Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multi-decadal timescales

Violette Zunz[1], Hugues Goosse[1], and Svetlana Dubinkina[2]

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

[2]Centrum Wiskunde & Informatica, Amsterdam, the Netherlands.

July 23, 2014

Abstract

In this study, we assess systematically the impact of different initialisation procedures on the predictability of the sea ice in the Southern Ocean. These initialisation strategies are based on three data assimilation methods: the nudging, the particle filter with sequential importance resampling and the nudging proposal particle filter. An Earth system model of intermediate complexity is used to perform hindcast simulations in a perfect model approach. The predictability of the Antarctic sea ice at interannual to multi-decadal timescales is estimated through two aspects: the spread of the hindcast ensemble, indicating the uncertainty of the ensemble, and the correlation between the ensemble mean and the pseudo-observations, used to assess the accuracy of the prediction. Our results show that at decadal timescales more sophisticated data assimilation methods as well as denser pseudo-observations used to initialise the hindcasts decrease the spread of the ensemble. However, our experiments did not clearly

demonstrate that one of the intialisation methods systematically provides with a more accurate prediction of the sea ice in the Southern Ocean than the others. Overall, the predictability at interannual timescales is limited to three years ahead at most. At multi-decadal timescales, the trends in sea ice extent computed over the time period just after the initialisation are clearly better correlated between the hindcasts and the pseudo-observations if the initialisation takes into account the pseudo-observations. The correlation reaches values larger than 0.5 in winter. This high correlation has likely its origin in the slow evolution of the ocean ensured by its strong thermal inertia, showing the importance of the quality of the initialisation below the sea ice.

1 Introduction

The last three decades have been characterised by an increase in sea ice extent in the Southern Ocean (e.g., *Comiso and Nishio*, 2008; *Parkinson and Cavalieri*, 2012). This recent expansion of the Antarctic sea ice has been attributed to different causes. Among them, a potential link with the stratospheric ozone depletion was pointed out (*Solomon*, 1999), but this hypothesis has not been confirmed in recent work (e.g., *Sigmond and Fyfe*, 2010; *Smith et al.*, 2012; *Bitz and Polvani*, 2012). Besides, *Mahlstein et al.* (2013); *Simpkins et al.* (2013); *Zunz et al.* (2013); *Polvani and Smith* (2013) drew attention to the fact that the internal variability of the climate system could also explain the positive trend in sea ice extent observed over the last decades. Other studies underlined the potential role of wind changes and of an enhanced stratification of the ocean (e.g., *Bitz et al.*, 2006; *Zhang*, 2007; *Lefebvre and Goosse*, 2008; *Stammerjohn et al.*, 2008; *Goosse et al.*, 2009; *Kirkman and Bitz*, 2010; *Landrum et al.*, 2012; *Holland and Kwok*, 2012; *Bintanja et al.*, 2013; *Goosse and Zunz*, 2014).

Nevertheless, no clear consensus on the processes responsible for this increase in sea ice extent has been reached yet. Understanding the evolution of the sea ice in the Southern Ocean is particularly difficult due to the lack of observations in this area, on the one hand, and the biases of climate models in the Southern Ocean, on the other hand. In particular, general circulation models involved in the 5th Coupled Model Intercomparison Project (CMIP5, *Taylor et al.* (2011)) generally overestimate the internal variability of the sea ice extent in the Southern Ocean and/or have a mean state that does not agree with the observations over the last 30 years, i.e. the time period for which reliable observations of the sea ice are available (e.g., *Turner et al.*, 2013; *Zunz et al.*, 2013).

A key issue is to determine whether the positive trend in sea ice extent would have been predictable if adequate observations and models were available some decades ago. In the same line, potential predictability in the Southern Ocean was pointed out (*Pohlmann et al.*, 2009) but the subject has been poorly studied so far (*Zunz et al.*, 2013; *Holland et al.*, 2013). In an idealised test case, *Holland et al.* (2013) described predictive capability for the position of the ice edge for several months if the system is initialised with nearly perfect observations. In a more realistic set up, *Zunz et al.* (2013) found that the skill of CMIP5 retrospective forecast simulations is generally weak for the Antarctic sea ice at interannual to multi-decadal timescales. The initialisation procedures used in the CMIP5 prediction simulations analysed by *Zunz et al.* (2013) are generally based on simple data assimilation methods, such as nudging, potentially reducing the skill of the predictions. Therefore, in parallel with an adjustment of the physical parameterisations included in the models that could reduce the biases in the Southern Ocean, more sophisticated initialisation methods deserve to be tested to check whether they improve the quality of the predictions of the sea ice in the Southern Ocean at interannual to multi-decadal timescales.

In the present study, we systematically examine how the predictability of Antarctic sea ice depends on the data assimilation method that is used to initialise the model simulation. In the recent study of *Pohlmann et al.* (2013), an analysis of the retrospective prediction skill of the Atlantic meridional overturning circulation was performed for different prediction systems using both different models and different initialisation procedures. In contrast to *Pohlmann et al.* (2013), we use only one climate model to analyse the Antarctic sea ice. This allows isolating more clearly the differences in the predictive skill that can be achieved due to the various initialisation procedures. Furthermore, unlike *Pohlmann et al.* (2013), the analyses proposed here were performed in an idealised framework. This approach consists of using pseudo-observations instead of actual observations for both the initialisation and the verification of the hindcats and has been used in many recent studies (e.g., *Holland et al.*, 2013; *Tietsche et al.*, 2013; *Servonnat et al.*, 2014). The pseudo-observations are obtained from a reference simulation performed with the same model as the one used in the hindcast and a noise is added to the pseudo-observations when they are included in the initialisation procedure. However, when comparing the results of initialisation methods with the pseudo-observations no noise is added providing the comparison with the truth. The use of pseudo-observations ensures that they have the same variability and mean state as the model results, since the incompatibilty in the mean state and variability between a model and observations may obscure the role of the initialisation method. Furthermore, working in an idealised framework allows testing the initialisation methods over longer time periods than if actual observations were used given that for the Antarctic sea ice reliable observations are available from the 1970s onwards only. Nevertheless, we have to keep in mind that the results discussed in this idealised framework correspond to an upper limit of predictability. For realistic prediction experiments, in which actual observations are simulated, model biases will tend to decrease the predictability.

The model used here is the Earth system model of intermediate complexity LOVECLIM1.2. It has a coarser resolution and a lower level of complexity than present-day general circulation models (GCMs), resulting in lower computational cost. Nevertheless, in the Southern Ocean it has a performance comparable to that of GCMs (*Goosse and Zunz*, 2014). It is thus an adequate tool to perform the large number of experiments required in our study. The skill of a prediction system is assessed here for the Antarctic sea ice through the analysis of hindcast simulations, i.e. simulations performed in the same conditions as if they were forecasts but spanning a past time period.

The climate model LOVECLIM1.2 is briefly described in Sect. 2.1 and the initialisation methods tested here are described in Sect. 2.2. Sect. 2.3 presents the scores that are used to assess the uncertainty and the accuracy of the hindcasts. The discussion of the results is divided into two parts: the interannual to decadal (Sect. 3.1) and the multi-decadal (Sect. 3.2) predictions. Finally, the main results are summarised and conclusions are proposed in Sect. 4.

2 Methodology

2.1 Model description

The three-dimensional model LOVECLIM1.2 (*Goosse et al.*, 2010) consists of the atmospheric component ECBilt2 (*Opsteegh et al.*, 1998), the oceanic component CLIO3 (*Goosse and Fichefet*, 1999) and the vegetation model VECODE (*Brovkin et al.*, 2002). The atmospheric model is a three-level quasi-geostrophic model with T21 horizontal resolution (corresponding to about $5.6^{\circ}\times5.6^{\circ}$). The stratosphere dynamics is not represented in ECBilt and the highest atmospheric level is at 200 hPa, preventing us to adequately take into account the influence of stratospheric ozone depletion. The oceanic model is an ocean general circulation model coupled to a sea ice model with horizontal resolution of $3^{\circ}\times3^{\circ}$ and 20 unevenly spaced vertical levels in the ocean. The vegetation component has the same horizontal resolution as ECBilt2 and simulates the evolution of the vegetation cover in terms of trees, grass and deserts. All the simulations performed with LOVECLIM1.2 over the 20th century are driven by anthropogenic and natural forcings (greenhouse gases increase, variations in volcanic activity, solar irradiance, orbital parameters and land use), corresponding to the ones adopted in the historical simulations performed in the framework of CMIP5 (*Taylor et al.*, 2011).

The model LOVECLIM1.2 simulates a realistic seasonal cycle of the sea ice extent in the Southern Hemisphere (*Goosse and Zunz*, 2014). It tends, however, to overestimate the sea ice extent during most of the year (not shown). The amplitude of those systematic biases in the sea ice simulated by LOVECLIM1.2 is comparable to the one of general circulation models involved in CMIP5 (e.g., *Turner et al.*, 2013; *Zunz et al.*, 2013). The too large sea ice extent simulated by LOVECLIM1.2 is the result of an overestimation of the sea ice concentration in the majority of the sectors of the Southern Ocean (Fig. 1). In summer (JFM), the averaged sea ice extent simulated by LOVECLIM1.2 between 1979 and 2009 reaches 6.1×10^{6} km^2 while it equals 4.2×10^{6} km^2 in the observations (*Fetterer et al.*, 2002, updated daily). In winter (JAS) over the period 1979-2009 the averaged sea ice extent reaches 19.8×10^{6} km^2 (17.8×10^{6} km^2) in LOVECLIM1.2 (in the observations).

2.2 Initialisation of the hindcasts

Eight initialisation methods are tested in this study. The methods are presented in this section and are summarised in Table 1. For each initialisation method, a hindcast is initialised on January 1 every 5 years between 1900 and 1990. One hindcast consists of an ensemble of 96 members and spans a period of 30 years. While the initialisation date slightly impacts the predictability of the Arctic sea ice (e.g., *Blanchard-Wrigglesworth et al.*, 2011; *Day et al.*, 2014), the influence of the initialisation date on the predictability of the Southern Ocean sea ice has not been firmly assessed yet (*Holland et al.*, 2013). However, this issue is out of the scope of the present study and is not addressed here.

In a first step, two extreme initialisation procedures are tested in the hindcast simulations. The first one does not take into account any pseudo-observations constraints. The corresponding non-initialised hindcasts, hereinafter referred to as HIND_noinit, do not require a specific procedure. They are simply taken from successive 30-yr time periods, separated by 5 years between 1900 and 1990, of a 96-member simulation driven by external forcing. Every three months, a perturbation is added to the surface air temperature of each member in order to be consistent with the experimental design of the simulations with data assimilation (see below). The hindcasts HIND_noinit are used to assess the part of the predictability that cannot be attributed to the initialisation with pseudo-observations. The second initialisation method is a nearly perfect initialisation. All the model variables of a perfectly initialised hindcast are initialised with values that are directly extracted from the pseudo-observations. A small perturbation is added to the surface air temperature of this initial state in order to generate different members of an ensemble. The perfectly initialised hindcasts, hereinafter referred to as HIND_perfect, allow assessing an upper limit of predictability.

In a second step, the hindcasts are initialised through different data assimilation (DA) methods. DA combines the model equations and available observations in order to estimate the state of the system as accurately as possible (*Talagrand*, 1997). In principle, a DA procedure allows updating the model solution not only for the variable that is assimilated but also for the other ones. Once a DA simulation has been run, the values of the state variables corresponding to different times at which we want to initialise a hindcast are extracted from the results of this simulation and are used as initial conditions. After the initialisation, no further information is provided by the pseudo-observations. A brief description of the DA methods used in the present study is given below and more detailed information is available in *Dubinkina and Goosse* (2013). These data assimilation methods demonstrated good performance in the Southern Ocean area in an idealised framework with pseudo-observations (*Dubinkina and Goosse*, 2013).

Here, pseudo-observations of monthly mean surface air temperature are assimilated. For initialisation of decadal prediction simulations assimilating subsurface oceanic data compared to assimilating only surface data can improve the performance of the forecast (e.g., *Dunstone and Smith*, 2010). However, in the Southern Ocean actual subsurface observations are even sparser in space and time than surface observations. Therefore, in order to test initialisation methods that could be easily transposed from this idealised study to a more realistic one, we assimilate only the surface air temperature data. Given the links between the surface air temperature and other climate variables, for instance the sea ice concentration, the reconstruction of the latter variables could be potentially improved by assimilating only the surface air temperature. The pseudo-observations correspond to the solution between the years 1850 and 2000 provided by a transient simulation driven by external forcing. This transient simulation starts in 850 from an equilibrium simulation. Four additional transient simulations spanning the period 850-1850, starting with perturbed initial conditions, were performed to provide the initial states for the simulations with data assimilation (for details see *Dubinkina and Goosse*, 2013). In order to mimic the instrumental errors, a Gaussian noise with standard deviation of 0.5°C is added to these pseudo-observations before they are assimilated in the model. Since we are working in an idealised framework in which the model and the pseudo-observations have the same climatology, there is no difference in assimilating anomalies or full-field variables (e.g., *Pierce et al.*, 2004; *Murphy et al.*, 2010; *Pohlmann et al.*, 2009; *Smith et al.*, 2013) and we choose to assimilate anomalies.

In addition, for each DA method used to generate the initial states of the hindcasts, two simulations were performed. In one simulation, dense pseudo-observations of the surface air temperature were assimilated, i.e. the pseudo-observations were available at every grid cell of the model. In the second simulation, sparse pseudo-observations were assimilated, i.e. the pseudo-observations were available only at the grid cells where observations from the HadCRUT3 dataset (*Brohan et al.*, 2006) are available between 1850-2000, the spatial coverage being displayed in Fig. 5 in *Dubinkina and Goosse* (2013). This allows assessing how the predictability decreases in a more realistic framework, compared to the idealised situation where pseudo-observations cover the whole model grid.

Nudging

Nudging is a DA method commonly used in decadal climate prediction studies (e.g., *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009; *Dunstone and Smith*, 2010; *Smith et al.*, 2010; *Kröger et al.*, 2012; *Swingedouw et al.*, 2012; *Matei et al.*, 2012a; *Servonnat et al.*, 2014). It consists of adding to the prognostic model equations a term that pulls the solution towards the (pseudo-) observations (e.g., *Kalnay*, 2007). In LOVECLIM1.2, the nudging term corresponds to an additional heat flux between the atmosphere and the ocean $Q = \gamma(T_{\text{mod}} - T_{\text{obs}})$. T_{mod} and T_{obs} are the monthly mean surface air temperature simulated by the model and from the (pseudo-) observations respectively. γ determines the relaxation time and equals 120 W m^{-2} K^{-1}. This value of γ stands between the values used in other studies (e.g., *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009; *Smith et al.*, 2010; *Matei et al.*, 2012a; *Swingedouw et al.*, 2012; *Servonnat et al.*, 2014). In addition, the nudging term is limited to a maximum value of 50 W m^{-2}.

The simulations that assimilate pseudo-observations through the nudging described above are 96-member ensembles. Each member of the ensemble is nudged every day towards the monthly mean pseudo-observations and every three months a perturbation is added to the surface air temperature of each member in order to work in the same experimental design for all three DA methods used here (see below). In both simulations assimilating dense and sparse pseudo-observations, the nudging is applied globally over the ocean where data are available, except the area covered by sea ice. Applying the nudging of surface temperature only on the grid cells free of sea ice is a common practice (e.g., *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009; *Matei et al.*, 2012a; *Servonnat et al.*, 2014). Excluding ice covered area from the nudging procedure prevents spurious forcing that would be introduced by the additional heat flux in the case when sea ice is present in the pseudo-observations but not in a simulation to be nudged. The hindcasts initialised with dense (sparse) pseudo-observations through the nudging are referred to as HIND_NUD$_{\mathrm{dense}}$ (HIND_NUD$_{\mathrm{sparse}}$) and summarised in Table 1.

Particle filter with sequential importance resampling

The particle filter with sequential importance resampling (SIR) is an ensemble DA method (e.g., *van Leeuwen*, 2009; *Dubinkina et al.*, 2011) and consists of the following steps. Starting from a set of different initial conditions, an ensemble of 96 simulations is propagated forward in time with the model for a period of three months. A realisation of the model (called *particle*) is different from another only due to different initial conditions. After the propagation step, a weight is assigned to each particle. This weight is computed based on the agreement between the surface air temperature estimated by the particle and the pseudo-observations (the better the agreement, the larger the weight). Then, particles are resampled: particles with small weights are eliminated while the ones with large weights are kept and duplicated in proportion to their weights, maintaining the total number of particles constant. A small perturbation is added to the duplicated particles in order to obtain initial conditions different from each other. The particles are then again propagated for three months using the model, and the whole procedure is repeated until the end of the period of interest.

Two sets of experiments were performed with SIR: one with dense pseudo-observations assimilated over the area covering the polar cap southward of 30°S and one with sparse pseudo-observations assimilated southward of 60°S. The choice of a smaller assimilation domain when sparse data were used has been made to avoid filter degeneracy (e.g., *van Leeuwen*, 2009, 2010; *Dubinkina and Goosse*, 2013). The hindcasts initialised with dense (sparse) pseudo-observations through the SIR are referred to as HIND_SIR$_{\mathrm{dense}}$ (HIND_SIR$_{\mathrm{sparse}}$).

Nudging proposal particle filter

The nudging proposal particle filter (NPPF) is a combination of the nudging and the particle filter with sequential importance resampling described above. During the propagation of the 96 particles using the model, a nudging term pulls the surface air temperature of the model towards the pseudo-observations. Then, the amplitude of the diagnosed nudging term is taken into account in the computation of the weight of each particle, as explained in *Dubinkina and Goosse* (2013). As for the particle filter with sequential importance resampling, two sets of experiments were performed with NPPF: one with dense pseudo-observations assimilated over the area covering the polar cap southward of 30°S and one with sparse pseudo-observations assimilated southward of 60°S. The nudging, in turn, is applied everywhere over the ocean, except the area covered by sea ice. Hereafter, the hindcasts whose initial conditions are extracted from a simulation that assimilated dense (sparse) pseudo-observations through the NPPF are referred to as HIND_NPPF$_{\mathrm{dense}}$ (HIND_NPPF$_{\mathrm{sparse}}$).

2.3 Assessment of the skill of the prediction system

On the one hand, the skill of the hindcasts is assessed through the spread of the solutions provided by the members belonging to the same ensemble. This spread is used here to quantify the uncertainty of the ensemble, but care must be taken while interpreting the spread as it does not systematically represent well the range of possibilities in a prediction (*Goddard et al.*, 2012). On the other hand, the ability of the ensemble mean to reproduce different characteristics of the sea ice present in the pseudo-observations provides a measure for the accuracy of the prediction. Both will thus be presented here.

In order to assess for how long the predictability of an initialised hindcast exceeds the one of a non-initialised experiment, the spread of the hindcast ensemble is generally compared to the spread computed from a reference simulation, the approach varying slightly from one study to another (e.g., *Pohlmann et al.*, 2004; *Phelps et al.*, 2004; *Koenigk and Mikolajewicz*, 2009; *Msadek et al.*, 2010; *Döscher et al.*, 2010; *Blanchard-Wrigglesworth et al.*, 2011; *Holland et al.*, 2013). We choose to use the prognostic potential predictability (PPP) introduced by *Pohlmann et al.* (2004) and applied in several recent studies (e.g., *Koenigk and Mikolajewicz*, 2009; *Msadek et al.*, 2010; *Holland et al.*, 2013). It consists of the ratio between the ensemble spread of the hindcasts and the variance of a control simulation. For a given variable x at lead time t,

$$\mathrm{PPP}(t) = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M-1}\sum_{j=1}^{M}\left[x_{ij}(t) - X_i(t)\right]^2}{\sigma_{\mathrm{clim}}^2}, \tag{1}$$

where i is the ensemble index (N ensembles initialised at different times), j is the member index within one ensemble (M members per ensemble), x_{ij} is the simulated variable x in the hindcast member j of the ensemble i, X_i is the ensemble mean of the ensemble i and σ_{clim}^2 is the variance of a 1000-yr control simulation with constant pre-industrial greenhouse gas levels taken from the year 1850. If the simulated variable corresponds to monthly or seasonal mean, as it is the case in Sect. 3.1, the variance σ_{clim}^2 is computed individually for each month or season of the year.

A value of the PPP close to 1 means that the ensemble spread is much smaller than the natural variability, indicating the existence of predictability arising from the knowledge of the initial state. On the contrary, when the PPP is close to 0 or negative, the ensemble spread equals or outstrips the natural variability, meaning that the potential predictability is lost. The significance of the PPP is assessed based on an F-test that takes into account the autocorrelation, as in *Pohlmann et al.* (2004).

We have to keep in mind that a high value of the PPP does not ensure that the ensemble mean constitutes an accurate prediction: an ensemble can display a small spread while disagreeing with the observed state. In the present study, we go a step further and compute the anomaly correlation coefficient (ACC) or the ordinary Pearson correlation, depending on the timescales considered, that tell us how well the hindcasts reproduce the year-to-year evolution of the pseudo-observations (Sect. 3.1) or the pseudo-observed trends spanning several decades (Sect. 3.2). For the ACC, we follow the formulation of *Pohlmann et al.* (2009):

$$ACC(t) = \frac{\sum_{i=1}^{N}\left[X_i(t) - \bar{X}\right]\left[o_i(t) - \bar{o}\right]}{\sqrt{\sum_{i=1}^{N}\left[X_i(t) - \bar{X}\right]^2 \sum_{i=1}^{N}\left[o_i(t) - \bar{o}\right]^2}}, \tag{2}$$

where t is the lead time (in years), X_i is the ensemble mean of the i^{th} hindcast for the simulated variable x, i is the ensemble index (different indices correspond to different times when the hindcast simulations are started). N is the number of ensembles. o_i is the pseudo-observation covering the time period spanned by the ensemble i. The overbar stands for the climatological mean computed from a reference simulation ($\bar{X}$) and of the pseudo-observations ($\bar{o}$) over the analysed period 1900-2000. The significance of the correlation is assessed thanks to a two-sided t-test.

To sum up, the predictive skill achieved thanks to different initialisation procedures at interannual to multi-decadal timescales is estimated here through the computation of the PPP and the correlation for different variables related to the sea ice and to the temperature in the Southern Ocean. These two skill measures complement each other since the PPP tells us about the scatter accross the solutions of a hindcast ensemble, while the correlation estimates the agreement between the pseudo-observations and the ensemble mean of the hindcast.

These skill scores are computed for the ice edge location—the latitude where the ice concentration in the Southern Ocean reaches 15%—, the sea ice extent (SIE)—the sum of the areas of all the model grid cells where the sea ice concentration is at least 15%—, and the ocean heat content in the upper 100m of the ocean. Analysing the predictive skill for the ice edge location provides an overview of the regional distribution of the predictability of the sea ice and allows an easy comparison with the results of *Holland et al.* (2013). The sea ice extent is a widely used metric in sea ice studies (e.g., *Turner et al.*, 2013; *Polvani and Smith*, 2013; *Germe et al.*, 2014) which provides an integrated view over the whole Southern Ocean. An alternative could be an analysis of the sea ice area which is often considered as a more natural measure of the total sea ice coverage. While the conclusions are generally most sensitive to the choice of ice extent or area when the results are compared to real observations (e.g., *Notz*, 2014), the choice of the metric for the sea ice cover should not have a large impact on our results since the analyses are here performed in a perfect model framework. Unless specified, the ocean heat content is computed over the area southward of 60°S, between 0 and 100m below the surface. This depth is close to the depth of convection reached in winter in LOVECLIM1.2 in most parts of the Southern Ocean, except in the area where deep convection occurs.

3 Results

3.1 Interannual to decadal predictability

In this section, we discuss the predictability of the Antarctic sea ice, from 1 month to 10 years ahead. The predictability of the ice edge location is computed for monthly means in order to compare our results to the recent study of *Holland et al.* (2013). Besides, we have chosen to discuss the interannual evolution of summer (average over January, February and March) and winter (average over July, August and September) sea ice extents separately rather than a month to month evolution in order to specifically investigate the difference in predicatbility between the seasons.

First, we analyse the PPP and the ACC for the hindcasts HIND_perfect (Sect. 3.1.1). In these hindcasts, all the model variables are initialised with values provided by the pseudo-observations (see Table 1). Second, the predictability provided through distinct initialisation methods is discussed for the sea ice extent (Sect. 3.1.2).

3.1.1 Predictability of the ice edge location

In the hindcasts HIND_perfect, the PPP of the ice edge location displays values between 0.7 and 1 during the first two months of integration everywhere around Antarctica (Fig. 2a). It then decreases everywhere in the Southern Ocean. Nevertheless, in the eastern Weddell Sea and in the western Indian Ocean (between 15˚E and 55˚E), in the Ross Sea (between 180˚E 210˚E) and in the Bellingshausen and Amundsen Seas (between 230˚E and 290˚E) potential predictability reemerges after a few months, in May or June of the first year, and persists until September or October. After 2 years the PPP becomes very low in all the sectors. Those results are in good agreement with the ones discussed in *Holland et al.* (2013), confirming the relevance of our study based on the model LOVECLIM1.2. *Holland et al.* (2013) have also highlighted an eastward propagation of the predictability related to the eastward flow of the ocean and sea ice. Such a propagation may play a role in our experiments in the Bellingshausen and Amundsen Seas, while the reemergence of the anomaly is the dominant feature in the other sectors.

The regions with high PPP are also characterised by high ACC of the ice edge location in the hindcasts HIND_perfect. It reaches values of at least 0.6 during the first three months of integration at all longitudes (Fig. 2b). In the eastern Weddell Sea and in the western Indian Ocean (between 0˚E and 40˚E) and in the Bellingshausen and Amundsen Seas (between 230˚E and 300˚E) the ACC remains higher than 0.6 until the end of the first year. This means that in these areas accurate prediction of the ice edge location potentially can be performed a year ahead. This is associated with high predictability of the sea ice concentration near the ice edge in those regions (not shown). In the western Pacific Ocean sector (between 90˚E and 160˚E) the ACC is close to 0 from April of the first year but higher values of the ACC reemerge between May and July of the first year. In December of the first year the ACC is lower than 0.4 at all longitudes. Nevertheless, in May of the second year the ACC reaches again values higher than 0.6 in the eastern Weddell Sea and in the western Indian Ocean (between 15˚E and 70˚E) as well as in the Bellingshausen and Amundsen Seas (between 250˚E and 300˚E).

The reemergence of predictability in winter is thus a dominant characteristic of the predictability of Antarctic sea ice. It cannot be accounted for by the memory of the sea ice itself. Indeed, the persistence of the Antarctic sea ice is very weak since it disappears almost entirely during the melting season. The memory is more likely provided by heat anomalies stored in the ocean, as proposed for instance by *Holland et al.* (2013). In LOVECLIM1.2 the high ACC of the ocean heat content (southward of 60˚S, in the first 100m below the surface) during the first two years of integration agrees well with this hypothesis (Fig. 3a). Note that the correlation between the ice edge location and the ocean heat content is particularly strong (in absolute value) between 0˚E and 50˚E and between 150˚E and 300˚E, especially during winter months, where the ACC of the ice edge is also high.

Given that the mixed layer is shallower in summer than in winter, the ocean surface is isolated from deeper levels during this season. In winter, thanks to the cooling and brine rejection during the formation of sea ice, the mixed layer is deeper and the interactions between the surface and the interior ocean are stronger. This leads to an enhanced heat flux from the ocean to the surface that plays a dominant role in the formation of sea ice, contributing to the significant negative correlation between the ice edge location and the ocean heat content below the sea ice (Fig. 3b). The changes in sea ice concentration also impact the ocean heat content, the two variables being linked by various feedback processes (e.g., *Martinson et al.*, 1981; *Goosse and Zunz*, 2014). Because of the multiple interactions it was not possible in the present framework to determine precisely to which extent the ocean drives the sea ice changes or if conversely, sea ice changes drive the ones in the ocean (not shown). Nevertheless, given the low persistence of the sea ice in the Southern Ocean, the results of Fig. 3 reasonably support the hypothesis that the high values reached by the ACC of the ice edge location during winter can be accounted for by the high ACC of the ocean heat content, achieved thanks to the strong thermal inertia of the ocean.

3.1.2 Predictability of the sea ice extent

The discussion of the predictive skill for the ice edge location presented above is focused on hindcast simulations initialised with perfect initial conditions. For this variable the values reached by the PPP and the ACC decrease rapidly after the initialisation and barely reach significant values, except in winter, after the first year of simulation. The skills of the simulations initialised with other methods display similar patterns but were even lower and were thus not presented for brevity. The PPP and the ACC may be better for a global variable such as the sea ice extent. Indeed, local errors of different signs in different sectors, which have a large effect on the PPP and the ACC of the ice edge location, may balance each other and lead to a better skill for the sea ice extent, in particular for the hindcasts that are not initialised with perfect initial conditions. It is thus instructive to analyse the PPP and the ACC of the sea ice extent for all the different initialisation methods tested here.

The PPP of summer SIE starts from a maximum in the first year of integration, gets close to the significance level or even falls below it in the second or in the third year of integration, except for the hindcasts HIND_NUD$_{\mathrm{sparse}}$ whose PPP never reaches statistically significant values. After three years of integration the PPP barely reaches significant values for any initialisation method (Fig. 4a). In the first year, whether the dataset used to initialise the hindcasts is dense or sparse strongly impacts the PPP, especially when the nudging proposal particle filter or the particle filter with sequential importance resampling are used to assimilate the data (blue and orange lines in Fig. 4a). Indeed, in the first year the values of the PPP of the hindcasts HIND_NPPF$_{\mathrm{dense}}$ and HIND_SIR$_{\mathrm{dense}}$ reach 0.70 and 0.65 respectively, while the PPP is much lower in the first year for the hindcasts HIND_NPPF$_{\mathrm{sparse}}$ and HIND_SIR$_{\mathrm{sparse}}$ (0.17 and 0.24 respectively). To a lesser extent, the hindcasts initialised through the nudging also provide a smaller PPP when sparse pseudo-observations are used but the PPP is always low when this DA method is used.

The low value of the PPP for the hindcasts HIND_NUD$_{\mathrm{dense}}$ and HIND_NUD$_{\mathrm{sparse}}$ indicates a large spread of the ensemble, already present at the initialisation of the hindcasts. In the simulation assimilating pseudo-observations through a nudging, the nudging term applied on each member of the ensemble tends to maintain the members close to each other. However, as explained in Sect. 2.2, the nudging is not applied over the sea ice covered area, obviously reducing the constraint from the pseudo-observations on the sea ice extent. On the contrary, sea ice covered area is included in the assimilation domain of the particle filters in SIR and NPPF, providing a stronger constraint from the pseudo-observations on the sea ice extent. ~~Note that, in the nudged simulations the members of the ensemble are also perturbed every three months in order to maintain the consistency with the experimental setup of the other simuations.~~

The fact that the PPP of the hindcasts initialised with sparse data is systematically lower than the one of the hindcasts initialised through the same method but with dense data can be accounted for by the weaker constraint applied on the initial state when sparse pseudo-observations are used. This results in a larger spread of the ensemble and, thus, in a lower PPP. The weaker constraint is first accounted for by the fewer amount of data available in the sparse pseudo-observations. In addition, when the particle filters (SIR or NPPF) are used, the domain over which the particle filters assimilate the data is smaller when sparse pseudo-observations are assimilated, also reducing the constraint on the ensemble.

As mentioned in Sect. 2.2, the reduction of the assimilation domain for the particle filters was required in order to avoid filter degeneracy in the simulation with assimilation. Fig. 4a and b illustrate that changing the dataset from dense pseudo-observations to sparse pseudo-observations decreases the constraint on the initial state much more than changing the DA method from the NPPF to the SIR. Consequently, particular attention has to be paid when interpreting the PPP because changes in the ensemble spread may be strongly related to the experimental design, such as the choice of the domain where the data are assimilated, and not on the predictability of the system itself.

The PPP of winter sea ice extent displays a slower decrease than the PPP for summer sea ice extent and stays above the significance level until the 7th year when initialised with perfect initial conditions (Fig. 4b). In the hindcasts HIND_NPPF$_{\mathrm{dense}}$ and HIND_SIR$_{\mathrm{dense}}$ the PPP remains significant during the first four years and in the hindcasts HIND_NPPF$_{\mathrm{sparse}}$ and HIND_SIR$_{\mathrm{sparse}}$ during the first three years. The hindcasts HIND_NUD$_{\mathrm{dense}}$ and HIND_NUD$_{\mathrm{sparse}}$ provide a PPP that is always below the 95% significance level. Overall, these results indicate that the winter sea ice extent is more predictable than the summer sea ice extent. The square of the autocorrelation (r^2) is used here to assess the predictability that is gained from the persistence. Fig. 4a and b indicate that the persistence of the winter sea ice extent is slightly higher than the one of summer sea ice extent during the first three years of integration but this persistence is too small to be responsible for the higher predictability obtained for winter sea ice extent. The higher PPP of the winter sea ice extent, compared to the summer sea ice extent, likely arises because of a stronger interaction between the sea ice and the interior ocean during winter, as discussed in the case of the reemergence of the predictability of the ice edge location in winter in Sect. 3.1.1.

Having assessed the PPP of the hindcasts of the Antarctic sea ice extent, we now focus on their accuracy through the computation of the ACC. The ACC for both summer and winter is always positive, even for the hindcasts HIND_noinit. This is due to the fact that all the hindcasts as well as the pseudo-observations are driven by the same external forcing that ensures at least a weak correlation between the hindcasts and the pseudo-observations.

In summer its value is, however, rarely above the significance level except in the first year for the hindcasts HIND_perfect and HIND_NPPF$_{sparse}$. Furthermore, the choice of the DA method and the use of dense or sparse pseudo-observations to initialise the hindcasts do not lead to a major difference in the ACC. In general, the ACC of sea ice extent in winter is larger than in summer (Fig. 4d), which is in agreement with the ACC of the ice edge location. For most of the initialisation methods, the ACC remains statistically significant during at least the first three years of integration. Moreover, during the first three years of integration the ACC of all hindcasts initialised with pseudo-observations, except HIND_SIR$_{sparse}$, outstrips the ACC of the hindcasts that were not initialised with pseudo-observations, though the improvement is weak and none of the tested intialisation methods systematically provides a higher ACC than the others.

The relative performance of each simulation with data assimilation is discussed in detail in *Dubinkina and Goosse* (2013). They demonstrated good performance of the reconstruction of atmosphere as well as ocean variables provided by the particle filters (SIR and NPPF). For the reconstructed sea ice concentration, a clear improvement is obtained with the NPPF compared to the SIR. *Dubinkina and Goosse* (2013) showed that the assimilation of pseudo-observations through the nudging provides satisfying reconstructions of surface air temperature and sea ice concentration in the Southern Hemisphere. However, because of the spurious impact of the nudging on the mixed layer dynamics in the model, the ocean temperature at depth and the surface salinity are not well simulated.

Despite those differences in the initial state, all the initialisation procedures tested here lead to relatively similar ACC. Firstly, this is due to the low ACC brought by the initialisation. The ACC of HIND_perfect, corresponding to the best possible initialisation, is generally not much higher than the one obtained in all the other simulations, except for the first year. Secondly, additional analyses performed on the initial states provided by the various assimilation methods have shown that they all have their own biases, either on the oceanic heat content or on the salt content. This potentially affects vertical heat fluxes and thus the oceanic heat flux at surface during the hindcast, reducing the prediction skill. The larger errors are found in HIND_NUD$_{dense}$ and HIND_NUD$_{sparse}$ for which the salt content in the first 100m in the initial state of the hindcast is negatively correlated with the one of the pseudo-observations, which is in agreement with *Dubinkina and Goosse* (2013).

3.2 Multi-decadal predictability

In this section, we discuss the predictability of the trends in the ice edge location and ice extent computed over time periods from 10 to 30 years. All the time periods over which the trends are computed start on the first year of the hindcasts, i.e. directly after the initialiation. Due to the large internal variability of the Antarctic sea ice, the trends in sea ice extent can substantially vary between the members of an ensemble performed with the same climate model (e.g., *Landrum et al.*, 2012; *Zunz et al.*, 2013). As a consequence, predictions may be very uncertain. However, an efficient intialisation of the hindcasts imposes a constraint on the initial state that could decrease the spread of the trend over the years following the initialisation. A comparison between the ensemble spread of the trends in hindcast simulations and a reference variance of the trends is thus required. As for the case of decadal predictions (Sect. 3.1), we use the prognostic potential predictability (Eq. (1)) as a measure of the uncertainty of the simulated trends in sea ice. Here, the variable x in Eq. 1 stands for the trend in ice edge location (sea ice extent) for each ensemble member and the time t in Eq. (1) represents the length of the time period over which the trend is computed. The climatological variance (σ^2_{clim} in Eq. (1)) is the variance of the trends computed over successive time periods, spaced by 5 years, of a 1000-yr control run simulation with constant pre-industrial greenhouse gas levels.

In order to assess the accuracy of the hindcasts, we focus on the ensemble mean of the trends in ice edge location (sea ice extent). We compute the ordinary Pearson correlation between the trends provided by the ensemble means of each hindcast and the corresponding trends computed from the pseudo-observations. On multi-decadal timescales, the external forcing has potentially a large impact on the trend in ice edge location (sea ice extent) and we have to disentangle this contribution from the one that can be attributed to the initialisation. For that purpose we take as a reference the correlation between the pseudo-observations and the hindcasts HIND_noinit. This hence represents the amplitude of the correlation of the trends that is provided by the external forcing.

For each initialisation method, the number of hindcasts used to compute the PPP and the correlation of the 10- to 30-yr trends is smaller than the number of hindcasts used in Sect. 3.1 (15 instead of 19). Indeed, the hindcasts initialised after 1970 are not considered in this section since the simulation providing the pseudo-observations ends in December 1999.

3.2.1 Predictability of the 10- to 30-yr trend in ice edge location

For the hindcasts HIND_perfect the PPP of the trend in ice edge location generally decreases as the lead time gets longer (Fig. 5). In summer the PPP is significant in the Ross Sea and in the Bellingshausen Sea (between 200°E and 300°E). It also reaches values higher than 0.5 in the eastern Weddell Sea and in the western Indian Ocean (between 0°E and 40°E) as well as in the western Pacific sector (between 110°E and 130°E, Fig. 5a). However, those relatively high values of the PPP appear in areas that are generally sea ice free during summer, i.e. where the interannual trend in the ice edge location is close to 0 and does not vary much from one simulation to another. The PPP is thus not meaningful in those conditions.

In winter the PPP of the trend in ice edge location barely exceeds 0.5 (Fig. 5b). For length of time periods up to 20 years, the largest values of the PPP (of at least 0.3) are found in the Indian Ocean (between 40°E and 90°E), in the western Pacific sector (between 90°E and 140°E) and in the Bellingshausen Sea (between 230°E and 280°E). In the eastern western Pacific sector (around 100°E) a PPP larger than 0.3 is found for any length of time periods. Overall, even when the hindcasts are perfectly initialised the trends in ice edge display a large spread, implying a rather uncertain prediction.

The correlation between the trends in ice edge location from the hindcasts HIND_perfect and the the corresponding trends from the pseudo-observations depends also on the longitude and on the length of the period over which the trend is computed (Fig. 6). In summer correlation higher than 0.6 is found in the western Weddell Sea and in the eastern Indian Ocean (between 10°E and 40°E), in the eastern western Pacific sector (between 130°E and 160°E), in the western and eastern Ross Sea (between 160°E and 220°E) and in the Bellingshausen and Amundsen Seas (between 230°E and 300°E), especially for time periods shorter than 20 years (Fig. 6a). In winter the correlation of the trend in ice edge location is generally higher than in summer, especially for time periods longer than 20 years (Fig. 6b). The correlation is generally higher than 0.6, except in the eastern Weddell Sea and in the western Indian Ocean (between 0˚E and 50˚E) for time period longer than 20 years, in the Ross Sea (between 160˚E and 200˚E) for time period longer than 14 years and in the Bellingshausen and Amundsen Seas (between 250˚E and 300˚E) for time periods longer than 25 years.

The values obtained for the PPP and the correlation of the trends in ice edge location for the hindcasts HIND_perfect confirm the relevance of these two scores to assess the skill of a prediction system. Indeed, the PPP and the ACC provide complementary information that can even appear contradictory, as for Fig. 5 and 6. On the one hand, the PPP of the ice edge location in the hindcasts HIND_perfect indicates that the spread of the ensemble is relatively large compared to the one of an uninitialised ensemble, i.e. that the potential predictability is low. On the other hand, the correlation reaches relatively high and statistically significant values, meaning that the trends computed from the hindcasts ensemble mean agree reasonably well with the corresponding trends in the pseudo-observations. This apparent disagreement between the low PPP and the high correlation could actually be accounted for by the fact that the ensemble mean of the trends is driven by the external forcing and the slowly varying components of the climate system such as the ocean, as discussed below. Besides, the trends provided by the individual members of the ensemble are widely scattered around the trend of the ensemble mean because these individual trends are influenced by the high frequency variability of the atmosphere, unpredictable at that timescale, that gives a range close to the one of an unitialised ensemble. Nevertheless, the overall decrease in PPP for increasing length of time periods over which the trend is computed does not necessarily imply an increase in the variance of the hindcast trends. Indeed, the climatological variance $\sigma^2_{\mathrm{clim}}(t)$ of the trend decreases with t, the length of the period over which the trend is computed. Consequently, although PPP decreases with t, the variance of the hindcast, giving a kind of measure of the uncertainty of the prediction, may still decrease.

High values of the correlation for summer and winter are much less widespread in the hindcasts HIND_noinit (Fig. 6c, d). In summer (winter) 8% (4%) of the values of the correlation shown on Fig. 6c (Fig. 6d) are statistically significant at the 95% level. For comparison, 39% (69%) of the correlations are statistically significant for the summer (winter) ice edge location of the hindcasts HIND_perfect (Fig. 6a,b). Therefore, statistically significant values of the correlation of the trends in the hindcasts HIND_noinit are rather marginal and are likely not meaningful as the percentages of statistically significant values are close to the p-value of 0.05 used to perform the statistical test.

Overall, the correlation of the trends in ice edge location for the hindcasts HIND_perfect and HIND_noinit shows that a large part of the predictability is likely provided by the initialisation and not by the response to the forcing. This indicates that even if the state of the sea ice in a given year is not predictable beyond three years ahead (see Sect. 3.1) the 10- to 30-yr trend in ice edge location is highly correlated between the hindcasts and the pseudo-observations.

As already proposed in Sect. 3.1 for interannual variations, the high correlation of the trend in ice edge location found in the hindcasts HIND_perfect is likely due to the thermal inertia of the ocean, which allows the anomalies characterising the initial state to impact the evolution of the system during years to decades with potential feedbacks between the oceanic heat content and the ice concentration. Indeed, in the hindcasts HIND_perfect the trend in annual mean ocean heat content (southward of 60°S, in the first 100m below the ocean surface) displays a correlation between the hindcasts and the pseudo-observations larger than 0.6 between 0°E and 75°E as well as between 190°E and 310°E (Fig. 7a). Furthermore, at these longitudes the correlation between the trend in ice edge location and the trend in ocean heat content reaches values close to -1 in winter (Fig. 7c) with less negative values in summer (Fig. 7b). The same analysis performed for the ocean heat content in the upper 300m provides results similar to the one shown here for the ocean heat content in the upper 100m. This high correlation of the trend in ocean heat content with pseudo-observations, combined with a large anti-correlation of the trend in ice edge location with the trend in ocean heat content, likely explains the large correlation of the trend in ice edge location between the hindcasts and the pseudo-observations, as shown in Fig. 6. However, between 75°E and 120°E the trend in winter ice edge location in the hindcasts is well correlated with the pseudo-observations (Fig. 6b), while the trend in ocean heat content does not display such a high correlation with the pseudo-observations. In some sectors of the Southern Ocean, additional mechanisms thus play a dominant role in the predictability of the ice edge location (e.g., transport from adjacent areas) and complementary studies are required to evaluate more precisely the origin of the predictability of the trend for each region.

3.2.2 Predictability of the 10- to 30-yr trend in sea ice extent

In this section, the predictability of the trends in summer and winter sea ice extents computed over increasing length of time period from 10 to 30 years is analysed. As for the 10- to 30-yr trend in ice edge location (Sect. 3.2.1), we have computed the prognostic potential predictability of the trend in sea ice extent (Fig. 8a, b) and the correlation between the ensemble means of the trends provided by hindcasts using different initialisation date and the corresponding trends in the pseudo-observations (Fig. 8c, d). These measures of the predictability are computed for the trends over 10- to 30-yr long time period (starting on the first year after the initialisation) shown on the x-axis of Fig. 8 and for different initialisation methods.

As discussed in Sect. 3.1.2 for the predictive skill of the sea ice extent at interannual to decadal timescales, the relative merits of the different initialisation methods tested here do not appear very clearly in the analysis of the trends in sea ice extent. Therefore for brevity, the results are shown here for the hindcasts HIND_perfect, HIND_noinit and HIND_SIR$_{\text{dense}}$. The skill of HIND_SIR$_{\text{dense}}$ is among the highest of all the experiments though HIND_SIR$_{\text{dense}}$ is not systematically better and its results are generally close to the ones obtained using other methods. Nevertheless, the results provided by the initialisation through the SIR appear more reliable to us since the SIR relies on a method that does not introduce any additional term in the model equations, ensuring that the model dynamics is preserved. Indeed, in our experimental design the nudging, when used alone, does not respect the ocean dynamics. Note, however, that this problem does not seem to occur when the nudging is combined with a particle filter in the NPPF (*Dubinkina and Goosse*, 2013).

The PPP of the trend in summer sea ice extent reaches at most 0.36. It is statistically significant at the 95% level up to 27-yr long time period for the hindcasts HIND_perfect (purple solid lines in Fig. 8a, b). The hindcasts HIND_SIR$_{dense}$ (orange solid line in Fig. 8a, b) is above the 95% significance level, for 10- to 13-yr long time periods. In winter the PPP is significant up to 20 years for the hindcasts HIND_perfect and up to 19 years for the hindcasts HIND_SIR$_{dense}$. For the other methods, the PPP is rarely significant for the trends in both summer and winter sea ice extent (not shown). As noticed for the PPP of the sea ice extent at interannual timescales (Sect. 3.1.2), for a given data assimilation method the PPP of the trend in sea ice extent of the hindcasts initialised with dense pseudo-observations is systematically higher than the PPP of the hindcasts initialised with sparse pseudo-observations.

For both seasons the correlation of the trend in sea ice extent between hindcasts and pseudo-observations is larger if pseudo-observations are taken into account at the initialisation, for any length of time period (Fig. 8c, d). In summer only the correlation computed from the hindcasts HIND_perfect reaches statistically significant values. The negative values obtained in summer for the hindcasts HIND_noinit (Fig. 8c) are not meaningful given that they are not statistically significant and they are not robust. Indeed, the same correlation computed with another set of pseudo-observations can provide slightly positive, still not statistically significant, values (not shown). In winter, however, both the hindcasts HIND_perfect and HIND_SIR$_{dense}$ provide statistically significant correlation of the trends in sea ice extent, while the correlation for the non-initialised hindcast is close to 0 and not statistically significant. This indicates again that a part of the predictability cannot be accounted for by the external forcing and arises from the initialisation.

The simulated trend in sea ice extent at multi-decadal timescales is due to a combination of the model internal variability and of the response to the external forcing. The external forcing is responsible for an overall decrease in sea ice extent between 1900 and 2000. The interannual variability is associated with positive and negative trends at multi-decadal timescales that are superimposed on this longer term trend, essentially externally driven. For sea ice extent in the Southern ocean those internally generated variations are much larger at decadal timescales than the externally driven decrease.

In the hindcasts HIND_noinit the internal variability of the sea ice extent present in the pseudo-observations is not captured in the ensemble mean and the trend in sea ice extent at multi-decadal timescales is essentially driven by the external forcing. As a consequence, the correlation between the pseudo-observed sea ice extent in the year preceding the initialisation and the trend in the hindcast sea ice extent is weakly positive (red dashed lines with circle markers in Fig. 9): since the model response under warming conditions is a melting of the sea ice in the Southern Ocean, the trend in sea ice extent tends to be more negative over time as mean ice extent is decreasing.

The goal of the initialisation of the hindcast is to reproduce the internally generated fluctuations in the pseudo-observations. First, as the trend in see ice extent is computed starting from the first year of the hindcast, it depends directly on the value of this extent in pseudo-observations the year preceding the initialisation of the hindcasts. Furthermore, because of the model dynamics (*Goosse and Zunz*, 2014), when initialised with a state that has a sea ice extent larger (smaller) than the climatological mean, a simulation generally provides a negative (positive) trend in sea ice extent during the following years. This results in a negative correlation between the sea ice extent in the pseudo-observations the year preceding the initialisation and the trend in sea ice extent in the hindcasts, especially in winter (purple dashed lines with circle markers in Fig. 9). Note that in contrast to *Goosse and Zunz* (2014), the simulations analysed here are transient simulations and the state averaged over several decades is thus not stationnary but slightly decreases in response to the external forcing.

This suggests that initialising a hindcast with a sea ice extent anomaly that fits the one of the pseudo-observations is a necessary condition to ensure an accurate prediction of the trend at multi-decadal timescales. This is, however, not sufficient since not only the state of the sea ice but also the state of the water column below the sea ice must be initialised with a state close to the pseudo-observations, given that the information provided to the ocean at the initialisation can impact the system over several decades, as already discussed at the end of Sect. 3.2.1. As expected, the trends in both summer and winter sea ice extents are strongly anti-correlated with the trend in ocean heat content southward of 60°S in the first 100m below the surface (solid lines in Fig. 9), similar results being obtained with the ocean heat content in the upper 300m (not shown). Consequently, it is essential to reproduce well this trend (Fig. 10) and thus the initial oceanic heat content as both are well anti-correlated (purple dotted lines with triangle markers in Fig. 9). As the data assimilation methods tested in the present study assimilate the surface air temperature only, the information provided by the pseudo-observations is not always adequately propagated in the ocean. The correlation of the trend in ocean heat content between the hindcasts and the pseudo-observations is maximum for the hindcasts HIND_perfect and is close to zero for the hindcasts HIND_noinit (Fig. 10).

To sum up, taking into account the pseudo-observations in the initialisation of the hindcasts straight from the pseudo-observations dataset (HIND_perfect) or from a DA simulation does not provide particularly high values for the PPP but clearly leads to more accurate ensemble mean at multi-decadal timescales. In particular, the higher correlation between the trends in sea ice extent of the hindcast and the corresponding ones of the pseudo-observations indicates that the initialisation with pseudo-observations triggers a shift of the ensemble mean of the trends towards the one of the pseudo-observations. Consequently, the initialisation with pseudo-observations potentially improves the correlation with pseudo-observations for the trends in ice edge location (sea ice extent) over several decades while the initialisation has only a weak impact on the predictive skill for the ice edge location (sea ice extent) at interannual to decadal timescales, as discussed in Sect. 3.1. This apparent disagreement may be accounted for by the fact that on timescales from months to several years the behaviour of the sea ice is strongly impacted by the quickly varying atmosphere. This high frequency variability tends to overwhelm the more predictable low frequency signal that could be provided by the ocean. Besides, on multi-decadal timescales it can be reasonably assumed that the variations in the ice edge location (sea ice extent) are mainly driven by the slowly varying ocean, limiting the impact of the unpredictable atmosphere.

4 Summary and conclusions

All the results discussed in the present study have been performed in a perfect model framework. Similar tests performed in a realistic framework, i.e. with the use of actual observations for both the initialisation and the verification of the hindcasts, would lead to a lower predictability, due to the models biases. Overall, even under such idealised conditions, the predictive skill of the model for the Antarctic sea ice is quite poor compared to other variables (e.g., *Kim et al.*, 2012; *Matei et al.*, 2012b). The analyses performed here have nevertheless highlighted interesting characteristics of the predictability and the forecast capability achieved thanks to different initialisation methods for the sea ice in the Southern Ocean at interannual to multidecadal timescales.

Firstly, in agreement with the recent study of *Holland et al.* (2013), the impact of the initialisation on the short-term predictability seems to be mainly driven by oceanic processes. More specifically, the predictability of the sea ice at interannual timescales is low in summer and increases in winter. This reemergence of the predictability in winter is provided by heat anomalies stored in the ocean. In summer these anomalies are isolated from the surface due to the weak vertical mixing in the ocean during this season. Conversely, in winter the enhanced vertical mixing allows these anomalies to reach the surface and impact the sea ice formation.

Secondly, the predictability of the Antarctic sea ice behaves very differently depending on the timescale considered. At interannual timescales, during the first three years of integration, the variance of the members within an ensemble is smaller than the climatological variance of the model. This suggests that the uncertainty of the ensemble mean is rather low. However, the scatter of the members depends on the perturbation method used to generate the ensemble and may underestimate the real uncertainty of the ensemble. Besides, at interannual timescales the signal that can be provided by the ocean is largely overwhelmed by the unpredictable variabiliy imposed by the atmosphere, resulting in a relatively weak correlation between the hindcasts ensemble mean and the corresponding pseudo-observations even during the first years of integration.

Although the predictability of the Antarctic sea ice during a particular year is limited to a few years ahead at best, there exists predictability of the trend in sea ice at multi-decadal timescales. Indeed, the correlation between the ensemble mean of the trends in hindcasts and the corresponding trend from the pseudo-observations easily reaches values greater than 0.5 in winter. Furthermore, the accuracy of the multidecadal trends in sea ice extent computed from hindcasts initialised with pseudo-observations is better compared to hindcasts initialised without taking into account any pseudo-observations. This indicates that at multi-decadal timescales the predictability is due to not only the external forcing but also the initialisation method as the initialisation with pseudo-observations clearly improves the accuracy of the prediction in our idealised framework. Besides, the spread of the trends within an ensemble roughly equals the variance of the model climatology, meaning that the uncertainty of the trends remains relatively large. In this framework, the ocean appears to impact the ensemble mean of the trend in sea ice while the atmosphere is responsible for the scatter of the members around this ensemble mean.

Thirdly, the method and the density of the pseudo-observations used to initialise the hindcasts influence the predictability of the sea ice. At both interannual and multi-decadal timescales the spread of the members within an ensemble is smaller if the hindcasts are initialised with dense pseudo-observations compared to an initialisation with sparse pseudo-observations. This is due to the stronger constraint applied at the initialisation, preventing the members to spread too quickly during the integration. The initialisation method also impacts the accuracy of the trend in hindcast simulations. The hindcasts initialised with perfect initial conditions display the highest correlation for the trend in sea ice extent as well as for the trend in ocean heat content. Therefore, we pointed out a clear link between the predictability of the sea ice and the quality of the initialisation of the ocean below it.

In our experiments, the relative skills at interannual to multi-decadal timescales of the initialisation methods based on data assimilation depend on the season, as well as on timescale investigated. None of the methods tested here has thus been clearly identified as the best suited for multi-decadal predictions of Antarctic sea ice. However, in the experimental setup employed here to assimilate pseudo-observations with a nudging, *Dubinkina and Goosse* (2013) have demonstrated that the nudging leads to a behaviour incompatible with the original model dynamics. Such incompatibility does not arise when the particle filters (SIR or NPPF) are used. If possible, the initialisation through a particle filter or a similar method should thus be preferred for studying the sea ice in the Southern Ocean.

We have applied here initialisation methods based on the assimilation of the surface air temperature only because relatively long time series are available and similar methods have been used in previous studies. Unfortunately, the persistence of surface variables is very low. Therefore, in order to provide contraints on longer term predictions, the data assimilation scheme has to propagate this information at depth in the ocean, where the persistence is much longer. Such a propagation is not always achieved with the methods used here. Therefore, in parallel to the models biases reduction efforts should be concentrated on better intialisation of the ocean. Furthermore, we have shown that the predictability is improved by the use of dense pseudo-observations in the initialisation procedure. Collecting observations in the Southern Ocean is thus crucial not only to improve the understanding of the processes occurring there but also to better initialise the simulations used to forecast the evolution of the sea ice around Antarctica.

Acknowledgements V. Zunz is Research Fellow with the Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agronomie (FRIA-Belgium). H. Goosse is Senior Research Associate with the Fonds National de la Recherche Scientifique (F.R.S. – FNRS-Belgium). This work is supported by the Belgian Federal Science Policy (Research Program on Science for a Sustainable Development). Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fond de la Recherche Scientifique de Belgique (FRS-FNRS). We thank three anonymous referees for their careful reading and constructive comments.

References

Bintanja, R., G. J. van Oldenborgh, S. S. Drijfhout, B. Wouters, and C. A. Katsman (2013), Important role for ocean warming and increased ice-shelf melt in Antarctic sea-ice expansion, *Nature Geosci*, *6*(5), 376–379.

Bitz, C. M., and L. M. Polvani (2012), Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model, *Geophys. Res. Lett.*, *39*(20), doi:10.1029/2012GL053393.

Bitz, C. M., P. R. Gent, R. A. Woodgate, M. M. Holland, and R. Lindsay (2006), The Influence of Sea Ice on Ocean Heat Uptake in Response to Increasing CO2, *Journal of Climate*, *19*(11), 2437–2450.

Blanchard-Wrigglesworth, E., C. M. Bitz, and M. M. Holland (2011), Influence of initial conditions and climate forcing on predicting Arctic sea ice, *Geophys. Res. Lett.*, *38*(18).

Brohan, P., J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones (2006), Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, *111*(D12), doi:10.1029/2005JD006548.

Brovkin, V., J. Bendtsen, M. Claussen, A. Ganopolski, C. Kubatzki, V. Petoukhov, and A. Andreev (2002), Carbon cycle, vegetation, and climate dynamics in the Holocene: Experiments with the CLIMBER-2 model, *Global Biogeochemical Cycles*, *16*(4), 1139, doi:10.1029/2001GB001662.

Comiso, J. (1999, updated daily), *Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS. Version 2, January 1980 to December 2009*, Boulder, Colorado USA: NASA DAAC at the National Snow and Ice Data Center.

Comiso, J. C., and F. Nishio (2008), Trends in the sea ice cover using enhanced and compatible AMSR-E, SSM/I, and SMMR data, *J. Geophys. Res.*, *113*(C02S07), doi:10.1029/2007JC004257.

Day, J. J., S. Tietsche, and E. Hawkins (2014), Pan-arctic and regional sea ice predictability: Initialization month dependence, *Journal of Climate*, *27*(12), 4371–4390.

Döscher, R., K. Wyser, H. E. M. Meier, M. Qian, and R. Redler (2010), Quantifying Arctic contributions to climate predictability in a regional coupled ocean-ice-atmosphere model, *Climate Dynamics*, *34*(7-8), 1157–1176, doi:10.1007/s00382-009-0567-y.

Dubinkina, S., and H. Goosse (2013), An assessment of particle filtering methods and nudging for climate state reconstructions, *Climate of the Past*, *9*(3), 1141–1152, doi:10.5194/cp-9-1141-2013.

Dubinkina, S., H. Goosse, Y. Sallaz-Damaz, E. Crespin, and M. Crucifix (2011), Testing a particle filter to reconstruct climate changes over the past centuries, *International Journal of Bifurcation and Chaos*, *21*(12), 3611–3618, doi:10.1142/S0218127411030763.

Dunstone, N. J., and D. M. Smith (2010), Impact of atmosphere and sub-surface ocean data on decadal climate prediction, *Geophys. Res. Lett.*, *37*(2), doi:10.1029/2009GL041609.

Fetterer, F., K. Knowles, W. Meier, and M. Savoie (2002, updated daily), *Sea Ice Index, January 1980 to December 2009*, Boulder, Colorado USA: National Snow and Ice Data Center, doi:http://dx.doi.org/10.7265/N5QJ7F7W.

Germe, A., M. Chevallier, D. Salas y Mélia, E. Sanchez-Gomez, and C. Cassou (2014), Interannual predictability of Arctic sea ice in a global climate model: regional contrasts and temporal evolution, *Climate Dynamics*, pp. 1–20, doi:10.1007/s00382-014-2071-2.

Goddard, L., et al. (2012), A verification framework for interannual-to-decadal predictions experiments, *Climate Dynamics*.

Goosse, H., and T. Fichefet (1999), Importance of ice-ocean interactions for the global ocean circulation: A model study, *Journal of Geophysical Research: Oceans*, *104*(C10), 23,337–23,355.

Goosse, H., and V. Zunz (2014), Decadal trends in the Antarctic sea ice extent ultimately controlled by ice–ocean feedback, *The Cryosphere*, *8*(2), 453–470, doi:10.5194/tc-8-453-2014.

Goosse, H., W. Lefebvre, A. de Montety, E. Crespin, and A. Orsi (2009), Consistent past half-century trends in the atmosphere, the sea ice and the ocean at high southern latitudes, *Climate Dynamics*, *33*(7), 999–1016–1016.

Goosse, H., et al. (2010), Description of the Earth system model of intermediate complexity LOVECLIM version 1.2, *Geoscientific Model Development*, *3*(2), 603–633, doi:10.5194/gmd-3-603-2010.

Holland, M. M., E. Blanchard-Wrigglesworth, J. Kay, and S. Vavrus (2013), Initial-value predictability of Antarctic sea ice in the Community Climate System Model 3, *Geophysical Research Letters*, *40*(10), 2121–2124, doi:10.1002/grl.50410.

Holland, P. R., and R. Kwok (2012), Wind-driven trends in Antarctic sea-ice drift, *Nature Geosci*, *5*(12), 872–875.

Kalnay, E. (2007), *Atmospheric Modeling, Data Assimilation and Predictability*, 4 ed., Cambridge University Press, Cambridge.

Keenlyside, N., M. Latif, J. H. Jungclaus, L. Kornbueh, and E. Roeckner (2008), Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, *453*, 84–88, doi:10.1038/nature06921.

Kim, H.-M., P. J. Webster, and J. A. Curry (2012), Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts, *Geophys. Res. Lett.*, *39*(10).

Kirkman, C. H., and C. M. Bitz (2010), 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*(9), 2224–2237.

Koenigk, T., and U. Mikolajewicz (2009), Seasonal to interannual climate predictability in mid and high northern latitudes in a global coupled model, *Climate Dynamics*, *32*(6), 783–798.

Kröger, J., W. Müller, and J.-S. von Storch (2012), Impact of different ocean reanalyses on decadal climate prediction, *Climate Dynamics*, *39*(3), 795–810.

Landrum, L., M. M. Holland, D. P. Schneider, and E. Hunke (2012), Antarctic Sea Ice Climatology, Variability, and Late Twentieth-Century Change in CCSM4, *Journal of Climate*, *25*(14), 4817–4838.

Lefebvre, W., and H. Goosse (2008), An analysis of the atmospheric processes driving the large-scale winter sea ice variability in the Southern Ocean, *J. Geophys. Res.*, *113*(C2), doi:10.1029/2006JC004032.

Mahlstein, I., P. R. Gent, and S. Solomon (2013), Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations, *Journal of Geophysical Research: Atmospheres*, *118*, 1–6, doi:10.1002/jgrd.50443.

Martinson, D. G., P. D. Killworth, and A. L. Gordon (1981), A Convective Model for the Weddell Polynya, *Journal of Physical Oceanography*, *11*(4), 466–488.

Matei, D., H. Pohlmann, J. Jungclaus, W. Müller, H. Haak, and J. Marotzke (2012a), Two Tales of Initializing Decadal Climate Prediction Experiments with the ECHAM5/MPI-OM Model, *Journal of Climate*, *25*(24), 8502–8523.

Matei, D., J. Baehr, J. H. Jungclaus, H. Haak, W. A. Müller, and J. Marotzke (2012b), Multiyear Prediction of Monthly Mean Atlantic Meridional Overturning Circulation at 26.5 °N, *Science*, *335*(6064), 76–79.

Msadek, R., K. W. Dixon, T. L. Delworth, and W. Hurlin (2010), Assessing the predictability of the Atlantic meridional overturning circulation and associated fingerprints, *Geophys. Res. Lett.*, *37*(19).

Murphy, J., V. Kattsov, N. Keenlyside, M. Kimoto, G. Meehl, V. Mehta, H. Pohlmann, A. Scaife, and D. Smith (2010), Towards Prediction of Decadal Climate Variability and Change, *Procedia Environmental Sciences*, *1*(0), 287–304.

Notz, D. (2014), Sea-ice extent and its trend provide limited metrics of model performance, *The Cryosphere*, *8*(1), 229–243.

Opsteegh, J. D., R. Haarsma, F. Selten, and A. Kattenberg (1998), ECBILT: a dynamic alternative to mixed boundary conditions in ocean models, *Tellus A*, *50*(3), 348–367.

Parkinson, C. L., and D. J. Cavalieri (2012), Antarctic sea ice variability and trends, 1979–2010, *The Cryosphere*, *6*(4), 871–880.

Phelps, M. W., A. Kumar, and J. J. O'Brien (2004), Potential Predictability in the NCEP CPC Dynamical Seasonal Forecast System, *Journal of Climate*, *17*(19), 3775–3785.

Pierce, D. W., T. P. Barnett, R. Tokmakian, A. Semtner, M. Maltrud, J. Lysne, and A. Craig (2004), The ACPI Project, Element 1: Initializing a Coupled Climate Model from Observed Conditions, *Climatic Change*, *62*(1), 13–28.

Pohlmann, H., M. Botzet, M. Latif, A. Roesch, M. Wild, and P. Tschuck (2004), Estimating the Decadal Predictability of a Coupled AOGCM, *Journal of Climate*, *17*(22), 4463–4472, doi:10.1175/3209.1.

Pohlmann, H., J. H. Jungclaus, A. Köhl, D. Stammer, and J. Marotzke (2009), Initializing Decadal Climate Predictions with the GECCO Oceanic Synthesis: Effects on the North Atlantic, *Journal of Climate*, *22*(14), 3926–3938.

Pohlmann, H., D. Smith, M. Balmaseda, N. Keenlyside, S. Masina, D. Matei, W. Müller, and P. Rogel (2013), Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system, *Climate Dynamics*, *41*(3-4), 775–785.

Polvani, L. M., and K. L. Smith (2013), Can natural variability explain observed Antarctic sea ice trends? New modeling evidence from CMIP5, *Geophysical Research Letters*, *40*(12), 3195–3199.

Servonnat, J., J. Mignot, E. Guilyardi, D. Swingedouw, R. Séférian, and S. Labetoulle (2014), Reconstructing the subsurface ocean decadal variability using surface nudging in a perfect model framework, *Climate Dynamics*, pp. 1–24, doi:10.1007/s00382-014-2184-7.

Sigmond, M., and J. C. Fyfe (2010), Has the ozone hole contributed to increased Antarctic sea ice extent?, *Geophys. Res. Lett.*, *37*(18), doi:10.1029/2010GL044301.

Simpkins, G. R., L. M. Ciasto, and M. H. England (2013), Observed variations in multidecadal Antarctic sea ice trends during 1979–2012, *Geophysical Research Letters*, pp. n/a–n/a.

Smith, D., R. Eade, and H. Pohlmann (2013), A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction, *Climate Dynamics*, *41*(11-12), 3325–3338.

Smith, D. M., R. Eade, N. J. Dunstone, D. Fereday, J. M. Murphy, H. Pohlmann, and A. A. Scaife (2010), Skilful multi-year predictions of Atlantic hurricane frequency, *Nature Geosci*, *3*(12), 846–849, doi:10.1038/NGEO1004.

Smith, K. L., L. M. Polvani, and D. R. Marsh (2012), Mitigation of 21st century Antarctic sea ice loss by stratospheric ozone recovery, *Geophys. Res. Lett.*, *39*(20), doi:10.1029/2012GL053325.

Solomon, S. (1999), Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, *37*(3), 275–316.

Stammerjohn, S. E., D. G. Martinson, R. C. Smith, X. Yuan, and D. Rind (2008), 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*(C3), doi:10.1029/2007JC004269.

Swingedouw, D., J. Mignot, S. Labetoulle, E. Guilyardi, and G. Madec (2012), Initialisation and predictability of the AMOC over the last 50 years in a climate model, *Climate Dynamics*, *40*(9-10), 2381–2399, doi:10.1007/s00382-012-1516-8.

Talagrand, O. (1997), Assimilation of Observations, an Introduction, *Journal of the Meteorological Society of Japan. Ser. II*, *75*(1B), 191–209.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2011), An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, *93*(4), 485–498.

Tietsche, S., D. Notz, J. Jungclaus, and J. Marotzke (2013), Predictability of large interannual Arctic sea-ice anomalies, *Climate Dynamics*, *41*(9-10), 2511–2526.

Turner, J., T. J. Bracegirdle, T. Phillips, G. J. Marshall, and J. S. Hosking (2013), An Initial Assessment of Antarctic Sea Ice Extent in the CMIP5 Models, *Journal of Climate*, *26*(5), 1473–1484, doi:10.1175/JCLI-D-12-00068.1.

van Leeuwen, P. J. (2009), Particle Filtering in Geophysical Systems, *Monthly Weather Review*, *137*(12), 4089–4114, doi:10.1175/2009MWR2835.1.

van Leeuwen, P. J. (2010), Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Quarterly Journal of the Royal Meteorological Society*, *136*(653), 1991–1999.

Zhang, J. (2007), Increasing Antarctic Sea Ice under Warming Atmospheric and Oceanic Conditions, *Journal of Climate*, *20*(11), 2515–2529.

Zunz, V., H. Goosse, and F. Massonnet (2013), How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent?, *The Cryosphere*, *7*(2), 451–468, doi:10.5194/tc-7-451-2013.

Hindcasts	Initial states	Assimilated pseudo-observations	Assimilation domain
HIND_noinit	Hindcasts initialised without taking into account any pseudo-observation.	-	-
HIND_perfect	The initial states are extracted directly from the pseudo-observations, to which a small perturbation is added in order to generate a 96-member ensemble.	Full model state from the pseudo-observations.	Whole model grid.
HIND_NUD$_{\mathrm{dense}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the nudging.	Dense surface air temperature.	Nudging applied everywhere over the ocean, except on the sea ice covered area.
HIND_NUD$_{\mathrm{sparse}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the nudging.	Sparse surface air temperature.	Nudging applied everywhere over the ocean, except on the sea ice covered area.
HIND_SIR$_{\mathrm{dense}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the particle filter with sequential importance resampling.	Dense surface air temperature.	Particle filter applied over the area southward of 30°S.
HIND_SIR$_{\mathrm{sparse}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the particle filter with sequential importance resampling.	Sparse surface air temperature.	Particle filter applied over the area southward of 60°S.
HIND_NPPF$_{\mathrm{dense}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the nudging proposal particle filter.	Dense surface air temperature.	Nudging applied everywhere over the ocean, except on the sea ice covered area. Particle filter applied over the area southward of 30°S.
HIND_NPPF$_{\mathrm{sparse}}$	The initial states are extracted from the 96 members of a simulation with data assimilation using the nudging proposal particle filter.	Sparse surface air temperature.	Nudging applied everywhere over the ocean, except on the sea ice covered area. Particle filter applied over the area southward of 60°S.

Table 1: Summary of the initialisation methods applied in the hindcasts analysed in this study. All the hindcasts are 96-member ensemble 30-yr long simulations.

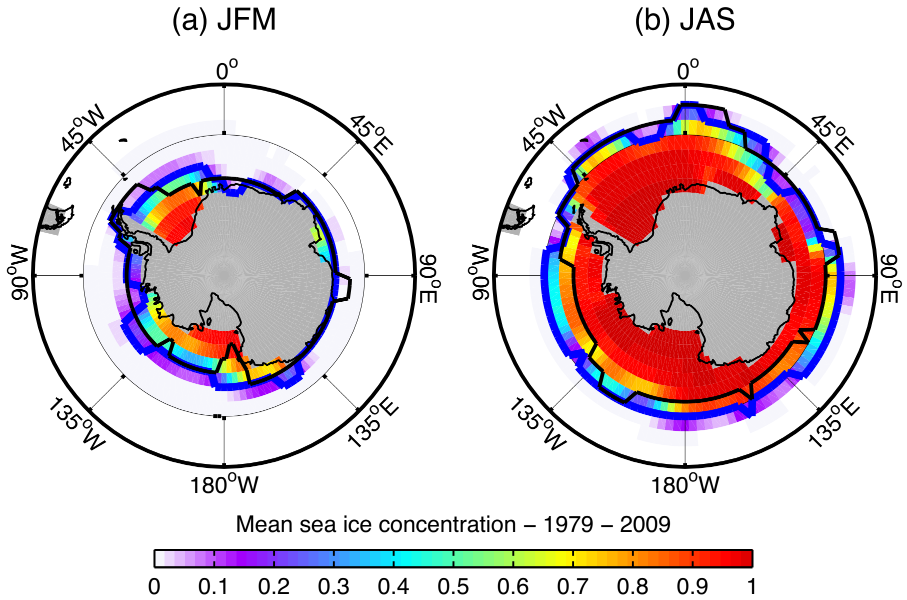

Figure 1: Mean sea ice concentration over the period 1979-2009 computed from a reference simulation performed with the model LOVECLIM1.2 driven by external forcing. Results are shown for (a) summer and (b) winter. The blue (black) line refers to the ice edge, i.e. the 15% concentration limit of the model simulation (observations interpolated on LOVECLIM1.2 ocean model grid, *Comiso*, 1999, updated daily).

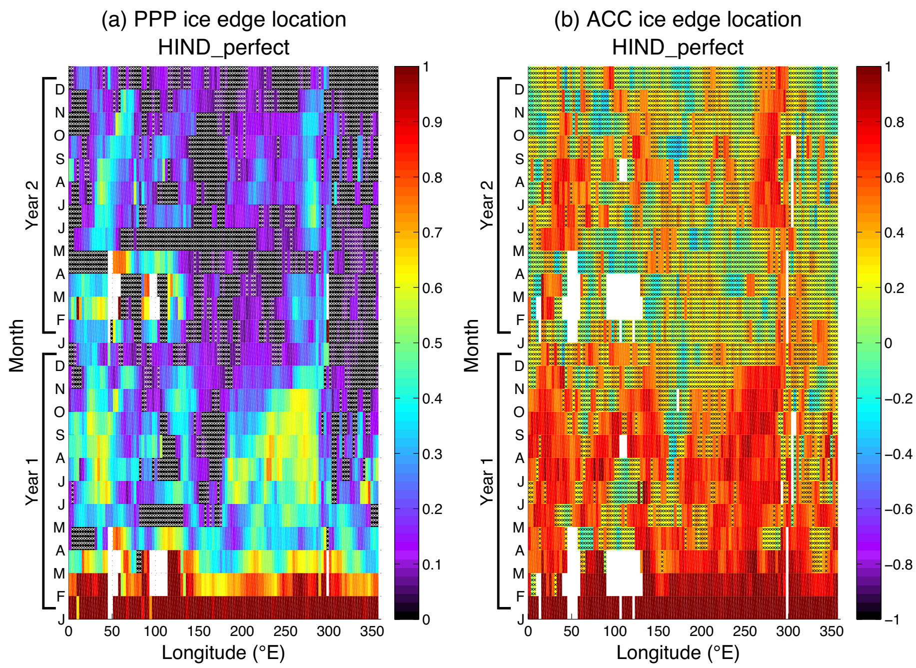

Figure 2: (a) Prognostic potential predictability and (b) anomaly correlation coefficient of the ice edge location computed from the hindcast HIND_perfect. The white (a) or black (b) crosses highlight the values that are not significant at the 95% level. The white areas correspond to undefined values coinciding with longitudes nearly free of sea ice during summer months.

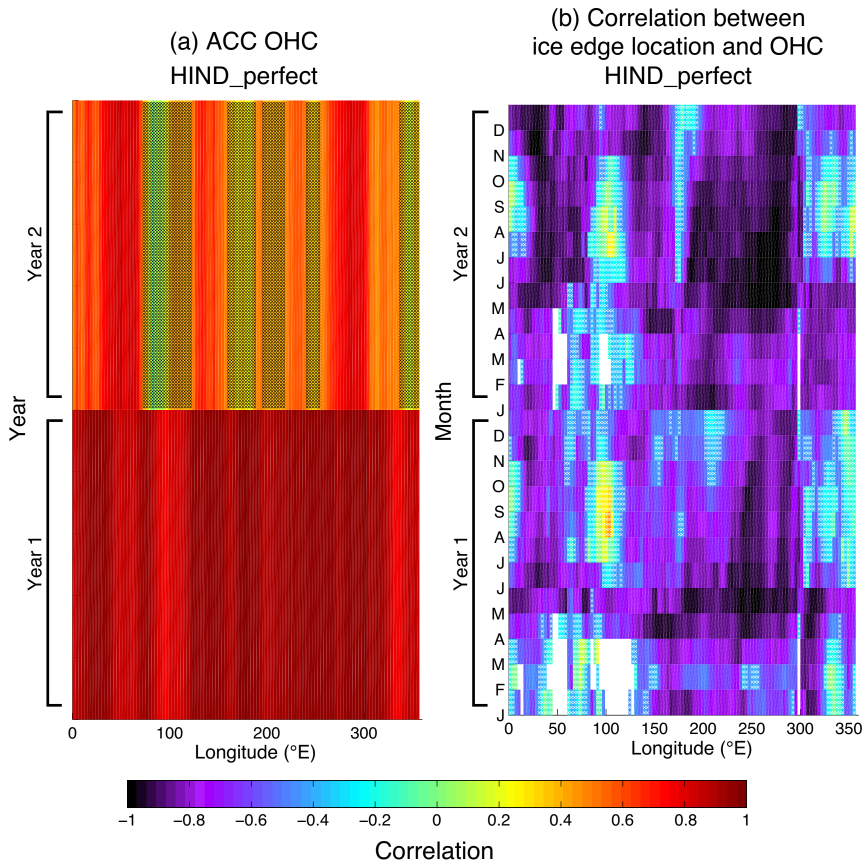

Figure 3: (a) ACC of the annual mean ocean heat content (OHC). (b) Correlation between the annual mean ocean heat content and the ice edge location. The ocean heat content is computed southward of 60°S and between 0 and 100m depth. The black (a) or white (b) crosses highlight the values that are not significant at the 95% level. The white areas correspond to undefined values coinciding with longitudes nearly free of sea ice during summer months.

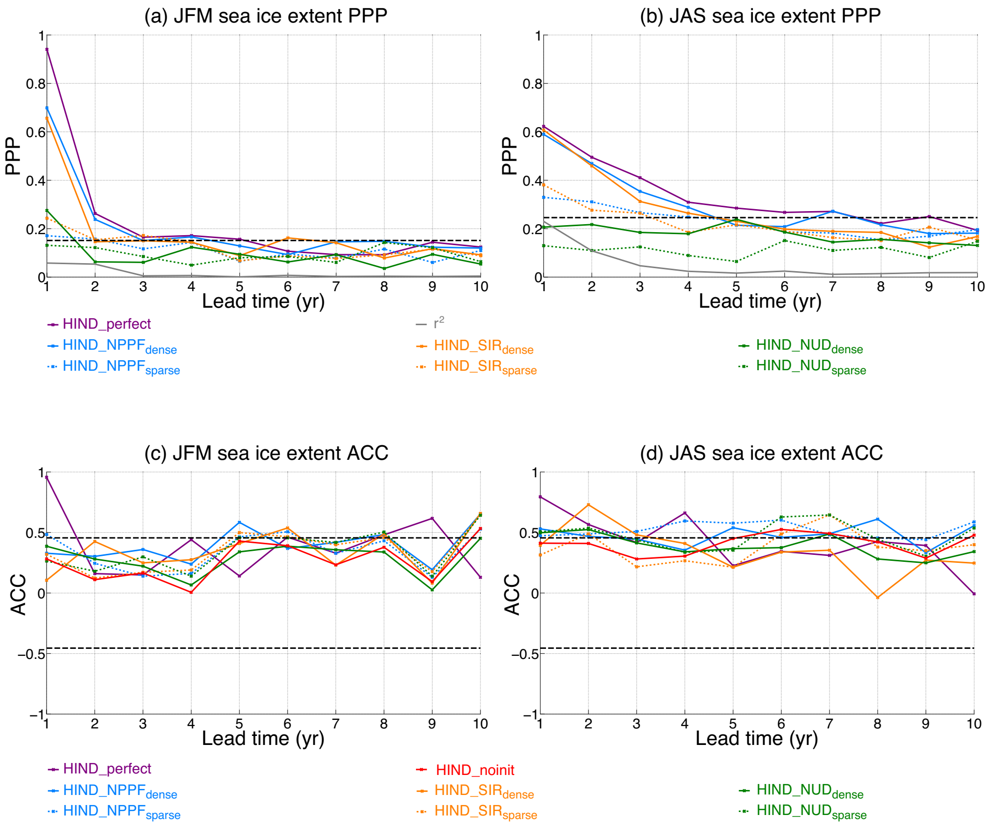

Figure 4: Prognostic potential predictability (a, b) and anomaly correlation coefficient (c, d) for summer (left column) and winter (right column) sea ice extent. The different colors correspond to different initialisation methods. Colored solid lines correspond to an initialisation with dense data, while colored dashed lines correspond to an initialisation with sparse data. The dashed black lines show the 95% significant level. For the PPP, the 95% significant level is higher for winter (b) than for summer (a) sea ice extent. This is due to the slightly larger persistence characterising winter sea ice extent leading to a fewer number of degrees of freedom used to perform the significance test. The grey line in (a) and (b) corresponds to the square of the autocorrelation that indicates the predictability arising from the persistence.

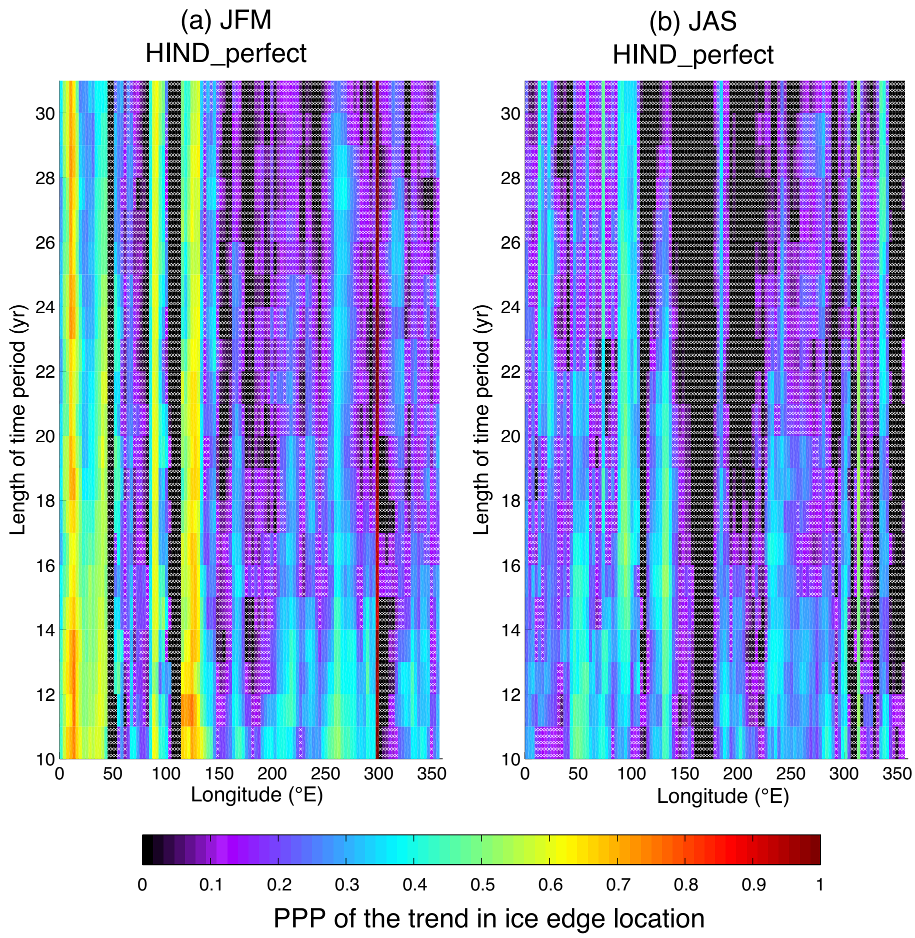

Figure 5: PPP of the trend for the hindcasts HIND_perfect in (a) summer and (b) winter ice edge location for increasing length of the time period over which the trends are computed. The white crosses highlight the values that are not significant at the 95% level.

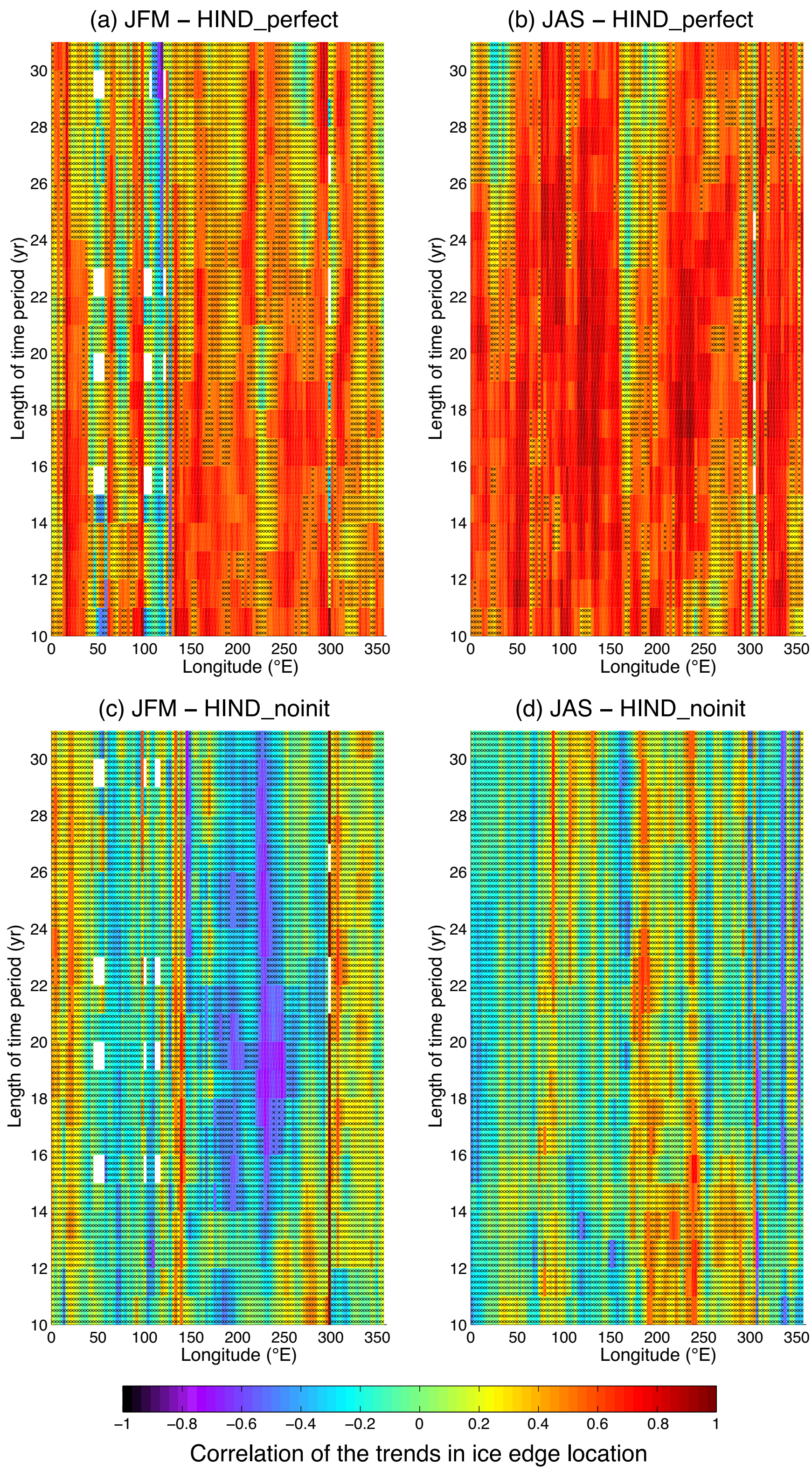

Figure 6: Correlation between the trend of the hindcasts ensemble mean and the trend of the pseudo-observations of the ice edge location in summer (left column) and in winter (right column) for the hindcast HIND_perfect (a, b) and HIND_noinit (c, d). The vertical axis refers to increasing length of the time period over which the trends are computed. The black crosses highlight the values that are not significant at the 95% level. The white areas correspond to undefined values.

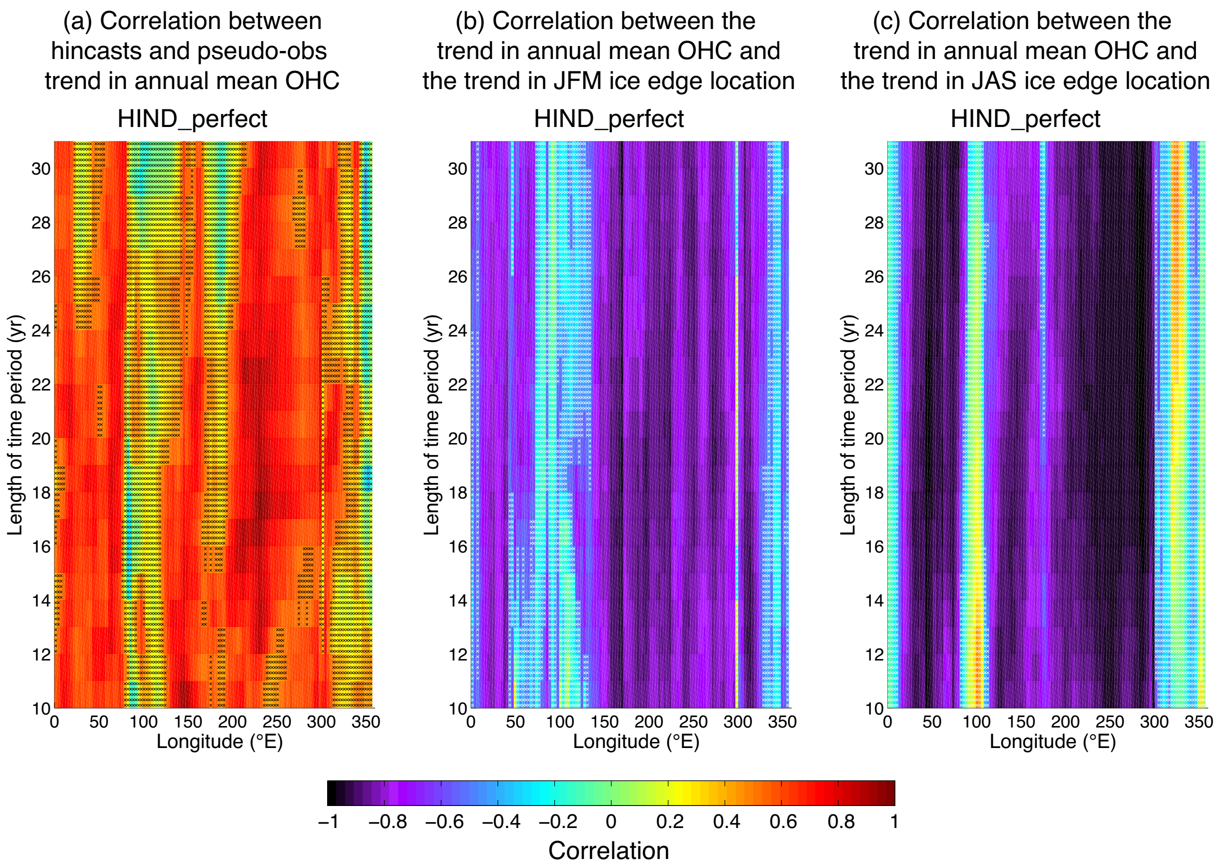

Figure 7: (a) Correlation between the trend in annual mean ocean heat content from the hindcasts and the one from the pseudo-observations. Correlation between the trend in summer (b) and winter (c) ice edge and the trend in annual mean ocean heat content. The vertical axis refers to increasing length of the time period over which the trends are computed. For each longitude, the ocean heat content is computed southward of 60°S and between 0 and -100 m in the ocean. The black (a) and white (b, c) crosses highlight the values that are not significant at the 95% level.

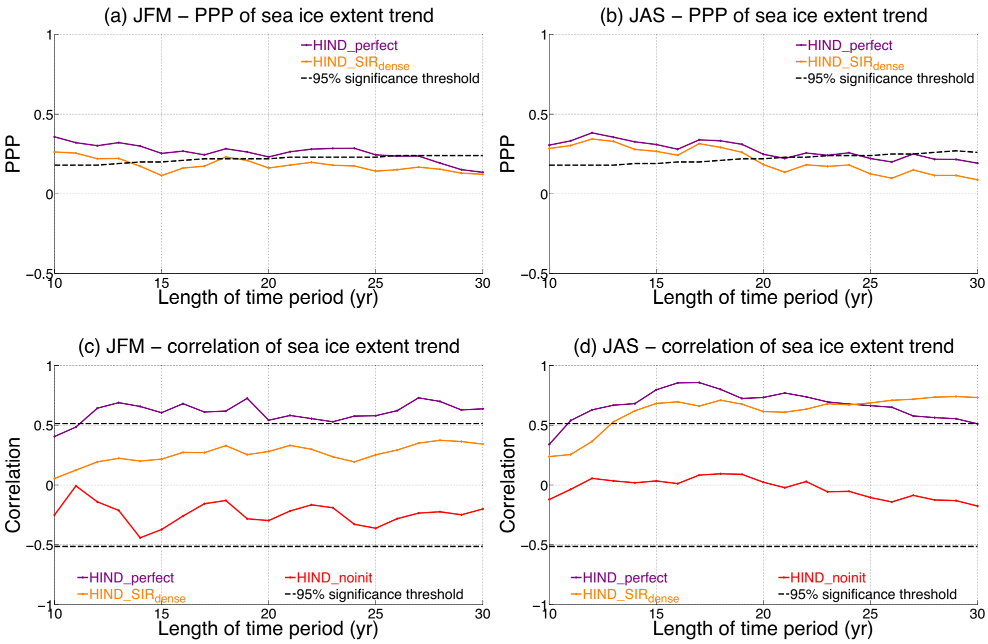

Figure 8: Prognostic potential predictability (a, b) and correlation with the pseudo-observations (c, d) of the trends in summer (left column) and winter (right column) sea ice extent, for increasing length of the time period over which the trends are computed. The different colors correspond to different initialisation methods. The dashed black lines show the 95% significance level. For the PPP (a, b), this significance level varies with the length of time period because it takes into account the autocorrelation of the trends computed over successive time periods used to compute the climatological variance of the trend (σ^2 in Eq. (1)). This autocorrelation depends on the length of time period used to compute the trends.

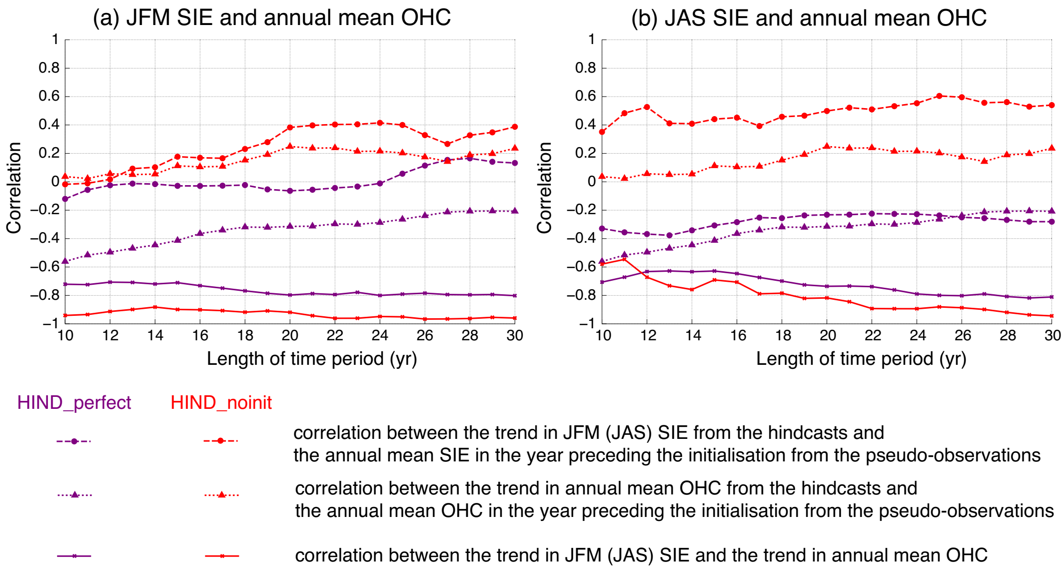

Figure 9: Correlation between the trends in (a) summer and (b) winter sea ice extents from the hindcasts and the annual mean sea ice extent from the pseudo-observations the year preceding the initialisation of the hindcasts (colored dashed lines with circle markers). Correlation between the trend in annual mean ocean heat content from the hindcasts and the annual mean ocean heat content from the pseudo-observations the year preceding the initialisation of the hindcasts (colored dotted lines with triangle markers). Correlation between the trend in sea ice extent from the hindcast and the trend in ocean heat content from the hindcasts (solid lines), for (a) summer and (b) winter sea ice extents. The ocean heat content is computed around Antarctica, southward of 60°S and between 0 and -100 m in the ocean. The x-axis refers the the increasing length of the time period over which the trends are computed.

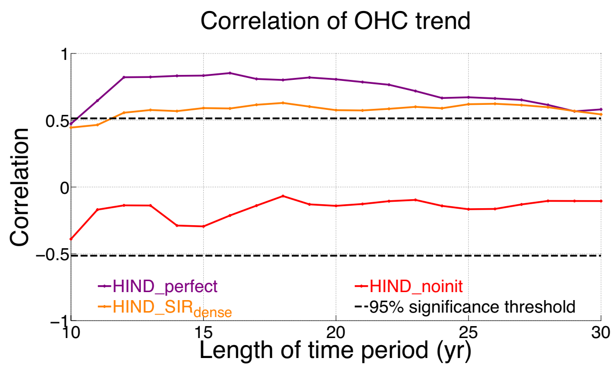

Figure 10: Correlation of the trends in annual mean ocean heat content between the hindcasts and the pseudo-observations, for increasing length of the time period over which the trends are computed. The ocean heat content is computed around Antarctica, southward of 60°S and between 0 and -100 m in the ocean. The dashed black lines show the 95% significance level.