Manuscript prepared for The Cryosphere
with version 4.2 of the LaTeX class copernicus.cls.
Date: 13 December 2012

How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent?

V. Zunz[1], **H. Goosse**[1], **and F. Massonnet**[1]

[1]Georges Lemaître Centre for Earth and Climate Research, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium

Abstract. Observations over the last 30 yr have shown that the sea ice extent in the Southern Ocean has slightly increased since 1979. Mechanisms responsible for this positive trend have not been well established yet. In this study, we tackle two related issues: is the observed positive trend compatible with the internal variability of the system and do the models agree with what we know about the observed internal variability? For that purpose, we analyze the evolution of sea ice around the Antarctic simulated by 24 different general circulation models involved in the 5th Coupled Model Intercomparison Project (CMIP5), using both historical and hindcast experiments. Our analyses show that CMIP5 models respond to the forcing, including the one induced by stratospheric ozone depletion, by reducing the sea ice cover in the Southern Ocean. Some simulations display an increase in sea ice extent similar to the observed one. According to models, the observed positive trend is compatible with internal variability. However, models strongly overestimate the variance of sea ice extent and the initialization methods currently used in models do not improve systematically the simulated trends in sea ice extent. On the basis of those results, a critical role of the internal variability in the observed increase in the sea ice extent in the Southern Ocean could not be ruled out but current models results appear inadequate to test more precisely this hypothesis.

1 Introduction

The way climate models reproduce the observed characteristics of sea ice has received a lot of attention (e.g. Flato, 2004; Arzel et al., 2006; Parkinson et al., 2006; Lefebvre and Goosse, 2008a; Sen Gupta et al., 2009). One conclusion of those studies is that the models skill is higher in the Northern Hemisphere than in the Southern Hemisphere. In particular, simulations performed for the 3rd Coupled Model Intercomparison Project (CMIP3) are generally able to reproduce relatively well the timing of the seasonal cycle of Southern Ocean sea ice extent but fail in simulating the observed amplitude (Parkinson et al., 2006). Furthermore, the models are usually unable to simulate the observed increase in Southern Ocean sea ice extent (e.g. Arzel et al., 2006; Parkinson et al., 2006), which is estimated to be of $11\,200 \pm 2680\,\mathrm{km}^2\,\mathrm{yr}^{-1}$ between 1979 and 2006 (Comiso and Nishio, 2008). At the regional scale, the 1979–2006 trend in observed sea ice extent is positive in all the sectors of the Southern Ocean, except in the Bellingshausen-Amundsen Seas sector, and the Ross Sea sector exhibits the largest positive trend (e.g. Cavalieri and Parkinson, 2008; Comiso and Nishio, 2008). Lefebvre and Goosse (2008a) have studied the trend simulated by several CMIP3 models in the different sectors of the Southern Ocean and they have shown that these models were not able to reproduce this observed spatial structure.

The observed increase in sea ice extent during the past decades is statistically significant at the 95% significant level (e.g. Cavalieri and Parkinson, 2008). However, its potential causes are still debated. We do not know the part of this trend that can be attributed to external forcing and the one that is due to natural variability. This issue has already been addressed for the Arctic sea ice extent (e.g. Kay et al., 2011) but remains poorly investigated for the Southern Ocean sea ice.

Several studies dealing with the potential role of the forced response have pointed out the relationship between stratospheric ozone depletion over the past few decades (Solomon, 1999) and changes in the atmospheric circulation at high latitudes (e.g. Turner et al., 2009; Thompson et al., 2011). Indeed, variations of sea ice extent in the Southern Ocean are strongly influenced by changes in the atmosphere circulation (e.g. Holland and Raphael, 2006; Goosse et al., 2009b). However, the link between atmospheric circulation and the

Correspondence to: Violette Zunz (violette.zunz@uclouvain.be)

sea ice extent integrated over the Southern Ocean is not straightforward (e.g. Lefebvre and Goosse, 2008b; Stammerjohn et al., 2008; Landrum et al., 2012) and several recent studies came to the conclusion that the stratospheric ozone depletion does not lead to an increase in the sea ice extent (e.g. Sigmond and Fyfe, 2010; Smith et al., 2012; Bitz and Polvani, 2012). A second potential cause of the observed expansion of sea ice cover relies on an enhanced stratification of the ocean which would inhibit the heat transfer to the surface. This strengthened stratification is mainly due to a freshening of the surface water, triggered by an increase in the precipitation over the Southern Ocean, the melting of the ice shelf and changes in the production and transport of sea ice (e.g. Bitz et al., 2006; Zhang, 2007; Goosse et al., 2009b; Kirkman and Bitz, 2010). Liu and Curry (2010) pointed out that an enhanced hydrological cycle may also increase the snowfalls at high latitudes in the Southern Ocean. In that case, the snow cover on thicker sea ice would raise the surface albedo, strengthen the insulation between the atmosphere and the ocean, and thus would protect the sea ice from melting. Nevertheless, this mechanism mainly impacts thick ice because for thin ice, the higher snow load leads to seawater flooding and to the formation of snow ice. This decreases the effect of the initial increase in snow thickness.

Another hypothesis suggests that the positive trend in the Southern Ocean sea ice extent could arise from the internal variability of the system that masks the warming signal in the Southern Ocean that should characterize the response to an increase in greenhouse gases concentration, according to climate models. In this framework, some recent studies have drawn the attention on the importance of distinguishing the lack of agreement between models from the lack of significant signal (e.g. Tebaldi et al., 2011; Deser et al., 2012). A trend can be significant from a statistical point of view, i.e. if it is above a threshold of significance computed through a statistical test. This does not imply that its value is outside of the range that can be reached by the internal variability. For instance, Landrum et al. (2012) have pointed out that large interannual variability in simulated sea ice concentration leads to late 20th Century trends in sea ice concentration that are not always statistically significant for individual members of an ensemble simulation. The observed positive trend of Southern Ocean sea ice extent is statistically significant at the 95% level for the last 30 yr (e.g. Cavalieri and Parkinson, 2008). However, this time period is too short to properly assess the multidecadal variability of the system. Consequently, we cannot estimate if this trend is exceptional or if similar conditions have already occurred many times in the recent past. The period spanning the last 30 yr during which sea ice cover slightly expanded in the Southern Ocean might follow a large melting that may have happened before 1979 (e.g. de la Mare, 1997, 2009; Cavalieri et al., 2003; Curran et al., 2003; Cotté and Guinet, 2007; Goosse et al., 2009b). This suggests that multidecadal variability in the Southern Ocean is large but the available data does not allow a quantitative estimate of its value. The results from model simulations appear thus to be crucial to balance this lack of observations. Provided that models are compatible with the available observations, they can help addressing the issue whether the observed positive trend in the Antarctic sea ice extent is due to external forcing or to internal variability, or to both of them.

The decreasing trend in many model simulations may be due to a misrepresentation of the response of the circulation and/or of the hydrological cycle to the forcing. Alternatively, the observed changes may belong to the range of the trends that can be attributed to the internal variability of the system. In this hypothesis, the positive trend observed over the last decades is just one particular realization among all the possible ones. A negative trend in one model simulations does not imply necessarily a disagreement between model and data as another simulation with the same model (another member of an ensemble, for instance) would likely display a positive one. Furthermore, if this is valid and if the internal variability is to some extent predictable, an adequate initialization of the system could lead to a better simulation of the evolution of the sea ice cover around the Antarctic.

In this paper, we examine outputs from general circulation models (GCMs) following the 5th Coupled Model Intercomparison Project (CMIP5) protocol. To further study the role of the internal variability in the increasing trend in sea ice extent in the Southern Ocean and in the apparent disagreements between models and observations, we deal with two kinds of simulations: historical and hindcast (or decadal) simulations. The first ones are driven by external forcing and are initialized without observational constraints. They are used to assess how well each model simulates the observed mean state, variability and trends in sea ice concentration and extent. The objective is to study the possible links between the internal variability of the system and the simulated trend in sea ice extent. Our purpose is, on the one hand, to test if the internal variability of the models agrees with the one of the observations. On the other hand, we check if the observed positive trend stands in the range of trends provided by models internal variability. Analyzing the mean state also appears to be important here because of its impact on the simulated variability (e.g. Goosse et al., 2009a). In addition to those points related to the variability of the system, the way stratospheric ozone is taken into account in models is also discussed to estimate if this has a significant impact on the simulated trends. However, it is out of the scope of this study to discuss specific mechanisms that link the sea ice extent and the stratospheric ozone variations.

The second kind of simulations, the hindcasts, are also driven by external forcing but, in contrast to the historical simulations, are initialized through data assimilation of observations. Consequently, these simulations allow us to assess how the state of the system in the early 80's impacts the variability of the models and their representation of the trend over the last 30 yr. Idealized model studies have shown high potential predictability at decadal time scales in the Southern Ocean (e.g. Latif et al., 2010), i.e. models have deterministic decadal variability, in particular for surface temperatures (Pohlmann et al., 2004). The predictive skill of the models at decadal time scales is also discussed here to see if this potential predictability is confirmed in real applications.

An initial investigation of the results of CMIP5 models has shown that, in agreement with previous studies related to CMIP3 models (e.g. Lefebvre and Goosse, 2008a), current GCMs do not simulate a spatial structure of the trend in sea ice extent similar to the observed one. This spatial structure might as well arise from the internal variability. In such a case, models would not have to fit the observed pattern, as discussed above. However, this remains a hypothesis and we have chosen to focus on the sea ice extent in the whole Southern Ocean rather than in the individual sectors to avoid the additional complexity associated with the spatial structure of the changes. Models and observation data are briefly presented in Sect. 2. The time period we analyze is limited by the available observations. For the Southern Ocean, validation data are quite sparse before 1979. We therefore examine outputs between 1979 and 2005. Results provided by models historical simulations are presented and discussed in Sect. 3. The analyses of hindcast simulations are described in Sect. 4. Finally, Sect. 5 summarizes our results and proposes conclusions.

2 Models and observation data

Models data were obtained from the CMIP5 (Taylor et al., 2011) multi-model ensemble: http://pcmdi3.llnl.gov/esgcet/home.htm. We have analysed results of historical simulations from 24 models which have the required data available. Among these models, 10 of them provide results for hindcast simulations. Both historical and hindcast simulations consist in ensemble simulations of various sizes. Models and their respective modeling groups are listed in Table 1, along with the number of members in each model historical and hindcast simulations. The models have different spatial resolution and representation of physical processes. The spatial resolution of models components is summarized in Table S1 of the Online Supplement Table of this paper. A reference is also given for more complete documentation.

We give specific information on the treatment of ozone in Table 2, as a basis for the discussion presented in Sect. 3.3. The AC&C/SPARC ozone database (Cionni et al., 2011) is used to prescribe ozone in most of the models without an interactive chemistry. In this database, stratospheric ozone for the period 1979–2009 is zonally and monthly averaged. It depends on the altitude and it takes solar variability into account. Whether they have interactive chemistry or prescribed stratospheric ozone, the 24 models analyzed in this study thus take into account the stratospheric ozone depletion in their historical simulations. This is an improvement since the CMIP3 simulations. Indeed, nearly half of the CMIP3 models prescribed a constant ozone climatology (Son et al., 2008). Nevertheless, some of the models have a coarse atmosphere resolution which sometimes does not encompass the whole stratosphere. In that case, processes related to the interaction between radiation and ozone as well as the exchange between the stratosphere and the troposphere may be represented rather crudely.

The hindcast simulations were initialized from a state that has been obtained through a data assimilation procedure, i.e. constrained to be close to some observed fields. There is a large panel of data assimilation methods but most of the models involved in CMIP5 assimilate observations through a nudging. This method consists in adding to the model equations a term that slightly pulls the solution towards the observations (Kalnay, 2007). MIROC4h and MIROC5 incorporate observations in their data assimilation experiments by an incremental analysis update (IAU). Details about this method can be found in Bloom et al. (1996). Table 3 summarizes the data assimilation method corresponding to each model, as well as the variable it assimilates. The relevant documentation was not available to us for CCSM4, FGOALS-g2 and MRI-CGCM3. All the models for which we have the adequate information, except BCC-CSM1.1 and CNRM-CM5, assimilate anomalies. Those anomalies are calculated for both model and observations by subtracting their respective climatology, computed over the same reference period. Working with anomalies does not prevent model biases but it avoids the initialization of the model with a state which is too far from its own climatology and thus limits model drift (e.g. Pierce et al., 2004; Smith et al., 2007; Troccoli and Palmer, 2007; Keenlyside et al., 2008; Pohlmann et al., 2009), as discussed in Sect. 4.

The model skill is measured through its representation of the sea ice concentration (the fraction of grid cell covered by sea ice) and sea ice extent (the sum of the areas of all grid cells having an ice concentration of at least 15 %). We consider the sea ice extent over the whole Southern Ocean and for models, it has been calculated on the original models grids. For each model providing an ensemble of simulations, the model mean is the average over the members belonging to the ensemble. The multi-model mean is then derived by computing the mean of the individual models means, without applying any weighting to the models. Sea ice concentration comes from the satellite observation of the National Snow and Ice Data Centre (NSIDC) (Comiso, 1999, updated 2008). The sea ice extent is then derived from this data set following the method described in Cavalieri et al. (1999) and applied by Cavalieri and Parkinson (2008) for the period 1979–2006.

3 Historical simulations

The historical simulations are driven by external forcing and are initialized without observational constraints. These simulations are here used to assess the mean state and the variability of the models using recent observations.

3.1 Mean state and variability

In a first step, we analyze the mean sea ice concentration over the period 1979–2005. Fig. 1 shows the multi-model mean of sea ice concentration in the Southern Ocean and compares the simulated sea ice edge to the observed one. Results are given for February (September), the month during which the observed sea ice extent reaches its minimum (maximum). In February, the multi-model mean underestimates the sea ice cover in the Belligshausen and Amundsen Seas as well as in the eastern part of the Ross Sea. In the Western Ross Sea and in small parts of the Weddell Sea and of the Indian Ocean sector, the multi-model mean overestimates the sea ice extent. In September, the shape of the sea ice edge computed from multi-model mean roughly fits the observations. However, the multi-model mean overestimates the sea ice cover everywhere except in the Indian Ocean sector and in the Eastern part of the Ross Sea sector.

This reasonable multi-model mean extent is the result of the average of a wide range of individual behaviors. To account for this variety of mean model states, we have plotted, for individual models, the mean of sea ice extent of each month of the year during the period 1979–2005. Figure 2a confirms that the multi-model mean fits quite well the observations, especially during winter months. However, the seasonal cycle of sea ice extent of the various models is largely spread around the observations and the timing of the minimum/maximum sea ice extent varies from one model to the other. In summer, 16 of the models underestimate the sea ice extent. In particular, CNRM-CM5 and MIROC5 are nearly sea ice free during summer. The latter strongly underestimates the ice extent all over the year and its winter sea ice extent is smaller than some models summer sea ice extent. On the contrary, CCSM4 and CSIRO-Mk3.6.0 overestimate the sea ice extent during the whole year, especially during summer. In winter, when the simulated sea ice cover reaches its maximum, the sea ice extent ranges from approximately 5×10^6 to $24\times10^6\ \mathrm{km}^2$ while the observations display a sea ice extent of about $17\times10^6\ \mathrm{km}^2$. 10 models underestimate the sea ice extent in September.

Since the internal variability of the climate system may also have played a role in the observed expansion of sea ice cover, we assess its representation in models by computing the standard deviation of the sea ice extent for each month of the year, over the period 1979–2005 (Fig. 2b). Here, to obtain both the ensemble mean of each model and the multi-model mean of standard deviations, an average of the individual standard deviations has been performed. We have chosen to detrend data before computing the standard deviation in order to suppress the direct impact of a trend on the standard deviation that could obscure our analysis of the potential links between those two variables discussed in Sect. 3.2. The monthly standard deviation indicates that the variability strongly differs between models. In February, 15 models have a standard deviation higher than the observed one and all of the 24 models overestimate the standard deviation during September. Consequently, the multi-model mean of standard deviations does not fit very well the observations. It overestimates the standard deviation all over the year, particularly during winter. The interannual variability in some models is significantly larger during winter months than during summer months. As a result, these models have a pronounced seasonal cycle of their standard deviation, in contrast to the observations, which display a relatively constant value throughout the year.

The analysis of Fig. 2b tells us two important things. On the one hand, it points out the inability of the majority of models to reproduce the observed interannual variability. In particular, they all overestimate the winter interannual variability. On the other hand, it highlights the fact that some models are characterized by a very different magnitude of the interannual variability from one season to the other. In order to avoid a loss of information, we have thus chosen in the following analysis to work with seasonal mean rather than with annual mean and to treat the summer and winter separately.

3.2 Trend over the period 1979–2005

For the historical simulations, we have computed for each member of the ensemble the trend from 1979 to 2005 of summer (average of January, February and March) and winter (average of July, August and September) sea ice extent. Each trend has been computed through a linear regression of the yearly values (between 1979 and 2005) of the summer or winter sea ice extent. In addition to a direct evaluation of model skill, one of our goals is to analyse if a relationship can be established between the mean state, the interannual variability simulated by the model and the ability to reproduce the observed trend.

Observations show that the summer sea ice extent expanded between 1979 and 2005, at a rate of approximately $149\,000\ \mathrm{km}^2$ per decade. This trend is significant at the 90% level. In Fig. 3a, it appears that almost all of the simulations performed with the 24 models fail in simulating the sign of this observed trend. Only three models (FGOALS-g2, GFDL-CM3 and GISS-E2-R) have an ensemble mean with a positive trend while most of them simulate a relatively large negative trend. For four additional models (CCSM4, CSIRO-Mk3.6.0, HadCM3 and MRI-CGCM3), some ensemble members display a positive trend. Among them, CSIRO-Mk3.6.0 and GFDL-CM3 are the only models displaying a positive trend significant at the 90% level in one of their members, as in the observations (see Table S2 of the Online Supplement Table of this paper). Nevertheless, CSIRO-Mk3.6.0 has a mean summer sea ice extent much larger than what is observed while GFDL-CM3 is well below the observations. Moreover, CSIRO-Mk3.6.0 has an interannual variability which is on average twice the one of the observations.

For summer sea ice extent, some given models display a standard deviation that could be quite different between members (Fig. 3b). Besides, the individual means of ensemble members performed with the same model are relatively similar (Fig. 3a). The range of values reached by the trends of the different members belonging to one model's simulation also differs strongly from one model to the other (Fig. 4a). We quantify the various ranges provided by the different models thanks to the ensemble standard deviation of the trends, for models that have at least 3 members in their historical simulations. This ensemble standard deviation of the trends stands between $26\,000\,\mathrm{km}^2$ per decade for MIROC-ESM and $470\,000\,\mathrm{km}^2$ per decade for BCC-CSM1.1 (see Table S2 of the Online Supplement Table of this paper). On average, the ensemble standard deviation of the trend equals $166\,000\,\mathrm{km}^2$ per decade. If we consider this average as an estimate of the range of the trend that can be associated with internal variability, the observed positive trend of $149\,000\,\mathrm{km}^2$ per decade is well among the values that could be due to natural processes alone and compatible with the available ensemble of model results. Nevertheless, given that many models have an interannual variability that is much larger than the one of the observations, it is not sure whether the range of the trends they provide is representative of the reality.

The comparison between the trend, the mean extent and standard deviation does not display any clear link in summer between those variables, some of the models that simulate an increase in the ice extent in at least one of their members overestimating the observed mean and variability, some underestimating it. Figure 3b also underlines the fact that models with little ice during summer often have a small interannual variability of summer sea ice extent, in agreement with results of Goosse et al. (2009a). Moreover, the spread of the sea ice extent trends and standard deviations of members belonging to one model ensemble grows with the mean summer sea ice extent.

Winter sea ice extent has also increased between 1979 and 2005, by approximately $86\,000\,\mathrm{km}^2$ per decade. This trend is not significant at the 90% level. Two models have an ensemble mean whose trend is positive: GFDL-CM3 and IPSL-CM5A-MR (Fig. 3c). The ensemble mean of GFDL-CM3 (5 members) has a positive trend which is close to the observed one but it strongly underestimates the mean winter sea ice extent. It is also an ensemble whose members are highly scattered along the trend axis, three having a positive trend (from approximately 470×10^3 to $1300 \times 10^3\,\mathrm{km}^2\,\mathrm{decade}^{-1}$) and two having a negative one (from approximately -290×10^3 to $-1120 \times 10^3\,\mathrm{km}^2\,\mathrm{decade}^{-1}$). The IPSL-CM5A-MR ensemble is made up of one member only. Its trend and its mean are both close to observations.

The 22 remaining models all have an ensemble mean showing a decrease in winter sea ice extent. However, as noticed for summer, a few of them have ensemble members displaying positive trends (BCC-CSM1.1, CSIRO-Mk3.6.0, IPSL-CM5A-LR and MRI-CGCM3). Two of three BCC-CSM1.1 historical simulation members present a positive trend. The last one has a very negative trend, reaching $-2520 \times 10^3\,\mathrm{km}^2\,\mathrm{decade}^{-1}$. Contrarily, the mean sea ice extent does not vary much between members of BCC-CSM1.1, all of them being larger than the observations. CSIRO-Mk3.6.0 ensemble contains 10 members. They all simulate a mean sea ice extent in winter relatively close to the observations. Only one member shows an increase in sea ice extent.

Figure 3d confirms that all the 24 models overestimate the interannual variability in winter. It also underlines the fact that simulations that have an ensemble mean of the trends close to the observed one have generally a standard deviation which is much larger than the one of the observations. IPSL-CM5A-MR single member, which has a trend and a mean state relatively close to the observations, has a standard deviation greater than $0.8\times10^6\,\mathrm{km}^2$ while the observed standard deviation stands around $0.25 \times 10^6\,\mathrm{km}^2$. GFDL-CM3 is a model that has a very high standard deviation (around 4 times the standard deviation of the observations). It is also a model with a large range of trends reached by its members (Fig. 4b).

For winter sea ice extent, considering again models that have at least 3 members in their historical simulations, the ensemble standard deviation of the trends varies between $100\times10^3\,\mathrm{km}^2\,\mathrm{decade}^{-1}$ for FGOALS-s2 and $1704\times10^3\,\mathrm{km}^2\,\mathrm{decade}^{-1}$ for BCC-CSM1.1 (see Table S3 of the Online Supplement Table of this paper). On average, this ensemble standard deviation of the trends equals $428\,000\,\mathrm{km}^2\,\mathrm{decade}^{-1}$. As for summer, if this value is representative of the range of trends due to internal variability, the observed trend of $86\,000\,\mathrm{km}^2$ per decade appears compatible with natural processes and the model ensemble. However, the model biases in their representation of the variance in winter during the last 30 yr is even larger than in summer, making this estimate of the uncertainty based on model results very questionable.

From this analysis of historical simulations, it appears that among all the simulations analyzed, only a few of them present a positive trend of the sea ice extent, for both summer and winter. 2 members over 85 have a statistically significant positive trend over the last 30 yr in summer (12 have a positive one) and 10 over 85 have a positive trend in winter. Those positive values appear thus as relatively rare events but are within the range of internal variability, according to model results. The important point here is that such positive trends are generally found in models that overestimate the interannual variability. Because of their high interannual variability, such models can provide a large range of possible trends, some of them agreeing with the observations.

3.3 Stratospheric ozone

CMIP5 models all take into account the stratospheric ozone depletion that occurred during the last 30 yr (see Table 2 for details). However, this improvement compared to CMIP3 brought to the stratospheric ozone does not lead to major changes in their representation of the trend in sea ice extent in the Southern Ocean.

To go a step further, we discuss if the way stratospheric ozone is treated has an influence on the results. The models with interactive chemistry (activated during the simulation or used in an offline simulation to compute the ozone dataset) and the ones whith higher atmospheric vertical resolution (≥ 35 layers) have on average a slightly smaller extent of sea ice in summer (Fig. 3a, respectively circle and triangle orange symbols). In winter, the models with high atmospheric resolution underestimate the sea ice extent while the ones with interactive chemistry overestimate it (Fig. 3c). The influence on the trend is hardly detected. This shows that, on average, the inclusion of an interactive chemistry or an increased vertical resolution do not make major differences compared to other models.

Looking now at individual models, we have seen in Sect. 3.2 that CSIRO-Mk3.6.0, GFDL-CM3 and IPSL-CM5A-MR provide results for sea ice extent trend in winter in relatively good agreement with observations but with much too high a standard deviation for GFDL-CM3 and IPSL-CM5A-MR. CSIRO-Mk3.6.0 has a quite coarse resolution in its atmosphere component (18 vertical layers) and prescribes the ozone from the AC&C/SPARC database. GFDL-CM3 and IPSL-CM5A-MR have a finer resolution (48 and 39 layers, respectively). They both have interactive chemistry but IPSL-CM5A-MR treats the interaction between ozone and climate through a semi offline approach. Again, from the available ensemble, the representation of ozone in models does not seem to be the dominant factor influencing the simulation of the trend in sea ice extent.

4 Hindcast simulations

We have shown in Sect. 3 that the lack of agreement between simulated and observed variance over the last 30 yr does not allow us to confidently establish the link between the internal variability and the positive trend found in observations of the sea ice extent. Nevertheless, if this link exists and if the internal variability in the Southern Ocean is in some way predictable, an adequate initialization of the system should improve the results of the simulated evolution of the sea ice extent. This hypothesis is tested in this section using the hindcast simulations performed in the framework of CMIP5. In contrast to the historical simulations, the hindcasts are initialized through data assimilation of observations. The data assimilation method and the variables assimilated vary from one model to the other, as summarized in Table 3.

4.1 Impact of the initialization on the simulated trends

The models used for the hindcast analysis have been chosen on the basis of the availability of their results. Fortunately, we see on Fig. 2 that these 10 models (dotted lines) constitute a subset which represents reasonably well the variety of general circulation models. In order to outline the effect of the initialization on the simulated trend in sea ice extent, for each model, we have computed the ensemble mean of the trends in hindcast simulations spanning the period 1981–2005, for winter and summer extent, and compared them to the ones from historical simulations (i.e. uninitialized), over the same time period. This period has been chosen as no hindcast was started in 1979. Here, the hindcasts were initialized in January 1981 for all the models except HadCM3, whose hindcast members were started in November 1980. On Fig. 5 showing the trend in sea ice extent computed from hindcast simulations against the one computed from historical simulations, a dot located on the line $y(x) = x$ means that the trend in hindcast simulation equals the one of historical simulation. If the trend simulated by hindcast is greater (smaller) than the one computed from historical simulation, then the dot will be above (below) the line $y(x) = x$.

Regarding summer sea ice extent (Fig. 5a), the initialization through a data assimilation procedure does not improve systematically the simulated trend. HadCM3, MIROC4h and MRI-CGCM3 hindcasts trends are closer to the observation than are their historical trends but they remain negative. BCC-CSM1.1, CNRM-CM5, IPSL-CM5A-LR and MPI-ESM-LR simulate a more negative trend in their hindcasts than in their historical runs. FGOALS-g2 has a largely positive trend in its hindcast while the trend in its historical simulation is slightly negative. CCSM4 hindcast displays a slightly positive trend while the one of its historical simulation is negative.

When initialized through data assimilation of observations, CCSM4, FGOALS-g2, CNRM-CM5 and BCC-CSM1.1 present a systematic drift (not shown). This drift is likely responsible for the high positive or negative trends found in the hindcasts of these models. Such a drift has its origin in the initialization of a model with a state that forces it to produce much more (or less) sea ice than has its climatological mean. After the initialization, the model does not have any constraint from observations anymore and the simulation tends to go back towards the model's climatology. We do not have information about the method used to initialize the models FGOALS-g2 and CCSM4. The use of raw data in the initialization procedures applied to BCC-CSM1.1 and to CNRM-CM5 may partly account for the drift occurring in their hindcast simulations.

Similarly, for winter sea ice extent, the initialization with observations does not systematically lead to a simulated trend in better agreement with observations. Figure 5b shows that hindcast simulations of MIROC4h, MIROC5 and MRI-CGCM3 have trends that are slightly closer to the observation than are the historical trends. The 7 other models perform worse or do not offer any improvement when they are initialized with observations. As in the case of summer sea ice extent (Fig. 5a), FGOALS-g2 simulates a large positive trend in its winter sea ice extent when it is initialized with observations and CNRM-CM5 has a more negative trend in its hindcast, for the same reasons as the one proposed above. For BCC-CSM1.1, the hindcast trend in winter sea ice extent does not differ significantly from the historical trend.

Results presented in Fig. 5 show that the initialization of models through data assimilation of observation does not bring significant improvement on the simulated trend. When raw data are used instead of anomalies, the initialization apparently deteriorates the trend in sea ice extent simulated by models. Corrections can be introduced to take into account that kind of biases (e.g. Troccoli and Palmer, 2007; Vannitsem and Nicolis, 2008). Nevertheless, such a procedure requires a larger amount of initialized simulations spanning several decades. Proposing such a method for sea ice and analyzing how it would impact the analysis of the trend is out of the scope of our study.

4.2 Correlation between models and observations

The forecast skill of the models can also be assessed by analyzing the predictions a few years ahead. To do so, for each model, we computed the anomaly correlation coefficient used in Pohlmann et al. (2009):

$$COR(t) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}[x_{ij}(t)-\bar{x}][o_i(t)-\bar{o}]}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}[x_{ij}(t)-\bar{x}]^2\sum_{i=1}^{N}M[o_i(t)-\bar{o}]^2}} \quad (1)$$

where t is the lead time (in years), x_{ij} are the hindcast simulations, i is the ensemble index (different indices correspond to different times when the hindcast simulations are started) and j is the index of the member belonging to the ensemble i. N is the number of ensembles and M is the number of members within each ensemble. o_i is the observation covering the time period spanned by the ensemble i. The overbar stands for the climatological mean of the uninitialized (historical) simulation and of the observations, over the analyzed period (here 1981–2005).

The correlation between hindcast simulations and observations is shown for summer (Fig. 6) and winter (Fig. 7) sea ice extent. This correlation has been computed from a series of 4 hindcasts ensemble simulations, initialized every 5 yr between January 1981 and January 1996 (every 5 yr between November 1980 and November 1995 for HadCM3). The 95 % significance level is computed using a t-test. This significance level varies from one model to another because of the different number of members in each model ensemble (see Table 1).

In summer, none of the 10 models analyzed here has a significant correlation for the first year after initialization (Fig. 6). HadCM3, IPSL-CM5A-LR and MIROC4h never outstrip the 95 % significant level. The 7 remaining models present one or two peaks of significant correlation several years after the initialization and almost all the models have a negative correlation during most of the 10 yr. The emergence of correlation later on in the simulation can occur randomly or it might still be a consequence of the initialization. Indeed, models might undergo an initial shock due to the initialization procedure before getting stabilized and benefit from the initialization. For winter sea ice extent (Fig. 7), the correlation is significantly positive during the first year for CCSM4, MIROC5 and MPI-ESM-LR models, indicating some predictive skill. Then the correlation decreases and reaches negative values. A negative correlation is also found in the other models. The significant correlation after one year in three models in winter likely arises from the initialization but the memory of the system is apparently not sufficient to keep a significant correlation during the following years. Unlike in the Arctic, sea ice around the Antarctic is relatively young. It disappears almost entirely during the melting season and recovers during winter months, preventing this sea ice to retain information from initialization. The ocean can keep the information over longer periods but in the available experiments, its role appears weak during the first year after initialization. Still, it may be responsible for the emergence of correlation several years after initialization, for both summer and winter sea ice extent, through local interactions or teleconnections with remote areas.

In any case, the skill of model predictions for Southern Ocean sea ice extent is quite poor compared to the one obtained for other variables. For instance, Kim et al. (2012) have analyzed hindcasts results from seven CMIP5 models and have shown that these models have a high skill in forecasting surface temperature anomalies over the Indian, North Atlantic and Western Pacific Ocean, up to 6–9 yr ahead. Matei et al. (2012a) have pointed out a significant correlation between hindcast and observations for the Atlantic Meridional Overturning Circulation (AMOC) strength at 26.5° N up to 4 yr ahead.

5 Summary and conclusions

From 24 CMIP5 models available to date, we have analyzed results of historical and hindcast simulations. This is still a small ensemble but we consider that it is diverse enough to constitute a reasonable sample to draw conclusions about current models behavior in the Southern Ocean.

The multi-model mean reproduces well the observed summer and winter sea ice edge as well as the annual cycle of sea ice extent. The skill of individual models is much lower. The majority of the biases in the simulated Southern Ocean sea ice highlighted for CMIP3 models persist for the CMIP5 ones. Furthermore, all the models analyzed here overestimate the variability of the sea ice extent in winter. In addition, we saw that, in contrast to observations, the variability in some models can vary significantly from one season to the other. We have thus chosen to analyze seasonal means rather than annual mean but the conclusions are similar whether we consider summer or winter sea ice extent.

The analyses performed in this paper aimed at better understand the role played by the internal variability in the observed increase of sea ice extent in the Southern Ocean. Our approach can be summarized in three questions that we can now partly answer.

Firstly, are the trend of winter and summer observed sea ice extent (statistically significant at the 90% level for summer but not for winter) compatible with a combination of the forced response and the internal variability according to model results? The models generally respond to the external forcing by a decrease in their sea ice extent. Our analysis of its representation in the different models has shown that the inclusion of stratospheric ozone depletion does not modify strongly the sign of the simulated trend in sea ice extent in the Southern Ocean compared to CMIP3, in which only half of the models took into account this forcing. Moreover, models with interactive chemistry or with higher atmospheric vertical resolution do not provide better results that the other ones. Nevertheless, natural variability can overwhelm the influence of the forced response, leading to a positive trend in some ensemble members. This case appears relatively rare among the available simulations. However, if we consider the wide range of trends each model provides because of its own dynamics only, the positive observed trend in sea ice extent can be accounted for by internal variability.

Secondly, does the models internal variability agree with the one of the observations? From our model analysis, positive trend in sea ice extent, such as the observed one, can arise from internal variability. Nevertheless, to have confidence in this conclusion, the models internal variability must fit the one of the observations. Unfortunately, we have shown that the models often have a climatological mean which is far from the observations or too high an interannual variability, or even both. None of the CMIP5 models provides thus a reasonable estimate of all the main characteristics of the sea ice cover over the last decades in the Southern Ocean, in contrast to the Arctic (e.g. Stroeve et al., 2012; Massonnet et al., 2012). Moreover, the few models that display an increase in sea ice extent have such a large variability that the sign of the trend is not robust. Because of those models biases, we cannot reasonably consider the results of these models as a good representation of the behavior of the Southern Ocean sea ice. As a consequence, even if the positive observed trend in sea ice extent is compatible with the models internal variability, the biases of these models prevent us from firmly assessing the link between the internal variability in the Southern Ocean and the observed increase in sea ice extent.

Thirdly, how does the initialization method impact the simulated evolution of sea ice extent in the Southern Ocean? If the internal variability is important, a correct initialization of the model state may lead to a better agreement with data. In this hypothesis, constraining the model with observations would put the system in a state that favors an increase in ice extent, for instance because of a more stratified or colder ocean. However, results from hindcast simulations have shown that there is no systematic improvement of the simulation of sea ice extent observed trend. Previous studies have demonstrated that models have a high potential predictability in the Southern Ocean region at decadal time scales (e.g. Latif et al., 2010), i.e. there exists in models deterministic decadal variability. The test in real conditions has not shown such predictability for sea ice extent. This may be due to some inadequate representation of physics and/or feedbacks in models but also to the initialization procedure. Indeed, observations required to initialize properly the system are quite sparse in that area and the time period they cover is relatively short. Furthermore, data assimilation methods used in general circulation models are essentially based on a nudging and improvement may be expected if more sophisticated methods are applied and systematically tested in the Southern Ocean.

Acknowledgements. We thank three anonymous reviewers for their insightful comments. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. We thank J. Jungclaus from the Max Planck Institute for Meteorology (Hamburg, Germany) for his useful advice and his help in the analysis of the MPI model results. We also thank S. Dubinkina for her help with this manuscript. V. Z. is Research Fellow with the Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agronomie (FRIA-Belgium). H. G. and F. M are Senior Research Associate and Research Fellow with the Fonds National de la Recherche Scientifique (F.R.S.-FNRS-Belgium), respectively. This work is supported by the F.R.S.-FNRS and by the Belgian Federal Science Policy Office (Research Program on Science for a Sustainable Development).

References

Arzel, O., Fichefet, T., and Goosse, H.: Sea ice evolution over the 20th and 21st centuries as simulated by current AOGCMs, Ocean Modelling, 12, 401–415, doi:10.1016/j.ocemod.2005.08.002, 2006.

Bitz, C. M. and Polvani, L. M.: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model, Geophys. Res. Lett., 39, 2012.

Bitz, C. M., Gent, P. R., Woodgate, R. A., Holland, M. M., and Lindsay, R.: The Influence of Sea Ice on Ocean Heat Uptake in Response to Increasing CO2, Journal of Climate, 19, 2437–2450, doi:10.1175/JCLI3756.1, 2006.

Bloom, S. C., Takacs, L. L., da Silva, A. M., and Ledvina, D.: Data Assimilation Using Incremental Analysis Updates, Monthly Weather Review, 124, 1256–1271, 1996.

Cavalieri, D. J. and Parkinson, C. L.: Antarctic sea ice variability and trends, 1979–2006, J. Geophys. Res., 113, doi:10.1029/2007JC004564, 2008.

Cavalieri, D. J., Parkinson, C. L., Gloersen, P., Comiso, J. C., and Zwally, H. J.: Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, J. Geophys. Res., 104, 15 803–15 814, doi:10.1029/1999JC900081, 1999.

Cavalieri, D. J., Parkinson, C. L., and Vinnikov, K. Y.: 30-Year satellite record reveals contrasting Arctic and Antarctic decadal sea ice variability, Geophys. Res. Lett., 30, doi:10.1029/2003GL018031, 2003.

Chikamoto, Y., Kimoto, M., Ishii, M., Mochizuki, T., Sakamoto, T., Tatebe, H., Komuro, Y., Watanabe, M., Nozawa, T., Shiogama, H., Mori, M., Yasunaka, S., and Imada, Y.: An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC, Climate Dynamics, pp. 1–22, doi:10.1007/s00382-012-1351-y, 2012.

Cionni, I., Eyring, V., Lamarque, J. F., Randel, W. J., Stevenson, D. S., Wu, F., Bodeker, G. E., Shepherd, T. G., Shindell, D. T., and Waugh, D. W.: Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, Atmos. Chem. Phys., 11, 11 267–11 292, doi:10.5194/acp-11-11267-2011, http://www.atmos-chem-phys.net/11/11267/2011/, 2011.

Comiso, J. C.: Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I, 1979-2007, Digital media, 1999, updated 2008.

Comiso, J. C. and Nishio, F.: Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data, J. Geophys. Res., 113, http://dx.doi.org/10.1029/2007JC004257, 2008.

Cotté, C. and Guinet, C.: Historical whaling records reveal major regional retreat of Antarctic sea ice, Deep Sea Research Part I: Oceanographic Research Papers, 54, 243–252, http://www.sciencedirect.com/science/article/pii/S0967063706003074, 2007.

Curran, M. A. J., van Ommen, T. D., Morgan, V. I., Phillips, K. L., and Palmer, A. S.: Ice Core Evidence for Antarctic Sea Ice Decline Since the 1950s, Science, 302, 1203–1206, http://www.sciencemag.org/content/302/5648/1203.abstract, 2003.

de la Mare, W. K.: Abrupt mid-twentieth-century decline in Antarctic sea-ice extent from whaling records, Nature, 389, 57–60, http://dx.doi.org/10.1038/37956, 1997.

de la Mare, W. K.: Changes in Antarctic sea-ice extent from direct historical observations and whaling records, Climatic Change, 92, 461–493, http://dx.doi.org/10.1007/s10584-008-9473-2, 2009.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S.: Uncertainty in the Backyard: Communicating the Role of Natural Variability in Future North American Climate, Perspective submitted to Nature Climate Change, 2012.

Donner, L. J., Wyman, B. L., Hemler, R. S., Horowitz, L. W., Ming, Y., Zhao, M., Golaz, J.-C., Ginoux, P., Lin, S. J., Schwarzkopf, M. D., Austin, J., Alaka, G., Cooke, W. F., Delworth, T. L., Freidenreich, S. M., Gordon, C. T., Griffies, S. M., Held, I. M., Hurlin, W. J., Klein, S. A., Knutson, T. R., Langenhorst, A. R., Lee, H.-C., Lin, Y., Magi, B. I., Malyshev, S. L., Milly, P. C. D., Naik, V., Nath, M. J., Pincus, R., Ploshay, J. J., Ramaswamy, V., Seman, C. J., Shevliakova, E., Sirutis, J. J., Stern, W. F., Stouffer, R. J., Wilson, R. J., Winton, M., Wittenberg, A. T., and Zeng, F.: The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3, Journal of Climate, 24, 3484–3519, doi:10.1175/2011JCLI3955.1, 2011.

Flato, G. M.: Sea-ice and its response to CO_2 forcing as simulated by global climate models, Climate Dynamics, 23, 229–241, doi:10.1007/s00382-004-0436-7, 2004.

Gao, F., Xin, X., and Wu, T.: Study on the Prediction of Regional and Global Temperature in Decadal Time Scale with BCC_CSM1.1, Chinese Journal of Atmospheric Sciences (in Chinese), 2012.

Goosse, H., Arzel, O., Bitz, C. M., de Montety, A., and Vancoppenolle, M.: Increased variability of the Arctic summer ice extent in a warmer climate, Geophys. Res. Lett., 36, doi:10.1029/2009GL040546, 2009a.

Goosse, H., Lefebvre, W., de Montety, A., Crespin, E., and Orsi, A.: Consistent past half-century trends in the atmosphere, the sea ice and the ocean at high southern latitudes, Climate Dynamics, 33, 999–1016–1016, doi:10.1007/s00382-008-0500-9, 2009b.

Holland, M. and Raphael, M.: Twentieth century simulation of the southern hemisphere climate in coupled models. Part II: sea ice conditions and variability, Climate Dynamics, 26, 229–245, doi:10.1007/s00382-005-0087-3, 2006.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, Cambridge, 4 edn., 2007.

Kawase, H., Nagashima, T., Sudo, K., and Nozawa, T.: Future changes in tropospheric ozone under Representative Concentration Pathways (RCPs), Geophys. Res. Lett., 38, doi:10.1029/2010GL046402, 2011.

Kay, J. E., Holland, M. M., and Jahn, A.: Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world, Geophys. Res. Lett., 38, doi:10.1029/2011GL048008, 2011.

Keenlyside, N., Latif, M., Jungclaus, J. H., Kornblueh, L., and Roeckner, E.: Advancing decadal-scale climate prediction in the North Atlantic sector, Nature, 453, doi:10.1038/nature06921, 2008.

Kim, H.-M., Webster, P. J., and Curry, J. A.: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts, Geophys. Res. Lett., 39, doi:10.1029/2012GL051644, 2012.

Kirkman, C. H. and Bitz, C. M.: The Effect of the Sea Ice Freshwater Flux on Southern Ocean Temperatures in CCSM3: Deep-Ocean Warming and Delayed Surface Warming, Journal of Climate, 24, 2224–2237, doi:10.1175/2010JCLI3625.1, 2010.

Landrum, L., Holland, M. M., Schneider, D. P., and Hunke, E.: Antarctic Sea Ice Climatology, Variability, and Late Twentieth-Century Change in CCSM4, Journal of Climate, 25, 4817–4838, doi:10.1175/JCLI-D-11-00289.1, 2012.

Latif, M., Delworrth, T., Dommenget, D., Drange, H., Hazeleger, W., Hurrell, J., Keenlyside, N., Meehl, G. A., and Sutton, R.: Dynamics of Decadal Climate Variability and Implications for its Prediction, in: Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society, vol. 2, Hall, J., Harrison D.E. & Stammer, D., Eds., ESA Publication WPP-306, Venice, Italy, 21-25 September 2009, 2010.

Lefebvre, W. and Goosse, H.: Analysis of the projected regional sea-ice changes in the Southern Ocean during the twenty-first century, Climate Dynamics, 30, 59–76–76, doi:10.1007/s00382-007-0273-6, 2008a.

Lefebvre, W. and Goosse, H.: An analysis of the atmospheric processes driving the large-scale winter sea ice variability in the Southern Ocean, J. Geophys. Res., 113, doi:10.1029/2006JC004032, 2008b.

Liu, J. and Curry, J. A.: Accelerated warming of the Southern Ocean and its impacts on the hydrological cycle and sea ice, Proceedings of the National Academy of Sciences, 107, 14 987–14 992, http://www.pnas.org/content/107/34/14987.abstract, 2010.

Massonnet, F., Fichefet, T., Goosse, H., Bitz, C. M., Philippon-Berthier, G., Holland, M. M., and Barriat, P. Y.: Constraining projections of summer Arctic sea ice, The Cryosphere, 6, 1383–1394, 2012.

Matei, D., Baehr, J., Jungclaus, J. H., Haak, H., Müller, W. A., and Marotzke, J.: Multiyear Prediction of Monthly Mean Atlantic Meridional Overturning Circulation at 26.5°N, Science, 335, 76–79, doi:10.1126/science.1210299, 2012a.

Matei, D., Pohlmann, H., Jungclaus, J., Müller, W., Haak, H., and Marotzke, J.: Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model, Journal of Climate, doi:10.1175/JCLI-D-11-00633.1, 2012b.

Parkinson, C. L., Vinnikov, K. Y., and Cavalieri, D. J.: Evaluation of the simulation of the annual cycle of Arctic and Antarctic sea ice coverages by 11 major global climate models, J. Geophys. Res., 111, doi:10.1029/2005JC003408, 2006.

Pierce, D. W., Barnett, T. P., Tokmakian, R., Semtner, A., Maltrud, M., Lysne, J., and Craig, A.: The ACPI Project, Element 1: Initializing a Coupled Climate Model from Observed Conditions, Climatic Change, 62, 13–28, http://dx.doi.org/10.1023/B:CLIM.0000013676.42672.23, 2004.

Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., and Tschuck, P.: Estimating the Decadal Predictability of a Coupled AOGCM, Journal of Climate, 17, 4463–4472, doi:10.1175/3209.1, 2004.

Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., and Marotzke, J.: Initializing Decadal Climate Predictions with the GECCO Oceanic Synthesis: Effects on the North Atlantic, Journal of Climate, 22, 3926–3938, doi:10.1175/2009JCLI2535.1, 2009.

Randel, W. J. and Wu, F.: A stratospheric ozone trends data set for global modeling studies, Geophys. Res. Lett., 26, 3089–3092, doi:10.1029/1999GL900615, 1999.

Sen Gupta, A., Santoso, A., Taschetto, A. S., Ummenhofer, C. C., Trevena, J., and England, M. H.: Projected Changes to the Southern Hemisphere Ocean and Sea Ice in the IPCC AR4 Climate Models, Journal of Climate, 22, 3047–3078, doi:10.1175/2008JCLI2827.1, 2009.

Sigmond, M. and Fyfe, J. C.: Has the ozone hole contributed to increased Antarctic sea ice extent?, Geophys. Res. Lett., 37, doi:10.1029/2010GL044301, 2010.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., and Murphy, J. M.: Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model, Science, 317, 796–799, doi:10.1126/science.1139540, 2007.

Smith, K. L., Polvani, L. M., and Marsh, D. R.: Mitigation of 21st century Antarctic sea ice loss by stratospheric ozone recovery, Geophys. Res. Lett., 39, 2012.

Solomon, S.: Stratospheric ozone depletion: A review of concepts and history, Rev. Geophys., 37, 275–316, http://dx.doi.org/10.1029/1999RG900008, 1999.

Son, S. W., Polvani, L. M., Waugh, D. W., Akiyoshi, H., Garcia, R., Kinnison, D., Pawson, S., Rozanov, E., Shepherd, T. G., and Shibata, K.: The Impact of Stratospheric Ozone Recovery on the Southern Hemisphere Westerly Jet, Science, 320, 1486–1489, 2008.

Stammerjohn, S. E., Martinson, D. G., Smith, R. C., Yuan, X., and Rind, D.: Trends in Antarctic annual sea ice retreat and advance and their relation to El Niño Southern Oscillation and Southern Annular Mode variability, J. Geophys. Res., 113, http://dx.doi.org/10.1029/2007JC004269, 2008.

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, Geophys. Res. Lett., 39, 2012.

Swingedouw, D., Mignot, J., Labtoulle, S., Guilyardi, E., and Madec, G.: Initialisation and predictability of the AMOC over the last 50 years in a climate model, Climate Dynamics, accepted, 2012.

Szopa, S., Balkanski, Y., Schulz, M., Bekki, S., Cugnet, D., Fortems-Cheiney, A., Turquety, S., Cozic, A., Deandreis, C., Hauglustaine, D., Idelkadi, A., Lathiere, J., Lefevre, F., Marchand, M., Vuolo, R., Yan, N., and Dufresne, J.-L.: Aerosol and Ozone changes as forcing for Climate Evolution between 1850 and 2100, Climate Dynamics, submitted, 2012.

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, doi:10.1175/BAMS-D-11-00094.1, 2011.

Tebaldi, C., Arblaster, J. M., and Knutti, R.: Mapping model agreement on future climate projections, Geophys. Res. Lett., 38, doi:10.1029/2011GL049863, 2011.

Thompson, D. W. J., Solomon, S., Kushner, P. J., England, M. H., Grise, K. M., and Karoly, D. J.: Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change, Nature Geosci, 4, 741–749, doi:10.1038/ngeo1296, http://dx.doi.org/10.1038/ngeo1296, 2011.

Troccoli, A. and Palmer, T. N.: Ensemble decadal predictions from analysed initial conditions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 2179–2191, doi:10.1098/rsta.2007.2079, 2007.

Turner, J., Comiso, J. C., Marshall, G. J., Lachlan-Cope, T. A., Bracegirdle, T., Maksym, T., Meredith, M. P., Wang, Z., and Orr, A.: Non-annular atmospheric circulation change induced by stratospheric ozone depletion and its role in the recent increase of Antarctic sea ice extent, Geophys. Res. Lett., 36, http://dx.doi.org/10.1029/2009GL037524, 2009.

Vannitsem, S. and Nicolis, C.: Dynamical Properties of Model Output Statistics Forecasts, Monthly Weather Review, 136, 405–419, doi:10.1175/2007MWR2104.1, 2008.

Voldoire, A., Sanchez-Gomez, E., Salas y Mélia, D., Decharme, B., Cassou, C., Sénési, S., Valcke, S., Beau, I., Alias, A., Chevallier, M., Déqué, M., Deshayes, J., Douville, H., Fernandez, E., Madec, G., Maisonnave, E., Moine, M. P., Planton, S., Saint-Martin, D., Szopa, S., Tyteca, S., Alkama, R., Belamari, S., Braun, A., Coquart, L., and Chauvin, F.: The CNRM-CM5.1 global climate model: description and basic evaluation, Climate Dynamics, doi:10.1007/s00382-011-1259-y, 2012.

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., Nozawa, T., Kawase, H., Abe, M., Yokohata, T., Ise, T., Sato, H., Kato, E., Takata, K., Emori, S., and Kawamiya, M.: MIROC-ESM: model description and basic results of CMIP5-20c3m experiments, Geosci. Model Dev. Discuss., 4, 1063–1128, doi:10.5194/gmdd-4-1063-2011, 2011.

Yukimoto, S., Yoshimura, H., Hosaka, M., Sakami, T., Tsujino, H., Hirabara, M., Tanaka, T. Y., Deushi, M., Obata, A., Nakano, H., Adachi, Y., Shindo, E., Yabu, S., Ose, T., and Kitoh, A.: Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1) - Model Description, Tech. Rep. 64, Meteorological Research Institute, Japan, 2011.

Zhang, J.: Increasing Antarctic Sea Ice under Warming Atmospheric and Oceanic Conditions, Journal of Climate, 20, 2515–2529, http://dx.doi.org/10.1175/JCLI4136.1, 2007.

Table 1. Model name, Institute and number of members in models historical and hindcast simulations.

Model name	Institute ID	Modeling center	Number of members in historical	Number of members in hindcasts
BCC-CSM1.1	BCC	Beijing Climate Center, China Meteorological Administration	3	4
CanESM2	CCCMA	Canadian Centre for Climate Modelling and Analysis	5	-
CCSM4	NCAR	National Center for Atmospheric Research	6	10
CNRM-CM5	CNRM-CERFACS	Centre National de Recherches Meteorologiques / Centre Europeen de Recherche et Formation Avancees en Calcul Scientifique	10	10
CSIRO-Mk3.6.0	CSIRO-QCCCE	Commonwealth Scientific and Industrial Research Organization in collaboration with Queensland Climate Change Centre of Excellence	10	-
EC-EARTH	EC-EARTH	EC-EARTH consortium	1	-
FGOALS-g2	LASG-CESS	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS,Tsinghua University	1	3
FGOALS-s2	LASG-IAP	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences	3	-
GFDL-CM3	NOAA GFDL	NOAA Geophysical Fluid Dynamics Laboratory	5	-
GFDL-ESM2M	NOAA GFDL	NOAA Geophysical Fluid Dynamics Laboratory	1	-
GISS-E2-R	NASA GISS	NASA Goddard Institute for Space Studies	5	-
HadCM3	MOHC	Met Office Hadley Centre	10	10
HadGEM2-CC	MOHC	Met Office Hadley Centre	1	-
HadGEM2-ES	MOHC	Met Office Hadley Centre	1	-
INM-CM4	INM	Institute for Numerical Mathematics	1	-
IPSL-CM5A-LR	IPSL	Institut Pierre-Simon Laplace	4	6
IPSL-CM5A-MR	IPSL	Institut Pierre-Simon Laplace	1	-
MIROC4h	MIROC	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	3	3
MIROC5	MIROC	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1	6
MIROC-ESM	MIROC	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	3	-
MIROC-ESM-CHEM	MIROC	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1	-
MPI-ESM-LR	MPI-M	Max Planck Institute for Meteorology	3	10 (3 in 30-year hindcast)
MRI-CGCM3	MRI	Meteorological Research Institute	3	3
NorESM1-M	NCC	Norwegian Climate Centre	3	-

Table 2. Summary of atmospheric vertical resolution and stratospheric ozone representation. Models in bold are the ones with an interactive chemistry, activated during the CMIP5 simulations or only activated in an offline simulation used to compute the ozone dataset prescribed in the CMIP5 simulations.

Model name	**Atmospheric vertical resolution**	**Stratospheric ozone**
BCC-CSM1.1	26 layers Top layer at 2.9 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
CanESM2	35 layers Top layer at 1 hPa	Prescribed AC&C/SPARC ozone database (Cionni et al., 2011)
CCSM4	26 layers	Prescribed; Data from an offline simulation of the CAM3.5 model with a fully interactive chemistry (Landrum et al., 2012)
CNRM-CM5	31 layers Top layer at 10 hPa	Interactive chemistry (Voldoire et al., 2012).
CSIRO-Mk3.6.0	18 layers	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011)
EC-EARTH	62 layers Top layer 5 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011)
FGOALS-g2	26 layers	No information available to us.
FGOALS-s2	26 layers Top layer at 2.19 hPa	No information available to us.
GFDL-CM3	48 layers	Interactive chemistry (Donner et al., 2011).
GFDL-ESM2M	24 layers	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
GISS-E2-R	40 layers Top layer at 0.1 hPa	Prescribed; Observational analyses of Randel and Wu (1999).
HadCM3	19 layers	Prescribed; Observational analyses of Randel and Wu (1999)
HadGEM2-CC	60 layers Top layer at 0.006 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
HadGEM2-ES	38 layers Top layer at 4 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
INM-CM4	21 layers Top layer at 10 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
IPSL-CM5A-LR	39 layers Top layer at 0.04 hPa	Prescribed; Data from an offline simulation of the LMDz-REPROBUS model (Szopa et al., 2012).
IPSL-CM5A-MR	39 layers Top layer at 0.04 hPa	Prescribed; Data from an offline simulation of the LMDz-REPROBUS model (Szopa et al., 2012).
MIROC4h	56 layers Top layer at 40 km	Prescribed; Data from an offline simulation of Kawase et al. (2011).
MIROC5	40 layers Top layer at 3 hPa	Prescribed; Data from an offline simulation of Kawase et al. (2011).
MIROC-ESM	80 layers Top layer at 0.003 hPa	Prescribed; Data from an offline simulation of Kawase et al. (2011).
MIROC-ESM-CHEM	80 layers Top layer at 0.003 hPa	Interactive chemistry (Watanabe et al., 2011).
MPI-ESM-LR	47 layers Top layer at 0.01 hPa	Prescribed; AC&C/SPARC ozone database (Cionni et al., 2011).
MRI-CGCM3	48 layers Top layer at 0.01 hPa	Interactive chemistry (Yukimoto et al., 2011).
NorESM1-M	26 layers Top layer at 2.9 hPa	No information available to us.

Table 3. Data assimilation methods used by the 10 models providing hindcast simulations.

Model name	Data assimilation method	References
BCC-CSM1.1	Nudging to 3D ocean temperature (raw data).	Gao et al. (2012)
CCSM4	Information not available to us	
CNRM-CM5	Nudging to 3D ocean temperature and salinity (raw data) as a function of depth and space, sea surface temperature and salinity nudging (raw data).	ftp://ftp.cerfacs.fr/pub/globc/exchanges/cassou/Michael/Aspen_CMIP5_wrkshop_cassou_2.ppt
FGOALS-g2	No information available to us.	
HadCM3	Nudging to 3D ocean temperature and salinity (anomalies), nudging to 3D atmosphere temperature and wind speed, nudging to surface pressure.	http://www.met.reading.ac.uk/~swr06jir/presentations/JIR_dept_seminar.pptx
IPSL-CM5A-LR	Nudging to sea surface temperature (anomalies).	Swingedouw et al. (2012)
MIROC4h	Incremental analysis update (IAU) of 3D ocean temperature and salinity (anomalies).	Chikamoto et al. (2012)
MIROC5	Incremental analysis update (IAU) of 3D ocean temperature and salinity (anomalies).	Chikamoto et al. (2012)
MPI-ESM-LR	Nudging to 3D ocean temperature and salinity (anomalies), except in the area covered by sea ice.	Matei et al. (2012b)
MRI-CGCM3	No information available to us.	

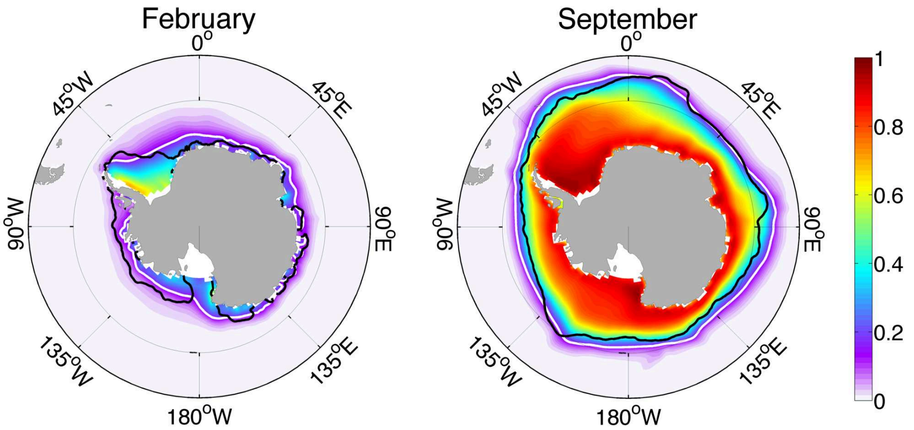

Fig. 1. Multi-model mean of sea ice concentration, computed from historical simulations over the period 1979–2005. White (black) line refers to the sea ice edge, i.e. the 15 % concentration limit of the multi-model ensemble mean (observations, Comiso, 1999, updated 2008).

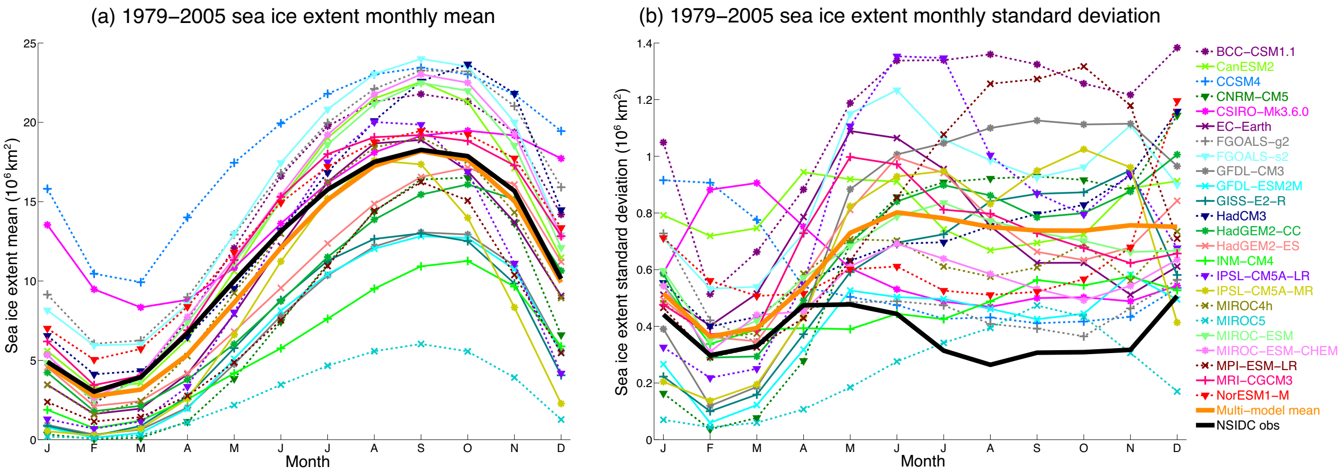

Fig. 2. **(a)** Monthly mean of Southern Ocean sea ice extent, computed over the period 1979–2005. **(b)** Standard deviation of detrended Southern Hemisphere sea ice extent, computed over the period 1979–2005 for each month of the year. Colors correspond to the ensemble mean of historical simulations from 24 different models. Dotted lines refer to models that provide both historical and hindcast simulations but here, results are only from historical simulations. Orange bold line is the multi-model mean. Black bold line refers to observations (Cavalieri and Parkinson, 2008).

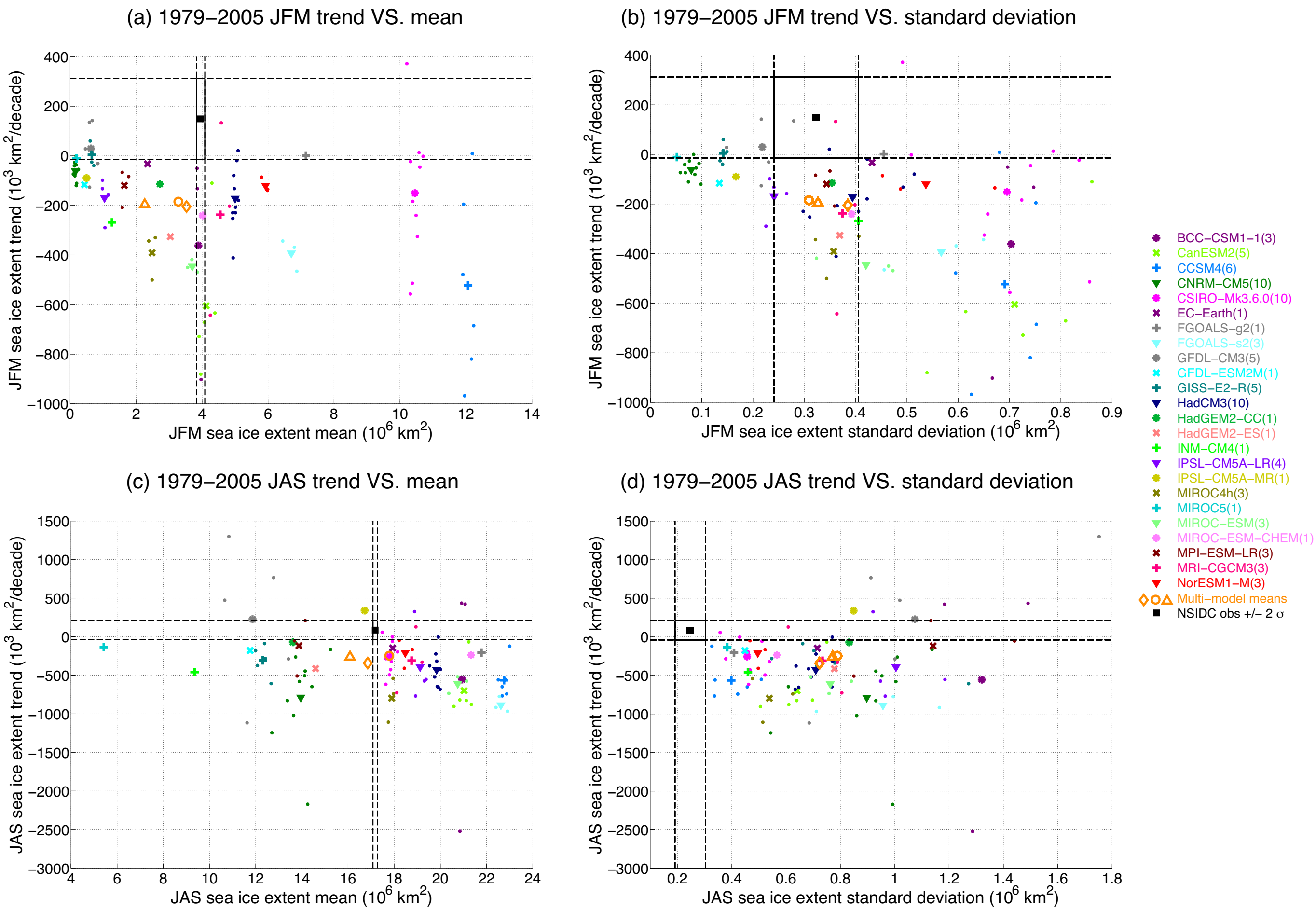

Fig. 3. Sea ice extent trend for the period 1979–2005 over the whole Southern Ocean vs. mean **(a, c)** and standard deviation **(b, d)**. The first row corresponds to summer (JFM), the second to winter (JAS). The different colors correspond to the historical simulations from 24 different models. For each color, the small dots refer to model individual members and the symbol specified in the legend is for the model ensemble mean. The number of members in each model is indicated in brackets in the legend. Orange refers to multi-model means: diamond sign is for the average over all the models, circle sign is for the mean of models with interactive chemistry (in bold in Table 2) and triangle sign is for the mean of models with 35 atmospheric levels or more on the vertical. Black square is for the observations (Cavalieri and Parkinson, 2008), surrounded by 2 standard deviations (black dashed lines).

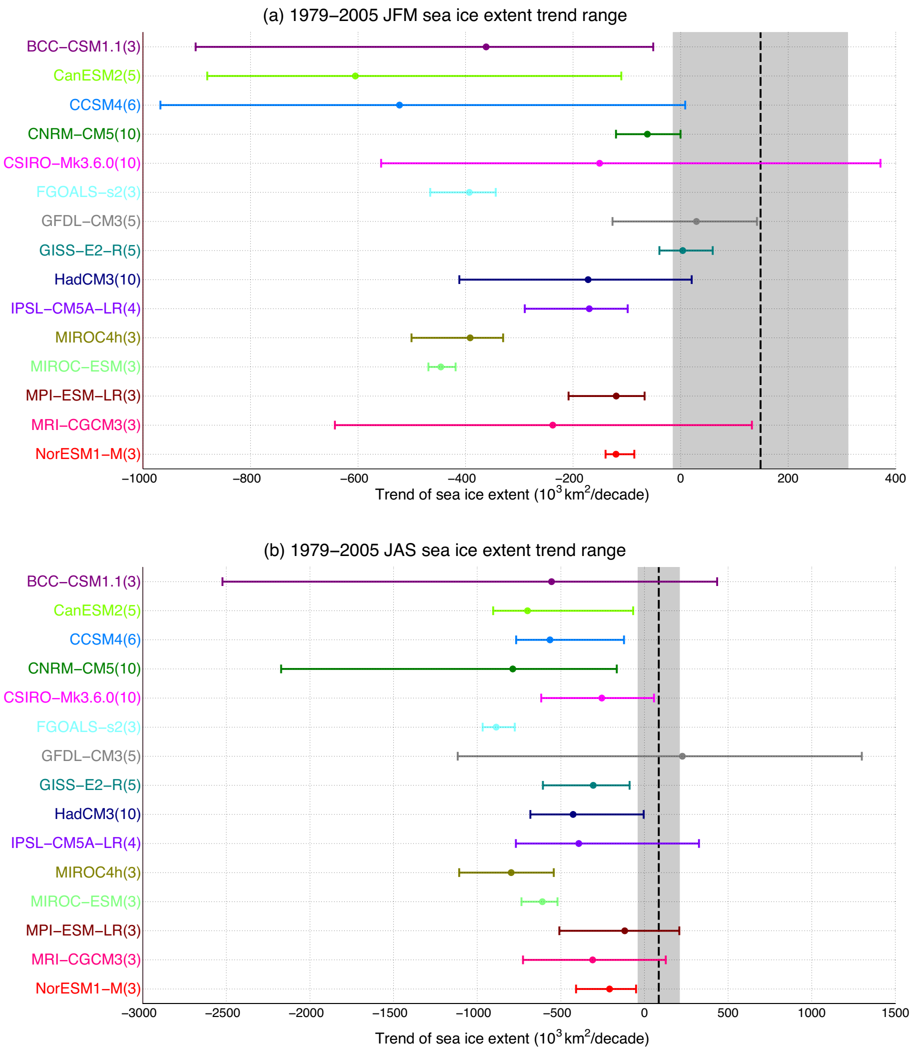

Fig. 4. Ensemble mean, minimum and maximum value of the sea ice extent trend for the period 1979–2005 over the whole Southern Ocean for summer **(a)** and winter **(b)**. The different colors correspond to the historical simulations from the 15 models that have at least 3 members in their ensemble. Dots refer to the ensemble means of the trends. Horizontal bars show the minimum and the maximum value of the trend reached by the members of one model ensemble. Black dashed line is for the trend of the observations (Cavalieri and Parkinson, 2008) surrounded by 2 standard deviations (grey shade).

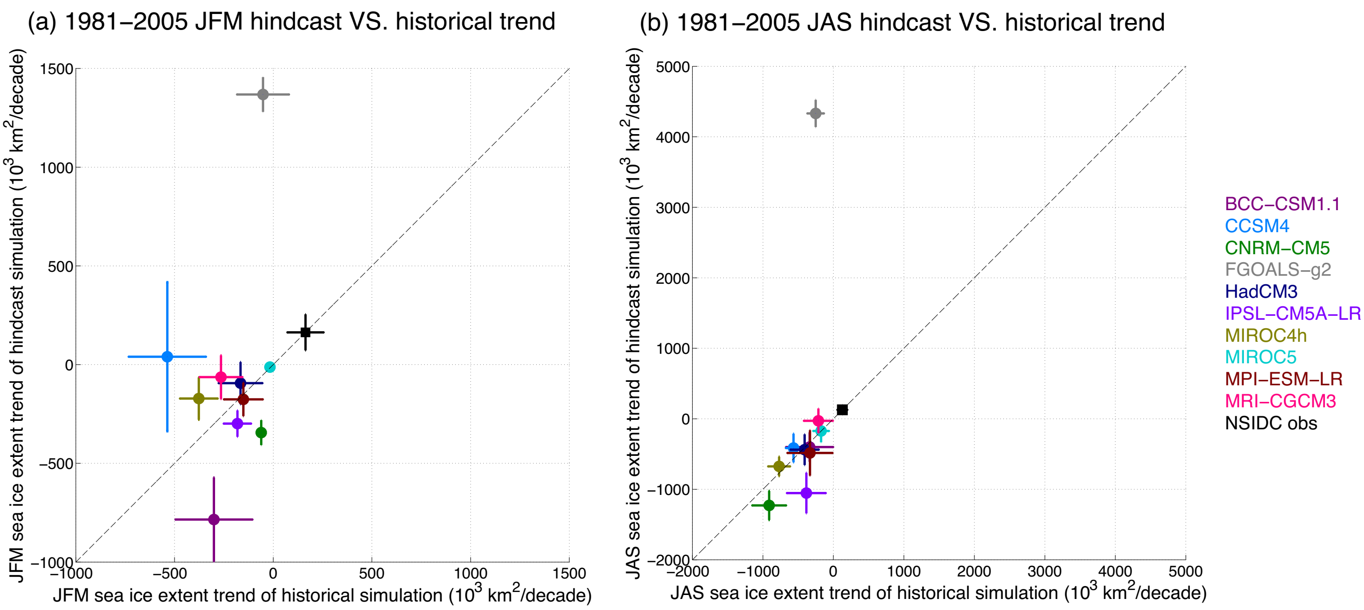

Fig. 5. Hindcast vs. historical Southern Ocean sea ice extent trend for summer **(a)** and winter **(b)**, computed over the period 1981–2005. The different colors refer to the different models. For each model, the dot refers to the ensemble mean of the trends and the horizontal (vertical) bar shows the ensemble mean of the standard deviations of the trends in the historical (hindcast) simulations. Black square is for the trend of the observations (Cavalieri and Parkinson, 2008). The vertical and the horizontal black bars are for the standard deviation of the observed trend. Dashed line represents the line $y(x) = x$.

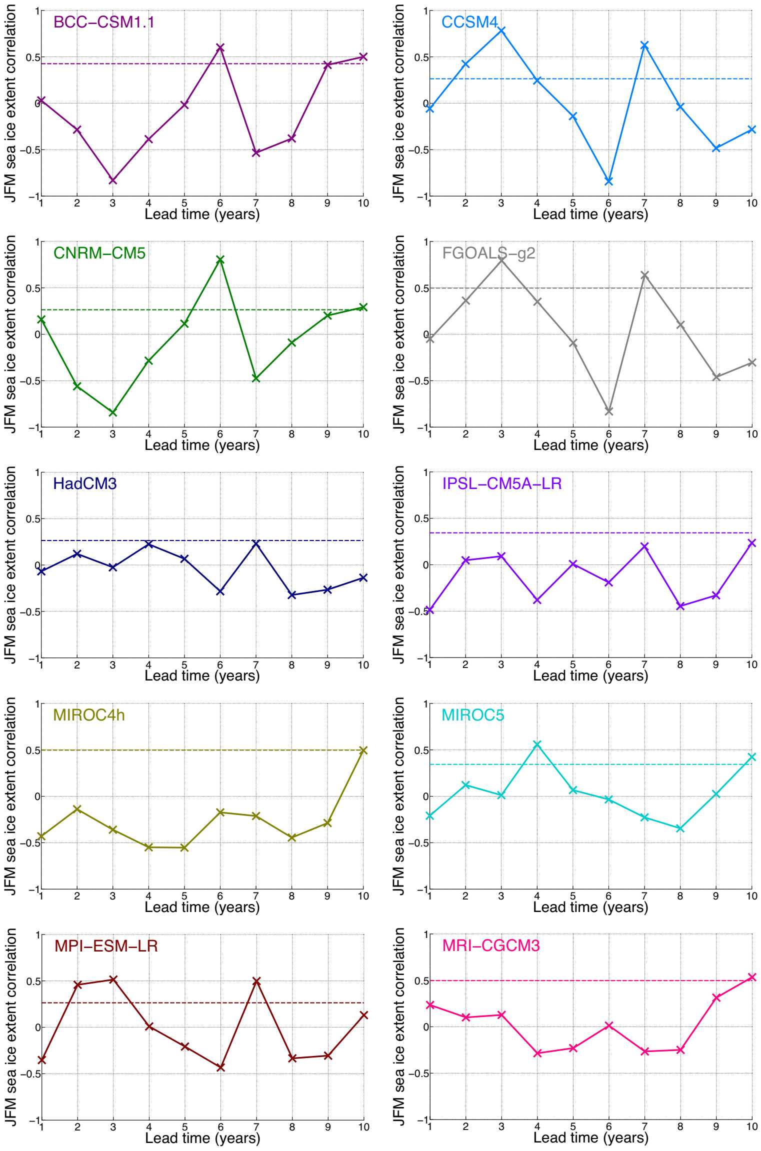

Fig. 6. Correlation between Southern Ocean summer (JFM) sea ice extent in models results and observations. For each model, the correlation is computed from a series of 4 hindcasts ensembles, initialized every 5 yr between January 1981 and January 1996 (between November 1980 and November 1995 for HadCM3). In each plot, the dashed line refers to the 95 % significance level.

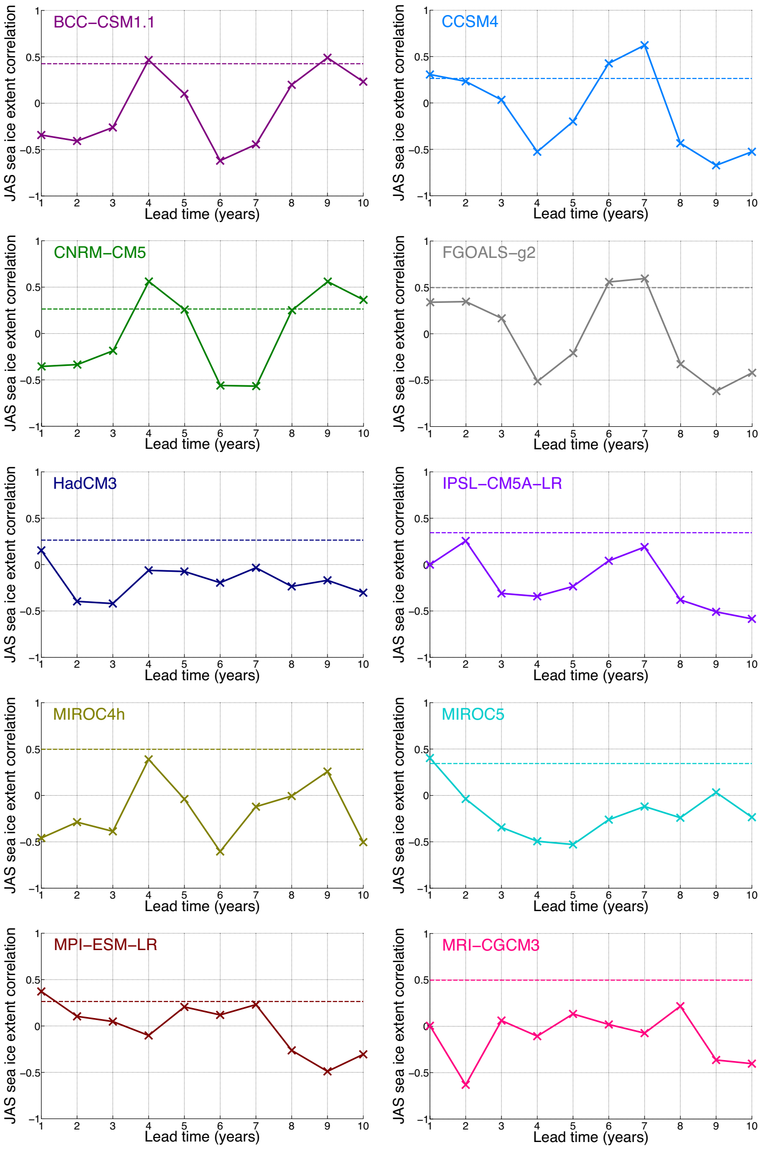

Fig. 7. Correlation between Southern Ocean winter (JAS) sea ice extent in models results and observations. For each model, the correlation is computed from a series of 4 hindcasts ensembles, initialized every 5 yr between January 1981 and January 1996 (between November 1980 and November 1995 for HadCM3). In each plot, the dashed line refers to the 95 % significance level.