slightly better fit for the two poles can be obtained with separate parameters sets. For the Arctic, it seems desirable to increase winter sea-ice thickness in the Chukchi Sea area (by increasing c_3 slightly) and to decrease the average sea-ice thickness over the Central Arctic (by decreasing c_2). Figures 14 and 15 show results for the Arctic with c_1=0.2m, c_2=2.4m and c_3=3. The spatial fit is slightly better, but~~ This method has been proposed and described in Ashfaq et al. (2011) It consists of adding, for each grid point and each calendar month's quantile in the observations, the ~~recent Arctic-mean decadal tendency towards decreased average sea-ice thickness is somewhat less well reproduced. For the Antarctic, the main feature to improve is the maximum ice thickness in the Weddell Sea, which can be decreased by decreasing c_2 to 2.0m. The Antarctic parameter set then becomes c_1=0.2m, c_2=2m~~ corresponding quantile change in the GCM data set, i.e. the difference between the maximum SST in the projected scenario and in the historical simulation, between the second highest SSTs in the two simulations, and



~~Observed Lindsay and Schweiger (2015) and parameterized Arctic sea-ice thickness (in m) for March and September, and difference between these (right), with the Arctic-specific parameter set.~~

Figure A1. Illustration of the quantile-quantile method for min. and max. of SST time series for a grid point in the Central Pacific : GCM historical simulation (blue, left), GCM projected scenario (red, left), observed SST(dashed, right), reconstructed future SST (thick, right)

~~c_3=2. The result~~ so on for each ranked SST quantile. However, unlike Ashfaq et al. (2011) , we did not create a new SST field for the present by replacing SST from the GCM in the historical period by its corresponding quantile in the observations, but we directly added the quantile change to the corresponding quantile of the observational time series (Figure ~~16)is indeed a decreased thickness of the perennial Weddell Sea ice with little impact elsewhere. In any case, these hemisphere-specific sea-ice parameter sets are not very different from each other and fairly similar to the original formulation.~~

~~Observed (black, after Lindsay and Schweiger (2015)) and diagnosed (red) 12-month moving average mean sea-ice thickness of the Arctic basin with the Arctic-specific parameter set.~~ A1). This conserves the chronology of the observations and their inter-annual variability in estimated SSTs for the future. In our results, we noticed a large fine-scale spatial variability of the constructed bias-corrected SSTs that was due to the large spatial variability of the climate change increments (quantile change) calculated individually for each pixel. To fix this, we applied a slight spatial filtering (3 grid point Hann box filter (Blackman and J.W., 1959)) of the quantile shifts in order to produce more consistent SST fields.

~~Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for March and September, and difference between these (right), with the Antarctic-specific parameter set.~~

A1 ~~Discussion and conclusion~~

~~Given the simplicity of the proposed diagnostic sea-ice thickness parameterization, the results are, at least in some aspects such as the predicted average Arctic sea-ice thinning, surprisingly good. The Central Arctic sea-ice thickness results are clearly adversely affected by the input sea-ice concentrations north of 82°N. Arctic winter sea-ice thickness in the marginal seas appears underestimated. In the Antarctic, the spatial pattern of SIT is very well represented. We think that in absence of pan-Arctic and pan-Antarctic satellite-based data before approximately 2000, this parameterization can serve as a surrogate for earlier periods, and that it can, because it seems to have predictive power, also serve for climate change experiments with AGCMs or RCMs. Because of its simplicity, implementing this parameterization should not be too complicated in any case provided the model does explicitly take into account sea-ice thickness in its computations of heat flow through sea ice. In that~~

~~case, sea-ice thickness can either be calculated online (with the need to keep track of annual minimum sea-ice thickness during the execution of the code) or be input as a daily boundary condition along with the sea-ice concentrations. Of course, another possibility would be to prescribe sea-ice thickness anomalies from coupled models. In this case, it would probably be wise to compute the prescribe SIT using relative sea-ice thickness changes. For example, in a climate change experiment, this would read~~ $\cancel{h_{presc}(t) = h_{obs,\,2003-2008} \cdot h_{sim}(t)\,/h_{sim,\,2003-2008}}$~~. In any case, it is very probable that Arctic sea ice thickness will further decrease as multi-year sea ice will be replaced by a predominantly seasonal sea-ice cover. This should probably be taken into account in future CORDEX- or HighResMip-style climate simulations, given the non-negligible impact of sea-ice thinning on winter heat fluxes in particular.~~

Competing interests. The authors have no competing interests.

Acknowledgements. This study was funded by the Agence Nationale de la Recherche through contract ANR-14-CE01-0001-01 (ASUMA). 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 participating to CMIP5 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.
R. Haarsma was funded through the PRIMAVERA project under Grant Agreement 641727 in the European Commission's Horizon 2020 research program.

References

Agosta, C., Favier, V., Krinner, G., Gallée, H., Fettweis, X., and Genthon, C.: High-resolution modelling of the Antarctic surface mass balance, application for the twentieth, twenty first and twenty second centuries, Climate Dynamics, 41, 3247–3260, https://doi.org/10.1007/s00382-013-1903-9, 2013.

Ashfaq, M., Skinner, C. B., and Diffenbaugh, N. S.: Influence of SST biases on future climate change projections, Climate Dynamics, 36, 1303–1319, https://doi.org/10.1007/s00382-010-0875-2, 2011.

Baumberger, C., Knutti, R., and Hirsch Hadorn, G.: Building confidence in climate model projections: an analysis of inferences from fit, Wiley Interdisciplinary Reviews: Climate Change, 8, e454–n/a, https://doi.org/10.1002/wcc.454, 2017.

Blackman, R. and J.W., T.: Particular pairs of windows, in: The measurement of power spectra, from the point of view of communications engineering., edited by Blackman, R. and J.W., T., chap. B5, pp. 95–100, Dover Publications, inc., New York, 1959.

Brekke, L. D., Dettinger, M. D., Maurer, E. P., and Anderson, M.: Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments, Climatic Change, 89, 371–394, https://doi.org/10.1007/s10584-007-9388-3, 2008.

Cannon, A. J.: Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure, Journal of Climate, 29, 7045–7064, 2016.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., et al.: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Sea level change, pp. 1137–1216, 2013.

de Elía, R., Laprise, R., and Denis, B.: Forecasting Skill Limits of Nested, Limited-Area Models: A Perfect-Model Approach, Monthly Weather Review, 130, 2006–2023, https://doi.org/10.1175/1520-0493(2002)130<2006:FSLONL>2.0.CO;2, 2006.

Denis, B., Laprise, R., Caya, D., and Côté, J.: Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment, Climate Dynamics, 18, 627–646, https://doi.org/10.1007/s00382-001-0201-0, 2002.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W. J., Cox, P., Driouech, F., Emori, S., Eyring, V., et al.: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2013, 5, 741–866, 2013.

Gerdes, R.: Atmospheric response to changes in Arctic sea ice thickness, Geophysical Research Letters, 33, n/a–n/a, https://doi.org/10.1029/2006GL027146, 2006.

Giorgi, F. and Gutowski, W. J.: Coordinated Experiments for Projections of Regional Climate Change, Current Climate Change Reports, 2, 202–210, https://doi.org/10.1007/s40641-016-0046-6, 2016.

Haarsma, R. J., Hazeleger, W., Severijns, C., de Vries, H., Sterl, A., Bintanja, R., van Oldenborgh, G. J., and van den Brink, H. W.: More hurricanes to hit western Europe due to global warming, Geophysical Research Letters, 40, 1783–1788, https://doi.org/10.1002/grl.50360, 2013.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fuckar, N. S., Guemas, V., et al.: High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6, Geoscientific Model Development, 9, 4185, 2016.

Haralick, R. M., Sternberg, S. R., and Zhuang, X.: Image Analysis Using Mathematical Morphology, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9, 532–550, https://doi.org/10.1109/TPAMI.1987.4767941, 1987.

Hawkins, E., Robson, J., Sutton, R., Smith, D., and Keenlyside, N.: Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach, Climate Dynamics, 37, 2495–2509, https://doi.org/10.1007/s00382-011-1023-3, 2011.

Hernández-Díaz, L., Laprise, R., Nikiéma, O., and Winger, K.: 3-Step dynamical downscaling with empirical correction of sea-surface conditions: application to a CORDEX Africa simulation, Climate Dynamics, 48, 2215–2233, https://doi.org/10.1007/s00382-016-3201-9, 2017.

Hewitson, B. C., Daron, J., Crane, R. G., Zermoglio, M. F., and Jack, C.: Interrogating empirical-statistical downscaling, Climatic Change, 122, 539–554, https://doi.org/10.1007/s10584-013-1021-z, 2014.

Holland, G., Done, J., Bruyere, C., Cooper, C., and Suzuki, A.: Model Investigations of the Effects of Climate Variability and Change on Future Gulf of Mexico Tropical Cyclone Activity, in: Proceeding of the Offshore Technology Conference, Houston, Texas, https://doi.org/10.4043/20690-MS, 2010.

Krinner, G., Genthon, C., Li, Z.-X., and Le Van, P.: Studies of the Antarctic climate with a stretched-grid general circulation model, Journal of Geophysical Research: Atmospheres, 102, 13 731–13 745, https://doi.org/10.1029/96JD03356, http://dx.doi.org/10.1029/96JD03356, 1997.

Krinner, G., Guicherd, B., Ox, K., Genthon, C., and Magand, O.: Influence of Oceanic Boundary Conditions in Simulations of Antarctic Climate and Surface Mass Balance Change during the Coming Century, Journal of Climate, 21, 938–962, https://doi.org/10.1175/2007JCLI1690.1, 2008.

Krinner, G., Rinke, A., Dethloff, K., and Gorodetskaya, I. V.: Impact of prescribed Arctic sea ice thickness in simulations of the present and future climate, Climate Dynamics, 35, 619–633, https://doi.org/10.1007/s00382-009-0587-7, 2010.

Krinner, G., Largeron, C., Ménégoz, M., Agosta, C., and Brutel-Vuilmet, C.: Oceanic Forcing of Antarctic Climate Change: A Study Using a Stretched-Grid Atmospheric General Circulation Model, Journal of Climate, 27, 5786–5800, https://doi.org/10.1175/JCLI-D-13-00367.1, 2014.

Kurtz, N. and Markus, T.: Satellite observations of Antarctic sea ice thickness and volume, Journal of Geophysical Research: Oceans, 117, 2012.

Lenaerts, J. T. M., Vizcaino, M., Fyke, J., van Kampenhout, L., and van den Broeke, M. R.: Present-day and future Antarctic ice sheet climate and surface mass balance in the Community Earth System Model, Climate Dynamics, 47, 1367–1381, https://doi.org/10.1007/s00382-015-2907-4, 2016.

Levine, R. C., Turner, A. G., Marathayil, D., and Martin, G. M.: The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall, Climate Dynamics, 41, 155–172, https://doi.org/10.1007/s00382-012-1656-x, 2013.

Li, G. and Xie, S.-P.: Tropical Biases in CMIP5 Multimodel Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems, Journal of Climate, 27, 1765–1780, https://doi.org/10.1175/JCLI-D-13-00337.1, 2014.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, 2015.

Maurer, E. P., Das, T., and Cayan, D. R.: Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction, Hydrology and Earth System Sciences, 17, 2147–2159, https://doi.org/10.5194/hess-17-2147-2013, 2013.

Mitchell, D., Achutarao, K., Allen, M., Bethke, I., Beyerle, U., Ciavarella, A., Forster, P., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, N., Massey, N., Fischer, E., Schleussner, C.-F., Scinocca, J., Seland, A., and Rashyd, Z.: Half a degree additional warming, prognosis and projected impacts (HAPPI): Background and experimental design, Geoscientific Model Development, 10, 571–583, 2017.

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., et al.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747, 2010.

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, https://doi.org/10.5194/tc-8-1871-2014, 2014.

Richter, I., Xie, S.-P., Behera, S. K., Doi, T., and Masumoto, Y.: Equatorial Atlantic variability and its relation to mean state biases in CMIP5, Climate Dynamics, 42, 171–188, https://doi.org/10.1007/s00382-012-1624-5, 2014.

Sato, T., Kimura, F., and Kitoh, A.: Projection of global warming onto regional precipitation over Mongolia using a regional climate model, Journal of Hydrology, 333, 144 – 154, https://doi.org/10.1016/j.jhydrol.2006.07.023, 2007.

Screen, J. A. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, Nature, 464, 1334, 2010.

Semmler, T., Jung, T., and Serrar, S.: Fast atmospheric response to a sudden thinning of Arctic sea ice, Climate Dynamics, 46, 1015–1025, 2016.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophysical Research Letters, 39, n/a–n/a, https://doi.org/10.1029/2012GL052676, 2012.

Taylor, K., Williamson, D., and Zwiers, F.: The sea surface temperature and sea-ice concentration boundary condition for AMIP II simulations, PCMDI Rep. 60, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, 25 pp., 2000.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, 2012.

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, Journal of Climate, 26, 1473–1484, https://doi.org/10.1175/JCLI-D-12-00068.1, 2013.

Williams, G., Maksym, T., Wilkinson, J., Kunz, C., Murphy, C., Kimball, P., and Singh, H.: Thick and deformed Antarctic sea ice mapped with autonomous underwater vehicles, Nature Geoscience, 8, 61–67, 2015.

Zhang, L. and Zhao, C.: Processes and mechanisms for the model SST biases in the North Atlantic and North Pacific: A link with the Atlantic meridional overturning circulation, Journal of Advances in Modeling Earth Systems, 7, 739–758, https://doi.org/10.1002/2014MS000415, 2015.

Interactive comment on "Assessing bias-corrections of oceanic surface conditions for atmospheric models" by Julien Beaumet et al.

F. Gallo (Referee)
florian.gallo@metoffice.gov.uk
Received and published: 10 January 2018

The authors thank the referee for accepting to review the paper and for the generally constructive remarks aiming at the improvement of the quality of the paper. The responses to the different comments are below :

****

General comments This paper proposes a way to evaluate bias-correction methods for SST and SIC for future climate projections, using a perfect model approach and a real-case application. There has clearly been a large amount of work in this study and this is clear when reading the paper. The analysis is thorough and the discussion hon-est, with the main caveats being highlighted and explained (at least, an explanation is proposed). The conclusion is clear and includes potential other methods to investigate. However, the presentation of methods and results might be a little bit confusing, given the amount of data.

1. Some extra introductory sentences explaining the point of using a perfect model approach would be welcome, as it might not be obvious to a reader that is not a specialist but wants to learn more. This could be done in section 2.4, where the description of the evaluation method (which is a main point of the paper) is a bit short. More generally, it would be interesting to provide some examples of the use of a perfect model approach in the literature.

Authors' response: The reviewer is right to state that some reader might not be familiar with the perfect model approach. We therefore added the following sentences that should explain the approach in a nutshell and refer to some examples: "A perfect model approach usually consists of using model data as a substitute for observations, and trying to predict projected model data from that model; this prediction can then be evaluated against the available model projections (e.g., Hawkins et al., 2011). In the real world, as observations of future climate are obviously not yet available, an equivalent approach is impossible if one cannot wait long enough for the future to become reality. Another type of perfect model approach are "Big Brother" experiments for evaluating downscaling techniques. In such studies, high-resolution model output is degraded in resolution and downscaling methods are then applied to these low-resolution data. The resulting synthetic high-resolution fields are then compared to the original high-resolution output (e.g. de Elía et al., 2002). Here, we consider SST and SIC…"

2. If the language is usually clear and understandable, the wording can be unusual and the authors are encouraged to have (another?) correction by a native speaker (which the reviewer is not. . .)

Authors' response: We had a native speaker correct the revised version.

3. On a more specific point, one might wonder why were the GCMs CNRM-CM5 and IPSL-CM5A-LR chosen? Was this a choice based on availability or were these models selected based on their respective performance for representing SST and SIC? It would be nice to have information on this point.

Authors response : "The search for suitable bias-corrections methods and their use was first motivated by the need to drive future scenarios climate experiments with atmosphere-only GCMs ARPEGE and LMDZ. Therefore, the work was started with SST and SIC coming from the *corresponding coupled model of these two atmosphere-only GCMs. HadGEM-ES was added later in order to verify if the results obtained were reproduced with this model, but we acknowledge that no criterion based on their respective performances has been used to select these models rather than another one. However, the fact that the results are very close for the three models investigated gives us confidence in the fact that they are robust and independent from the AOGCM chosen as initial material.*

Finally some caveats and issues are treated too lightly and would require a more thorough description and explanation (see specific comments). Overall, the proposed paper describes an interesting and detailed work that should be of interest to many users in the climate modelling community. I therefore propose this manuscript to be accepted after the minor changes described in this review document.

Specific comments.

Specific comments

P1 L12: the part about RCPs is not needed, isn't it? The sentence is a bit long
Authors reponse : « Comment taken into account : the RCP acronym and the corresponding reference is now introduced in the Data section : "Only the first ensemble members of the historical, and of the Representative Concentration Pathways (RCPs; \citet{Moss2010}) 4.5 and 8.5 simulations have been considered"».

L17 bracket missing somewhere

Authors response : « Ok, comment taken into account. »

L19: Would it be possible to have some other examples from the literature? Surely a list of 4 or 5 references should be easy to find

Authors responses : "Some additional references were add as examples of the considerable literature on the bias of CMIP5 models, especially on SST. References demonstrating the added value of bias-corrected SST have been included as well : "The absence of the Pacific cold tongue bias and the reduction of the double ITCZ problem in AMIP experiments with respect to the CMIP5 model experiments \citep{Li2014} shows the importance of forcing atmospheric model by SST close to the observations. For instance, improvements in the modelling of the tropical cyclone activity in the Gulf of Mexico \citep{Holland2010} and of summer precipitation in Mongolia \citep{Sato2007} were obtained by bias-correcting SST and other AOGCM outputs before using them as forcing for RCMs."

L19 "For example, it has. . ."
Authors response : « Ok, comment taken into account. »

L20 the seasonal cycle and the trend
Authors response : « Ok, comment taken into account. »

P2 L24 describe "AMIP" as "Atmospheric Model Intercomparison Project" if it is not done any-where else
Authors response : « Ok, comment taken into account. »

P3 L25 Is there a reference for the Hann box filter? Why did you choose this filter?

Authors response : "The first reference to Hann function is in "Particular pairs of windows" in "The measurement of power spectra, from the point of view of communications engineering" by R.B. Blackman and J. Tukey, 1959. We have chosen Haan filter because it is the lightest filter amongst the commonly used box filter."

P5 L13 Any information on the number / proportion of GCMs that were
dismissed? What does "poorly" mean for the selection process?

Authors response : « We built our library by selecting AOGCMs who have a reasonable representation of the sea-ice extent annual cycle, its maximum and minimum, in present climate following the literature (e.g. Turner et al., 2013, Stroeve et al., 2012). For instance, our list for the « real-case » application of the method contains historical simulation and future scenarios of the following AOGCMs : MIROC-ESM, EC-EARTH, NorESM1-M, CCSM4 and IPSL-CM5A. »

L13 AOGCMs and remove"overly"

Authors response : « Ok, comment taken into account. »

P7 L12 "We assume that an ideal bias correction method should reproduce the same change in mean and variance between the observations and the estimated future SST and SIC as between the used coupled GCM historical simulation and the climate change experiment." That seems obvious but is there any reference regarding this issue? Is there any discussion among the scientific community?

Authors response : "There is indeed debate about this issue and so far, probably no consensus. For the bias-correction of future scenarios, one usually makes the hypothesis that one can rely on the climate signal coming from a model even if this model has bias in the reproduction of present climate. There are indeed reasons to believe that model biases are time invariant (e.g. Maurer et al. ,2013 (www.hydrol-earth-syst-sci.net/17/2147/2013) although whether we should correct the climate change signal remains an open question (see Ehret et al, 2012 (https://www.hydrol-earth-syst-sci.net/16/3391/2012/)).

L23 What is the point of applying the perfect model approach for SST, as we use only "regular" bias correction? You highlight this issue, but you might want to shrink this section a bit.

Authors response : "Indeed, the part on bias correction of SST is less novel. Following this remark and remarks of the second reviewer, the part on the methodology for the bias-correction of SST as well as the part on the perfect model test have been shrunk."

P8 Fig4 Are you sure about the color? There seems to be a very large initial bias between the obs and the historical simulation for North Atlantic, is that expected? Moreover the RCP4.5 looks quite cold compared to the corrected values. If this is correct, can you highlight and explain that in the text?

Author response : « The colors are right. The North Atlantic is a region where coupled GCMs often exhibit large biases (usually cold biases) because of their poor skills in representing correctly the Atlantic Meridional Overturning Circulation (AMOC). This example is indeed another argument for the bias correction of SSTs .»

L11 "methods"

Authors response : « Ok, comment taken into account. »

L12 delete "in"
Authors response : « Ok, comment taken into account. »

P9 L9 This comment is valid for the whole paper, but is the use of "biases" valid when describing the results of the perfect model experiment? It is a bit confusing with the original bias that we are trying to correct. Again, if it has been used previously in the literature in that context, I'm ok, but maybe "difference" or "error" would be clearer, as it is a bias created by the method, and not a bias originally in the data

Authors response : « Comment taken into account, the term "error" or mean "error" is now used throughout the text in order to make it less confusing.»

P10 L12 "more or less" – can we find a more scientific term please?
Authors response : « Ok, comment taken into account. »

L14 "is easy to explain" – Is it? Can you develop, please?
Authors response : « Ok, comment taken into account, some explanations are given in the text : "The presence of such peaks is easy to explain by taking into account the structure of the LUT as i) for a given month, the SIC does not always increase monotonically with decreasing SST, ii) the discrete nature of LUT is not in favour of a continuous SIC frequency distribution »

L29 Should an ideal method apply the same statistical changes? It sounds right, but what about skewed distribution (precipitation) where the BC would change the distribution, therefore changing the distribution of changes? I think there is quite a discussion about that topic, so, if I agree with you, I would change to "We consider here that an ideal method. . ."
Authors response : « Ok, the sentence has been modified following your recommendation.»

P11 All text – Would it be possible to have some correlation value in order to quantify the error among the different methods? Maybe a correlation coefficient, or the value of the minimum, maximum and mean error for each graph?

"*Authors response :"Mean errors and root mean square errors for each graph are added on the plots. In the text, we now discuss the average mean error or average RMSE for every scenarios and for the Arctic and Antarctic combined in order to quantify and compare more objectively the errors between the three methods."*

Fig8 and 9 It is difficult to see which point correspond to what – Maybe adding a letter to each of them to point to the region would help – Please try but it might make the figure impossible to read. It would be nice to be able to navigate alone within the points

Author response : "We changed the legend of the figure so that we can now distinguish the different regions with the help of different colors. Different signs (crosses and circles) are used to distinguish scenarios from CNRM-CM5 and IPSL-CM5A-LR. The more important for these figures is first to distinguish the regions, then the models. The distinction between rcp4.5 and rcp8.5 is less essential for the interpretation of the results and the connections with the text."

Interactive comment on "Assessing bias-corrections of oceanic surface conditions for atmospheric models" by Julien Beaumet et al.
Anonymous Referee #2
Received and published: 2 February 2018

The authors thank the referee for accepting to review the paper and for the generally constructive remarks aiming at the improvement of the quality of the paper. The responses to the different comments are below:

****

This manuscript discusses bias correction of sea-surface temperature using the anomaly method and the quantile-quantile method, and bias correction of sea ice concentration using the look-up table method, the iterative relative anomaly method, and the analog method. These bias correction methodologies are evaluated using a perfect model test (i.e. evaluated using the given model as "observations") and a real-case application in which the bias correction methods are compared to observations.
It is assumed that ideal bias correction will reproduce changes in the mean and variance between observations and projected climate as between historical simulations and projected climate. The authors determine that the presented methods for bias correcting SST are reliable. The methods presented for sea-ice concentration are less reliable, however, the analog method showed promising results and improvement over other bias correction methods. Additionally, the authors provide an appendix with a proposed method to parameterize sea-ice thickness, with potential for use in climate modeling applications.
I have a number of major and minor comments for the authors to address. Some of the manuscript was unclearly written, making the arguments difficult to follow. I also question the inclusion of SST bias correction evaluation. In regards to the review criteria, the manuscript does present relevant information that is related to modeling questions, particularly for sea-ice concentration, rendering it suitable for publication. However, much of the methodology for sea-surface temperature bias correction has been noted in other manuscripts.

My comments are below.
General Comments:
1. While the presented results for SIC are novel and will be very helpful for future modeling studies, the presented results for SST are somewhat less of an advancement. SST bias correction has been studied previously. In fact, there is much less discussion surrounding SST bias correction, and the results are almost glossed over by the authors in comparison. While the results are helpful in a summary sense for an interested reader, the concept seems less novel. This section may be able to be reduced even more, or eliminated completely.

Authors response : "We agree that this part of the paper is less novel and that these issues have been addressed in previous papers. Its presence in the manuscript is justified by the need to highlight the consistency with the work done for sea-ice, and the consistency between the response for the two variables, and to show the possibility to generalize the evaluation methods. However, in order to avoid redundant results, and emphasize the parts of the paper that are innovative, some parts of the result section were cut and the presentation of the methods have been mostly sent to the Appendix section".

2. The Appendix describes a methodology for parameterizing sea-ice thickness, which was noted in Section 4.3 as a strong influence. While you state that an in-detail evaluation of sea ice thickness

prescription is beyond the scope of this paper, you evaluate and further refine one of the methods for parameterization in the Appendix. This seems like an important contribution to the field that has been studied comparatively less than, for example, SST bias correction methodologies. I'm concerned that this contribution will be lost due to its presence in supplementary material, and would potentially warrant a separate manuscript that delves more deeply into the topic.

Authors response : « In some way, the work done on sea-ice thickness is not entirely innovative either, as it was already presented by one of the author (Krinner, 1997) and used in another study (Krinner, 2010). The innovation here is that the parameterization is further refined with parameters set for the Arctic and the Antarctic and that the results are objectively evaluated with sea-ice thickness measurements which so far were seldom, particularly in the Southern Ocean. However, we think that the current material on this topic is not sufficient to deserve a separate manuscript and it seems complicated to delves more deeply into the topic far enough to be able to produce a second manuscript. However, in order to avoid this contribution to be lost and in order to improve the manuscript consistency, we introduced the work on sea-ice thickness in the main part of the paper.

3. I am curious why the CNRM-CM5, IPSL-CM5A-LR, and HadGEM-ES coupled GCM data were explicitly chosen for this study. In addition, you note that HadGEM-ES was used in Section 2.1 near line 25, but never mention results from this model.

Authors response : "The search for suitable bias-corrections methods and their use was first motivated by the need to drive future scenarios climate experiments with atmosphere-only GCMs ARPEGE and LMDZ. Therefore, the work was started with SST and SIC coming from the corresponding coupled model of the latter two atmosphere-only GCMs. HadGEM-ES was added later, in order to verify if the results obtained were reproduced with this model, but we acknowledge that criterion based on model performances was used to select this model rather than another one. However, the fact that the results are very close for the three models investigated gives us some confidence in the fact that they are robust and independent from the AOGCM chosen as initial material. Results/figures with HadGEM-ES are not presented in order to limit the length of the paper, nevertheless some of the results with HadGEM-ES could be included in the appendix section for some transparency purposes."

4. I am also curious why you selected the given bias correction methods for SST and SIC, are these arguably the most popular methods in use? If so, it would be helpful to note this as a motivation for the work.

Authors response : "Absolute anomaly and quantile-quantile methods are likely amongst the most popular methods for bias-corrections, especially for SST. Absolute anomaly for SST and iterative relative anomaly method for SIC have been introduced and used by one of the co-author in a previous work (Krinner et al., 2008). Evaluating the Look-Up Table method was motivated by the fact that this method is, so far, the method recommended in the frame of the HighResMIP for the production of bias-corrected SIC boundary conditions for atmospheric models. In the light of our results, this should be changed in favour of the analog method in a near future. We developed and introduced the analog method as we weren't satisfied by the results for the bias correction of SIC with the first two methods.

5. Because you are using a perfect model test, can these results be generalized to other models,
or are these results specific to the models used?

Authors response : "We applied the perfect model test to rcp4.5 and rcp8.5 scenarios from three AOGCM of the CMIP5 experiments, and the results are very similar for each scenarios .From the *perfect model test perspective, the results are not dependent on the models used. Relative performances of the three bias correction methods in the "real-case" application also corroborate the results from the perfect model experiment which gives us confidence in the fact that results are essentially not model-dependent."*

6. The introduction could benefit from additional discussion on SST biases, as it is written the focus is on SIC biases.

Authors responses : "Some additional references were add as examples of the considerable literature on the bias of CMIP5 models, especially on SST. References demonstrating the added value of bias-corrected SST have been included as well in the introduction : "The absence of the Pacific cold tongue bias and the reduction of the double ITCZ problem in AMIP experiments with respect to the CMIP5 model experiments \citep{Li2014} shows the importance of forcing atmospheric model by SST close to the observations. For instance, improvements in the modelling of the tropical cyclone activity in the Gulf of Mexico \citep{Holland2010} and of summer precipitation in Mongolia \citep{Sato2007} were obtained by bias-correcting SST and other AOGCM outputs before using them as forcing for RCMs."

7. Figure 6 and resulting discussion: How does one determine what is a "reasonable" and "very small" error? To me, these look like large errors overall, but perhaps they are reasonable and very small with respect to the relative anomaly method?

Authors response : "The use of terms such as "reasonable" or "very small" has been reduced. Now, mean errors and root mean square errors for each graph were added on the plots. In the text, we now discuss the average mean error or average RMSE for every scenarios and for the Arctic and Antarctic combined in order to quantify and compare more objectively the errors between the three methods."

8. In Section 4.2, page 18, last sentence on the page: Preferentially selecting output of reasonably "well behaving" AOGCMs is perhaps too simplistically stated here. There are a variety of issues in selecting which models are "well behaving". Though the following reference focuses on selecting models for regional hydrological studies, some of the general comments will still hold true for model selection: Brekke LD, Dettinger MD, Maurer EP, Anderson M (2008) Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments. Clim Change 89:371–394 . doi: 10.1007/s10584-007-9388-3

Authors response : "Indeed, the selection of "well behaving" models for climate change applications is a complex issue extremely dependent on the processes and the region of interest. Further in the general discussion, we highlight this issue and add two references dealing with it : "The selection of climate models based on their credibility for climate change scenario is a complex issue \citep[e.,g.]{Brekke2008,Baumberger2017}, dependent on the purposes, the processes and the region of study. Whether the climate change signal should be corrected remains on open question \citep{Ehret2012}, even though there are good reasons to believe that model biases are time invariant \citep{Maurer2013}.". In the discussion for the bias correction of SIC, we make clear that in the light of our results, it is preferable to avoid to use AOGCMs that have large or persistent negative bias on sea-ice in present climate as initial material for the analog or the iterative, relative anomaly method."

9. Is the main result for SST bias correction that either method is appropriate for use due to your evaluation of the reliability of these methods? How does this result differ from other work on SST bias correction?

Authors response : "Following our evaluation, this is indeed the main result for SST bias correction methods. This can be a little surprising as one can expect that the quantile-quantile method is more appropriate to reproduce change in variance and correct biases that are quantile-specifics. For the absolute anomaly, the fact that we use the complete time series of the AOGCM scenario rather than the climatological mean allows for taking into account the projected change in variance present in this scenario. However, given the fact that the quantile-quantile method is widely applied for bias correction of climate variables and has proven to be appropriate, at least for variables that have no skewed distribution such as temperatures, we would recommend the use of the quantile-quantile method. This is however not the main point of the paper."

Technical Comments: The following comments should be easy to address, but will substantially improve the readability of the manuscript.

1. Please confirm that all acronyms are clearly defined, I have not listed all instances, but a few examples follow: CMIP5, AOGCM, AMIP, PCDMI, AGCM, etc.

Authors response : "Ok, comment taken into account"

2. Please confirm that all acronyms are consistent throughout the manuscript. I have not listed all instance, but a few examples follow: a. You define sea-ice concentration (SIC) in the beginning of the abstract, but spell it out in other places. b. You defined sea-ice area (SIA) twice.

Authors response : "Ok, comment taken into account"

3. Some of your terminology is inconsistent throughout the manuscript. For example, sometimes you say "future SST and SIC", other times you say "projected SST and SIC", etc., which makes the manuscript difficult to follow.

Authors response : "Ok, comment taken into account"

4. There are a number of grammatical and spelling errors, for example "The presence of SIC maps from futures AOGCM projections. . ." should read "The presence of SIC maps from future AOGCM projections. . .", please double-check the manuscript for grammar and spelling.

Authors response : "Comment taken into account, we had the manuscript read by a native speaker."

5. Figure 1 (right): It is difficult to determine which line is thick and which is thin, I suggest using a dashed line or adding more thickness.

Author response : "Ok, comment taken into account"

6. Figure 3 caption: Where should the reader go to "see text"?

Authors response : "Comment taken into account, there is now reference towards the corresponding section."

7. Figure 7: Including a key for the lines such as in Figure 4 would be helpful for clarity

Authors response : "Ok, comment taken into account"

8. Figure 8 and 9: The text refers to specific regions, for example the Weddell sea, but I'm not sure how to determine the regions from this Figure.

Authors response : "We changed the legend of the figure so that we can now distinguish the different regions with the help of different colors. Different signs (crosses and circles) are used to distinguish scenarios from CNRM-CM5 from IPSL-CM5A-LR. The more important for these figures is first to distinguish the regions, then the models. The distinction between rcp4.5 and rcp8.5 is less essential for the interpretation of the results and the connections with the text."

9. Figure A2 and A5: Including a key for the lines such as in Figure 4 would be helpful for clarity

Authors response : "Ok, comment taken into account"

10.Equation 1: As some of the parts of the equation refer to a climatological mean, and some to monthly data, adding in summations or "bar" notation would be very helpful.

Authors response : "Ok, comment taken into account"

11. Table 1 and resulting discussion: I may have missed something, but the labeling of this table confuses me, as well as the discussion in the text. In Section 3, below line 25, you that when comparing corrected RCP SST using the perfect model test and original SST from CNRM-CM5 RCP8.5 you obtain a null bias for the entire globe. Yet in this table you show CNRM-CM5 rcp8.5 – CNRM-CM5 hist has a mean difference of +3.04 degrees C. I assume something is written incorrectly here, but I'm not sure what. In addition, this table is referenced in only in Section 3.1.2, which references the IPSL-CM5A-LR data. I'm confused why you're changing models here.

Authors response : "The goal of the real case application is to show that the mean and standard deviation shift due to the climate change is similar between the observations in the historical periods and the bias corrected scenario than between the historical simulation and the original scenario of the AOGCM used as initial material. Perhaps, the term mean difference was confusing here, we propose to replace it by change in mean and change in standard deviation "

12. As SIC is bias corrected independently of SSTA as noted in the first sentence of Section 3.3, this should also be mentioned somewhere in the methods section, providing context for the examination of physical consistency in Section 3.3.

Authors response : "We now refer to the examination of the physical consistency analysis in the introduction and the method section of the paper : "As SST and SIC are bias-corrected separately, section~\ref{sec3.3} presents a few considerations about SST and SIC consistency after performing bias corrections. The effects of the corrections applied \textit{a posteriori} in order to ensure the physical consistency between the two variables are evaluated within the framework of the perfect model test."

Assessing bias-corrections of oceanic surface conditions for atmospheric models

Julien Beaumet[1], Gerhard Krinner[1], Michel Déqué[2], Rein Haarsma[3], and Laurent Li[4]

[1]Univ. Grenoble Alpes, CNRS, IRD, Institut des Géosciences de l'Environnement, F-38000, Grenoble, France
[2]Météo-France, Centre National de Recherche Météorologiques, Toulouse, France
[3]Royal Netherlands Meteorological Institute (KNMI), De Bilt, Netherlands
[4]Laboratoire de Météorologie Dynamique, Université Pierre-Marie Curie, CNRS, Paris, France

Correspondence to: Julien Beaumet (Julien.Beaumet@univ-grenoble-alpes.fr)

Abstract. Future sea-surface temperature and sea-ice concentration from coupled ocean-atmosphere general circulation models such as those from the CMIP5 experiment are often used as boundary ~~forcing~~ forcings for the downscaling of future climate ~~experiment~~experiments. Yet, these models show some considerable biases when compared to the observations over present climate. In this paper, existing methods such as an absolute anomaly method and a quantile-quantile method for sea surface temperature (SST) as well as a look-up table and a relative anomaly method for sea-ice concentration (SIC) are presented. For SIC, we also propose a new analog method. Each method is objectively evaluated with a perfect model test using CMIP5 model ~~experiment~~ experiments and some real-case applications using observations. ~~With~~ We find that with respect to other previously existing methods~~for SIC~~, the analog method is a substantial improvement for the bias correction of future SIC. Consistency between the constructed SST and SIC fields is an important constraint to consider, as is consistency between the prescribed sea-ice ~~concentrations~~concentration and thickness; we show that the latter can be ensured by using a simple parameterization of sea-ice thickness as a function of instantaneous and annual minimum SIC.

1 Introduction-Context

Coupled climate models are the most reliable tools that we have today for large-scale climate projections, such as in the Coupled Model Intercomparison Project ~~,~~ Phase 5 (CMIP5 ~~, project (Taylor et al., 2012)), in which these projections were based on Representative Concentration Pathways (RCPs; Moss et al. (2010))~~(Taylor et al., 2012)). Regional-scale information is obtained by using these global simulations as a basis for downscaling exercises. Dynamical downscaling, as opposed to empirical-statistical downscaling (e.g., Hewitson et al., 2014), is carried out either with Regional Climate Models (RCM) (e.g., Giorgi and Gutowski, 2016) or with high-resolution ~~global atmospheric general~~ atmospheric global circulation models (Haarsma et al., 2016). In both cases, information about the projected changes of sea-surface conditions, such as Sea Surface Temperatures (SST), Sea-Ice Concentration (SIC) and Sea-Ice ~~Thickness~~ thickness (SIT), is required as a lower boundary condition for the higher-resolution models. However, SST and SIC conditions modelled by coupled Atmosphere-Ocean ~~General Circulation Model (AOGCMs pr~~ Global Circulation Models (AOGCMs or CGCMs) show important biases for the present climate ~~(Flato et al., 2013) . It has, for example,~~ (Flato et al., 2013; Li and Xie, 2014; Richter et al., 2014; Levine et al., 2013; Zhang and Zha For example, it has been highlighted that most of the CMIP5 models had difficulties in reliably modelling the seasonal cycle and the trend of sea-ice extent in the Antarctic over the historical period (Turner et al., 2013). Therefore, the validity and reliability of such coupled simulations is questionable for future climate projections (e.g. end of the ~~21st century)~~ 21^{st} century), and so is their use as boundary conditions when performing dynamical downscaling of future climate projections.

Prescribing correct SST is crucial for atmospheric modelling because SST determines heat and moisture exchanges with the atmosphere (Ashfaq et al., 2011; Hernández-Díaz et al., 2017). ~~In high latitude~~ The absence of the Pacific cold tongue bias and the reduction of the double ITCZ problem in AMIP experiments with respect to the CMIP5 model experiments (Li and Xie, 2014) shows the importance of forcing atmospheric model by SST close to the observations. For instance, improvements in the modelling of tropical cyclone activity in the Gulf of Mexico (Holland et al., 2010) and of summer precipitation in Mongolia (Sato et al., 2007) were obtained by bias-correcting SST and other AOGCM outputs before using them as forcing for RCMs. At high latitudes, SIC (Krinner et al., 2008; Screen and Simmonds, 2010; Noël et al., 2014) and, in some cases, SIT (Gerdes, 2006; Krinner et al., 2010) are two additional ~~required and~~ crucial boundary conditions for atmospheric ~~modelling of recent and future climate change~~ models. Krinner et al. (2014) demonstrated that for the Antarctic climate as simulated by an atmospheric model, prescribed SST and sea-ice changes have greater influence than prescribed greenhouse gas concentration changes. ~~Integrated~~ Large-scale average winter sea-ice extent and summer SST have been identified among the key boundary forcings for regional modelling of the Antarctic surface mass balance (Agosta et al., 2013), which is the only potentially significant negative contributor to the global eustatic sea level change ~~in~~ over the course of the ~~21st~~ 21^{st} century (Agosta et al., 2013; Church et al., 2013; Lenaerts et al., 2016). We note that while there is a considerable body of scientific literature on the effect of varying SST and SIC on simulated climate, very few studies focused on the role of varying SIT in atmosphere-only simulations (Gerdes, 2006; Krinner et al., 2010; Semmler et al., 2016) , although air-sea fluxes in the presence of sea ice are strongly influenced by the thickness of the sea ice and the overlying snow cover. Gerdes (2006) and Krinner et al. (2010) have shown that the atmospheric response to changes in Arctic SIT can induce atmospheric signals that are of similar magnitude as those due to changes in sea ice cover. In most atmosphere-only General Circulation Models (AGCMs), SIT will therefore also need to be prescribed along with SST and SIC. When SST and SIC from a coupled climate model are directly used, SIT from that same run should of course be used; however, in case SST and SIC from the coupled model run are bias-corrected, as we strongly suggest here, we argue that SIT should be prescribed in a physically consistent manner in the atmosphere-only simulation.

In this study, we describe, evaluate and discuss different existing and new methods for the construction of bias-corrected future SST~~and SIC~~ , SIC and SIT. These methods generally take into account observed oceanic boundary conditions as well as the climate change signal coming from CMIP5 AOGCM scenarios to build more reliable SST and SIC conditions for future climate, which should reduce the uncertainties when used to force future climate projections. The different methods have been evaluated using a perfect model ~~test~~ approach, and by carrying out real-case applications on observations. Applied changes in mean and variances have been investigated as well as the coherence of SIC and SST after applying bias correction methods. The analysis of the results focuses on methods for sea-ice, as bias correction of SIC is ~~a more complicated~~ more complicated an issue to deal with. For SIT, we propose a diagnostic using SIC following Krinner et al. (1997) . Because there were no reliable observational data sets available until recently (Lindsay and Schweiger, 2015; Kurtz and Markus, 2012, e.g.) , we evaluate here directly diagnosed SIT against new observations. In the following, we present the bias-correction methods, the data and the evaluation methods in section~~2.~~ 2.1. The results of the evaluation are shown in section~~3 and are~~ 3. Because SST and SIC are bias-corrected separately, section 3.3 presents a few considerations about SST and SIC consistency after performing bias corrections. The results are then discussed together with general considerations on bias correction of oceanic surface conditions in section~~4. Finally, our findings are summed up and we~~ 4. Finally, we sum up our findings and draw conclusions in section~~5.~~ 5.

2 Data and methods

2.1 Data

Application and validation of the methods for bias correction have been achieved using ~~observationnal~~ observational SST and SIC data from the Program for Climate Model Diagnosis and Intercomparison (PCMDI) that are generally used as boundary conditions for Atmospheric Model Intercomparion Project (AMIP) experiments (Taylor et al., 2000), called “PCMDI obs.” or “observations” in this paper. The AOGCM’s historical and ~~future simulated~~ projected sea-surface conditions come from CMIP5 simulations (Taylor et al., 2012). Only the first ensemble members of the historical, ~~rcp4.5 and rcp8.5~~ and of the Representative Concentration Pathways (RCPs; Moss et al. (2010)) 4.5 and 8.5 simulations have been considered. Most methods have been tested using CNRM-CM5, IPSL-CM5A-LR and HadGEM-ES coupled GCM. Data from NorESM1-M~~and the~~ , MIROC-ESM, EC-EARTH, CCSM4 models have also been used as analog candidates in the analog method for sea-ice. Prior to any application of the bias correction methods, AOGCMs data have been ~~bilinearly~~ bi-linearly regridded onto a common regular 1°x1° grid. For the evaluation of the diagnosed SIT, we used the Lindsay and Schweiger (2015) data for the Arctic. For the Antarctic, in spite of recent observations with autonomous underwater vehicles by Williams et al. (2015) which tend to suggest occurrence of thicker Antarctic sea-ice than previously acknowledged, we will use the Kurtz and Markus (2012) data because of their large spatial coverage.

2.2 Sea Surface Temperature methods

The bias correction of simulated SST is a ~~fairly~~ relatively easy and a straightforward issue to deal with. ~~Nevertheless, different~~ Different methods have been developed ~~. In this section, we describe an anomaly-based method and a quantile-quantile method. Results from their application are presented in section 3.~~

2.2.1 ~~Anomaly method~~

~~This frequently used method (e.g., Krinner et al., 2008) simply consists~~ and presented in the literature. Here we re-evaluate two different frequently used methods. The first is an absolute anomaly method (Krinner et al., 2008, e.g.,) , which consists of simply adding the SST ~~anomaly coming from the difference between a coupled AOGCM projection and the corresponding historical simulation to the present-day observations. In practice, for each grid point, the difference between the SST~~ difference for a given month ~~in the future from a climate change simulation and the climatological mean SST in the corresponding historical simulation from the same coupled AOGCM is added to the observed climatological mean SST (e.g . PCMDI, 1971-2000):~~

$$SST_{Fut,est} = \overline{SST_{obs}} + \left(SST_{Fut,AOGCM} - \overline{SST_{Hist,AOGCM}}\right)$$

~~In (A1),~~ ~~$SST_{Fut,est}$~~ ~~is the estimated future SST for a given month,~~ ~~$\overline{SST_{obs}}$~~ ~~the observed climatological monthly mean,~~ ~~$SST_{Fut,AOGCM}$~~ ~~the model future SST for a given month in the future AOGCM scenario and~~ ~~$\overline{SST_{Hist,AOGCM}}$~~ ~~the model climatological monthly~~ from an AOGCM scenario to the climatological mean in the ~~AOGCM historical simulation for the same reference period as for the observed climatology. As a result, the reconstructed SST time series has the chronology of the AOGCM projected scenario.~~

2.2.1 ~~Quantile-quantile method~~

~~This method has been proposed and described in Ashfaq et al. (2011) It consists in adding, for each grid point and each calendar month's quantile in the observations, the corresponding quantile change in the GCM data set, i.e. the difference between the maximum SST in the projected scenario and in the historical simulation, between the second highest SSTs in the two simulations, and so on for each ranked SST quantile. However, unlike Ashfaq et al. (2011), we did not create a new SST field for the present by replacing SST from the GCM in the historical period by its corresponding quantile in the observations, but we directly added the quantile change~~ observations. The second is a quantile-quantile method presented in Ashfaq et al. (2011) , where for each quantile and each month, the climate change signal coming from the AOGCM scenario is added to the corresponding quantile ~~of the observational time series (Figure A1). This allows keeping the observations chronology and their inter-annual variability in estimated SSTs for the future. In our results, we noticed a large fine-scale spatial variability of~~ in the ~~constructed bias-corrected SSTs that was due to the large spatial variability of the climate change increments (quantile change) calculated individually for each pixel. To fix this , we applied a slight spatial filtering (3 grid point Hann box filter) of the quantile shifts in order to produce more consistent SST fields~~observations. Presenting these well-known methods in detail is of limited interest for the main part of this paper. However, interested readers can find a more complete description of the methods in Appendix A.

~~Illustration of the quantile-quantile method for min. and max. of SST time series for a grid point in the Central Pacific : GCM historical simulation (blue, left), GCM projected scenario (red, left), observed SST(thin, right), reconstructed future SST (thick, right)~~

2.3 Sea-ice Concentration methods

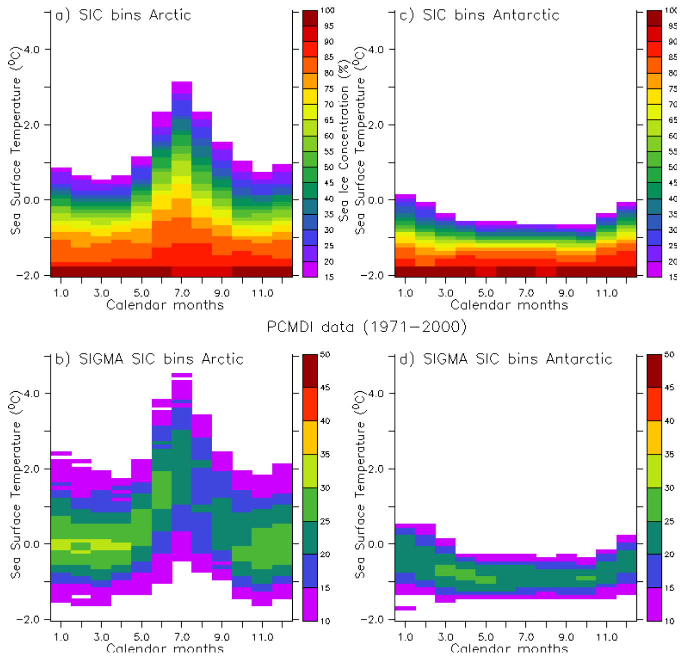


Figure 1. Look Up Tables (top) linking SST and SIC for the Arctic (left) and the Antarctic (right) built using 1971-2000 PCMDI observations and the associated uncertainty (root mean square error) on the computed SIC average (bottom).

~~Sea-ice concentration~~ SIC is more difficult to bias correct because it is a relative quantity that must be strictly bounded between 0 and 100 %. This difficulty led some authors to neglect SIC bias correction altogether in studies with prescribed corrected future SSTs that did not specifically focus on polar regions (e.g., Hernández-Díaz et al., 2017) . In this section, we present three methods: a look-up table, ~~a~~ an iterative relative anomaly and an analog method.

2.3.1 Look-up Table method

This method has been developed at *the Royal Netherlands Meteorological Institute* (KNMI). It is used in Haarsma et al. (2013) and within the framework of the High Resolution Model Intercomparison Project (HighResMIP) (Haarsma et al., 2016). ~~It is based on the assumption~~ A regression of SIC as a function of SST is also used in the HAPPI project (Mitchell et al., 2017) . In this method,the assumption is made that SIC is a function of SST. Therefore, SST are ranked per 0.1 K bins and the corresponding average SIC for each temperature bin between -2 and +5°C is calculated. Relations between SST and SIC have been found to be dependent on seasons and hemispheres. Therefore, using monthly mean values of SST and SIC from historical observations, look-up tables are built, separately for the Arctic and the Antarctic, for each calendar month (Figure 1). Then, with the help of future SSTs, these ~~look-up tables~~ Look-up Tables (LUT) are used to retrieve future SIC.

2.3.2 Iterative relative anomaly method

Here we follow a method described by Krinner et al. (2008). It is based on relative regional sea-ice area (SIA) changes ~~which~~ and is essentially an iterative scheme of mathematical morphology for image erosion and dilation (Haralick et al., 1987). The Arctic and the Antarctic are divided into sectors of equal longitude. In each sector, the average SIA is calculated by spatially integrating SIC. With respect to the method introduced in Krinner et al. (2008), we introduce the use of a quantile-quantile method to determine the targeted SIA in the bias-corrected projection. This targeted SIA is then calculated for each sector and each quantile, with the help of the following equation:

$$SIA_{Fut,est} = SIA_{obs} \cdot \left(\frac{SIA_{Fut,AOGCM}}{SIA_{Hist,AOGCM}} \right) \tag{1}$$

In (2), ~~$SIA_{\mathrm{Fut,est}}$~~ $SIA_{Fut,est}$ is the estimated projected SIA for the current month and sector, ~~SIA_{Obs}~~ SIA_{Obs} the SIA from the observations, and ~~$SIA_{\mathrm{Fut,AOGCM}}$ and $SIA_{\mathrm{Hist,AOGCM}}$~~ $SIA_{Fut,AOGCM}$ and $SIA_{Hist,AOGCM}$ are respectively computed SIA for the corresponding quantile to the observations, using SIC from a future scenario and a historical AOGCM's simulation. Starting from an observed present SIC map and using the computed relative SIA change for a given sector, the decrease (increase) in SIC is then realized using an iterative process: SIC in each grid box is replaced by the minimum (maximum) SIC of all adjacent pixels (Figure 2); the new spatially integrated SIA is calculated and the operation is repeated until the obtained change converges towards the computed targeted SIA retrieved from AOGCM's simulation sea-ice data and observations. Afterwards, the decrease/increase process is repeated on the hemisphere scale in order to ensure that the change in SIC reproduces the total hemispheric SIA change.

2.3.3 Analog method

In this method, we divide the Arctic and the Antarctic into ~~n_{s}~~ n_s geographical sectors that correspond to different seas of the Arctic and the Southern Oceans; we defined ~~n_{s}~~ n_s = 12 sectors for the Arctic and ~~n_{s}~~ n_s = 7 sectors for the Antarctic. For each sector and each month, the quantiles of the sea-ice extent (SIE: total area with SIC above 15%) and the SIA are computed from SIC observations over the AMIP period. Corresponding quantile changes in SIE and SIA are computed using SICs from a CMIP5 AOGCM ~~'s~~ historical simulation and a projected scenario run. Computed quantile changes are then applied to the corresponding quantiles in the observations in order to obtain ~~target~~ targeted future SIA and SIE for each month, quantile and sectors. Then, a library of future SIC fields is built by collecting SIC observations from the AMIP period as well as SIC from CMIP5 projections. The presence of SIC maps from ~~futures~~ AOGCM projections in this library is justified by the need to take into account physically plausible future SIC distributions outside of the current observed range. However, ~~AOGCM that overly~~ AOGCMs that poorly represent sea-ice ~~distribution~~ annual cycle in present-day climate are preferably dismissed from this library. Future SIC is then finally reconstructed by searching the analog for each quantile *q*, sector *s* and month *m* in the library, that is to say the SIC field that minimizes the cost function C expressed by:

$$C_{q,m,s} = \sqrt{\left(\frac{SIA_s - SIA_{T_{(q,m,s)}}}{SIA_{max_{(q,m,s)}}} \right)^2 + \left(\frac{SIE_s - SIE_{T_{(q,m,s)}}}{SIE_{max_{(q,m,s)}}} \right)^2} \tag{2}$$

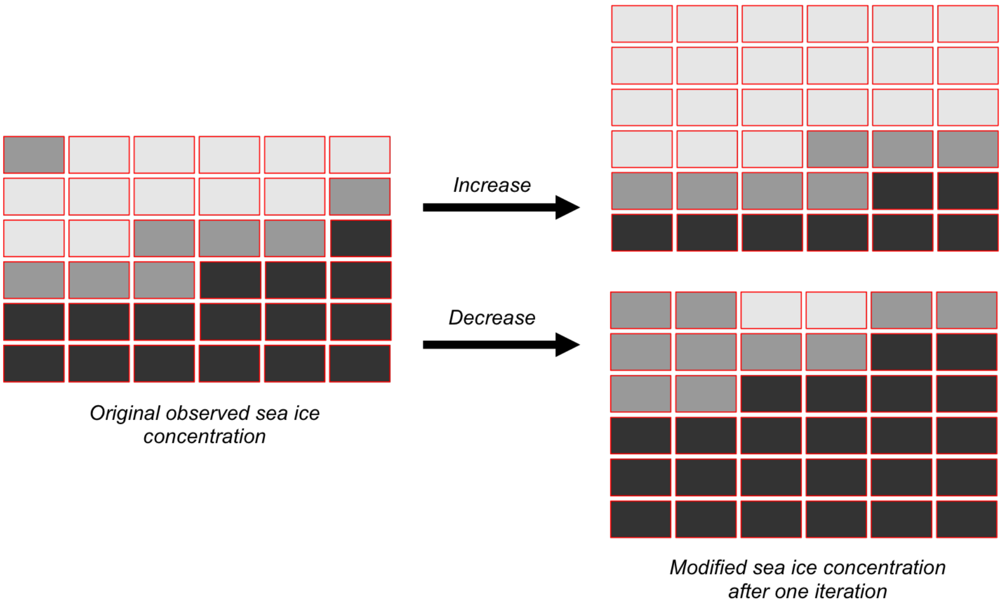


Figure 2. Iteratively constructing a "corrected" future SIC field using the iterative relative anomaly method (see ~~text~~section 2.3.2).

where SIA_s and SIE_s are the SIA and SIE of the processed sectors of the analog candidate from the library, ~~$SIA_{T_{(q,m,s)}}$ and $SIE_{T_{(q,m,s)}}$~~ $SIA_{T_{(q,m,s)}}$ and $SIE_{T_{(q,m,s)}}$ are the targeted ~~future sea-ice area and extent~~ projected SIE and SIA computed using the quantile-quantile method, and ~~$SIA_{max_{(q,m,s)}}$ and $SIE_{max_{(q,m,s)}}$~~ $SIA_{max_{(q,m,s)}}$ and $SIE_{max_{(q,m,s)}}$ are the maximum SIA and SIE of the processed sector. The double criterion on both SIE and SIA was introduced in order to be able to distinguish cases in which the total SIE in a sector is similar but the average SIC is very different (and vice versa). In order to avoid issues introduced by different land masks between AOGCMs and PCMDI data, we filled land grid points with sea-ice using a nearest neighbour method and masked all the grid points with the same land mask built with land fraction from PCMDI data in order to compute SIEs and SIAs for each region with the same reference. Analogs are attributed without taking into account the month of the analog candidate in the library. This allows for instance attributing a summer sea-ice map from present observations for a future winter month reconstructed sea-ice field. For each quantile q, month m and sector s, this procedure yields an hemispheric SIC field ~~$SIC_{opt_{(i,q,m,s)}}$~~ $SIC_{opt_{(i,q,m,s)}}$ that minimizes the cost function for the given sector, month and quantile. For a given month and quantile, there are thus ~~n_s~~ n_s hemispheric SIC fields ~~$SIC_{opt_{(i,q,m,s)}}$~~ $SIC_{opt_{(i,q,m,s)}}$. At each grid point i, the corresponding ~~n_s~~ n_s SIC values are then blended using a weight function ~~$w_{(i,s)}$~~ $w_{(i,s)}$ depending on the distance ~~$d_{(i,s)}$~~ $d_{(i,s)}$ of that grid point to the center of each of the sectors in order to obtain the final reconstructed SIC, ~~$SIC_{(i,q,m)}$~~ $SIC_{(i,q,m)}$, for a given quantile q and month m:

$$SIC_{(i,q,m)} = \sum_{s=1}^{n_s} \left(w_{(i,s)} \times SIC_{opt_{(i,q,m,s)}} \right) \tag{3}$$

with

$$w_{(i,s)} = \left(1 + \left(\frac{d_{(i,s)}}{d_r}\right)^4\right)^{-1} \tag{4}$$

Here, $\cancel{d_r}$ d_r is a reference distance of 500 km, yielding a smooth transition at the boundaries between two adjacent sectors. At the center of a sector, this yields a weight that is very close to 1 for the relevant field that was identified as optimal for that sector and that is close to 0 for the fields identified as optimal for the other sectors; at the boundary between two sectors, the weights are typically 0.5 for the two relevant sectors and close to 0 for the others.

2.4 Sea Ice Thickness method

2.4.1 Diagnosing sea-ice thickness from sea-ice concentration

As described by Krinner et al. (2010) , the parameterization of sea-ice thickness SIT (denoted h_S in the following) as a function of the local instantaneous SIC f and annual-minimum SIC f_{min} is designed such as to yield h_S of the order of 3 meters for multi-year sea ice (deemed to be dominant when the local annual minimum fraction $f_{min} \gg 0$) and h_S below 60cm (with a stronger annual cycle) in regions where sea-ice completely disappears in summer (that is, $f_{min} = 0$), and intermediate values for intermediate cases:

$$h_S = (c_1 + c_2 f_{min}^2) \cdot (1 + c_3(f - f_{min})) \tag{5}$$

with c_1=0.2m, c_2=2.8m and c3=2. This corresponds to the observed characteristics of Arctic and Antarctic sea ice, with multi-year sea ice being generally much thicker than first year ice. The parameter c_3 introduces a seasonal ice thickness variation in areas where there is a concomitant seasonal cycle of SIC. A more parsimonious formulation using only two parameters could have been designed to comply with these constraints. However, for the sake of consistency with previous work, we used the equation proposed by Krinner et al. (1997) who designed the parameterization such as to allow for a fairly strong seasonal cycle of SIT also in regions with intermediate values of f_{min}.

2.5 Evaluation

Evaluation of the above methods is mainly achieved with a perfect model approach. ~~In this test~~A perfect model approach usually consists of using model data as a substitute for observations, and trying to predict projected model data from that model; this prediction can then be evaluated against the available model projections (e.g., Hawkins et al., 2011) . In the real world, as observations of future climate are obviously not yet available, an equivalent approach is impossible if one cannot wait long enough for the future to become reality. Another type of perfect model approach are "Big Brother" experiments for evaluating downscaling techniques. In such studies, high-resolution model output is degraded in resolution and downscaling methods are then applied to these low-resolution data. The resulting synthetic high-resolution fields are then compared to the original high-resolution output (e.g., Denis et al., 2002; de Elía et al., 2006) . Here, we consider SST and SIC from the historical simulation of one coupled AOGCM as being the observations. Then, we apply the different bias correction methods using

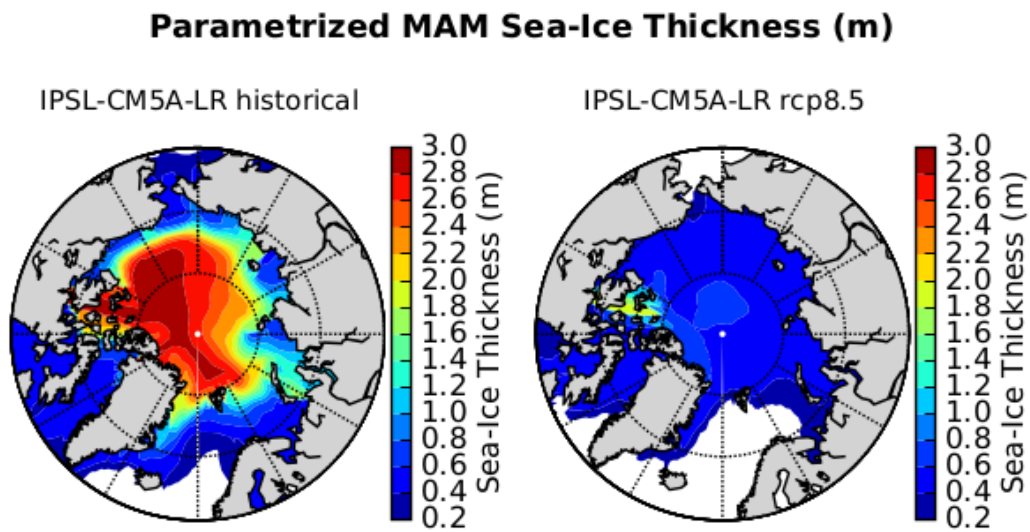


Figure 3. Spring (MAM) estimated mean SIT(m) using parametrization from (Krinner et al., 1997) and IPSL-CM5A-LR SIC data from the historical run (1971-2000, left) and the rcp8.5 scenario (2071-2100, right).

the climate change signal coming from a scenario of the same ~~AOGCM~~model. Obtained projected SST and SIC using this perfect model test are finally compared with original SST and SIC from the AOGCM climate change experiment. Additionally, we also performed an assessment of real case applications using observations and climate change signals coming from AOGCM projections. Changes in mean and variance in the coupled model projection with respect to the historical simulation are compared to the introduced change in mean and variance in the estimated future SST and SIC using bias correction methods with respect to the observed climatological data. We ~~assume~~ consider here that an ideal bias correction method should reproduce the same change in mean and variance between the observations and the bias-corrected projected SST and SIC as between the used coupled GCM historical simulation and its climate change scenario. For SIT, since the method is a diagnostic using SIC in order to ensure the consistency between these two variables, the evaluation of the method is achieved by comparing estimated SIT with observations that were not available until recently (Lindsay and Schweiger, 2015; Kurtz and Markus, 2012) .

As SST and SIC are bias-corrected separately, section 3.3 presents a few considerations about SST and SIC consistency after performing bias corrections. The effects of the corrections applied *a posteriori* in order to ensure the physical consistency between the two variables are evaluated within the framework of the perfect model test.

3 Results

3.1 Sea Surface Temperatures

3.1.1 Perfect model test

~~In this section, we discuss the application of the perfect model test for both the anomaly and the quantile-quantile method. To apply this test, we used CNRM-CM5 data from the historical simulation over the 1971-2000 period and from the rcp8.5 projection for the 2071-2100 period. Corrected rcp8.5 SST have been compared with the original SST projection. For the anomaly method~~ Absolute anomaly or quantile-quantile methods have been used for SST in previous bias-correction applications cited before in this paper. As a consequence, the utility of a perfect model test here is limited for SSTs, and it was only applied in order to be consistent with the evaluation of the method for SIC. For both methods, the relation between the ~~anomaly-corrected~~ bias-corrected projected SST and the SST directly obtained from the AOGCM projection is trivial when we replace observed SST by ~~SST~~ the one from the AOGCM historical simulation, as for instance in (1). As a result, ~~when comparing corrected rcp SST using the perfect model test and original SST from CNRM-CM5 rcp8.5 scenario, we obtain, by construction, a null bias all over the world (*figure not shown*). For the quantile-quantile method, the bias is also null in most regions. However, since we applied a very slight spatial filtering of the quantile increment, some slight biases (positive or negative) appear in regions of steep SST gradients (i.e. regions with major oceanic currents). Nevertheless, these biases are negligible (a few tenths of degrees Celsius; *figure not shown*)~~ the resulting errors were null or close to zero, and the results are therefore not presented or discussed.

3.1.2 Real-case application

Here, we present the application of the anomaly and the quantile-quantile methods in a ~~real case~~ real-case application. For this application, we use SST ~~data~~ from PCMDI observations data set over 1971-2000, from ~~the~~ IPSL-CM5A-LR and CNRM-CM5 historical simulation over the same period, as well as the rcp8.5 scenario over 2071-2100. Histograms of frequency distribution of SST for different regions of the world (Weddell Sea, Central Pacific and North Atlantic) have been plotted in order to compare frequency distributions in the observations, in the GCM historical and future simulations, as well as in the estimated bias-corrected future SST using the quantile-quantile and the anomaly ~~method~~ methods (Figure 4). In this figure, we can appreciate the change in mean and variance between the GCM historical simulation and the GCM future scenario and between the PCMDI observations and ~~in the estimated~~ the bias-corrected SST scenario. ~~This~~ In Figure 4 (bottom), we can see the large cold bias of the AOGCM with respect to the observations in the North Atlantic, as coupled models usually struggle to correctly represent the Atlantic Meridional Overturning Circulation (AMOC). The change in mean and variance due to the climate change signal is more explicitly ~~calculated and~~ presented for the North Atlantic for the application with CNRM-CM5 model in Table 1. Results from the anomaly method and from the quantile-quantile method are very similar, and both methods succeed in applying the same change in mean and variance coming from the AOGCM scenario to the observations when producing bias-corrected SST.

Table 1. Mean and standard deviation ~~difference~~ change between present and future SST data sets for North Atlantic (45°N to 58°N, 105°W to 85°W)

	Mean ~~difference~~ change (°C)	STD ~~difference~~ change (°C)
CNRM-CM5 rcp8.5 - CNRM-CM5 hist	+3.04	+0.59
Anomaly meth. app. - PCMDI obs	+3.06	+0.66
Quantile-quantile meth. app. - PCMDI obs	+3.04	+0.68

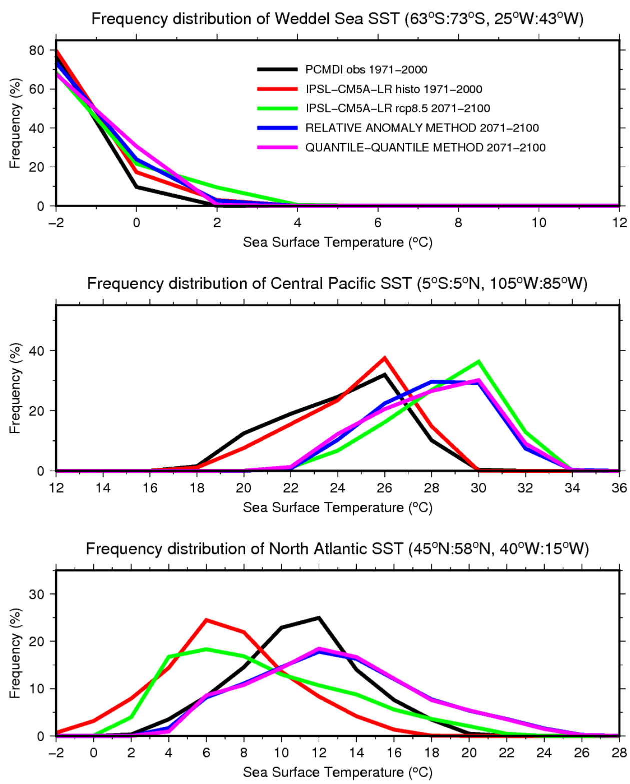


Figure 4. Frequency distribution of SST for PCMDI observations (black), IPSL-CM5A-LR historical (red) over 1971-2000 and rcp8.5 (green), quantile-quantile method (pink) and anomaly method (blue) applications over 2071-2100 for Weddell Sea (top), Central Pacific (center) and North Atlantic (bottom)

3.2 Sea-Ice Concentration

3.2.1 Perfect model test

In this section, we present the results of the application of the perfect model test for the three methods for bias correction of SIC. The term "perfect model test" is not absolutely pertinent for the evaluation of the Look-up Table method, as we first ~~computing look-up tables~~ computed LUTs using SST and SIC from an AOGCM historical simulation. Then, we used the SST of the climate change projection from the same AOGCM and retrieved SIC with the help of the previously computed LUT. An example of computed LUT using data of the historical simulation of CNRM-CM5 can be seen in Figure 5. It is noteworthy that this new ~~look-up table~~ LUT is significantly different from the one using PCMDI observations (Figure 1). Even though, the use of this LUT for the perfect model test instead of LUTs computed using observed SST and SIC over the AMIP period can be discussed, the use of LUT computed using observations would necessarily produce poorer result for the reconstruction of SIC ~~the AOGCM future~~ of the AOGCM's scenario in a perfect model test. Using AOGCM data, inconsistent or missing results were found for most of SST bins at or below the freezing point of sea water (-1.8°C). In order to fill the LUT, we therefore fixed SIC=99% for SST=-2.0°C and linearly interpolated SIC between -1.7°C and -2.0°C.

The perfect model test is more rigorously applied for the evaluation of the relative anomaly and the analog method, as we simply replaced time series of the observed SIC by the one from the AOGCM historical ~~scenario~~ simulation before applying the method without any specific modification or calibration. For the analog method, ~~we mention that~~ the tested AOGCM projection has been excluded from the possible analog candidates before applying the method and the perfect model test.

~~Mean biases~~ Errors (%) after applying the perfect model test are shown for the three methods for the rcp4.5 and rcp8.5 scenarios of the IPSL-CM5A-LR and CNRM-CM5 AOGCM (Figure 6). ~~One can see that the mean bias~~ These errors are generally lower for the LUT method : the mean Root Mean Square Error (RMSE) on the estimation ~~of sea-ice concentration remains reasonable for most of~~ for each scenarios for the Arctic and the Antarctic ~~for the analog method and very small for the look-up table method~~ is 4,8%. The mean error (ME) using this method tends to be positive in the Arctic and negative in the Southern Oceans. Errors using the relative anomaly method exhibits some larger values (mean RMSE = 8%). The errors using the analog method have intermediate values with respect to the first two methods (mean RMSE = 5.9%). Some of the ~~biases~~ errors of the analog method for regions with very complex coastal geography, such as the Canadian Archipelago, are due to the differences in land mask between the tested and the chosen AOGCM as analog candidate, despite the care taken for this issue. ~~Mean bias for the relative anomaly method exhibits some larger values.~~ The pattern of the ~~biases using this method~~ errors using the iterative relative anomaly seems robust between the different AOGCM scenarios. It is also noteworthy that the pattern of the ~~biases~~ errors is also similar between different methods, especially if we consider the results in the Arctic for the scenarios of the CNRM-CM5 model.

With the results of the perfect model test, we also performed a comparison between the frequency distribution of the mean SIC in the AOGCM future scenario (here CNRM-CM5, rcp8.5) and in the corresponding estimation using the bias correction methods (Figure 7). In these plots, we represented the histogram of frequency of ~~sea-ice concentration~~ SIC for four regions: Ross Sea (72°S:77°S; 174°E:163°W), Weddell Sea (63°S:73°S; 45°W:25°W), Arctic Basin (80°N:90°N; 180°W:180°E), and the Canadian Archipelago (66°N:80°N; 130°W:80°W). These regions have been chosen because they are the principal regions where there remains a significant amount of ~~sea-ice~~ sea ice by the end of the 21st century under the rcp8.5 scenario. With the ~~look-up table~~ LUT method (blue lines in Figure 7), the distribution of ~~sea-ice concentration is more or less~~ SIC is quite well reproduced in the Arctic (Figure 7 c and d), whereas in the Antarctic seas the distribution (Figure 7 a and b) exhibits well-marked peaks that we do not find in the GCM data set (black lines). The presence of such peaks is easy to explain by taking into account the structure of the ~~look-up tables~~LUT as i) for a given month, the SIC does not always increase monotonically with decreasing SST, ii) the discrete nature of LUT is not in favour of a continuous SIC frequency distribution. Moreover, using this method, we find a large underestimation of ~~the sea-ice concentrations~~ SIC above 90%, mainly in the Southern Hemisphere, with almost no occurrence of these high ~~sea-ice concentrations~~ SIC values in the estimations using the LUT method for the Ross and Weddell Seas. The frequency distribution of the sea-ice using the relative anomaly method (green lines in Figure 7) ~~seems more reasonable~~is closer to the distribution in the AOGCM, even if there is a slight overestimation of the frequency for concentrations between 70 and 90% and an underestimation for very high ~~sea-ice concentrations~~ SICs (above 90%). Finally, the distribution obtained using the analog method (red lines on Figure 7) is very close to the distribution of the original AOGCM ~~future~~ scenario. The results are robust because differences of sea-ice frequency distribution between ~~future estimation and future AOGCM future scenario~~ bias-corrected projections and AOGCMs scenarios are very similar for ~~rcp4.5 from CNRM-CM5 as well as for both scenarios from IPSL-CM5A-LR~~ other scenarios and coupled models (*figures not shown*).

3.2.2 Real-case application

~~In this section, we~~ We applied the three bias correction methods using PCMDI SIC ~~observations~~ data from the 1971-2000 period, as well as the IPSL-CM5A-LR and CNRM-CM5 historical data over the same period and ~~the~~ data from the rcp4.5 and rcp8.5 ~~future~~ scenarios from 2071-2100 in order to obtain future bias-corrected ~~sea-ice corrections~~SIC. The reliability of the methods is evaluated by comparing the change in mean and variance between the observations ~~over present climate and future estimated sea-ice concentrations to~~ and the bias-corrected projected SICs with the corresponding changes in the ~~climate change simulation~~ original AOGCM scenario with respect to the historical simulation. ~~An~~ We consider here that an ideal method should apply the same statistical changes to observed sea-ice as the one present in the climate change projection used to derive climate change signal.

In Figure 8, the bias-corrected mean ~~sea-ice concentration~~ SIC change is plotted against the corresponding change in mean SIC in the AOGCM ~~future~~ scenario used to determine the climate change signal. All points in the plot are obtained by the same four AOGCM ~~future~~ scenarios as well as the same four "test regions" as in previous section (Ross and Weddell Seas, Arctic Basin, Canadian Archipelago). Similarly, in Figure 9, applied changes in standard deviation for the ~~future estimated~~ bias-corrected projected SIC are plotted against corresponding standard deviation change in the AOGCM climate change experiment.

For the ~~look-up table~~ LUT method (Figure 8a), future SSTs have been bias-corrected using the quantile-quantile method before using computed LUT for the retrieval of future SIC. Using this method, there seem to be no systematic ~~errors~~ error in the applied change in mean SIC. ~~However, the~~ The mean error on the estimation of the change in mean SIC for every regions and scenarios

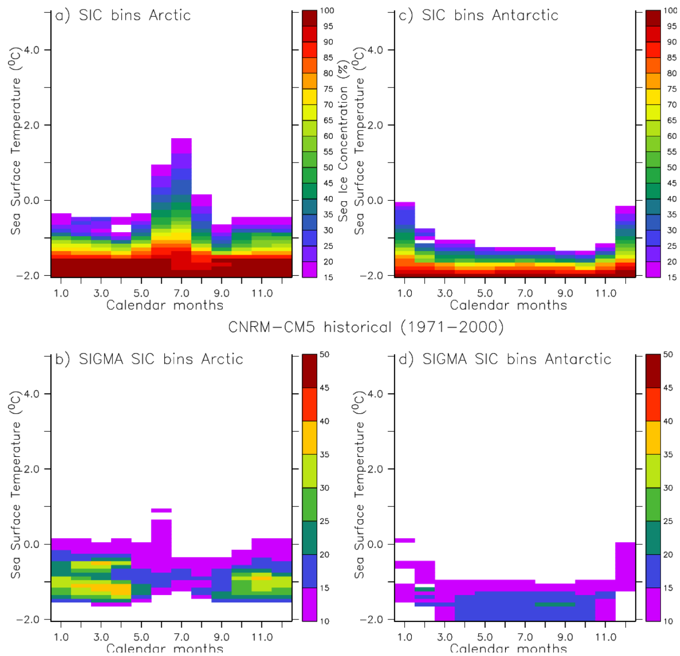


Figure 5. Look-up tables linking SST and SIC for the Arctic (a) and the Antarctic (c) built using 1971-2000 CNRM-CM5 historical simulation data and the associated uncertainty (root mean square error) on the computed SIC average (b,d)

is -2.2% and the RMSE is 42%. The spread of the points seems to increase for stronger decreases ~~of~~ in sea-ice. Main outliers with a high overestimation of the decrease in SIC are ~~constituted by~~ points representing the evolution of sea-ice in the Weddell Sea, mainly for CNRM-CM5 scenarios. If we consider change in SIC variability (Figure 9a), ~~there is a strong systematic bias and the~~ systematic error (-14.9%) and RMSE (69.3%) are strong. The decrease in SIC variability in the ~~future~~ Antarctic seas in the projection is strongly overestimated. Indeed, due to the structure of the ~~look-up table itself~~LUTs themselves, the variability of SIC in ~~future estimations~~ the bias-corrected projections is much lower than in the observations or in the original scenarios. The application of the relative anomaly method shows a more general overestimation (ME = -11.6% ; RMSE = 52.2%) of the decrease in mean SIC (Figure 8b). This overestimation is more pronounced for the Weddell Sea area and for the scenarios of the CNRM-CM5 model. Only the decrease in mean SIC in the Arctic Basin is correctly reproduced with respect to the AOGCMs ~~future~~ scenarios. Concerning the change in SIC variability (Figure 9b), the scores are comparable to the application of the LUT method (ME = -11,6% ; RMSE = 64.7%). The increase in variability in the Arctic Basin and in the Canadian Archipelago is correctly reproduced whereas for the Antarctic seas and particularly the Weddell sector, the decrease in SIC variability is once again ~~massively~~ dramatically overestimated.

Finally, the application of the analog method ~~is able to reproduce a great part~~ gives intermediate scores (ME = -8% ; RMSE = 48.7%) with respect to the two previous methods for the estimation of the change in mean SIC (Figure 8c). ~~Nevertheless,~~ These scores are greatly deteriorated by distinct outliers corresponding to the Weddell Sea sector ~~are once again present~~ for each AOGCM scenario, with ~~a strong~~ an overestimation of the decrease in sea-ice. As for the relative anomaly method, the

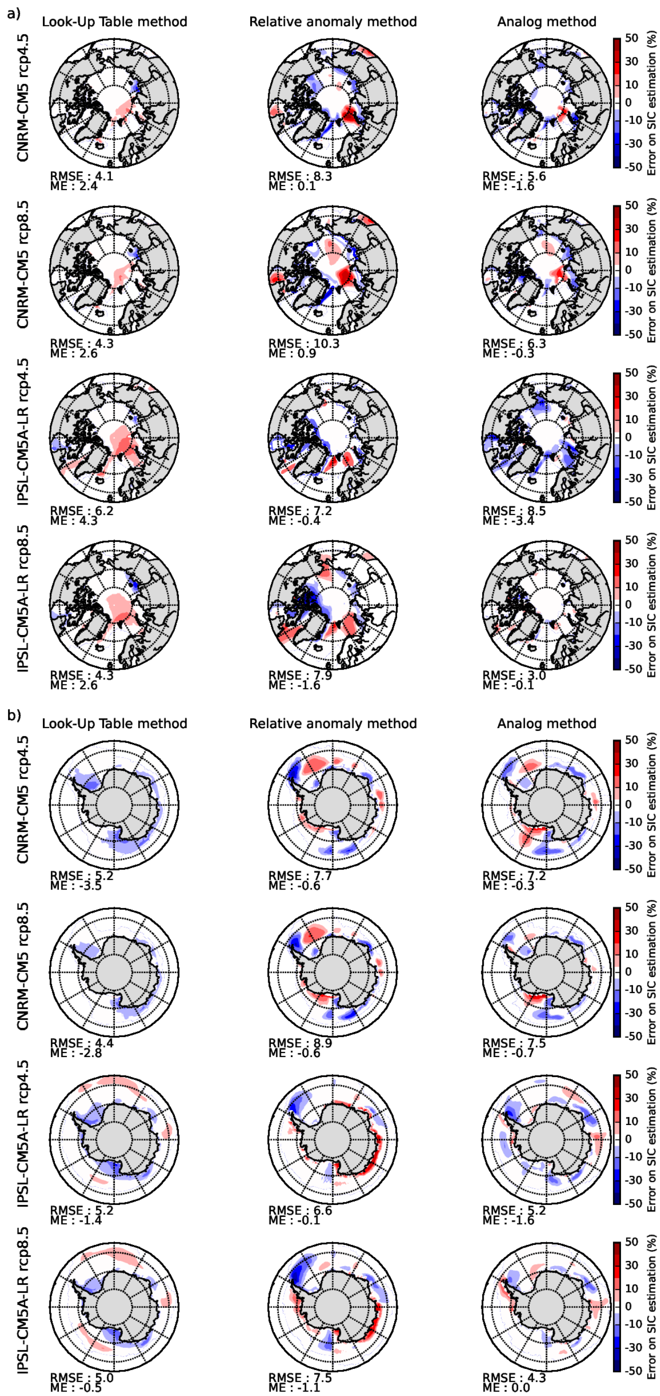


Figure 6. Mean ~~bias~~ error on the estimation of SIC with respect to the original AOGCM future scenario for the LUT, iterative relative anomaly and analog ~~method~~ methods with CNRM-CM5 and IPSL-CM5A-LR rcp4.5 and rcp8.5 scenarios for the Arctic (a) and the Antarctic (b)

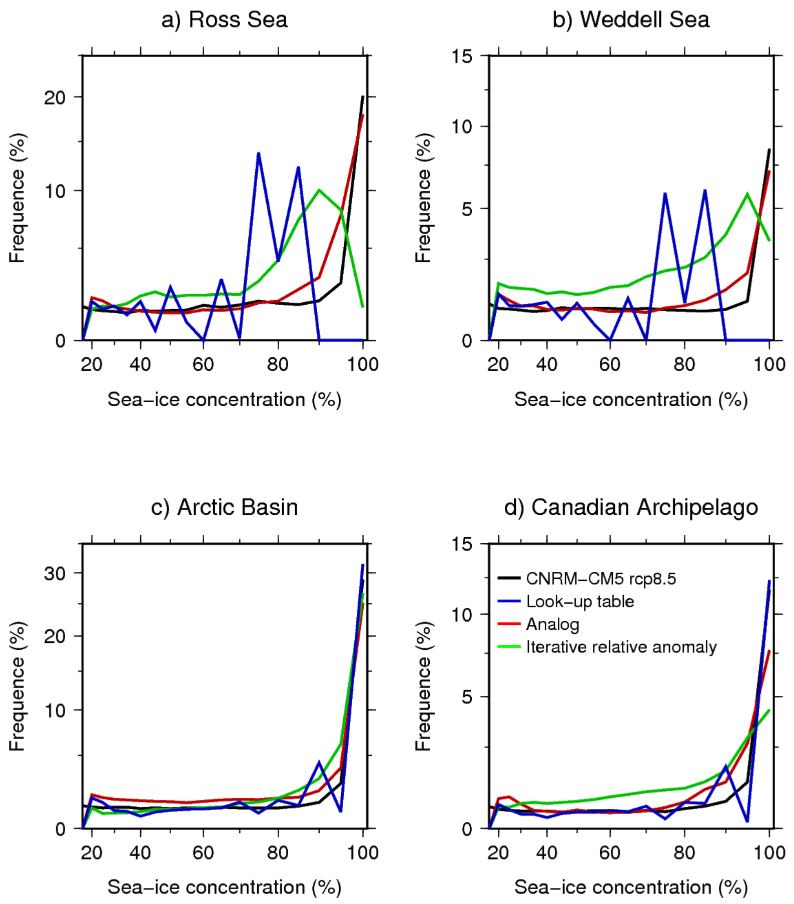


Figure 7. Frequency distribution of SIC in CNRM-CM5 rcp8.5 scenario (black) and in estimation using different methods in a perfect model test: Look-up table (blue), analog (red), and iterative relative anomaly (green). Regions are: a) Ross Sea (72°S:77°S, 174°E:163°W); b) Weddell Sea (63°S:73°S, 43°W:25°W); c) Arctic Basin (80°N:90°N, 180°W:180°E); d) Canadian Archipelago (66°N:88°N, 130°W:80°W)

change in SIC variability (Figure 9c) is correctly reproduced (ME = -9.3% ; RMSE = 60.3%), especially in the Arctic, while there is ~~a strong~~ an overestimation of the decrease in variability around Antarctica, particularly for the Weddell Sea.

3.3 Consistency between Sea Surface Temperature and Sea-ice Concentration~~consistencyAs bias correction~~

As bias corrections of SST and sea-ice are performed separately, the physical consistency between the two variables ~~is assessed~~ needs to be ensured a posteriori. To do so, three different issues are examined:

- There is a considerable amount of sea-ice (>15%) in the corrected scenario where the SST is above fresh water freezing point (273.15K). In this case, we set SST equal to the sea water freezing point (271.35K) for any SIC equal or greater than 50%. If the future calculated SIC is between 15 and 50%, the future SST is obtained by linearly interpolating between the sea water freezing point and the freshwater freezing point.

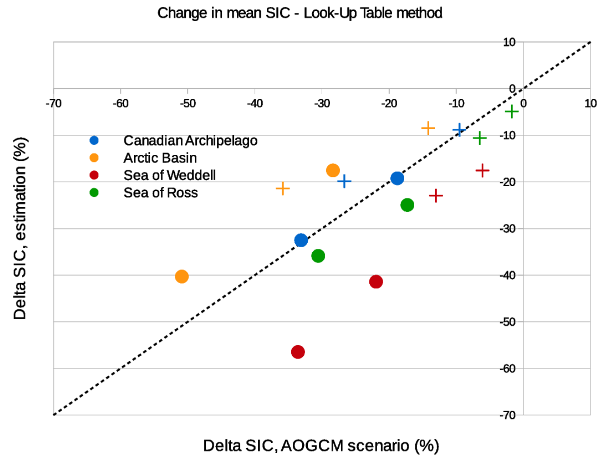
- The future corrected ~~sea-surface temperature~~ SST is below the fresh water freezing point but there is no significant (<15%) SIC in the bias-corrected scenario. In this case, we put the SST of the concerned grid point equal to the fresh water freezing point.

- SST has been used to remove very localized suspicious presence of sea-ice (no-ice) in the Arctic in summer. Any sea-ice for SST above 276.15K has been removed, this temperature being the highest temperature at which significant amount of sea-ice (15%) is found is the Arctic in the computed ~~look-up table~~ LUT using PCMDI data.

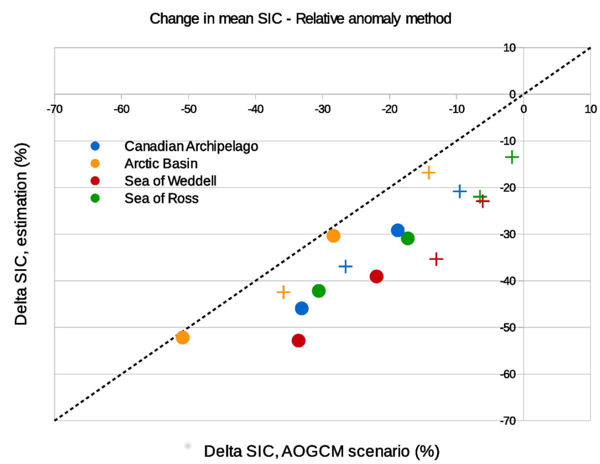
The impact of these modifications has been evaluated using the framework of the perfect model test. After applying the analog method for SIC and the quantile-quantile method for SST in a perfect model approach, we applied the correction for SST and SIC consistency and compared obtained SSTs to the original AOCGM future scenario used to carry out the experiment. The ~~biases~~ error can be seen in Figure 10 for the application of the method with IPSL-CM5A-LR and CNRM-CM5 scenarios. ~~It~~ Error is negligible in most regions. Very locally, it can reach up to 1°C. These regions generally correspond to regions where the analog method has shown some ~~biases~~ errors for the reconstruction of sea-ice especially for CNRM-CM5 scenarios. The occurrences of the three cases mentioned above have been assessed for both the perfect method test and the real-case application. ~~First~~ The first and third cases are very ~~seldom~~ rare and about 1% or less of the global oceanic surfaces experience at least one case during a 30 years experiment. The second case is more frequent, more than 20% of the global oceanic surfaces experience at least one occurrence during a 30 year experiment, while the mean occurrence at each time step is about 1 to 2% of the global oceanic surfaces. This case is responsible for the small (0.25 to 0.5K) but widespread warm bias on SST that can be seen in the Antarctic seas for the reconstruction of IPSL model scenarios in Figure~~10.~~ 10. Nevertheless, this slight decrease in the quality of the reconstruction of SST is worth considering in order to ensure physical consistency between SST and SIC.

4 ~~Discussion~~

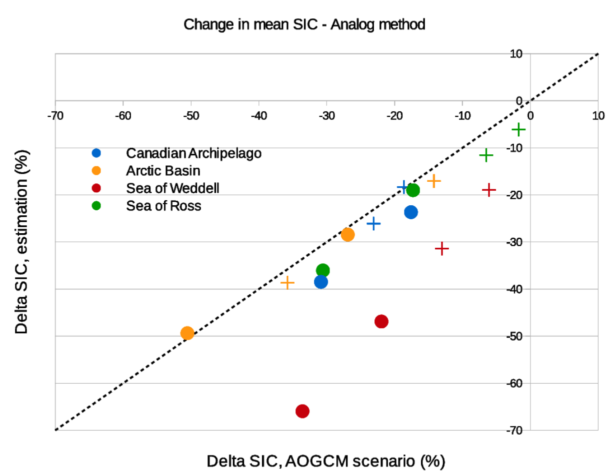
3.1 Sea Ice Thickness



(a)

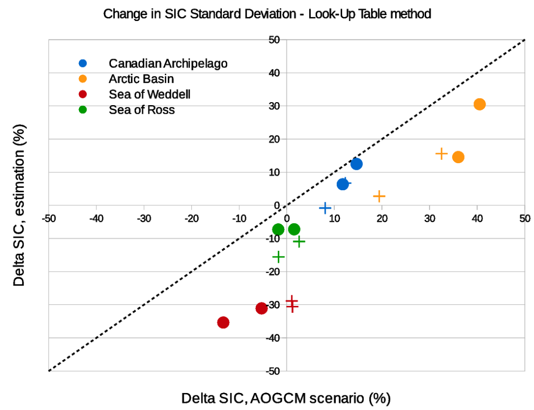


(b)

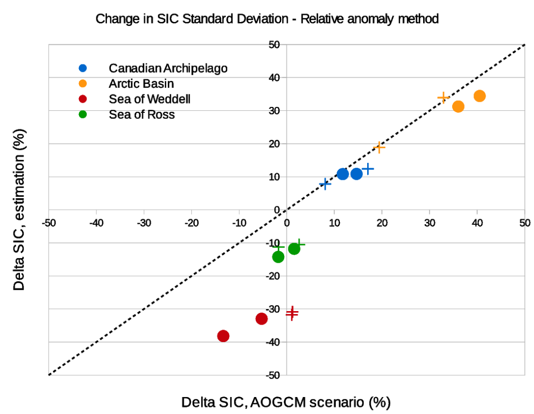


(c)

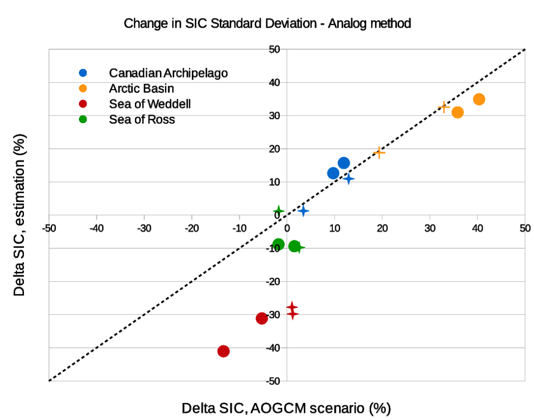
Figure 8. Change in mean ~~estimated future~~ bias-corrected SIC projections using a) ~~look-up table~~Look-Up Table, b) iterative relative anomaly, c) analog ~~method~~ methods against corresponding mean change in the AOGCM ~~future~~ scenario for the four test regions ~~(~~: Canadian Archipelago (blue), Arctic Basin (orange), Weddell Sea (red) and Ross Sea (green). Circles represent scenarios (rcp4.5 and rcp8.5) of



(a)



(b)



(c)

Figure 9. Change in ~~estimated future~~ bias-corrected SIC projections standard deviation using a) ~~look-up table~~Look-Up Table, b) iterative relative anomaly, c) analog ~~method~~ methods against corresponding mean change in the AOGCM ~~future~~ scenario for the four test regions ~~(~~: Canadian Archipelago (blue), Arctic Basin (orange), Weddell Sea (red) and Ross Sea (green). Circles represent scenarios (rcp4.5 and rcp8.5) of CNRM-CM5 and crosses, scenarios of IPSL-CM5A-LR

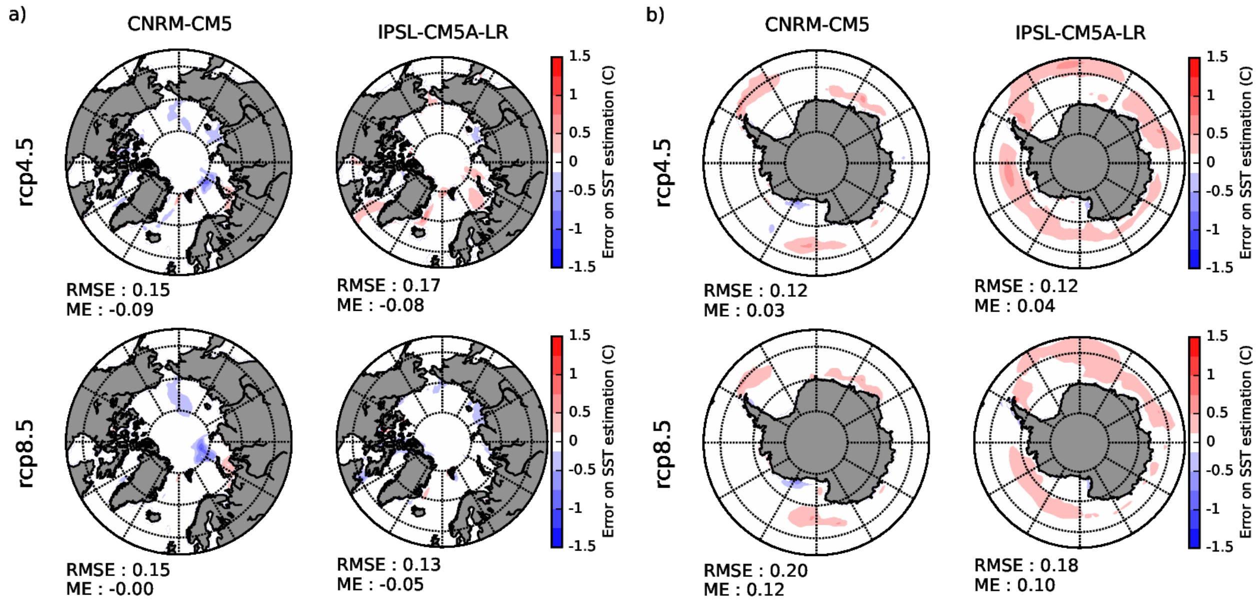


Figure 10. Mean ~~bias~~ error on the estimation of ~~the sea surface temperature~~ SST with respect to the corresponding original AOGCM ~~future~~ scenario after applying the analog method for sea-ice, the quantile-quantile method for SST and the correction for SST and SIC consistency for the Arctic (a) and the Southern Oceans (b)

The original formulation by Krinner et al. (1997) was parameterized for both hemispheres. We will therefore first present results for the original unique parameter set $c_{1,2,3}$ applied to both hemispheres. In a second step, we will present results for separate Arctic and Antarctic parameter sets, yielding a better fit to the observations. The reasoning is that, at the expense of generality of the diagnostic parameterization, one could argue that the strong difference between the Arctic and Antarctic geographic configuration — a closed small ocean favouring ice ridging and thus thicker sea ice in the Arctic, and large open ocean favouring thinner sea ice around Antarctica — justifies choosing different parameter sets for the two hemispheres. As changes of the position of the continents will be irrelevant over the time scales of interest here, climate change experiments will not be adversely affected by this loss of generality.

3.1.1 Option 1: Global parameter set

A comparison between the observed (Lindsay and Schweiger, 2015) and our diagnosed evolution of the Arctic mean SIT is given in Figure 11. The geographical patterns of the observed (in fact, observation-regressed) and parameterized Arctic ice thickness for March and September over the observation period 2000-2013 (Figure 12) do bear some resemblance, but they also show some clear deficiencies of the diagnostic parameterization. The diagnostic parameterization reproduces high SIT north of Greenland and the Canadian Archipelago, linked to persistent strong ice cover, but underestimates maximum ice thickness (due in part to compression caused by the ocean surface current configuration). Thinner sea ice over the seasonally

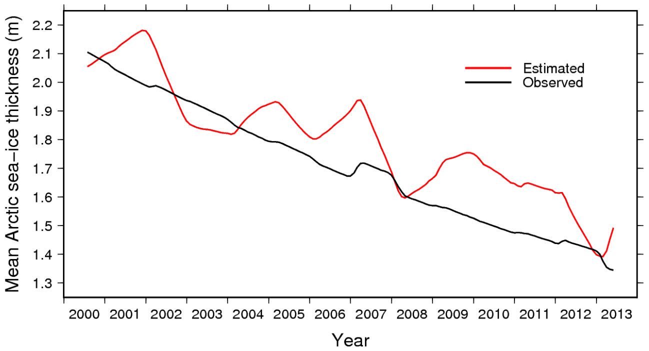


Figure 11. Observed (black, after (Lindsay and Schweiger, 2015)) and diagnosed (red) 12-month moving average mean sea-ice SIT of the Arctic basin (see Figure 12). The global parameter set is used here. Slight differences to Figure 13 of Lindsay and Schweiger (2015) appear because here we mask ice-free (SIC < 15%) areas that have a finite, non-zero ice thickness in the regression proposed by Lindsay and Schweiger (2015) who extend their regression to the entire Arctic Basin at all seasons.

ice-free parts of the basin is reproduced, but it is actually too thin, particularly in winter (for example in the Chukchi Sea). Obvious artifacts appear in September north of about 82°N where the SIC in the ERA-Interim data set clearly bears the signs of limitations due to the absence of satellite data.

Both for spring (Oct-Nov) and fall (May-Jun), our diagnosed SIT (Figure 13) compares generally well with the ICESat data except for an overestimate in the Weddell Sea, at both seasons. The geographical pattern of alternating regions with thin and thick sea ice is remarkably well reproduced.

3.1.2 Option 2: Separate Arctic and Antarctic parameter sets

A slightly better fit for the two poles can be obtained with separate parameters sets. For the Arctic, it seems desirable to increase winter SIT in the Chukchi Sea area (by increasing c_3 slightly) and to decrease the average SIT over the Central Arctic (by decreasing c_2). Figures 14 and 15 show results for the Arctic with c_1=0.2m, c_2=2.4m and c_3=3. The spatial fit is slightly better, but the recent Arctic-mean decadal trend towards decreased average SIT is somewhat less well reproduced. For the Antarctic, the main feature to improve is the maximum ice thickness in the Weddell Sea, which can be decreased by lowering c_2 to 2.0m. The Antarctic parameter set then becomes c_1=0.2m, c_2=2m and c_3=2. The result (Figure 16) is indeed a decreased thickness of the perennial Weddell Sea ice with little impact elsewhere.

In any case, these hemisphere-specific sea-ice parameter sets are not very different from each other and fairly similar to the original formulation.

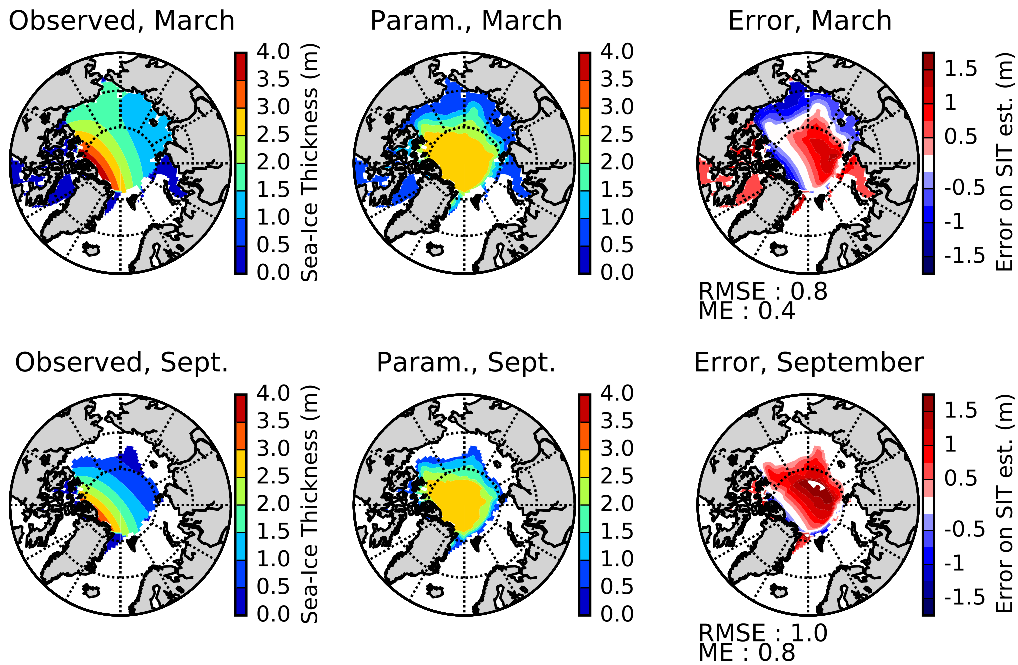


Figure 12. Observed (regressed, Lindsay and Schweiger (2015)) and parameterized Arctic SIT (in m) for March and September, and difference between these (right), with the global parameter set.

4 Discussion

4.1 Sea Surface Temperatures

The bias correction of projected ~~SSTs~~ SST coming from AOGCM scenarios is ~~an issue~~ fairly easy to deal with, and different appropriate solutions have already been proposed in the literature ~~(e.g., Krinner et al., 2008; Ashfaq et al., 2011; Hernández-Díaz et al., 2017)~~ In these papers, it has been demonstrated that the use of bias-corrected SSTs has considerable influences on the modeled climate and its response in projected scenarios for regions and processes as different as precipitation and temperature in the ~~Tropics and~~ tropics, the West African Monsoon ~~as well as for~~ and the climate of Antarctica.

In this paper, we reviewed two existing bias-correction methods and propose a validation that allows objectively evaluating the efficiency of these methods with the use of a perfect model test and a real-case application. Since both methods show no ~~bias~~ biases in the perfect model test and succeed in reproducing the change in mean and variability coming from the AOGCM future scenarios, we can be confident in the use of these methods for bias-correction of future AOGCM scenarios.

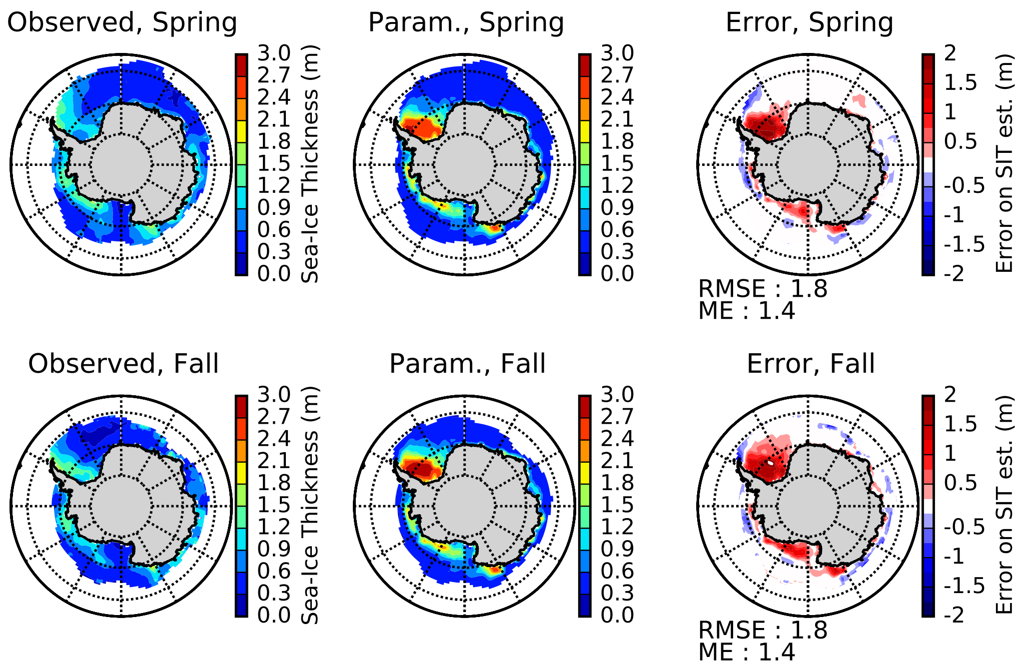


Figure 13. Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for Spring and Fall, and difference between these (right), with the global parameter set.

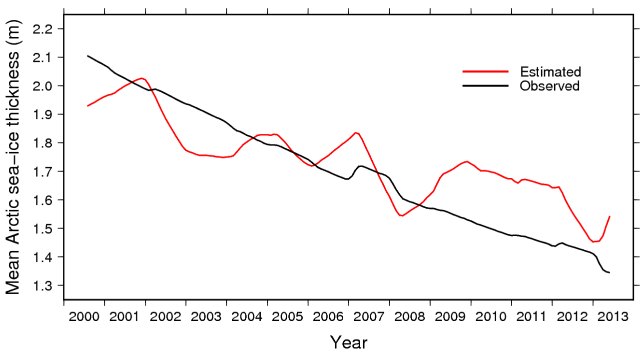


Figure 14. Observed (black, after Lindsay and Schweiger (2015)) and diagnosed (red) 12-month moving average mean SIT of the Arctic basin with the Arctic-specific parameter set.

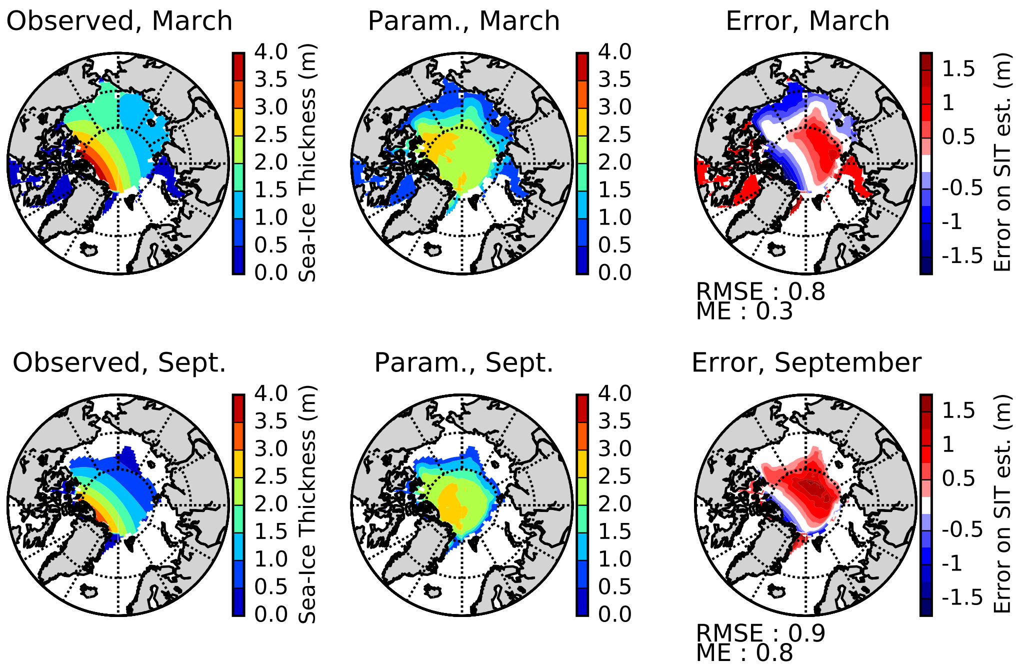


Figure 15. Observed Lindsay and Schweiger (2015) and parameterized Arctic SIT (in m) for March and September, and difference between these (right), with the Arctic-specific parameter set.

4.2 Sea-Ice Concentration

SIC is a quantity that has to remain strictly bounded between 0 and 100%, exhibits some sharp gradients and has to remain physically consistent with SST. Therefore the empirical bias correction of future SIC from coupled models scenarios is a much more complex issue to deal with than the bias correction of SSTs. The absence of satisfying solution proposals for this issue in the literature has led to incorrect bias-correction of future SIC in a recent study (Hernández-Díaz et al., 2017). Yet, the proposal of convenient solutions for the bias correction of sea-ice for ~~future~~ projected scenarios is crucial for the community interested in the downscaling of ~~future climate scenarios~~ climate scenarios experiments for polar regions.

In the perfect model test, we have seen that the ~~look-up table~~ LUT method shows some reduced ~~mean bias~~ errors over most regions (Figure 6). However, we have seen that the frequency distribution of future SIC obtained using this method is ~~different from~~ very different than the original distribution in the AOGCM and unavoidably exhibits some peaks due to the structure of LUT (Figure 7). Moreover, the absence of SIC above 90% in the Antarctic is also a considerable limitation to the method considering the large differences in terms of heat and moisture exchanges in winter between an ocean fully covered by sea-ice and an ocean that exhibits some ice-free channels (Krinner et al., 2010). In addition, the use of SST as a proxy for SIC is physically questionable, as we should expect a large SIC gradient around the freezing point. The fact that both SST and SIC

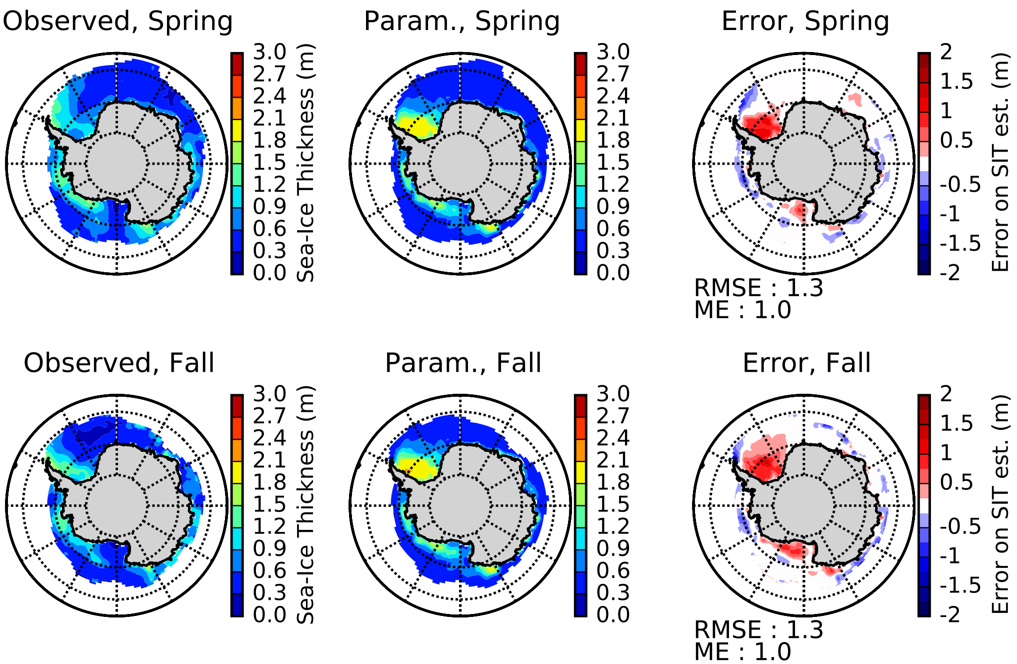


Figure 16. Observed Kurtz and Markus (2012) and parameterized Antarctic SIT (in m) for Spring and Fall, and difference between these (right), with the Antarctic-specific parameter set.

are averaged over a long period (one month) and over a considerable area (1°x1°) is probably the main reason why we find nevertheless a relation between the two variables. The ~~real case~~ real-case application of the method also shows some difficulties for the reconstruction of large decreases in mean SIC (Figure 8a) as well as a poor reconstruction of the change in variability in future SIC (Figure 9a).

The relative anomaly method (Krinner et al., 2008) shows the largest spatial mean ~~biases~~ errors in the perfect model test (Figure 6). The structure of some ~~biases~~ errors seems to be constant across the reconstruction of different climate scenarios used in the perfect model test. The empirical reduction of SIC by an iterative "erosion" from the edges of the sea-ice covered regions has most likely the tendency to overestimate the decrease of sea-ice for some coastal regions, while it probably fails to reproduce some processes involved in the disappearance of sea-ice in the future such as for example the inflow of warmer waters through the Barents Sea or the Bering Strait in the Arctic. The "real-case" application of the relative anomaly method has shown some systematic negative ~~bias~~ errors in the reconstruction of the decrease in mean SIC (Figure 8b) and ~~some important~~ a substantial overestimation of the decrease in variability in the Antarctic seas (Figure 9b).

The evaluation of the analog method with the perfect model test ~~allows to highlight some mean biases locally slightly bigger~~ shows that the mean error can be locally slightly higher than for the ~~look-up table~~ LUT method (Figure 6). However, the frequency distribution of the ~~future estimated~~ bias-corrected SIC perfectly reproduces the frequency distribution of the sea-ice in the original AOGCM ~~future~~ scenario (Figure 7). The real-case application of the method succeeds in reproducing the change in mean and variability of SIC for most of the tested regions and scenarios (Figure 8c). However, the decrease in mean (Figure 8c) and variability (Figure 9c) of the sea-ice in the Antarctic, particularly the Weddell Sea, is also largely overestimated using this method. With respect to the relative anomaly method, the fact that we use observed or ~~AOGCMs modeled~~ AOGCM-simulated sea-ice maps to reconstruct estimated future ~~sea-ice~~ sea ice, and that we use a criterion for both ~~sea-ice area and sea-ice extent~~ SIA and SIE, allows us to better reproduce some critical features of future sea-ice cover, and to obtain a more realistic frequency distribution. It should be noted that in the perfect model test as well as in the real-case application, the original AOGCM is not present among the possible analog candidates. If this is done, the results are even better using this method.

The fact that the analog method and the relative anomaly method share the same ~~bias~~ errors in the real-case application with a strong overestimation of the decrease in mean and variability of the sea-ice in the Weddell Sea in particularly for the scenarios of the CNRM-CM5 model is not a coincidence. For both methods, the targeted future SIE (or SIA) for a given sector is a product of the division of the integrated SIE (SIA) in the AOGCM ~~future~~ scenario by the corresponding quantity in the historical simulation. As a consequence, the targeted projected SIE (SIA) for a given sector and a given month is null when the integrated SIE (SIA) is null in the future AOGCM scenario. Therefore, the bias in the ~~future~~ scenario is not corrected in that case. The fact that both methods overestimate the decrease in sea ice mainly for CNRM-CM5 scenarios is to be linked to the fact that the historical simulation of this AOGCM shows some considerable negative biases for the sea-ice in the Weddell Sea with respect to the observations. Consequently, SIC in the Weddell Sea in CNRM-CM5 ~~future~~ rcp8.5 scenario is low and the number of months with a complete disappearance of sea ice is large. For these months, SIC in these sectors is not bias-corrected with the latter two methods. This means that although the methods described here are in principle applicable to any AOGCM output, it seems to be wise to ~~preferentially select output of reasonably "well-behaving" AOGCMs~~ exclude AOGCMs with large negative bias on sea-ice in their historical simulation as initial material for the bias-correction.

4.3 ~~A note on sea-ice thickness~~Sea-ice Thickness

~~Air-sea fluxes in the presence of sea ice arestrongly influenced by the thickness of the sea ice and the overlying snow cover. Gerdes (2006) and Krinner et al. (2010) have shown that the atmospheric response to changes in Arctic sea-ice thickness is substantial. In most AGCMs, sea-ice thickness will also need to be prescribed along with sea-surface temperature and~~ Given the simplicity of the proposed diagnostic SIT parameterization, the results are, at least in some aspects such as the predicted average Arctic sea-ice ~~concentration. When SST and SICfrom a coupled climate model are directly used, sea-ice thickness from that same run should of course be used; however, in case SST and SIC from~~ thinning, surprisingly good. The Central Arctic SITs results are clearly adversely affected by the input SICs North of 82°N. Arctic winter SIT in the marginal seas appears underestimated. In the Antarctic, the spatial pattern of SIT is very well represented.

We think that in the absence of pan-Arctic and pan-Antarctic satellite-based data before approximately 2000, this parameterization can serve as a surrogate, and that it can, because it seems to have predictive power, also serve for climate change experiments with AGCMs or RCMs. Because of its simplicity, implementing this parameterization should not be too complicated in any case provided the model does explicitly take into account SIT in its computations of heat flow through sea ice. In that case, SIT can either be calculated online (with the need to keep track of annual minimum SIC during the execution of the code) or be input as a daily boundary condition along with the SIC.

Of course, another possibility would be to prescribe SIT anomalies from coupled models. In this case, it would probably be wise to compute the prescribe SIT using its relative thickness changes. For example, in a climate change experiment, this would read $h_{presc}(t) = h_{obs,2003-2008}\, h_{sim}(t)/h_{sim,2003-2008}$. Problems could of course occur in areas where the coupled model ~~run are bias-corrected, as we strongly suggest here, we argue that sea-ice thickness should be prescribed in a physically consistent manner in the atmosphere-only simulation. An in-detail evaluation of sea-ice thickness prescription methods is beyond the scope of the main part of this paper. Therefore, an evaluation and further refinement of a simple parameterization of~~ simulates no sea-ice ~~thickness~~ cover at present. A physically consistent diagnostic parameterization of SIT as a function of ~~instantaneous and annual minimum SIC, initially suggested by Krinner et al. (1997) and used by Krinner et al. (2010) , is presented in the supplementary material of this paper.~~ constructed SIC, as proposed here, would not suffer from such problems.

In any case, it is very probable that Arctic SIT will further decrease as multi-year sea ice will be replaced by a predominantly seasonal sea-ice cover. This should probably be taken into account in future modeling exercises similar to CORDEX or HighResMip, given the non-negligible impact of sea-ice thinning on winter heat fluxes in particular.

4.4 General considerations on bias correction of oceanic forcings

As already mentioned~~before~~, one may doubt whether it is possible to bias-correct a GCM that has overly ~~strong~~ large biases in present-day climate. Indeed, most of the bias-correction methods rely on the hypothesis than the climate change signal coming ~~of~~ from an AOGCM scenario is not dependent on the bias in the historical simulations. This hypothesis can largely be questioned in a non-linear system (formed by SIC and SST). For example, in a model with a ~~strong~~ large negative bias in sea-ice for present-day climate, most of the additional energy due to an enhanced greenhouse effect will be used to heat the ocean, while it would be primarily used to melt sea-ice in a model with a correct initial sea-ice state. For such a model, the reliability of the climate change signal in SST is thus necessarily questionable. The selection of climate models based on their credibility for climate change scenario is a complex issue (Brekke et al., 2008; Baumberger et al., 2017, e.,g.) , dependent on the purposes, the processes and the region of study. Whether the climate change signal should be corrected remains on open question (Ehret et al., 2012) , even though there are good reasons to believe that model biases are time invariant (Maurer et al., 2013) .

Skills of coupled GCMs in reproducing the observed climate and its variability for a region of interest are often evaluated in order to use the GCM output as forcing for downscaling experiments. However, skills of atmospheric ~~GCM~~ GCMs are generally better when forced by observed oceanic boundary conditions ~~(Krinner et al., 2008; Ashfaq et al., 2011; Hernández-Díaz et al., 2017)~~ (Krinn

Similarly, even though bias correction methods have some limitations, for future climate experiments, there are good reasons to believe that simulations produced using bias-corrected oceanic forcings bear reduced uncertainties with respect to simulations realized with “raw” oceanic forcings from coupled model scenarios such as those from the CMIP5 experiments.

Bias-corrected oceanic forcings can be used to force a regional climate model (RCM), but in this case an additional modelling step has to be carried out, as bias-corrected oceanic forcings should be used to force an atmosphere only GCM that will provide atmospheric lateral boundary conditions for the RCM in order to ensure the consistency between oceanic and atmospheric forcings, such as in Hernández-Díaz et al. (2017). In this framework, the use of a variable resolution GCM which allows to directly use bias-corrected oceanic forcings and downscale ~~future climate experiments~~ climate scenarios is an alternative worth considering, as it also allows two-way interactions between the downscaled regions and the general atmospheric circulation.

5 Conclusions

In this paper, we reviewed existing methods for bias correction of SST and SIC and proposed new ones, such as the analog method for sea-ice. We also proposed validation methods that allow objectively evaluating bias-correction methods with the use of a perfect model test and real-case applications.

The bias-correction of SST is an issue that has already been widely addressed in recent papers and its importance for the modeling and downscaling of future climate scenarios has been demonstrated for multiple regions of the world. In our analysis, we were able to demonstrate the reliability and the suitability of absolute anomaly and quantile-quantile methods for the bias correction of future SST scenarios.

The bias correction of SIC is a more difficult issue to address. With the analog method, we propose a method that shows promising results in most cases and that allows reconstructing future SIC with a realistic frequency distribution in the future. However, the fact that the relative anomaly between an AOGCM ~~future scenario and the~~ scenario and its historical simulation is also used in this method in order to determine future targeted sea-ice extent and area, prevent from bias-correcting cases where sea-ice disappears entirely in a given sector or even an hemisphere. Despite the absence of a perfect and definite answer to this issue, we propose a new and improved method as well as a convenient, objective way to evaluate bias correction methods for ~~future~~ climate scenarios. ~~We draw the attention on the bias-correction of sea-ice that~~ The bias correction of sea ice is currently somewhat overlooked by the community. The application of a multivariate bias correction method (Cannon, 2016) is also a perspective that could help with the bias correction of SST and SIC ~~future~~ projected scenarios at the same time. Nevertheless, corrected SIC using the analog method ~~represent~~ represents a substantial improvement with respect to other previously existing bias-correction methods for sea-ice scenarios and will therefore be made available to anyone willing to use them as forcing for bias-corrected downscaling experiments.

Code and data availability. FORTRAN code enabling the generation of bias-corrected future SST and SIC using CMIP5 scenarios and PCMDI data as input are publicly available for each method via *https* transfer (https://mycore.core-cloud.net/index.php/s/3Lo3Tlr9wsyUGjk) or *ftp* transfer (ftp://ftp.lthe.fr/pub/beaumet/Sourcecode_SSTSICmethods.tar.gz). Bias-corrected future CMIP5 scenarios (rcp4.5 and 8.5) realized within the frame of this study (IPSL-CM5A-LR and CNRM-CM5) are available as well (https://mycore.core-cloud.net/index.php/s/Q1cIsS71Mo4vC or ftp://ftp.lthe.fr/pub/beaumet/Data_BCSST-SIC.tar.gz).

Appendix A: A ~~simple diagnostic parameterization of sea-ice thickness for AGCM simulations~~Bias correction methods : Sea Surface Temperatures

A1 ~~Introduction - general remarks~~

~~Atmospheric circulation models (AGCMs or regional climate models) require information about the state of~~

A0.1 Anomaly method

This frequently used method (e.g., Krinner et al., 2008) simply consists of adding the SST anomaly coming from the difference between a coupled AOGCM projection and the ~~sea surface as a lower boundary condition. While much attention has been paid to sea-surface temperature (SST) and sea-ice concentration (SIC) in that respect, the issue of prescribing correct (or at least reasonable) sea-ice thickness (SIT) has been somewhat neglected historically. While there is a considerable body of scientific literature on the effect of varying SST and SIC on simulated climate, only very few studies focused on the role of varying SIT in atmosphere-only simulations. The authors are aware of three such studies (Gerdes, 2006; Krinner et al., 2010; Semmler et al., 2016) . Gerdes (2006) concluded that "realistic sea ice thickness changes can induce atmospheric signals that are of similar magnitude as those due to changes in sea ice cover", while Krinner et al. (2010) show that the impact of a variable sea-ice thickness compared to a uniform value is essentially limited to the cold seasons and the lower troposphere, and that sea-ice thickness changes have a significant impact also in the context of climate change simulations. Near-surface temperature changes of the order of a few °C are observed in response to the replacement of a uniform thick Arctic sea-ice cover by variable sea-ice thickness. In this note, a simple diagnostic parameterization initially developed by Krinner et al. (1997) is discussed and evaluated against new Arctic and Antarctic sea-ice thickness data that were not available in the mid-90s.The idea is to propose a simple parameterization of sea-ice thickness that can be used in a variety of climate modelling applications, in particular for AGCM or RCM simulations of climate conditions different than today, from palaeoclimate studies to climate projections.In these applications, this parameterization can be particularly useful in cases where future sea-surface conditions (SST, SIC and SIT)are not directly prescribed from a coupled ESM run, but rather obtained using a bias correction method.~~

A1 ~~Methods~~

A0.1 ~~Diagnosing sea-ice thickness from sea-ice concentration~~

~~As described by Krinner et al. (2010) , the parameterization of sea-ice thickness h_S as a function of the local instantaneous sea-ice fraction f is designed such as to yield h_S of the order of 3 meters for multi-year sea ice (deemed to be dominant when the local annual minimum fraction $f_{min} \gg 0$) and h_S below 60cm (with a stronger annual cycle) in regions where sea-ice completely disappears in summer (that is, $f_{\text{min}} = 0$), and intermediate values for intermediate cases:~~

$$h_S = (c_1 + c_2 f_{min}^2) \cdot (1 + c_3(f - f_{min}))$$

~~with c_1=0.2m, c_2=2.8m and c_3=2. This corresponds to the observed characteristics of Arctic and Antarctic sea ice, with multi-year sea ice being generally much thicker than first year ice. The parameter c_3 introduces a seasonal ice thickness variation in areas where there is a concomitant seasonal cycle of sea-ice concentration. A more parsimonious, simply bilinear formulation could have been designed to comply with these constraints. However, for the sake of consistency with previous work, we used the equation proposed by Krinner et al. (1997) who designed the parameterization such as to allow for a fairly strong seasonal cycle of sea-ice thickness also in regions with intermediate values of f_{min}. Figure 3 (from Krinner et al. (2010)) illustrates diagnosed Arctic sea-ice thickness for the present and for the end of the 21st century (2081-2100) using bias-correction applied to sea-ice concentrations from a coupled ESM SRES-A1B simulation (Krinner et al., 2008) .~~

~~Prescribed annual mean Arctic sea-ice thickness (in m) in an AGCM climate change experience with bias-corrected sea-surface conditions, using the proposed diagnostic parameterization (figure from Krinner et al. (2010)). Left: present (1981-2000), right: SRES-A1B for 2081-2100. Bias correction for SST and SIC after Krinner et al. (2008)~~ corresponding historical simulation to the present-day observations. In practice, for each grid point, the difference between the SST for a given month in the future from a climate change simulation and the climatological mean SST in the corresponding historical simulation from the same coupled AOGCM is added to the observed climatological mean SST (e.g . PCMDI, 1971-2000):

$$SST_{Fut,est} = \overline{SST_{obs}} + \left(SST_{Fut,AOGCM} - \overline{SST_{Hist,AOGCM}}\right) \qquad \text{(A1)}$$

In (A1), $SST_{Fut,est}$ is the estimated future SST for a given month, $\overline{SST_{obs}}$ the observed climatological monthly mean, $SST_{Fut,AOGCM}$ the model future SST for a given month in the future AOGCM scenario and $\overline{SST_{Hist,AOGCM}}$ the model climatological monthly mean in the AOGCM historical simulation for the same reference period as for the observed climatology. As a result, the reconstructed SST time series has the chronology of the AOGCM projected scenario.

A0.1 ~~Data~~

~~In the following, we used sea-ice concentration data extracted from the ERA-Interim output; this is typically the kind of data that would be used in AGCM or RCM simulations.Lindsay and Schweiger (2015) recently proposed a 15-parameter spatial and temporal regression of Arctic sea-ice thickness observations from submarines, aircraft and satellites. We will use these observations here. Kurtz and Markus (2012) have deduced Antarctic SIT from ICESat data for the period 2003-2008. Although observations with autonomous underwater vehicles by Williams et al. (2015) tend to suggest occurrence of thicker Antarctic sea-ice than previously acknowledged, we will use the Kurtz and Markus (2012) data because of their large spatial coverage.~~

A1 ~~Results~~

~~The original formulation by Krinner et al. (1997) was parameterized for both hemispheres. We will therefore first present results for the original unique parameter set $c_{1,2,3}$ applied to both hemispheres. In a second step, we will present results for separate Arctic and Antarctic parameter sets, yielding a better fit to the observations. The reasoning is that, at the expense of generality of the diagnostic parameterization, one could argue that the strong difference between the Arctic and Antarctic geographic configuration — a closed small ocean favouring ice ridging and thus thicker sea ice in the Arctic, and large open~~

~~ocean favouring thinner sea ice around Antarctica — justifies choosing different parameter sets for the two hemispheres. As the position of the continents will not change over the time scales of interest here, climate change experiments will not be adversely affected by this loss of generality.~~

A0.1 ~~Option 1: Global parameter set~~Quantile-quantile method

~~A comparison between the observed (Lindsay and Schweiger, 2015) and our diagnosed evolution of the Arctic mean sea-ice thickness is given in Figure 11. The geographical patterns of the observed (in fact, observation-regressed) and parameterized Arctic ice thickness for March and September over the observation period 2000-2013 (Figure 12) do bear some resemblance, but they also show some clear deficiencies of the diagnostic parameterization. The diagnostic parameterization reproduces high sea-ice thickness north of Greenland and the Canadian Archipelago, linked to persistent strong ice cover, but underestimates maximum ice thickness (due in part to compression caused by the ocean surface current configuration). Thinner sea ice over the seasonally ice-free parts of the basin is reproduced, but it is actually too thin, particularly in winter (for example in the Chukchi Sea). Obvious artifacts appear in September north of about 82°N where the SIC in the ERA-Interim data set clearly bears the signs of limitations due to the absence of satellite data. Both for spring (Oct-Nov) and fall (May-Jun), our diagnosed SIT (Figure 13) compares generally well with the ICESat data except for an overestimate in the Weddell Sea, at both seasons. The geographical pattern of alternating regions with thin and thick sea ice is remarkably well reproduced.~~

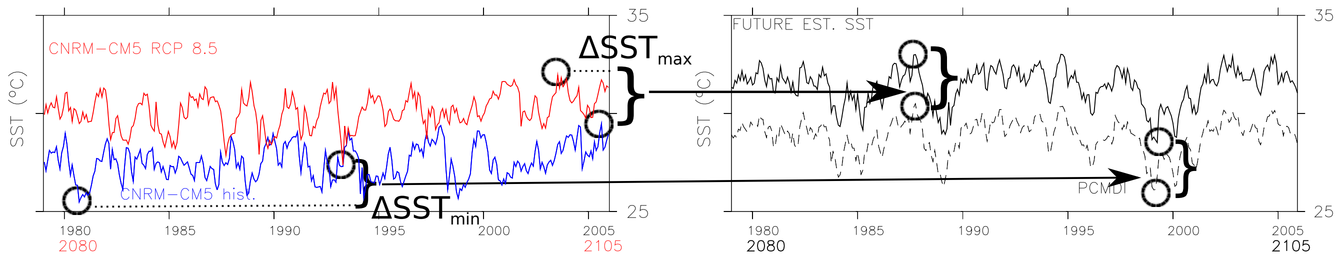
~~Observed (black, after Lindsay and Schweiger, 2015) and diagnosed (red) 12-month moving average mean sea-ice thickness of the Arctic basin (see Figure 12). The global parameter set is used here. Slight differences to Figure 13 of Lindsay and Schweiger (2015) ap because here we mask ice-free (SIC < 15%) areas that have a finite, non-zero ice thickness in the regression proposed by Lindsay and Schweiger (2015) who extend their regression to the entire Arctic Basin at all seasons.~~

~~Observed (regressed, Lindsay and Schweiger (2015)) and parameterized Arctic sea-ice thickness (in m) for March and September, and difference between these (right), with the global parameter set.~~

~~Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for March and September, and difference between these (right), with the global parameter set.~~

A0.2 ~~Option 2: Separate Arctic and Antarctic parameter sets~~

~~A slightly better fit for the two poles can be obtained with separate parameters sets. For the Arctic, it seems desirable to increase winter sea-ice thickness in the Chukchi Sea area (by increasing c_3 slightly) and to decrease the average sea-ice thickness over the Central Arctic (by decreasing c_2). Figures 14 and 15 show results for the Arctic with c_1=0.2m, c_2=2.4m and c_3=3. The spatial fit is slightly better, but~~ This method has been proposed and described in Ashfaq et al. (2011) It consists of adding, for each grid point and each calendar month's quantile in the observations, the ~~recent Arctic-mean decadal tendency towards decreased average sea-ice thickness is somewhat less well reproduced. For the Antarctic, the main feature to improve is the maximum ice thickness in the Weddell Sea, which can be decreased by decreasing c_2 to 2.0m. The Antarctic parameter set then becomes c_1=0.2m, c_2=2m~~ corresponding quantile change in the GCM data set, i.e. the difference between the maximum SST in the projected scenario and in the historical simulation, between the second highest SSTs in the two simulations, and



~~Observed Lindsay and Schweiger (2015) and parameterized Arctic sea-ice thickness (in m) for March and September, and difference between these (right), with the Arctic-specific parameter set.~~

Figure A1. Illustration of the quantile-quantile method for min. and max. of SST time series for a grid point in the Central Pacific : GCM historical simulation (blue, left), GCM projected scenario (red, left), observed SST(dashed, right), reconstructed future SST (thick, right)

~~c_3=2. The result~~ so on for each ranked SST quantile. However, unlike Ashfaq et al. (2011) , we did not create a new SST field for the present by replacing SST from the GCM in the historical period by its corresponding quantile in the observations, but we directly added the quantile change to the corresponding quantile of the observational time series (Figure ~~16)is indeed a decreased thickness of the perennial Weddell Sea ice with little impact elsewhere. In any case, these hemisphere-specific sea-ice parameter sets are not very different from each other and fairly similar to the original formulation.~~

~~Observed (black, after Lindsay and Schweiger (2015)) and diagnosed (red) 12-month moving average mean sea-ice thickness of the Arctic basin with the Arctic-specific parameter set.~~ A1). This conserves the chronology of the observations and their inter-annual variability in estimated SSTs for the future. In our results, we noticed a large fine-scale spatial variability of the constructed bias-corrected SSTs that was due to the large spatial variability of the climate change increments (quantile change) calculated individually for each pixel. To fix this, we applied a slight spatial filtering (3 grid point Hann box filter (Blackman and J.W., 1959)) of the quantile shifts in order to produce more consistent SST fields.

~~Observed Kurtz and Markus (2012) and parameterized Antarctic sea-ice thickness (in m) for March and September, and difference between these (right), with the Antarctic-specific parameter set.~~

A1 ~~Discussion and conclusion~~

~~Given the simplicity of the proposed diagnostic sea-ice thickness parameterization, the results are, at least in some aspects such as the predicted average Arctic sea-ice thinning, surprisingly good. The Central Arctic sea-ice thickness results are clearly adversely affected by the input sea-ice concentrations north of 82°N. Arctic winter sea-ice thickness in the marginal seas appears underestimated. In the Antarctic, the spatial pattern of SIT is very well represented. We think that in absence of pan-Arctic and pan-Antarctic satellite-based data before approximately 2000, this parameterization can serve as a surrogate for earlier periods, and that it can, because it seems to have predictive power, also serve for climate change experiments with AGCMs or RCMs. Because of its simplicity, implementing this parameterization should not be too complicated in any case provided the model does explicitly take into account sea-ice thickness in its computations of heat flow through sea ice. In that~~

~~case, sea-ice thickness can either be calculated online (with the need to keep track of annual minimum sea-ice thickness during the execution of the code) or be input as a daily boundary condition along with the sea-ice concentrations. Of course, another possibility would be to prescribe sea-ice thickness anomalies from coupled models. In this case, it would probably be wise to compute the prescribe SIT using relative sea-ice thickness changes. For example, in a climate change experiment, this would read $h_{presc}(t) = h_{obs,\,2003-2008} \cdot h_{sim}(t) / h_{sim,\,2003-2008}$. In any case, it is very probable that Arctic sea ice thickness will further decrease as multi-year sea ice will be replaced by a predominantly seasonal sea-ice cover. This should probably be taken into account in future CORDEX- or HighResMip-style climate simulations, given the non-negligible impact of sea-ice thinning on winter heat fluxes in particular.~~

Competing interests. The authors have no competing interests.

Acknowledgements. This study was funded by the Agence Nationale de la Recherche through contract ANR-14-CE01-0001-01 (ASUMA). 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 participating to CMIP5 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.
R. Haarsma was funded through the PRIMAVERA project under Grant Agreement 641727 in the European Commission's Horizon 2020 research program.

References

Agosta, C., Favier, V., Krinner, G., Gallée, H., Fettweis, X., and Genthon, C.: High-resolution modelling of the Antarctic surface mass balance, application for the twentieth, twenty first and twenty second centuries, Climate Dynamics, 41, 3247–3260, https://doi.org/10.1007/s00382-013-1903-9, 2013.

Ashfaq, M., Skinner, C. B., and Diffenbaugh, N. S.: Influence of SST biases on future climate change projections, Climate Dynamics, 36, 1303–1319, https://doi.org/10.1007/s00382-010-0875-2, 2011.

Baumberger, C., Knutti, R., and Hirsch Hadorn, G.: Building confidence in climate model projections: an analysis of inferences from fit, Wiley Interdisciplinary Reviews: Climate Change, 8, e454–n/a, https://doi.org/10.1002/wcc.454, 2017.

Blackman, R. and J.W., T.: Particular pairs of windows, in: The measurement of power spectra, from the point of view of communications engineering., edited by Blackman, R. and J.W., T., chap. B5, pp. 95–100, Dover Publications, inc., New York, 1959.

Brekke, L. D., Dettinger, M. D., Maurer, E. P., and Anderson, M.: Significance of model credibility in estimating climate projection distributions for regional hydroclimatological risk assessments, Climatic Change, 89, 371–394, https://doi.org/10.1007/s10584-007-9388-3, 2008.

Cannon, A. J.: Multivariate bias correction of climate model output: Matching marginal distributions and intervariable dependence structure, Journal of Climate, 29, 7045–7064, 2016.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., et al.: Climate change 2013: the physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Sea level change, pp. 1137–1216, 2013.

de Elía, R., Laprise, R., and Denis, B.: Forecasting Skill Limits of Nested, Limited-Area Models: A Perfect-Model Approach, Monthly Weather Review, 130, 2006–2023, https://doi.org/10.1175/1520-0493(2002)130<2006:FSLONL>2.0.CO;2, 2006.

Denis, B., Laprise, R., Caya, D., and Côté, J.: Downscaling ability of one-way nested regional climate models: the Big-Brother Experiment, Climate Dynamics, 18, 627–646, https://doi.org/10.1007/s00382-001-0201-0, 2002.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrology and Earth System Sciences, 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W. J., Cox, P., Driouech, F., Emori, S., Eyring, V., et al.: Evaluation of Climate Models. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Climate Change 2013, 5, 741–866, 2013.

Gerdes, R.: Atmospheric response to changes in Arctic sea ice thickness, Geophysical Research Letters, 33, n/a–n/a, https://doi.org/10.1029/2006GL027146, 2006.

Giorgi, F. and Gutowski, W. J.: Coordinated Experiments for Projections of Regional Climate Change, Current Climate Change Reports, 2, 202–210, https://doi.org/10.1007/s40641-016-0046-6, 2016.

Haarsma, R. J., Hazeleger, W., Severijns, C., de Vries, H., Sterl, A., Bintanja, R., van Oldenborgh, G. J., and van den Brink, H. W.: More hurricanes to hit western Europe due to global warming, Geophysical Research Letters, 40, 1783–1788, https://doi.org/10.1002/grl.50360, 2013.

Haarsma, R. J., Roberts, M. J., Vidale, P. L., Senior, C. A., Bellucci, A., Bao, Q., Chang, P., Corti, S., Fuckar, N. S., Guemas, V., et al.: High resolution model intercomparison project (HighResMIP v1. 0) for CMIP6, Geoscientific Model Development, 9, 4185, 2016.

Haralick, R. M., Sternberg, S. R., and Zhuang, X.: Image Analysis Using Mathematical Morphology, IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-9, 532–550, https://doi.org/10.1109/TPAMI.1987.4767941, 1987.

Hawkins, E., Robson, J., Sutton, R., Smith, D., and Keenlyside, N.: Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach, Climate Dynamics, 37, 2495–2509, https://doi.org/10.1007/s00382-011-1023-3, 2011.

Hernández-Díaz, L., Laprise, R., Nikiéma, O., and Winger, K.: 3-Step dynamical downscaling with empirical correction of sea-surface conditions: application to a CORDEX Africa simulation, Climate Dynamics, 48, 2215–2233, https://doi.org/10.1007/s00382-016-3201-9, 2017.

Hewitson, B. C., Daron, J., Crane, R. G., Zermoglio, M. F., and Jack, C.: Interrogating empirical-statistical downscaling, Climatic Change, 122, 539–554, https://doi.org/10.1007/s10584-013-1021-z, 2014.

Holland, G., Done, J., Bruyere, C., Cooper, C., and Suzuki, A.: Model Investigations of the Effects of Climate Variability and Change on Future Gulf of Mexico Tropical Cyclone Activity, in: Proceeding of the Offshore Technology Conference, Houston, Texas, https://doi.org/10.4043/20690-MS, 2010.

Krinner, G., Genthon, C., Li, Z.-X., and Le Van, P.: Studies of the Antarctic climate with a stretched-grid general circulation model, Journal of Geophysical Research: Atmospheres, 102, 13 731–13 745, https://doi.org/10.1029/96JD03356, http://dx.doi.org/10.1029/96JD03356, 1997.

Krinner, G., Guicherd, B., Ox, K., Genthon, C., and Magand, O.: Influence of Oceanic Boundary Conditions in Simulations of Antarctic Climate and Surface Mass Balance Change during the Coming Century, Journal of Climate, 21, 938–962, https://doi.org/10.1175/2007JCLI1690.1, 2008.

Krinner, G., Rinke, A., Dethloff, K., and Gorodetskaya, I. V.: Impact of prescribed Arctic sea ice thickness in simulations of the present and future climate, Climate Dynamics, 35, 619–633, https://doi.org/10.1007/s00382-009-0587-7, 2010.

Krinner, G., Largeron, C., Ménégoz, M., Agosta, C., and Brutel-Vuilmet, C.: Oceanic Forcing of Antarctic Climate Change: A Study Using a Stretched-Grid Atmospheric General Circulation Model, Journal of Climate, 27, 5786–5800, https://doi.org/10.1175/JCLI-D-13-00367.1, 2014.

Kurtz, N. and Markus, T.: Satellite observations of Antarctic sea ice thickness and volume, Journal of Geophysical Research: Oceans, 117, 2012.

Lenaerts, J. T. M., Vizcaino, M., Fyke, J., van Kampenhout, L., and van den Broeke, M. R.: Present-day and future Antarctic ice sheet climate and surface mass balance in the Community Earth System Model, Climate Dynamics, 47, 1367–1381, https://doi.org/10.1007/s00382-015-2907-4, 2016.

Levine, R. C., Turner, A. G., Marathayil, D., and Martin, G. M.: The role of northern Arabian Sea surface temperature biases in CMIP5 model simulations and future projections of Indian summer monsoon rainfall, Climate Dynamics, 41, 155–172, https://doi.org/10.1007/s00382-012-1656-x, 2013.

Li, G. and Xie, S.-P.: Tropical Biases in CMIP5 Multimodel Ensemble: The Excessive Equatorial Pacific Cold Tongue and Double ITCZ Problems, Journal of Climate, 27, 1765–1780, https://doi.org/10.1175/JCLI-D-13-00337.1, 2014.

Lindsay, R. and Schweiger, A.: Arctic sea ice thickness loss determined using subsurface, aircraft, and satellite observations, The Cryosphere, 9, 269–283, 2015.

Maurer, E. P., Das, T., and Cayan, D. R.: Errors in climate model daily precipitation and temperature output: time invariance and implications for bias correction, Hydrology and Earth System Sciences, 17, 2147–2159, https://doi.org/10.5194/hess-17-2147-2013, 2013.

Mitchell, D., Achutarao, K., Allen, M., Bethke, I., Beyerle, U., Ciavarella, A., Forster, P., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, N., Massey, N., Fischer, E., Schleussner, C.-F., Scinocca, J., Seland, A., and Rashyd, Z.: Half a degree additional warming, prognosis and projected impacts (HAPPI): Background and experimental design, Geoscientific Model Development, 10, 571–583, 2017.

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., et al.: The next generation of scenarios for climate change research and assessment, Nature, 463, 747, 2010.

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, https://doi.org/10.5194/tc-8-1871-2014, 2014.

Richter, I., Xie, S.-P., Behera, S. K., Doi, T., and Masumoto, Y.: Equatorial Atlantic variability and its relation to mean state biases in CMIP5, Climate Dynamics, 42, 171–188, https://doi.org/10.1007/s00382-012-1624-5, 2014.

Sato, T., Kimura, F., and Kitoh, A.: Projection of global warming onto regional precipitation over Mongolia using a regional climate model, Journal of Hydrology, 333, 144 – 154, https://doi.org/10.1016/j.jhydrol.2006.07.023, 2007.

Screen, J. A. and Simmonds, I.: The central role of diminishing sea ice in recent Arctic temperature amplification, Nature, 464, 1334, 2010.

Semmler, T., Jung, T., and Serrar, S.: Fast atmospheric response to a sudden thinning of Arctic sea ice, Climate Dynamics, 46, 1015–1025, 2016.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M., and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, Geophysical Research Letters, 39, n/a–n/a, https://doi.org/10.1029/2012GL052676, 2012.

Taylor, K., Williamson, D., and Zwiers, F.: The sea surface temperature and sea-ice concentration boundary condition for AMIP II simulations, PCMDI Rep. 60, Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, Livermore, CA, 25 pp., 2000.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498, 2012.

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, Journal of Climate, 26, 1473–1484, https://doi.org/10.1175/JCLI-D-12-00068.1, 2013.

Williams, G., Maksym, T., Wilkinson, J., Kunz, C., Murphy, C., Kimball, P., and Singh, H.: Thick and deformed Antarctic sea ice mapped with autonomous underwater vehicles, Nature Geoscience, 8, 61–67, 2015.

Zhang, L. and Zhao, C.: Processes and mechanisms for the model SST biases in the North Atlantic and North Pacific: A link with the Atlantic meridional overturning circulation, Journal of Advances in Modeling Earth Systems, 7, 739–758, https://doi.org/10.1002/2014MS000415, 2015.