Faculté des Sciences, Ecole de Physique
Earth and Life Institute
Georges Lemaître Centre for Earth and Climate Research

Antarctic sea ice variability and predictability at decadal timescales

Doctoral Dissertation presented by

Violette Zunz

in fulfillment of the requirements
for the degree of Doctor in Sciences

Thesis committee:

Prof. Hugues Goosse (Supervisor)	Université catholique de Louvain
Prof. Thierry Fichefet	Université catholique de Louvain
Prof. Sophie Vanwambeke	Université catholique de Louvain
Dr. Daniela Matei	Max Planck Institute, Hamburg
Dr. Christophe Cassou	Centre national de la recherche scientifique – CERFACS, Toulouse
Prof. Marie-Laurence De Keersmaecker (Chair)	Université catholique de Louvain

Louvain-la-Neuve
August 2014

Les quelques phrases alignées sur cette page marquent la fin d'une belle aventure, celle du doctorat. Ce fut une expérience très enrichissante, tant d'un point de vue scientifique que humain. J'ai bien sûr traversé des moments de doutes et j'ai dû surmonter des difficultés. Heureusement, la réalisation de ma thèse de doctorat ne fut pas une aventure en solitaire, loin de là ! Vous êtes nombreux à m'avoir guidée, encouragée et soutenue au cours de ces quatre années de thèse. C'est avec beaucoup de plaisir que je tiens à remercier chacune et chacun d'entre vous.

Je remercie tout d'abord le Professeur Hugues Goosse, promoteur de ma thèse. Merci, Hugues, pour la confiance que tu m'as témoignée en me proposant de réaliser une thèse de doctorat au sujet passionnant. Merci aussi pour ta disponibilité infaillible, ton optimisme et ton sens de l'humour. Merci pour les nombreuses discussions que nous avons eues, pour tes conseils très justes et tes suggestions toujours pertinentes.

Je remercie également les autre membres de mon jury pour l'intérêt qu'ils ont porté à mon sujet de thèse, pour leurs relectures attentives et pour leurs commentaires constructifs et encourageants. Thierry, merci pour la justesse et la précision de tes remarques. Sophie, merci pour ta curiosité et pour le regard pertinent que tu as porté sur mon travail de recherche. Daniela, thank you very much for your relevant comments during the review process of my thesis and for the stimulating discussions we had during my short stay at the Max Planck Institute in 2011. Christophe, merci pour ton enthousiasme et pour les pistes de réflexion judicieuses que tu m'as suggérées.

Je tiens à remercier l'Université catholique de Louvain pour m'avoir accordé un mandat UCL-FSR durant les premiers mois de mon doctorat. Je remercie le Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA) pour m'avoir octroyé une bourse de recherche durant mes quatre années de doctorat. Je suis également reconnaissante envers la Politique scientifique fédérale (BELSPO) et le centre de calcul intensif et de stockage de masse (CISM) pour avoir mis à ma disposition les moyens logistiques nécessaires au bon déroulement de cette thèse.

Je voudrais maintenant remercier mes collègues, dont nombreux d'entre eux sont devenus au fil des années des amis. Merci à toute l'équipe pour ces bons moments partagés autour d'un verre, d'un barbecue ou à l'occasion d'un match de badminton. Merci aussi pour tous les coups de mains reçus par-ci, par-là, qui m'ont souvent facilité la tâche. Je remercie en particulier Elisabeth, compagne idéale de bureau, de fin de thèse et de visite des stands commerciaux de l'EGU... Quatre années de « co-habitation », ça crée des liens ! Merci pour ton écoute toujours attentive, ta disponibilité et ta serviabilité. Merci aussi

spécialement à Oli, fondateur du centre de loisirs Sérigny-plage et coach de la cuvée de docteurs 2014, pour ta gentillesse sans pareille et tes encouragements continus. Longue vie aux week-ends festifs à Sérigny ! Lise, merci pour ton amitié et pour ton aide chaque fois que j'en avais besoin. Merci également à Dominique et à Monique pour leur travail discret et efficace réalisé dans la bonne humeur.

Je profite de l'occasion pour remercier Monsieur Canard et Monsieur Dartevelle, enseignants à l'Institut Notre-Dame de Thuin, pour avoir éveillé en moi cette passion pour les sciences physiques et mathématiques.

Je remercie également chaleureusement ma famille. Merci à mes parents. Partis bien trop tôt, ils m'ont imprégnée de leurs valeurs qui guident mes choix. Je remercie en particulier ma mère pour son courage et pour m'avoir toujours offert ce qu'il y a de meilleur. Tata, Tonton, merci d'avoir assuré la relève, de m'entourer de votre tendresse et pour votre soutien inconditionnel. Manolo, mon grand frère –Kurt– adoré, merci d'être si attentionné et toujours disponible. Merci également à Catherine, Marie, Anne et toute la clique. Je remercie aussi ma belle-famille. Bernadette, Thierry, merci pour votre gentillesse, vos encouragements et votre confiance.

Je remercie bien sûr tous mes amis pour leurs précieux encouragements et leur compréhension face à mon manque de disponibilité ces derniers mois. Merci pour tous ces bons moments partagés qui sont souvent l'occasion de discussions passionnantes.

Enfin, tout ceci n'aurait sans doute pas été possible sans toi, Bertrand. Ta patience, ta compassion et ton soutien indéfectibles m'ont permis de ne pas perdre pied. Tu es toujours le premier à m'encourager à relever de nouveaux défis et à rétablir ma confiance en moi quand je la perds. Tes délicates attentions à mon égard, en particulier durant ces derniers mois, m'ont permis de vivre cette fin de doctorat beaucoup plus sereinement. Bertrand, merci pour tout, merci d'être toujours là pour moi.

Contents

Contents **iii**

1 Introduction **1**
1.1 Decadal climate predictions 2
1.2 Antarctic sea ice 15
1.3 Contribution of the thesis 23

2 Variability of the Antarctic sea ice in GCMs **33**
2.1 Introduction 34
2.2 Models and observational data 37
2.3 Historical simulations 44
2.4 Hindcast simulations 52
2.5 Summary and conclusion 59

3 Impact of initialisation on the sea ice predictability **61**
3.1 Introduction 62
3.2 Methodology 64
3.3 Results 72
3.4 Summary and conclusions 90

4 Influence of a meltwater input on the sea ice forecast skill **93**
4.1 Introduction 94
4.2 Methodology 97
4.3 Results 103
4.4 Summary and conclusions 120

5 Conclusions **123**
5.1 Main achievements 123
5.2 Predicting sea ice in the Southern Ocean 125
5.3 Perspectives 127

Appendices **131**

A Acronyms **131**

B Supplementary material for Chapter 2 **133**

References **141**

CHAPTER 1

Introduction

The sea ice cover present in the polar regions significantly affects the exchanges between the atmosphere and the ocean, reflects a large part of the solar radiation back to space and is involved in several feedback processes that strongly link it to the other components of the climate system. Consequently, the changes in the sea ice cover around Antarctica are generally associated to changes in the ocean and in the atmosphere of adjacent areas. Furthermore, the Southern Ocean has also an influence on more remote areas, in particular through the oceanic overturning circulation, illustrating the important role of the sea ice surrounding Antarctica in the climate system.

While the Earth's climate is currently undergoing a global warming, the Antarctic sea ice has been slightly expanding since the beginning of satellite observations in the late 1970's, following a trend opposite to the one displayed by its Arctic counterpart (e.g., *Turner and Overland*, 2009). Several hypotheses have been proposed to explain how this increase in sea ice extent could have been occurring in a global warming context but no consensus have been reached yet. Given the lack of understanding of the processes that drive the evolution of the Antarctic sea ice, simulating and predicting its evolution remains a major challenge. In this framework, the research carried out during this thesis tackled two main questions: (1) is the sea ice in the Southern Ocean predictable at the decadal timescale and (2) how can these predictions be improved?

1.1 Decadal climate predictions

Decadal climate predictions refer to predictions that are performed a few years to several decades in advance. They are thus particularly helpful for decision makers, in many sectors of activities such as agriculture, food production, energy management, etc. (e.g., *Vera et al.*, 2010). Decadal predictions bridge the gap between seasonal-to-interannual predictions and longer-term ones that aim at forecasting the evolution of the climate for decades to centuries ahead (Fig. 1.1).

Performing a climate prediction requires a climate model that is integrated forward in time from slightly different initial conditions. This leads to what is called an ensemble simulation, which provides a set of estimates of the future evolution of the climate. Performing such an ensemble is necessary in decadal prediction studies because it allows assessing the sensitivity of a simulation to initial conditions, which is a major issue at decadal timescale, as discussed in Sect. 1.1.1. Different methods are used to generate an ensemble, for instance by adding a random perturbation to the initial conditions or by using initial states corresponding to slightly different dates (e.g., *Pohlmann et al.*, 2004; *Doblas-Reyes et al.*, 2011; *Du et al.*, 2012; *Yeager et al.*, 2012). Additionally, some techniques allow determining the optimal spatial structure of the perturbations, supposed to lead to a rapid spread of the ensemble (e.g., *Toth and Kalnay*, 1997; *Zanna and Tziperman*, 2008; *Ham et al.*, 2014).

Decadal predictions are generally performed with general circulation models (GCMs). These models simulate the evolution of the climate over the whole Earth, with the highest resolution and complexity available up to now. The fifth phase of the Coupled Model Intercomparison Project (CMIP5, *Taylor et al.*, 2011) has proposed a coordinated framework to perform decadal prediction simulations with GCMs from different research centres. However, a simulation performed with a GCM requires large computational resources, limiting the number of simulations that can be carried out. If systematic tests with different experimental designs are needed, the use of an Earth system model of intermediate complexity (EMIC) may be preferred. An EMIC has a lower level of complexity and a coarser resolution than a GCM. This offers the noticeable advantage of reducing the computational cost of a simulation with the drawback of misrepresenting or not representing at all some important mechanisms taking place in the climate system, as briefly discussed in Sect. 1.3.2 for the LOVECLIM model that will be used in Chapters 3 and 4. The ability of an EMIC to simulate specific processes, in particular those occurring at small spatial scales and short timescales, can thus be lower than the one of a GCM.

1.1.1 Uncertainties in decadal climate predictions

Lorenz (1975) has distinguished two kinds of climate predictions. The first one consists in an initial value problem. The uncertainty in predictions of the first kind is due to the chaotic nature of the climate, implying that small differences in the initial conditions can lead to different evolutions of the system (*Lorenz*, 1963). In particular, a set of initial conditions based on observations is subject to uncertainties due to measurement errors. Even if these uncertainties are initially small, they grow during the model integration. After a certain time of simulation, which depends on the variable of interest, any solution of an ensemble initialised with observations cannot be distinguished from the ones of a randomly initialised ensemble simulation. At this stage, the information provided by the observations at the initialisation is lost and the limit of predictability is reached. Daily weather forecasts or seasonal-to-interannual climate predictions generally consist in predictions of the first kind (Fig. 1.1).

Lorenz (1975) has identified the predictions of the *second kind* as a boundary condition problem. These predictions depend on the evolution of the external forcings, arising from natural and anthropogenic sources. The uncertainty in predictions of the second kind is thus inherent to the scenario of future emissions of greenhouse gases and aerosols that impact the radiative forcing. This scenario depends on human behaviour and on the decisions of policymakers. Additionally, natural external forcings such as the volcanic activity or the variations in solar irradiance will also impact the future climate but these phenomenons are difficult to anticipate. Long-term climate change predictions, spanning decades to centuries, are referred to as projections (*Kirtman et al.*, 2013) and consist in predictions of the second kind (Fig. 1.1).

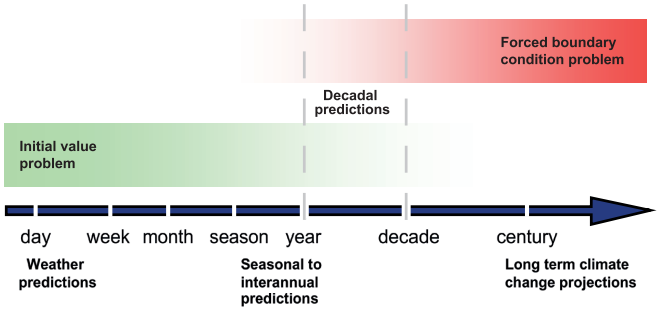

Figure 1.1: A schematic illustrating the progression from an initial-value based prediction at short timescales to the forced boundary-value problem of climate projection at long timescales. Decadal prediction occupies the middle ground between the two. Based on *Meehl et al.* (2009). Reproduction of Box 11.1, Figure 2 of *Kirtman et al.* (2013).

Additionally, both kinds of predictions are subject to uncertainties arising from the model errors. Indeed, a climate model includes approximations and is thus unable to simulate perfectly all the processes occurring in the real world. The mean state and the internal variability (i.e. the natural fluctuations produced without the intervention of an external source) of the different climate variables are thus not systematically well represented by a model. Consequently, simulations initialised from the same initial conditions but performed with different climate models may provide different solutions for the evolution of the system. Similarly, different models may simulate different responses to the same radiative forcing.

Regarding their timescales, decadal predictions lie between seasonal-to-interannual predictions and long-term projections (e.g., *Collins and Allen*, 2002; *Meehl et al.*, 2009; *Hurrell et al.*, 2010; *Latif et al.*, 2010). Decadal predictions are thus both an initial-value problem and a boundary-condition one (Fig. 1.1). While it appears difficult to reduce the uncertainty arising from the emissions scenarios, performing skillful decadal predictions relies on the ability of the model to reproduce the observed mean state and internal variability as well as on a synchronisation between the variability of the prediction and the one of the observations. Consequently, in a decadal prediction framework, it is essential to constrain the initial state of the forecast simulation to be as close as possible to the observed one.

1.1.2 Initialisation through data assimilation

Observations of climate variables are often scattered in space and time, and thus cannot be used directly to initialise a simulation. In addition, the observed state is not necessarily compatible with the model physics. Therefore, if raw observations are used to initialise a prediction, the latter may undergo a transient phase, generally characterised by a spurious signal, whose duration may be longer than the forecast period. These two issues can be partially solved through the use of data assimilation methods which combine the model equations with the observations to provide an estimate of the state of the system as accurate as possible, and compatible with both the observations and the model (*Talagrand*, 1997). In other words, given an incomplete set of observations and a climate model, a data assimilation method computes values for each model variable, at each model grid point, based on the information provided by both the solutions of the model equations and the available observations (Fig. 1.2).

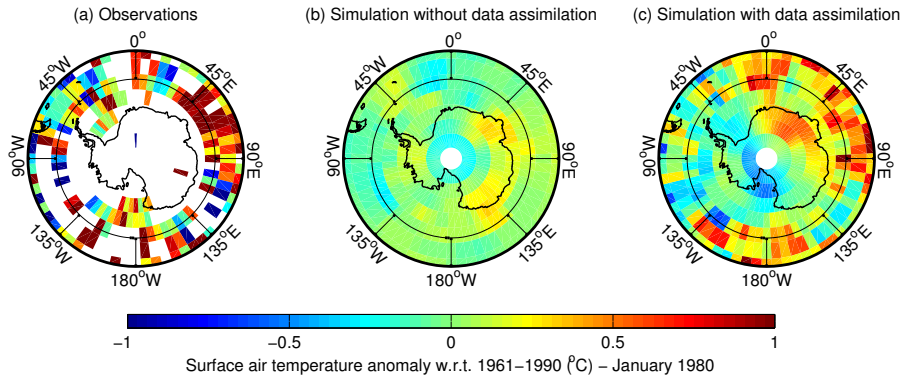

Figure 1.2: Example of the impact of assimilating anomalies of surface air temperature with the climate model LOVECLIM (*Goosse et al.*, 2010a). Results show the monthly mean of January 1980 surface air temperature anomalies. The anomalies are computed with regard to the period 1961-1990, for (a) the observations (HadCRUT3 dataset, *Brohan et al.*, 2006), (b) the mean of an ensemble simulation driven by the external forcings only and (c) the mean of an ensemble simulation driven by the external forcings that assimilates anomalies of surface air temperature through a nudging proposal particle filter (see the text for details).

The data assimilation methods all share a common objective: for a set of n model variables contained in the vector $\mathbf{x} \in \mathbb{R}^{n\times 1}$, given a set of p observations contained in the vector $\mathbf{y}_o \in \mathbb{R}^{p\times 1}$ and a background knowledge $\mathbf{x}_b \in \mathbb{R}^{n\times 1}$ of the vector $\mathbf{x}$ (provided, for instance, by a model integration), the data assimilation aims at finding an optimal analysis $\mathbf{x}_a \in \mathbb{R}^{n\times 1}$ for the set of variables in $\mathbf{x}$. From an algorithmic point of view, in a decadal prediction context, the analysis is often computed through a sequential approach. It consists in integrating a model forward in time and, at some specific times, updating the solution of the model using available observations before continuing the integration (Fig. 1.3). Data assimilation allows reconstructing the evolution of the system during a time interval. The state of the system at a given date can then be extracted from this reconstruction to initialise a prediction simulation. Many data assimilation methods, based on different implementations, can be used to estimate the state of the system. The most standard ones are summarised below.

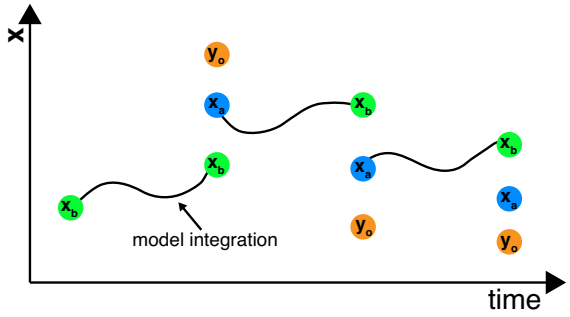

Figure 1.3: Illustration of the basic principles of the data assimilation. The background vectors $\mathbf{x}_b$ are shown as the green bullets, the analysis vectors $\mathbf{x}_a$ are shown as the blue bullets and the observation vectors $\mathbf{y}_o$, displayed in the model space for simplicity, are shown as the orange bullets.

The nudging

The nudging, or Newtonian relaxation, is a very simple and widely used data assimilation method that consists in adding to the model equations a term that pulls the solution towards the observations (*Kalnay*, 2007). The nudging is generally implemented as follows:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{G}\mathbf{H}^T \left[\mathbf{y}_o - \mathbf{H}\mathbf{x}_b\right]. \tag{1.1}$$

$\mathbf{H} \in \mathbb{R}^{p \times n}$ is a matrix that projects the variables in the model space onto variables in the observation space. To be more general, the matrix $\mathbf{H}$ should be replaced by an operator but, here, the assumption of a linear observation operator is made. $\mathbf{G} \in \mathbb{R}^{n \times n}$ is a diagonal matrix whose elements determines the relaxation timescale of the nudging term applied on the model variables corresponding to the observed ones. These relaxation timescales are chosen empirically. One disadvantage of this method is that it introduces ad-hoc terms in the equations without any physical justification. Therefore, the amplitude of the elements in $\mathbf{G}$ has to be chosen carefully. If this amplitude is too large, the solution converges quickly towards the observations, preventing an adjustment through the model dynamics. When used to initialise a prediction simulation, this technique may trigger an initialisation shock that results in abrupt changes at the beginning of the simulation. On the contrary, if the amplitude of the elements in $\mathbf{G}$ is too small, the impact on the model solution is weak and the nudging may be useless. This data assimilation method is rather simple to implement in climate models and it does not increase the computational needs. It has often been used to assimilate observations in general circulation models in order to estimate an initial state for decadal prediction simulations

(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).

The optimal interpolation and the Kalman filtering

While the nudging is based on empirical choices of the magnitudes of the relaxation coefficients, other data assimilation methods, such as the optimal interpolation or the Kalman filtering, rely on the solution of an optimisation problem. Additionally, these methods allow a multivariate data assimilation, i.e. they provide an update for the whole solution of the model and not only for the variables corresponding to the observed ones, preserving the co-variability between the model variables. The optimal interpolation is based on the following equation:

$$\mathbf{x}_a = \mathbf{x}_b + \mathbf{K}\left[\mathbf{y}_o - \mathbf{H}\mathbf{x}_b\right]. \tag{1.2}$$

$\mathbf{K} \in \mathbb{R}^{n\times p}$ is the gain matrix, also called weight matrix. The gain matrix $\mathbf{K}$ is chosen to minimise the mean squared analysis error $\mathbf{A} = E(\epsilon_a \epsilon_a^T)$, where $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\epsilon_a = \mathbf{x}_a - \mathbf{x}_t$ with $\mathbf{x}_t \in \mathbb{R}^{n\times 1}$ the (unknown) truth in the model space. $E()$ refers to the expected value, i.e. the average over an infinite number of cases. Under the assumptions that the background $\mathbf{x}_b$ and the observations $\mathbf{y}_o$ are unbiased and that their errors are uncorrelated, the optimal matrix $\mathbf{K}$ is

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T(\mathbf{R}+\mathbf{H}\mathbf{B}\mathbf{H}^T)^{-1}. \tag{1.3}$$

$\mathbf{B} = E(\epsilon_b \epsilon_b^T)$ is the error covariance matrix ($\mathbf{B} \in \mathbb{R}^{n\times n}$) for the background ($\epsilon_b = \mathbf{x}_b - \mathbf{x}_t$), $\mathbf{R} = E(\epsilon_o \epsilon_o^T)$ is the error covariance matrix ($\mathbf{R} \in \mathbb{R}^{p\times p}$) for the observations ($\epsilon_o = \mathbf{y}_o - \mathbf{y}_t$ where $\mathbf{y}_t \in \mathbb{R}^{p\times 1}$ is the truth in the observation space), see for instance *Kalnay* (2007) for details. The optimal interpolation does not only provide an optimal estimate of $\mathbf{x}_a$ but it also produces an error covariance matrix for the analysis (e.g., *Talagrand*, 1997; *Kalnay*, 2007).

In the optimal interpolation approach, the background error covariance matrix $\mathbf{B}$ is assumed to be constant in time. The Kalman filter relies on the same principles as the optimal interpolation, but with one major difference: the background error covariance matrix, called here the forecast error matrix $\mathbf{P}^f(t)$, is not stationary anymore. The propagation of the error covariance matrix is achieved thanks to the linear tangent model $\mathbf{L}$, i.e. the Jacobian matrix of the model operator M (e.g., *Kalnay*, 2007). The operator M_{i-1} allows propagating the analysis solution $\mathbf{x}_a$ at time t_{i-1} to the forecast solution $\mathbf{x}_f$ at time t_i, $\mathbf{x}_f(t_i) = M_{i-1}[\mathbf{x}_a(t_{i-1})]$. This apparently small change between the optimal interpolation and the Kalman filter actually implies the multiplication of very large matrices and is thus very expensive from a computational point of view.

The ensemble version of the Kalman filter (e.g., *Evensen*, 2003) allows computing a time evolving forecast error matrix without the use of the linear tangent model. The matrix $\mathbf{P}^f(t)$ is here estimated from the solutions provided

by the N members of an ensemble simulation at time t, as for instance

$$\mathbf{P}^f(t) \approx \frac{1}{N-1}\sum_{j=1}^{N}(\mathbf{x}_f^j(t) - \mathbf{X}_f(t))(\mathbf{x}_f^j - \mathbf{X}_f(t))^T, \tag{1.4}$$

where the j index refers to the member j and $\mathbf{X}_f$ is the ensemble mean. All the members of the ensemble assimilate the same observations. Attention has thus to be paid in order to maintain a sufficiently large variance in the ensemble, through for instance the addition of a random perturbation to the observations assimilated in each member.

Data assimilation methods based on the optimal interpolation or on the Kalman filter have been used in a few decadal prediction simulations performed in the framework of CMIP5 (e.g., GFDL-CM2.1 model, *Zhang et al.*, 2007). Besides, *Massonnet et al.* (2013) have used an ensemble Kalman filter to reconstruct the thickness of the sea ice in the Southern Ocean through the assimilation of sea ice concentration data in an ocean-sea ice model. Additionally, *Massonnet et al.* (2014) have assessed the impact of the initialisation of the hindcasts through the assimilation of sea ice concentration data with the ensemble Kalman filter in an ocean-sea ice model. According to their results, this initialisation method could improve the forecast skill of the sea ice in the Arctic but the benefits for the Antarctic sea ice are negligible.

Variational data assimilation

The optimal analysis field $\mathbf{x}_a$ can also be found by minimising a cost function (e.g., *Talagrand*, 1997; *Kalnay*, 2007). This is the basic principle of the variational data assimilation. The variational data assimilation can be time independent, referred to as the 3D-Var. In this case, the cost function J is described as

$$J(\mathbf{x}) = \frac{1}{2}\left[[\mathbf{x}-\mathbf{x}_b]^T\mathbf{B}^{-1}[\mathbf{x}-\mathbf{x}_b] + [\mathbf{y}_o - \mathbf{H}\mathbf{x}]^T\mathbf{R}^{-1}[\mathbf{y}_o - \mathbf{H}\mathbf{x}]\right]. \tag{1.5}$$

The vector $\mathbf{x}$ that minimises the cost function J provides the analysis vector $\mathbf{x}_a$. While in the optimal interpolation, the control variable for the minimisation is the gain matrix $\mathbf{K}$, in the variational approach, this control variable is the vector $\mathbf{x}_a$. However, under adequate assumptions, the 3D-Var and the optimal interpolation are equivalent and provide an identical optimal analysis field $\mathbf{x}_a$ (see, e.g., *Kalnay*, 2007, for details).

The time dependent variational data assimilation, referred to as the 4D-Var, aims at minimising the following cost function:

$$\begin{aligned} J(\mathbf{x}(t_0)) \quad = \quad & \frac{1}{2}\left[\mathbf{x}(t_0) - \mathbf{x}_b(t_0)\right]^T \mathbf{B}_0^{-1}\left[\mathbf{x}(t_0) - \mathbf{x}_b(t_0)\right] \\ & + \frac{1}{2}\sum_{m=0}^{K}\left[\mathbf{H}\mathbf{x}\left(t_m\right) - \mathbf{y}_o(t_m)\right]^T \mathbf{R}(t_m)^{-1}\left[\mathbf{H}\mathbf{x}(t_m) - \mathbf{y}_o(t_m)\right]. \end{aligned}$$

4D-Var is thus used to adjust globally a model solution to the observations distributed within a time interval. In this case, the model is used as a strong constraint given that the analysis solution has to satisfy the model equations. It is equivalent to identifying an initial state that would provide a solution, through the model integration, that best fits the observations. 4D-Var differs from a sequential approach since it propagates the information from the observations both forward and backward in time. This method thus requires the use of the linear tangent model $\mathbf{L}$ and its transpose, the adjoint model $\mathbf{L}^T$ (see, e.g., *Kalnay*, 2007).

The variational data assimilation is often used for meteorological applications. 3D-Var and 4D-Var have been implemented in several weather prediction systems as, for instance, the one developed at the European Centre for Medium-Range Weather Forecasts (ECMWF, *Bouttier and Rabier*, 1997) or the one of Météo-France (*Thépaut et al.*, 1996). A 3D-Var data assimilation method has also been used by the National Centers for Environmental Prediction (NCEP) to perform a reanalysis of the atmosphere, the NCEP ocean reanalysis having been developed thanks to an optimal interpolation (*Saha et al.*, 2010).

Particle filtering

As the ensemble Kalman filter, the particle filtering (PF) is an ensemble data assimilation method (e.g., *van Leeuwen*, 2009). Nevertheless, unlike the ensemble Kalman filter, the principle of the particle filter does not imply a modification of the solution provided by each member of the ensemble. The particle filtering approach rather consists in a resampling of the probability distribution of the model to improve the agreement between the latter and the observations. The idea is to approximate the evolution of the probability distribution of the model states by propagating forward in time an ensemble of simulations, called particles. The particles differ from each other because they are initialised from slightly different states. Then, at a time step when observations are available, the ensemble simulation is stopped and a weight is attributed to each particle, based on its agreement with the observations. The ensemble is resampled according to the weights of the particles and is then again propagated with the model. The entire procedure is repeated until the end of the period of interest.

In a degenerated form of the particle filtering, as the one used for instance by *Goosse et al.* (2010b), the weight corresponds to the Euclidian distance between the model solution and the observations. The particle with the largest weight is chosen and duplicated as many times as the total number of particles within the ensemble, while the other particles are eliminated. The number of particles is thus maintained constant through the simulation. A small perturbation is then added to the copied particles in order to generate different solutions for the following period. The particles are then again propagated forward in time by the model and the different steps are repeated. This simplified particle filter

has demonstrated good skill to reconstruct the changes in the atmosphere, the ocean and the sea ice at high southern latitudes in the past half century (*Goosse et al.*, 2009a), the warming in the Arctic during the 15th century (*Crespin et al.*, 2009) as well as the evolution of the climate over the last 600 years (*Goosse et al.*, 2010b).

Besides, more advanced versions of the particle filtering exist, as for instance the particle filter with sequential importance resampling (e.g., *van Leeuwen*, 2009). *Dubinkina et al.* (2011) have implemented this data assimilation method in the LOVECLIM climate model (*Goosse et al.*, 2010a). In this approach, the weight attributed to each particle is computed assuming that the conditional probability of the observations given a model solution $p(\mathbf{y}_o|\mathbf{x}^j)$, called the likelihood of the particle j, is Gaussian (*van Leeuwen*, 2009), i.e.

$$p(\mathbf{y}_o|\mathbf{x}^j) = A \exp\left\{-\frac{1}{2}\left[\mathbf{y}_0 - \mathbf{H}\mathbf{x}^j\right]^T \mathbf{R}^{-1}\left[\mathbf{y}_0 - \mathbf{H}\mathbf{x}^j\right]\right\}, \tag{1.6}$$

where A is a normalisation factor. The weight w_j of the particle j is then defined as

$$w_j = \frac{p(\mathbf{y}_o|\mathbf{x}^j)}{\sum_{k=1}^{M} p(\mathbf{y}_o|\mathbf{x}^k)}. \tag{1.7}$$

The particles with a large weight are duplicated in proportion to their weight, while the ones with a small weight are eliminated (Fig. 1.4). As in the degenerated form of the particle filter, the total number of particles is maintained constant and a small perturbation is applied on the duplicated particles before they are propagated with the model. The particle filter with sequential importance resampling, in combination with the climate model LOVECLIM, has been applied in various climate studies, such as the investigation of the causes of the cooling that occurred in the Southern Hemisphere from 10 to 8 ka BP (*Mathiot et al.*, 2013), the reconstruction of the evolution of the climate during the mid-Holocene (*Mairesse et al.*, 2013) and the prediction of the Southern Ocean sea ice (*Zunz et al.*, 2014).

Dubinkina and Goosse (2013) have implemented in LOVECLIM a combination of a particle filter with sequential importance resampling and a nudging, referred to as the nudging proposal particle filter (NPPF). Through this approach, each particle is nudged towards the observations during the model integration. The amplitude of the nudging that has been applied on each particle is taken into account in the computation of the weight. This method has been used in *Goosse and Zunz* (2014) to study the decadal trends in Antarctic sea ice as well as in *Zunz et al.* (2014) and *Zunz and Goosse* (2014) to analyse the predictability of the sea ice in the Southern Ocean.

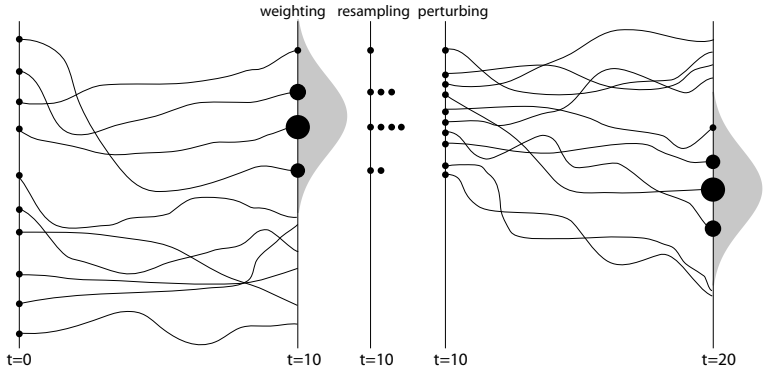

Figure 1.4: The particle filter with sequential importance resampling. The curved lines show the trajectories followed by the particles and the size of the bullets correspond to the weight of each particle. Figure adapted from *van Leeuwen* (2009).

Assimilated field

The choice of the fields to be assimilated is obviously based on the availability and on the reliability of the observations, but the geographic areas and the timescales of interest have to be taken into account as well. In particular, in a decadal prediction framework, the initialisation of slowly varying components, as for instance the ocean, can increase the predictive skill of a climate model (e.g., *Dunstone and Smith*, 2010). This initialisation of the ocean state can be reached directly through the assimilation of the corresponding variables, but subsurface data are sparse in space and time. However, the initialisation can be indirect, through for instance the assimilation of surface variables, assuming a propagation of the information from the surface to the interior ocean.

Systematic biases present in climate models generally result in a model climatology that does not fit the observed one. These biases thus raise the issue whether anomaly or full-field variables should be assimilated (e.g., *Murphy et al.*, 2010; *Smith et al.*, 2013a; *Magnusson et al.*, 2013a,b). Full-field data assimilation can reduce the model biases by constraining the model solution to get closer to the observations. Nevertheless, when a simulation is initialised through full-field data assimilation, if the solution starts with a mean state that differs from the model climatology, it will likely drift towards this latter once the constraints from observations are removed. Anomalies are constructed by subtracting from the observations their climatological mean. The model climatology is then added to these observation anomalies before they are assimilated by the model. Assimilating anomalies thus allows preserving the model mean state and can prevent a model drift when used to initialise predictions. Nevertheless, there can remain inconsistencies between the structure of the anomalies

and the model dynamics that could still give rise to an initialisation shock (e.g., *Magnusson et al.*, 2013b). The relative merits of full-field or anomaly data assimilation have been discussed in *Magnusson et al.* (2013a), *Magnusson et al.* (2013b) and *Smith et al.* (2013a). The benefits of each method depend on the area and on the timescale considered as well as on the model used to perform the hindcasts. *Magnusson et al.* (2013b) and *Smith et al.* (2013a) have shown that the variability in the El Niño-Southern Oscillation and its related teleconnections are better reproduced in simulations initialised with full-field data. At seasonal timescales, according to *Smith et al.* (2013a), the temperature and precipitation in simulations initialised with full-field data are generally better simulated than in the case where anomalies are used for the initialisation. At multi-year timescales, these two methods do not lead to significant differences in the predictive skill, except if the ensemble size is small ($\lesssim$ 20 members), in which case the anomaly initialisation could improve the simulation results (*Smith et al.*, 2013a).

1.1.3 Assessment of the prediction system skill

To justify the interest of any prediction system, its skill has first to be measured. This requires performing hindcast simulations, also called retrospective forecasts. A hindcast is a simulation carried out in the same conditions as if it was a forecast, except that it spans a past period of time. For decadal predictions, the hindcast consists of an ensemble of simulations. The analysis of the hindcast results generally focuses on two criteria: (1) the agreement between the different members of the ensemble simulation, referred to as the predictability in *Kirtman et al.* (2013), and (2) the agreement between the ensemble mean and the observations, referred to as the forecast quality or forecast skill in *Kirtman et al.* (2013).

The predictability allows assessing how fast small errors on the knowledge of the initial conditions grow. If states which are initially close to each other scatter rapidly, traducing a rapid loss of the information provided at the initialisation, the predictability is weak and vice versa. The predictability is a property of the prediction system itself and depends on both the perturbation applied to generate the ensemble and the way the model dynamics propagates this perturbation forward in time. The forecast quality, in turn, depends on the ability of the climate model to reproduce the observed evolution of the system.

The metrics used to assess the predictability and the forecast quality vary from one study to the other. However, they rely on common principles that are introduced in the following paragraphs.

Potential predictability

The potential predictability can arise from the externally or internally generated variability (e.g., *Griffies and Bryan*, 1997; *Phelps et al.*, 2004; *Pohlmann et al.*, 2004; *Latif et al.*, 2010; *Kirtman et al.*, 2013). Indeed, if the external forcing and the response of the system to this forcing are known, the potential predictability may increase. However, predictable components of the internal variability can also lead to higher potential predictability of the variables characterising this component.

In many studies (e.g., *Griffies and Bryan*, 1997; *Pohlmann et al.*, 2004; *Koenigk and Mikolajewicz*, 2009; *Döscher et al.*, 2010; *Msadek et al.*, 2010; *Holland et al.*, 2013; *Germe et al.*, 2014; *Zunz et al.*, 2014), the potential predictability is estimated on the basis of the variance of the ensemble (i.e. the mean of the squared distance between the ensemble members and the ensemble mean) normalised by a variance of reference, generally the climatological variance. This approach has been formalised in the concept of prognostic potential predictability (PPP), introduced by *Phelps et al.* (2004) and *Pohlmann et al.* (2004):

$$\mathrm{PPP(t)} = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M-1}\sum_{j=1}^{M}\left[x_{ij}(t) - X_i(t)\right]^2}{\sigma^2_{\mathrm{clim}}}, \tag{1.8}$$

where i is the ensemble index (N ensembles initialised at different times), j is the member index within one ensemble (M members in each ensemble), x_{ij} is the variable x in the hindcast member j of the ensemble i, X_i is the ensemble mean of the ith ensemble. σ^2_{clim} is the climatological variance, generally computed from a reference simulation. A value of the PPP close to 1 (0) means a high (low) potential predictability. Note that his approach does not distinguish the predictability that can be accounted for by the external forcing and the one arising from the internal variability. *Boer* (2011) and *Boer et al.* (2013), for instance, propose a method to disentangle the contribution of the internal and external components to the total potential predictability (not discussed here).

Using the ensemble variance to estimate the potential predictability assumes that the ensemble spread is representative of the range of possible solutions. However, this is not systematically the case given that the ensemble can be too strongly (weakly) scattered, leading to an underestimate (overestimate) of the potential predictability. For instance, this could be due to an inadequate sample of initial states or a wrong amplification of the initial errors by the model. The ensemble spread is considered to appropriately express the forecast uncertainty if the ensemble probability distribution is consistent with the one of the observations (e.g., *Wilks*, 2006; *Goddard et al.*, 2012; *Corti et al.*, 2012). This consistency between the ensemble and observations probability distribution can be assessed through different methods, as for instance the reliability diagram, the rank histogram (also called Talagrand diagram), the Brier score

or the (continuous) ranked probability score, that are detailed in *Talagrand and Vautard* (1997), *Wilks* (2006) and *Goddard et al.* (2012).

Forecast quality

The potential predictability introduced above corresponds to an upper limit of predictability in the analysed model and does not consider the agreement between the hindcast results and the corresponding observations, i.e. the forecast quality. Assessing the forecast quality is generally achieved through the computation of the skill scores based on the mean squared error (MSE) or the anomaly correlation coefficient (e.g., *Pohlmann et al.*, 2009; *Matei et al.*, 2012b,a; *Kröger et al.*, 2012; *Goddard et al.*, 2012; *Pohlmann et al.*, 2013) between the hindcast ensemble mean and the observations.

$$\mathrm{MSE}(t) = \frac{1}{N}\sum_{i=1}^{N}\left(X_i(t) - o_i(t)\right)^2, \tag{1.9}$$

where X_i is the ensemble mean of the ith ensemble, o_i is the observations spanning the same time period as X_i and N is the total number of ensembles. The mean squared skill score (MSSS) compares the MSE computed for the initialised hindcast ($\mathrm{MSE}_{\mathrm{hindcast}}$) with the MSE of a reference forecast such as the climatological mean ($\mathrm{MSE}_{\mathrm{clim}}$) or an uninitialised simulation,

$$\mathrm{MSSS}(t) = 1 - \frac{\mathrm{MSE}_{\mathrm{hindcast}}(t)}{\mathrm{MSE}_{\mathrm{clim}}(t)}. \tag{1.10}$$

In the case where X_i and o_i in Eq. 1.9 are anomalies computed with regard to the climatology of the model and of the observations, respectively, the forecast based on the model climatology corresponds to a zero anomaly and Eq. 1.10 can be expanded (e.g., *Murphy*, 1988; *Wilks*, 2006) as

$$\mathrm{MSSS}(t) = r^2 - \left[r - \frac{s_X}{s_o}\right]^2, \tag{1.11}$$

where $s_X^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i)^2$, $s_o^2 = \frac{1}{N}\sum_{i=1}^{N}(o_i)^2$ and r is the anomaly correlation coefficient (ACC), defined as

$$r(t) = \frac{\sum_{i=1}^{N}\left[X_i(t)o_i(t)\right]}{\sqrt{\sum_{i=1}^{N}\left[X_i(t)\right]^2\sum_{i=1}^{N}\left[o_i(t)\right]^2}}. \tag{1.12}$$

The MSSS thus combines the square of the ACC (first term in the right hand side of Eq. 1.11 and, in the second term in the right hand side of Eq. 1.11, what is called the conditional bias of the forecast (e.g., *Murphy*, 1988; *Wilks*, 2006). The ACC measures the relative association between the hindcasts and the observations but is independent of the magnitude of the changes. As a

consequence, the ACC will likely reach high values in the presence of a trend in both the hindcasts and the observations. The conditional prediction bias, in turn, takes into account the relative amplitude of the time series and is thus sensitive to outliers.

1.1.4 Skill of current decadal prediction systems

The assessment of the skill of decadal predictions and its dependance on the experimental setup have been the subject of many recent studies (e.g., *Fyfe et al.*, 2011; *Matei et al.*, 2012a; *Kröger et al.*, 2012; *Kim et al.*, 2012; *Smith et al.*, 2013b). The CMIP5 coordinated framework (*Taylor et al.*, 2011) has much facilitated the comparison between the results provided by different general circulation models. Currently, the research often focus on the North Atlantic oceanic circulation and the surrounding areas as well as on the Arctic. This is motivated by the high potential predictability reached in these regions at decadal timescales (*Pohlmann et al.*, 2004; *Koenigk and Mikolajewicz*, 2009). Several studies using different models (e.g., *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009; *Matei et al.*, 2012b; *Swingedouw et al.*, 2012; *Pohlmann et al.*, 2013) have demonstrated that skillful decadal predictions of surface temperature over large parts of the North Atlantic, Europe, North America and northern Africa as well as of the meridional overturning circulation in the North Atlantic could be expected if the predictions are adequately initialised.

The potential predictability of the Arctic sea ice extent is generally higher in winter than in summer and is higher in the Atlantic sector than in the Pacific sector of the Arctic (e.g., *Koenigk and Mikolajewicz*, 2009; *Germe et al.*, 2014). While the quality of the predictions of the sea ice volume and extent over the whole Arctic seems to be mainly driven by the external forcing, the initialisation with observations can regionally improve these predictions up to 6 years in advance (*Germe et al.*, 2014). Besides, according to *Tietsche et al.* (2013), the magnitude of strong negative sea ice anomalies are difficult to predict and the initialisation with observations does not clearly improve these predictions.

High potential predictability has also been pointed out in other areas, in particular in the Southern Ocean (e.g., *Pohlmann et al.*, 2004; *Boer and Lambert*, 2008). However, the skill of decadal predictions around Antarctica has been poorly studied so far (e.g., *Holland et al.*, 2013; *Zunz et al.*, 2014; *Zunz and Goosse*, 2014) and is the object of the present thesis.

1.2 Antarctic sea ice

The main concepts relative to decadal predictions have been presented in the previous section. The present section aims at giving an overview on what is known about sea ice in the Southern Ocean: its general characteristics (Sect. 1.2.1), the main processes ruling its dynamics (Sect. 1.2.2) and its evolution

as deduced from models and observations (Sect. 1.2.3). Through the present thesis, unless specified, summer and winter refer to the austral seasons, and we consider that austral summer (winter) spans January, February and March (July, August and September).

1.2.1 How does it look like?

The Southern Ocean is a circumpolar ocean that connects together the Atlantic, Indian and Pacific Oceans. It is usually divided into five sectors: the Weddell Sea, the Indian Ocean, the Western Pacific Ocean, the Ross Sea and the Bellingshausen and Amundsen Seas (Fig. 1.5). At high southern latitudes, this ocean is covered by a blanket of sea ice (Fig. 1.6). The sea ice consists of frozen seawater that forms when the ocean surface temperature reaches the freezing point ($\sim$ -1.9 °C). Among other properties, it is characterised by a high albedo (e.g., *Brandt et al.*, 2005) that prevents a large part of the incoming solar radiation to penetrate into the ocean. Thanks to its low thermal conductivity (e.g., *Pringle et al.*, 2007), the sea ice efficiently insulates the ocean from the atmosphere.

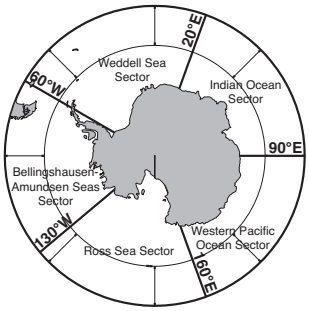

Figure 1.5: Sectors of the Southern Ocean generally used in sea ice analyses.

On average between 1979 and 2010, the sea ice around Antarctica ranges from a minimum extent of 3.1×10^6 km^2 in February and a maximum extent of 18.5×10^6 km^2 in September (*Parkinson and Cavalieri*, 2012). The fraction of the Southern Hemisphere covered by sea ice thus oscillates between $\sim$1% and $\sim$7%. The sea ice cover is not uniform as it is made up of different categories of ice, of varying thickness and composition. Ice-free areas, such as leads[1] and polynyas[2], are also present within the ice pack. Given that most of the sea ice in the Southern Ocean disappears during the melting season, only a small

[1]Lead: narrow, roughly linear crack in the sea ice resulting from the divergence or the shear of ice velocity field.

[2]Polynya: area of persistent open water within the sea ice pack.

fraction of the sea ice persists during more than one year. This results in relatively thin sea ice, mostly less than one meter thick (*Worby et al.*, 2008). The thicker sea ice is found in the western Weddell Sea, around 45°W (*Worby et al.*, 2008), where a part of the sea ice survives to the melting season (Fig. 1.6).

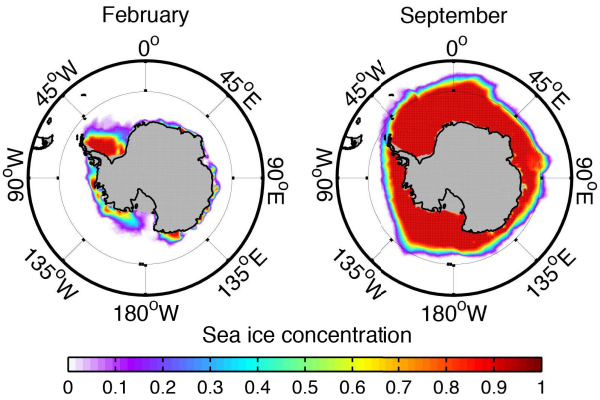

Figure 1.6: Antarctic sea ice concentration derived from the Nimbus-7 SMMR and DMSP SSM/I-SSMIS satellite observations (*Comiso*, 1999, updated daily), on average over the period 1979-2012.

1.2.2 How does it interact with the climate system?

Albeit the relatively small dimensions of the sea ice cover in the Southern Ocean, it strongly affects the Antarctic ecosystem and the evolution of the climate, at both local and global scales. Examples of interaction between the sea ice and the other climate components are presented below.

Thermodynamic feedbacks

The sea ice is involved in several feedback processes that rely on the thermodynamic properties of the sea ice (e.g., *Ebert and Curry*, 1993; *Bitz and Roe*, 2004; *Flanner et al.*, 2011). These feedbacks do not only impact the sea ice itself but also play an important role in the energy balance of the atmosphere above and of the ocean below the sea ice. One of the most important feedbacks implying sea ice is the ice albedo feedback. This positive retroaction loop is associated to the high albedo of sea ice. When sea ice melts, a larger fraction of the incoming solar radiation is absorbed by the ocean that warms up, enhancing the melting of sea ice. This positive feedback is particularly intense during the melting season, i.e. when the incoming solar radiation is strong. However, because sea ice in the Southern Ocean is farther from the South Pole

than is the Arctic sea ice from the North Pole, it receives less incoming solar radiation in austral summer compared to the one received by the Arctic sea ice in the boreal summer, resulting in a weaker impact of the ice albedo feedback on the Antarctic sea ice than on its Arctic counterpart (*Turner and Overland*, 2009). Sea ice is also involved in negative feedback processes that prevent a variation in sea ice to be amplified. For instance, the ice thickness growth feedback occurs during the growing season because of the dependency of the insulation capacity of the sea ice on its thickness. Where sea ice is thin, more heat can escape from the ocean surface and sea ice grows faster.

Ocean stratification and circulation

When sea ice is formed, it rejects brine into the ocean. On the contrary, the melting of sea ice results in a freshwater input at the ocean surface. As a consequence, the melting and the formation of sea ice strongly affects the ocean stratification. In particular, in parts of the continental shelves where strong sea ice formation occurs, salty water sinks along the continental slope and gives birth to the Antarctic Bottom Water (AABW, e.g., *Gordon et al.*, 1991). The latter lies at the bottom of the World Ocean and is an important contributor to the global ocean circulation. The Southern Ocean being overall weakly stratified, especially the Weddell Sea (e.g., *Martinson et al.*, 1981), it is strongly sensitive to changes in density of its different layers, associated to variations in temperature and salinity (e.g., *Gordon et al.*, 1991). Consequently, the modification of the ocean stratification arising from sea ice formation and melting is involved in different feedback processes (e.g., *Martinson*, 1990; *Zhang*, 2007; *Goosse and Zunz*, 2014). For instance, sea ice formation induces a weaker stratification of the water column, allowing warm water from the interior ocean to be entrained in the surface layer through convection, and then limiting the sea ice growth (and vice-versa, *Martinson*, 1990). This retroaction loop thus stabilises the evolution of sea ice. The formation and melting cycle of sea ice and the resulting changes in ocean stratification can also induce positive feedbacks, as discussed for instance in *Zhang* (2007) and *Goosse and Zunz* (2014), that potentially contributes to the observed changes in sea ice extent over the last decades (see Sect. 1.2.3).

Internal modes of variability

The atmospheric variability in the Southern Hemisphere is dominated by the Southern Annular Mode (SAM, e.g., *Thompson and Wallace*, 2000), also known as the Antarctic oscillation. The SAM emerges as the leading Empirical Orthogonal Function (EOF) in many atmospheric fields, including surface pressure, geopotential height and zonal wind (see *Thompson and Wallace*, 2000, and references therein). EOF analysis, also known as Principal Component Analysis (PCA), provides a decomposition of the total variance of a climate variable into independent modes, or patterns, of variability. The fraction of the total variability explained by each EOF mode is the highest for the first

mode and decreases with the following modes. The SAM is characterised by a zonally symmetric structure that involves exchanges of mass between the mid- and high latitudes. The difference between zonal mean pressure at low and high southern latitudes is commonly used as an index to quantify the strength of the SAM (e.g., *Gong and Wang*, 1999). Alternatively, the time series of the principal component associated to the first EOF mode is also often employed as a SAM index (e.g., *Thompson and Wallace*, 2000; *Lefebvre and Goosse*, 2008a).

A positive phase of the SAM consists of low pressure over the continent, high pressure in the extra-tropics and stronger circumpolar westerly winds (and vice-versa). During a positive phase of the SAM, the westerly winds encircling Antarctica isolates the latter from warmer area. The subsequent cooling could lead to an increase in sea ice formation. Additionally, the westerly winds result in a northward Ekman current in the ocean surface that pushes sea ice away from the continent, as well as an upwelling due to the Ekman pumping in the ocean at high latitudes. The northward drift of sea ice could also contribute to an expansion of the ice cover (e.g., *Hall and Visbeck*, 2002). A link between the variations in sea ice around Antarctica and the SAM has been confirmed on the basis of observations as well as through model studies (e.g., *Lefebvre et al.*, 2004; *Lefebvre and Goosse*, 2005; *Holland and Raphael*, 2006; *Stammerjohn et al.*, 2008; *Goosse et al.*, 2009a). Non-annular components of the SAM likely contribute to the regional pattern of the sea ice concentration. Indeed, when the SAM index is high, the pressure pattern in the Southern Ocean area is not perfectly annular and results in southward surface wind stress in the Weddell, Bellingshausen and Amundsen Seas, while the surface wind stress is directed northward in the Ross Sea. This induces an increase in sea ice in the Ross Sea and a decrease in the Weddell Sea associated to high SAM index (e.g., *Lefebvre et al.*, 2004; *Lefebvre and Goosse*, 2005).

Additionally, studies have highlighted teleconnections between the high latitudes variability in the Southern Ocean and the variability occurring in remote areas, in particular the El Niño-Southern Oscillation (ENSO) (e.g., *Yuan and Martinson*, 2001; *Liu et al.*, 2002; *Yuan*, 2004; *Holland and Raphael*, 2006; *Stammerjohn et al.*, 2008). The mechanisms linking ENSO to the variations in sea ice concentration imply a mode of variability called the Antarctic Dipole (ADP), occurring at high southern latitudes. The ADP consists of a quasi-stationary wave characterised by opposite signals between the Pacific and Atlantic sectors of the Southern Ocean (e.g., *Yuan and Martinson*, 2001). This dipole is essentially present in the surface temperature and sea ice fields. During warm ENSO events, the surface temperature in the Pacific sector of the Southern Ocean generally increases, while the surface temperature decreases in the Atlantic sector. Warm ENSO events are thus associated with a growth of sea ice in the Atlantic sector of the Southern Ocean and a melting of sea ice in the Pacific sector. The opposite behaviour emerges during cold ENSO events.

1.2.3 How has it evolved over past decades?

The Southern Ocean is an isolated and harsh area. Observations in this region are thus relatively sparse before the satellite era. Thanks to satellite passive-microwave observations, records of the sea ice in the Southern Hemisphere are available from December 1972 (e.g., *Comiso*, 1999, updated daily; *Cavalieri et al.*, 2003; *Parkinson and Cavalieri*, 2012). However, data derived from satellite observations prior to October 1978 are generally excluded from sea ice analyses because of the lack of intercalibration with more recent data. Prior to the 1970's, a few data exist that could extent the length of the period covered by observations. For instance, *Meier et al.* (2013) have initiated a work to process data from satellites operating in the late 1960's and early 1970's. Their results indicate that the 1964 sea ice extent was higher than the 1979-present estimates. Ice core and whaling records are also used to estimate indirectly the state of the sea ice before the 1970's (e.g., *de la Mare*, 1997; *Curran et al.*, 2003; *de la Mare*, 2009). These measurements are in agreement with the study of *Meier et al.* (2013) and suggest that the slight increase in sea ice extent characterising the last three decades (see below) might follow a large melting that may have happened before the late 1970's. However, these indirect observations provide limited spatial and temporal information and have to be interpreted cautiously. The presumed sea ice decline between 1950 and the late 1970's is consistent with the observations of atmospheric variables at high southern latitudes which indicate that the 1950-1978 period was warmer and characterised by weaker westerlies and higher sea level pressure than the 1979-2011 period (*Fan et al.*, 2014). Data set for the ice thickness or for its snow cover are even more limited in their spatial and temporal coverage than sea ice concentration data (e.g., *Giles et al.*, 2008).

Between November 1978 and December 2012, the sea ice extent (SIE) in the Southern Ocean has increased at a rate of 0.13 to 0.2 million km^2 per decade (*Kirtman et al.*, 2013). Over this period, the trends in annual mean sea ice concentration are characterised by a regional pattern, displaying a decrease in the Bellingshausen and Amundsen Seas associated to a warming of the Antarctic Peninsula (e.g., *Vaughan et al.*, 2003), and an increase elsewhere, which is particularly strong in the Ross Sea (Fig. 1.7). The behaviour of the sea ice in the Southern Hemisphere thus clearly contrasts with the one in the Arctic that decreases in response to the global warming (e.g., *Turner and Overland*, 2009). Different processes have been pointed out as possible causes for this observed evolution of the sea ice in the Southern Ocean. These are summarised below.

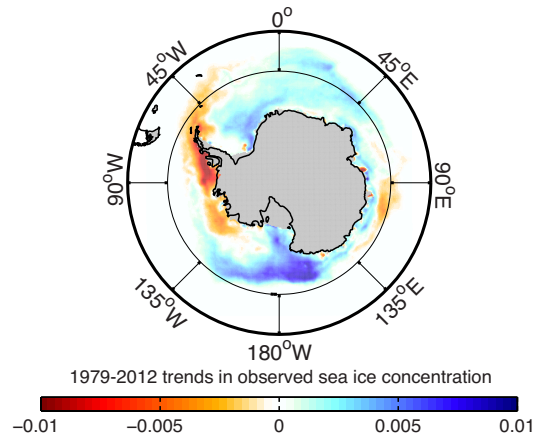

Figure 1.7: 1979-2012 trends in annual mean sea ice concentration in the Southern Ocean, from data derived from the Nimbus-7 SMMR and DMSP SSM/I-SSMIS satellite observations (*Comiso*, 1999, updated daily).

The recent trend in SIE and the atmospheric circulation

Recent decades are characterised by a positive phase of the SAM that strengthens with time (e.g., *Thompson and Solomon*, 2002; *Marshall*, 2003). This increase likely occurs in response to the emissions of greenhouse gases coupled with the stratospheric ozone depletion, although the internal variability of the climate system might also be involved (e.g., *Fyfe et al.*, 1999; *Kushner et al.*, 2001; *Turner et al.*, 2009; *Thompson et al.*, 2011; *Swart and Fyfe*, 2012). Albeit the links between the variability in sea ice and the atmospheric modes of variability (See Sect. 1.2.2), the relation between the trend in the SAM and the trends in the observed sea ice extent has not been confirmed (e.g., *Liu et al.*, 2004; *Lefebvre et al.*, 2004; *Lefebvre and Goosse*, 2005; *Stammerjohn et al.*, 2008; *Lefebvre and Goosse*, 2008a; *Simpkins et al.*, 2012). Besides, the hypothesis that the observed recent increase in sea ice extent could be linked to the stratospheric ozone depletion is not in agreement with recent studies (e.g., *Sigmond and Fyfe*, 2010; *Bitz and Polvani*, 2012; *Smith et al.*, 2012; *Sigmond and Fyfe*, 2013). The changes in the sea ice cover over the last 30 years thus do not seem to be strongly influenced by one single large-scale process. However, the regional distribution of the trends in sea ice concentration could still arise from atmospheric patterns that are not systematically related to a known mode of internal variability and that act at regional scale (e.g., *Lefebvre and Goosse*, 2008a; *Holland and Kwok*, 2012).

The recent trend in SIE and the freshening of the ocean surface

In parallel to the increase in sea ice extent, the last decades are also characterised by a freshening of the surface of the Southern Ocean (e.g., *Jacobs et al.*, 2002). Such a freshening generally leads to a stabilisation of the water column that results in a reduced vertical mixing in the ocean. Consequently, the heat contained by the ocean is trapped in the intermediate layer, the surface cools down and the observed freshening could thus also be responsible for the observed increase in sea ice extent. This freshening has been attributed to different causes.

First, *Liu and Curry* (2010) have pointed out a link between the increase in sea ice extent and the enhanced hydrological cycle associated to the large-scale warming. This enhanced hydrological cycle results in a net freshwater loss at low southern latitudes and a net freshwater gain at high southern latitudes (e.g., *Fyfe et al.*, 2012). In addition, given that precipitation at high latitudes is mainly in the form of snow, it increases the ice albedo and, according to *Liu and Curry* (2010), it stimulates the growth of sea ice through the ice albedo feedback. Besides, the young and thin sea ice that constitutes most of the ice cover in the Southern Ocean is subject to the formation of snow ice, in response to the higher snow load, that directly contributes to an increase in the sea ice thickness. The thickening of the sea ice strengthen the thermal insulation between the ocean and the atmosphere, resulting in a weakening of the basal growth of sea ice. However, the increase in sea ice volume due to the formation of snow ice may compensate for or even outstrip the reduced basal sea ice growth (e.g., *Fichefet and Maqueda*, 1999; *Lecomte et al.*, 2013), leading to a stable or growing sea ice volume in response to enhanced snowfalls.

Second, the melting of the Antarctic ice sheet and icebergs also contributes to the freshwater budget of the Southern Ocean surface (e.g., *Silva et al.*, 2006; *Rignot et al.*, 2008; *Velicogna*, 2009; *Pritchard et al.*, 2012; *Shepherd et al.*, 2012). The model results of *Bintanja et al.* (2013) seems to confirm a link between the increase in sea ice extent and the Antarctic ice sheet retreat. On the contrary, *Swart and Fyfe* (2013) have suggested that including a freshwater flux compatible with the Antarctic ice sheet mass imbalance has little impact on the simulated sea ice.

Third, feedbacks associated to the growth and melting cycle of the sea ice could also result in a net freshening of the ocean surface (e.g., *Zhang*, 2007; *Goosse and Zunz*, 2014). For instance, in their model study, *Goosse and Zunz* (2014) have shown that, while the brine rejected during ice growth is exported in the intermediate layers thanks to the strong vertical mixing occurring in winter, the freshwater provided by the melting of sea ice during summer is included at the ocean surface only. This induces a net downward transport of salt and a freshening of the surface on an annual average because of the ice formation, leading to a positive feedback.

The recent trend in SIE and the internal variability

Several studies have discussed the link between the observed increase in sea ice extent and the internal variability of the climate system (e.g., *Turner et al.*, 2013; *Polvani and Smith*, 2013; *Mahlstein et al.*, 2013; *Zunz et al.*, 2013; *Swart and Fyfe*, 2013). For instance, the magnitude of the multi-decadal trends in observed sea ice extent and sea ice concentration is influenced by decadal variability and is thus sensitive to the time period chosen for the analysis (e.g., *Simpkins et al.*, 2013). In addition, as mentioned at the beginning of Sect. 1.2.3, the few observations before the 1970's indicate that, at the end of the 1960's, the sea ice extent in the Southern Ocean was likely larger than the 1979-2010 average (e.g. *de la Mare*, 1997, 2009; *Meier et al.*, 2013). The period spanning the last 30 years could have thus been preceded by a large melting of the sea ice, consistent with the behaviour of atmospheric variables, as pointed out in several model and observation based studies (e.g., *Fichefet et al.*, 2003; *Curran et al.*, 2003; *Cavalieri et al.*, 2003; *Cotté and Guinet*, 2007; *Goosse et al.*, 2009a; *Fan et al.*, 2014), suggesting a large interdecadal variability of the Antarctic sea ice.

Due to the limited time period over which observations are available, the internal variability of the sea ice at multi-decadal time scales has to be analysed through a model approach. Although the amplitude of the observed trend in sea ice extent during the last 30 years lies within the range of the internal variability simulated by current general circulation models, it is a rather rare event among the simulations performed with these models over the period 1979-2005 (e.g., *Turner et al.*, 2013; *Polvani and Smith*, 2013; *Mahlstein et al.*, 2013; *Zunz et al.*, 2013; *Swart and Fyfe*, 2013). The link between the recent trend in sea ice extent and the internal variability is thus plausible according to a model analysis, but current GCMs generally simulate an interannual variability of the sea ice extent that is not in agreement with the observed one (*Turner et al.*, 2013; *Zunz et al.*, 2013).

1.3 Contribution of the thesis

The first two sections of the present Introduction have allowed identifying two weakly understood aspects of the climate system: (1) the recent variability (Sect. 1.2) and (2) the predictability (Sect. 1.1) of the Antarctic sea ice at decadal timescales. These have defined the direction of the research carried out during this thesis. The results of this work, presented in details in the following chapters, have shed some light on these issues and are summarised below.

1.3.1 Variability of the Antarctic sea ice in GCMs (Chapter 2, p. 33)

As mentioned in Sect. 1.2.3, reliable observations of the whole sea ice cover in the Southern Ocean are available from the late 1970's and indicate an increase in sea ice extent between November 1978 and December 2012 (*Kirtman et al.*, 2013). Possible causes for this positive trend have been presented in Sect. 1.2.3, but the respective contributions of the proposed mechanisms are still debated (*Kirtman et al.*, 2013). Given the short time period spanned by the observations, the results provided by climate models are particularly useful to investigate the relevancy of the hypotheses proposed to account for the recent expansion of the sea ice.

In Chapter 2, the accent is put on the internal variability of sea ice. Two main questions have been addressed. (1) Is the observed positive trend in sea ice extent compatible with a combination of the forced response and the internal variability of the climate system as simulated by current general circulation models? (2) How does the initialisation with observations improve the simulated trend?

The sea ice concentration simulated by 24 general circulation models have been collected from the CMIP5 database. Historical simulations of all of these 24 models have been analysed. Each historical simulation consists of an ensemble whose size varies from one model to the other and generally spans the period 1850-2005. No observation is taken into account at the initialisation, but these simulations are driven by anthropogenic and natural external forcings. On average over all the historical simulations, the seasonal cycle of the sea ice concentration and extent over the period 1979-2005 are in good agreement with the observations. However, the models considered individually provide a large range of solutions (Fig. 1.8a). The timing as well as the amplitude of the extrema of the seasonal cycle of the sea ice extent vary much between the models and barely fit the observations. The monthly standard deviation of the sea ice extent also differs strongly from one model to the other and is generally larger in winter than in summer (Fig. 1.8b). During winter months, all of the 24 GCMs analysed here display a standard deviation larger than the observations. This strong and robust bias thus raises the question of whether current GCMs can be reasonably used to investigate the internal variability of the sea ice in the Southern Ocean.

The trends in sea ice extent between 1979 and 2005, computed from the results of these historical simulations, display a large range. Furthermore, these trends may strongly differ between the members of the same model ensemble. Considering all the members of the 24 model ensembles, 14% (12%) of the trends in summer (winter) sea ice extent are positive. This points out the rarity of the positive trend in sea ice extent, such as the observed one, among the CMIP5 ensemble. Anyway, the magnitude of the observed trend ranges

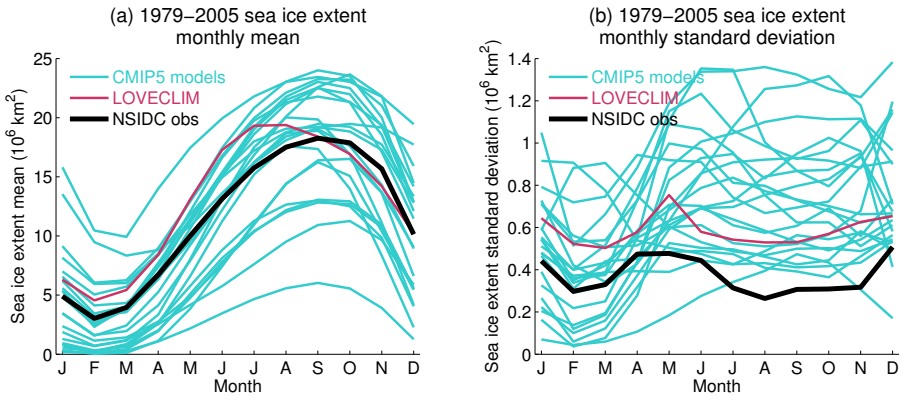

Figure 1.8: (a) Monthly mean of the Antarctic sea ice extent, computed over the period 1979-2005. (b) Standard deviation of the detrended Antarctic sea ice extent, computed over the period 1979-2005 for each month of the year. Black bold line refers to observations (Cavalieri and Parkinson, 2008). Blue lines refer to CMIP5 models. Purple line corresponds to the LOVECLIM model whose results are discussed in Chapters 3 and 4. Figure adapted from *Zunz et al.* (2013).

among the values of the trends provided by the individual members of the analysed simulations. It can thus be concluded that the observed trend in sea ice extent is compatible with the internal variability of general circulation models. However, this finding must be considered cautiously given the strong bias in the internal variability of the simulated sea ice extent.

The decadal simulations of 10 CMIP5 models, spanning the period 1980-2009, have also been analysed. Contrarily to the historical simulations, the initialisation of the decadal simulations takes into account the observations through a data assimilation procedure (generally based on a nudging). According to our analysis, this initialisation with observations does not clearly improve the simulated trends in Antarctic sea ice extent. In some cases, it can even introduce spurious trends at the beginning of the simulation that obscure the longer-term trend. This lack of improvement might be due to the fact that sea ice and atmospheric data are barely taken into account in the initialisation procedure of CMIP5 models. Initialising sea ice and atmospheric variables might improve the simulated sea ice as the atmospheric variability potentially drives the evolution of the latter, as suggested by the reconstructions of the Antarctic sea ice cover obtained with global ice ocean models driven by atmospheric reanalyses (e.g., *Fichefet et al.*, 2003; *Zhang*, 2014). The lack of improvement of the simulated trends in decadal CMIP5 simulations might also be accounted for by the rather simple data assimilation methods implemented

in these models and the effect of more sophisticated data assimilation needs to be investigated. This encouraged us to orientate the next steps of this research towards systematic tests of varying initialisation procedures, based on different data assimilation methods. The results of this work are presented in Chapter 3 and summarised in the following section.

1.3.2 Impact of the initialisation on the Antarctic sea ice predictability (Chapter 3, p. 61)

Having pointed out the systematic biases in the mean state and internal variability of the Antarctic sea ice simulated by current GCMs, it appeared relevant to first test the impact of different initialisation strategies in a perfect model framework. This approach consists in using a reference simulation performed with the model as pseudo-observations, instead of actual observations. This allows getting rid of the model biases given that the model and the pseudo-observations have, by construction, identical climatological mean and variance. Furthermore, it allows performing tests over several time periods, the availability in space and time of the pseudo-observations being unlimited.

Systematic tests require performing a large amount of simulations that cannot be afforded with a GCM. In that context, working with an Earth system model of intermediate complexity, whose computational cost is lower than a GCM, appears to be useful. Our choice turned to the model LOVECLIM (*Goosse et al.*, 2010a). Its sea ice extent mean state and variability ranging well among the ones of the CMIP5 GCMs (Fig. 1.8), LOVECLIM constitutes an adequate tool to study the impact of the initialisation on the simulated sea ice in the Southern Ocean.

LOVECLIM is made up of 5 interactive components (see Fig. 1 in *Goosse et al.* (2010a)). In the present study, only the atmospheric (ECBilt, *Opsteegh et al.*, 1998), the oceanic-sea ice (CLIO, *Goosse and Fichefet*, 1999) and the vegetation (VECODE, *Brovkin et al.*, 2002) components are activated. The two other components (the ice sheet model AGISM (*Huybrechts*, 2002) and the ocean carbon cycle model LOCH (*Mouchet and François*, 1996)) evolve too slowly to have a significant impact on multi-decadal timescales and are thus deactivated. ECBilt is a T21 (corresponding to a horizontal resolution of $\sim$ 5.6°$\times$ 5.6 °) three-level quasi-geostrophic model. CLIO is an ocean general circulation model coupled to a comprehensive thermodynamic-dynamic sea ice model with a horizontal resolution of 3°$\times$ 3° and 20 unequally spaced vertical levels in the ocean. VECODE simulates the evolution of the fraction of trees, grass and desert at the same horizontal resolution as ECBilt.

The coarse resolution of LOVECLIM and its lower level of complexity compared to GCMs offer a considerable advantage in terms of computational cost, but these have the drawback of misrepresenting some important processes. For instance, the atmospheric component of LOVECLIM has only three vertical

levels, the upper one is at 200 hPa and the changes in the vertical profiles of temperature and humidity are rather crudely represented (*Goosse et al.*, 2010a). While LOVECLIM reproduces reasonably well the spatial distribution of the observed surface temperature, it overestimates the surface temperature in the tropics (*Goosse et al.*, 2010a). Overall, LOVECLIM does not simulate correctly the atmospheric circulation in the tropics and strongly underestimates the variance of El Niño-Southern Oscillation (*Sriver et al.*, 2014), one of the dominant mode of variability in this area that likely impacts the variability at high southern latitudes, as discussed in Sect. 1.2.2.

The stratospheric dynamics is not included in ECBilt and the response of the atmospheric circulation to the stratospheric ozone depletion cannot thus be properly simulated. However, in the extra-tropics, LOVECLIM performs relatively well and is able to reproduce the main modes of atmospheric variability characterising the North Atlantic and the Arctic areas, in particular the North Atlantic and the Arctic Oscillations (e.g., *Goosse et al.*, 2001, 2005). In the Southern Ocean, an EOF analysis performed on the monthly mean anomalies of the geopotential height at 800 hPa, southward of 20°S, from the last 1000 years of a 5000-year control simulation carried out with LOVECLIM provides a first EOF mode that displays an annular pattern similar to the Southern Annular Mode presented in Sect. 1.2.2 (Fig. 1.9a). Positive values extend from mid- to high latitudes, while negative values cover the Antarctic region (between approximately 90°S and 50°S). The latitude where the sign of the first EOF mode reverses (black contour in Fig. 1.9) roughly corresponds to the one derived from reanalyses data (e.g., *Thompson and Wallace*, 2000; *Lefebvre et al.*, 2004), except between 45°W and 180°W where this sign shift is apparently located too close to the Antarctic continent in LOVECLIM results. The second and the third EOF modes of the geopotential height in LOVECLIM are not well separated (Fig. 1.9b, c) and likely correspond to the combination of several modes of variability identified in previous studies (e.g., *Mo and White*, 1985; *Raphael*, 2004)

The correlation between the SAM index derived from atmospheric reanalyses and the sea ice concentration provided by satellite observations displays a spatial pattern characterised by negative values in the Weddell Sea sector and positive values in the Ross Sea sector, especially during winter (Fig. 1.10a, c), in agreement with *Lefebvre et al.* (2004). As mentioned in Sect. 1.2.2, this regional structure of the response of the sea ice to the SAM can be accounted for by the non-annular components of the SAM. This spatial distribution of the correlations is not well reproduced in the results of a control simulation performed with the model LOVECLIM (Fig. 1.10b, d), suggesting that the response of the sea ice to the SAM may be not well represented in this model.

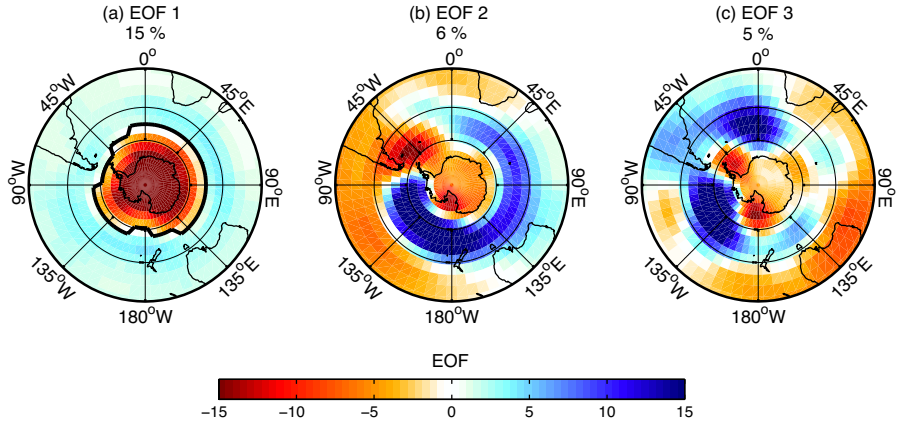

Figure 1.9: Patterns associated with the first three EOF modes of the monthly mean geopotential height at 800 hPa, provided by the last 1000 years of the 5000-year control simulation performed with LOVECLIM analysed in *Goosse and Zunz* (2014). The latitudes southward of 20°S are included in the computation of the EOFs. The percentages indicate the fraction of the total variance that is accounted for by each EOF mode. The black contour in (a) corresponds to 0 values contour and indicates the latitude where the values of the first EOF mode reverses sign.

To assess the impact of the initialisation on the predictability of sea ice in the Southern Ocean, hindcast simulations made up of 96 members were launched every 5 years between 1900 and 1990. These hindcasts were initialised through different methods: without taking into account any observations, from a perfect initial state, i.e. a state directly extracted from the pseudo-observations, or from a state obtained through the data assimilation of pseudo-observations.

Different simulations with data assimilation, based on a nudging, a particle filter with sequential importance resampling or a nudging proposal particle filter (see Sect. 1.1.2), were tested. The results of those experiments have been discussed in *Dubinkina and Goosse* (2013). They assimilate pseudo-observations of surface air temperature, to which noise has been added in order to mimic the observation errors.

The skill of the prediction system was assessed through the prognostic potential predictability and the correlation with the pseudo-observations (see Sect. 1.1.3) for the ice edge location and for the sea ice extent. Two types of timescales were investigated: the interannual ones, i.e. between one month and 10 years, and the multi-decadal ones, through the analysis of the trends over 10- to 30- year long time periods.

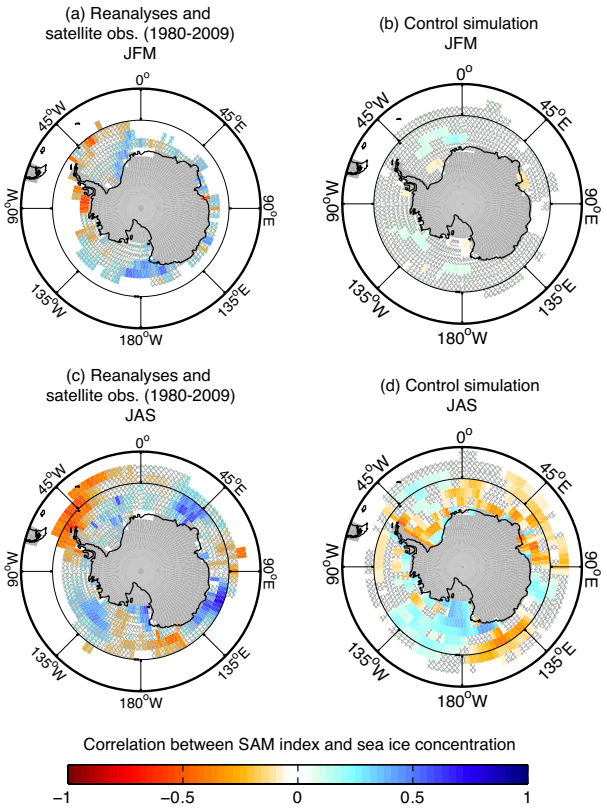

Figure 1.10: Correlation between the sea ice concentration and the SAM index (a,b) in summer and (c,d) in winter. The first column correspond to the correlation, computed over the period 1980-2009, between the sea ice concentration provided by satellite observations (*Comiso*, 1999, updated daily) and the SAM index derived from NCEP-NCAR atmospheric reanalyses (*Kalnay et al.*, 1996), available at `http://www.jisao.washington.edu/aao/slp`. The right column display the results obtained from the last 1000 years of the 5000-year control simulation performed with LOVECLIM analysed in *Goosse and Zunz* (2014). Hatched areas highlight values that are not statistically significant at the 95% level and the autocorrelation of the residuals is taken into account to compute the significance threshold.

At interannual timescales, the PPP of the Antarctic sea ice extent is larger in winter than in summer, because of the enhanced interaction between the surface and the interior ocean during winter. Except in perfectly initialised hindcasts, the PPP of the sea ice extent is statistically significant during the first 4 years of the hindcasts at best. The correlation between the sea ice

extent provided by the hindcasts and the one of the pseudo-observations reaches very similar values whether pseudo-observations are taken into account at the initialisation or not. This suggests that, at interannual timescales, most of the correlation arises from the external forcings and that the improvement brought by the initialisation is rather slight, except for lead times up to two years in the eastern Weddell Sea, in the western Indian Ocean and in the Bellingshausen and Amundsen Seas during winter.

At multi-decadal timescales, the correlation between the trends in Antarctic sea ice extent provided by the hindcasts and the one of the pseudo-observations easily reaches values higher than 0.5 and is generally higher in winter than in summer. In addition, the hindcasts whose initial state takes into account the pseudo-observations systematically provide a correlation with pseudo-observations that is greater than the one of the hindcasts initialised without taking into account any observation. This indicates that, at multi-decadal timescales, the correlation is not provided entirely by the external forcings, the initialisation with pseudo-observations also playing a role. The analysis of the heat contained in the first 100 m of the ocean highlights a link between the quality of the initialisation of the ocean below sea ice and the simulated trends in sea ice extent.

The rather weak predictive skill of the Antarctic sea ice at interannual timescales obtained in a perfect model framework strongly reduces the hope of performing good predictions of the sea ice in realistic conditions at these timescales. On the contrary, in this idealised framework, the predictive skill at multi-decadal timescales reaches satisfying values. An improvement of the simulated trends in sea ice extent can thus be expected in realistic conditions if the hindcasts are adequately initialised. This issue is specifically addressed in Chapter 4, whose results are summarised in the following section.

1.3.3 Influence of a meltwater input on the Antarctic sea ice forecast skill (Chapter 4, p. 93)

The last part of this thesis is devoted to the analysis of the trends in sea ice extent and concentration between 1980 and 2009 simulated by the model LOVECLIM under different experimental setups. The model results were compared to actual observations. Two kinds of simulations have been carried out: simulations spanning the period 1850-2009, that assimilate actual observations of surface air temperature anomalies (HadCRUT3, *Brohan et al.*, 2006) through a nudging proposal particle filter, and hindcast simulations, initialised in 1980 from a state provided by a simulation with data assimilation and running until 2009. All the simulations were 96-member ensembles and were driven by natural and anthropogenic external forcings. The sea ice extent and concentration computed by these simulations are compared to data derived from satellite observations (*Comiso*, 1999, updated daily; *Fetterer et al.*, 2002, updated daily).

Given the potential link between the freshening of the surface of the Southern Ocean and the observed expansion of the sea ice (see Sect. 1.2.3), it appeared relevant to check the impact of an additional freshwater flux on the results of the hindcast simulations. However, reliable estimate of the freshwater input derived from the Antarctic ice sheet mass imbalance over the last century are lacking. In this study, this flux was reconstructed through the assimilation process: a freshwater flux that follows an autoregressive process was added to each member of the ensemble; thanks to the particle filter, the simulations in which the random freshwater flux adequately pulls the model solution towards the observations were kept and duplicated. The mean freshwater flux that brought the system close to the observations was then reconstructed a posteriori and applied in other simulations.

If no additional freshwater flux is included during the simulation with data assimilation, the simulated trend in sea ice extent between 1980 and 2009, on average over the ensemble, is negative and is thus in disagreement with the observations. Including an additional freshwater flux increases the number of degrees of freedom of the system and allows the simulation with data assimilation to provide a trend in sea ice extent averaged over the ensemble that fits the observed one. Furthermore, the regional distribution of the trends in sea ice concentration also agrees well with the observations. The additional freshwater flux is negative on average over the period 1850-2009, indicating that a saltier surface ocean provides a better agreement between the model and the observations. This mean freshwater flux actually allows compensating for model deficiencies, but the experimental design implemented here does not allow determining the origin of these deficiencies (e.g., misrepresentation of the hydrological cycle or lack of interactive ice sheet component). The additional freshwater flux also induces a shift of the mean state of the sea ice extent towards lower values.

The hindcast simulations performed in this study brought three main conclusions. (1) The initialisation through data assimilation of observations impacts the simulated Antarctic sea ice at multi-decadal timescales, confirming the results obtained in a perfect model framework. Furthermore, an adequate initialisation of the hindcast allows reproducing the recent increase in sea ice extent, supporting the hypothesis that this increase could arise from the internal variability of the system. (2) The experimental design implemented for a hindcast simulation, in particular here the amplitude of the additional freshwater flux, has to be consistent with the experimental setup used in the simulation with data assimilation providing the initial state of the hindcast. If this is not the case, a drift of the model may produce unrealistic trends at the beginning of the hindcast that can blur the longer term trend. (3) An increase in the additional freshwater flux is not required to simulate a growth in sea ice extent such as the observed one. Overall, the results of both simulations with data assimilation and hindcasts confirm the link between the freshwater budget at

the surface of the ocean and the state of the sea ice above it. However, they do not prove any relationship between a recent melting of the Antarctic ice sheet and the increase in Antarctic sea ice extent observed over 1980-2009.

CHAPTER 2

Variability of the Antarctic sea ice in GCMs

This Chapter is based on the following paper: 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.

Abstract

Observations over the last 30 years have shown that the sea ice extent in the Southern Ocean has slightly increased since 1979. Mechanisms responsible for this positive trend have not been well established yet. In this study, we tackle two related issues: is the observed positive trend compatible with the internal variability of the system, and do the models agree with what we know about the observed internal variability? For that purpose, we analyse the evolution of sea ice around Antarctica simulated by 24 different general circulation models involved in the 5th Coupled Model Intercomparison Project (CMIP5), using both historical and hindcast experiments. Our analyses show that most of the

CMIP5 models respond to the forcings, including the one induced by stratospheric ozone depletion, by reducing the Antarctic sea ice cover, with only a few simulations displaying an increase in sea ice extent similar to the observed one. The observed positive trend ranging among the solutions provided by CMIP5 models, it is concluded that this trend is compatible with the models internal variability. However, models strongly overestimate the variance of sea ice extent and the internal variability they simulate may thus be not realistic. Furthermore, the initialisation methods currently used in models do not improve systematically the simulated trends in sea ice extent. On the basis of those results, a critical role of the internal variability in the observed increase in sea ice extent in the Southern Ocean could not be ruled out, but current models results, because of the model biases, appear inadequate to test more precisely this hypothesis.

2.1 Introduction

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

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

Several studies dealing with the potential role of the forced response have pointed out the relationship between stratospheric ozone depletion over the past few decades (*Solomon*, 1999) and changes in the atmospheric circulation at high latitudes (e.g., *Turner et al.*, 2009; *Thompson et al.*, 2011). Indeed, variations of sea ice extent in the Southern Ocean are strongly influenced by changes in the atmosphere circulation (e.g., *Holland and Raphael*, 2006; *Goosse et al.*, 2009a). However, the link between atmospheric circulation and the sea ice extent integrated over the Southern Ocean is not straightforward (e.g., *Lefebvre and Goosse*, 2008a; *Stammerjohn et al.*, 2008; *Landrum et al.*, 2012) and several recent studies came to the conclusion that the stratospheric ozone depletion does not lead to an increase in sea ice extent (e.g., *Sigmond and Fyfe*, 2010; *Smith et al.*, 2012; *Bitz and Polvani*, 2012). A second potential cause of the observed expansion of the sea ice cover relies on an enhanced stratification of the ocean which would inhibit the heat transfer to the surface. This strengthened stratification is mainly due to a freshening of the surface water, triggered by an increase in precipitation over the Southern Ocean, the melting of the ice shelves, and changes in the production and transport of sea ice (e.g., *Bitz et al.*, 2006; *Zhang*, 2007; *Goosse et al.*, 2009a; *Kirkman and Bitz*, 2010). *Liu and Curry* (2010) pointed out that an enhanced hydrological cycle may also increase the snowfalls at high latitudes in the Southern Ocean. In that case, the snow cover on thicker sea ice would raise the surface albedo, strengthen the insulation between the atmosphere and the ocean, and thus would protect the sea ice from melting. Nevertheless, this mechanism mainly impacts thick ice because, for thin ice, the higher snow load leads to seawater flooding and to the formation of snow ice. This decreases the effect of the initial increase in snow thickness.

Another hypothesis suggests that the positive trend in the Southern Ocean sea ice extent could arise from the internal variability of the system that masks the warming signal in the Southern Ocean that should characterise the response to an increase in greenhouse gases concentrations, according to climate models. In this framework, some recent studies have drawn attention to the importance of distinguishing the lack of agreement between models from the lack of significant signal (e.g., *Tebaldi et al.*, 2011; *Deser et al.*, 2012). A trend can be significant from a statistical point of view, i.e. if it is above a threshold of significance computed through a statistical test. This does not imply that its value is outside the range that can be reached by the internal variability. For instance, *Landrum et al.* (2012) have pointed out that large interannual variability in simulated sea ice concentration leads to late 20th Century trends in sea ice concentration that are not always statistically significant for individual members of an ensemble simulation. The observed positive trend of Antarctic sea ice extent is statistically significant at the 95 % level for the last 30 years (e.g., *Cavalieri and Parkinson*, 2008). However, this time period is too short to properly assess the multi-decadal variability of the system. Consequently,

we cannot estimate if this trend is exceptional or if similar conditions have already occurred many times in the recent past. The period spanning the last 30 years during which sea ice cover slightly expanded in the Southern Ocean might follow a large melting that may have happened before 1979 (e.g., *de la Mare*, 1997, 2009; *Cavalieri et al.*, 2003; *Curran et al.*, 2003; *Cotté and Guinet*, 2007; *Goosse et al.*, 2009a). This suggests that multi-decadal climate variability in the Southern Ocean area is large, but the available data do not allow a quantitative estimation of its amplitude. Sparse data from the 1960s are currently being processed (e.g., *Meier et al.*, 2013), making observations of the sea ice extent available over a longer time period. Further analyses based on these prolonged time series might therefore improve our knowledge of the internal variability of the sea ice extent. Nevertheless, until longer continuous time series are available, the results from model simulations appear to be crucial to balance the lack of observations. Provided that models are compatible with the available observations, they can help addressing the issue whether the observed positive trend in the Antarctic sea ice extent is due to external forcings or to internal variability, or to both of them.

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

In this chapter, we examine outputs from general circulation models (GCMs) following the 5th Coupled Model Intercomparison Project (CMIP5) protocol. To further study the role of the internal variability in the increasing trend in sea ice extent in the Southern Ocean and in the apparent disagreements between models and observations, we deal with two kinds of simulations: historical and hindcast (or decadal) simulations. The first ones are driven by external forcings and are initialised without observational constraints. They are used to assess how well each model simulates the observed mean state, variability and trends in sea ice concentration and extent. The objective is to study the possible links between the internal variability of the system and the simulated trend in sea ice extent. Our purpose is, on the one hand, to test if the internal variability of the models agrees with the one of the observations. On the other hand, we check if the observed positive trend stands in the range of trends provided by the model internal variability. Analysing

the mean state also appears to be important here because of its impact on the simulated variability. In particular, according to *Goosse et al.* (2009b), a smaller Antarctic sea ice extent is generally associated to a weaker variance of the extent. In addition to those points related to the variability of the system, the way stratospheric ozone is taken into account in models is also discussed to estimate if this has a significant impact on the simulated trends. However, it is out of the scope of this study to discuss specific mechanisms that link the sea ice extent and the stratospheric ozone variations.

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

An initial investigation of the results of CMIP5 models has shown that, in agreement with previous studies related to CMIP3 models (e.g., *Lefebvre and Goosse*, 2008b), current GCMs do not simulate a spatial structure of the trend in sea ice extent similar to the observed one. This spatial structure might as well arise from the internal variability. In such a case, models would not have to fit the observed pattern as discussed above. However, this remains a hypothesis and we have chosen to focus on the sea ice extent in the whole Southern Ocean rather than in the individual sectors to avoid the additional complexity associated with the spatial structure of the changes. Models and observation data are briefly presented in Sect. 2.2. The time period we analyse is limited by the available observations. For the Southern Ocean, validation data are quite sparse before 1979. We therefore examine outputs between 1979 and 2005. Results provided by models' historical simulations are presented and discussed in Sect. 2.3. The analyses of hindcast simulations are described in Sect. 2.4. Finally, Sect. 2.5 summarises our results and proposes conclusions.

2.2 Models and observational data

The models' data were obtained from the CMIP5 (*Taylor et al.*, 2011) multi-model ensemble: `http://pcmdi3.llnl.gov/esgcet/home.htm`. We have analysed results of historical simulations from 24 models which have the required data available. Among these models, 10 provide results for hindcast simulations. Both historical and hindcast simulations consist in ensemble simulations

of various sizes. Historical runs finish in 2005 and we have decided not to prolong them with the RCP (Representative Concentration Pathways) simulations. Given that these latter contain less members, it would have made the analysis of the internal variability less reliable. Models and their respective modelling groups are listed in Table 2.1, along with the number of members in each model historical and hindcast simulation. The models have different spatial resolution and representation of physical processes. The spatial resolution of models' components is summarised in Table B.1 of the Appendix B. A reference is also given for more complete documentation.

We give specific information on the treatment of ozone in Table 2.2 as a basis for the discussion presented in Sect. 2.3.3. The AC&C/SPARC ozone database (*Cionni et al.*, 2011) is used to prescribe ozone in most of the models without interactive chemistry. In this database, stratospheric ozone for the period 1979–2009 is zonally and monthly averaged. It depends on the altitude and it takes solar variability into account. Whether they have interactive chemistry or prescribed stratospheric ozone, the 24 models analysed in this study thus take into account the stratospheric ozone depletion in their historical simulations. This is an improvement since the CMIP3 simulations. Indeed, nearly half of the CMIP3 models prescribed a constant ozone climatology (*Son et al.*, 2008). Nevertheless, some of the models have a coarse atmosphere resolution which sometimes does not encompass the whole stratosphere. In that case, processes related to the interaction between radiation and ozone as well as the exchange between the stratosphere and the troposphere may be represented rather crudely.

The hindcast simulations were initialised from a state that has been obtained through a data assimilation procedure, i.e. constrained to be close to some observed fields. There is a large panel of data assimilation methods (see Chapter 1), but most of the models involved in CMIP5 assimilate observations through a nudging. This method consists in adding to the model equations a term that slightly pulls the solution towards the observations (*Kalnay*, 2007). MIROC4h and MIROC5 incorporate observations in their data assimilation experiments by an incremental analysis update (IAU). Details about this method can be found in *Bloom et al.* (1996). Table 2.3 summarises the data assimilation method corresponding to each model as well as the variables it assimilates. The relevant documentation was not available to us for CCSM4, FGOALS-g2 and MRI-CGCM3. All the models for which we have the adequate information, except BCC-CSM1.1 and CNRM-CM5, assimilate anomalies. Those anomalies are calculated for both model and observations by subtracting their respective climatology, computed over the same reference period. Working with anomalies does not prevent model biases, but it avoids the initialisation of the model with a state which is too far from its own climatology and thus limits model drift (e.g., *Pierce et al.*, 2004; *Smith et al.*, 2007; *Troccoli and Palmer*, 2007; *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009), as discussed in Sect. 2.4.

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

Table 2.1: Model name, institute and number of members in model historical and hindcast simulations.

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

Table 2.1: Continued.

Model name	Institute ID	Modelling center	Number of members in historical	Number of members in hindcasts
MIROC5	MIROC	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology	1	6
MIROC-ESM	MIROC	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	3	–
MIROC-ESM-CHEM	MIROC	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (The University of Tokyo), and National Institute for Environmental Studies	1	–
MPI-ESM-LR	MPI-M	Max Planck Institute for Meteorology	3	10 (3 in 30-year hindcast)
MRI-CGCM3	MRI	Meteorological Research Institute	3	3
NorESM1-M	NCC	Norwegian Climate Centre	3	–

Table 2.2: Summary of atmosphere vertical resolution and stratospheric ozone representation. Models in bold are the ones with interactive chemistry, activated during the CMIP5 simulations or only activated in an offline simulation used to compute the ozone dataset prescribed in the CMIP5 simulations.

Model name	Atmosphere vertical resolution	Stratospheric ozone
BCC-CSM1.1	26 layers Top layer at 2.9 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
CanESM2	35 layers Top layer at 1 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
CCSM4	26 layers	Prescribed; Data from an offline simulation of the CAM3.5 model with a fully interactive chemistry (*Landrum et al.*, 2012).
CNRM-CM5	31 layers Top layer at 10 hPa	Interactive chemistry (*Voldoire et al.*, 2012).
CSIRO-Mk3.6.0	18 layers	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
EC-EARTH	62 layers Top layer 5 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
FGOALS-g2	26 layers	No information available to us.
FGOALS-s2	26 layers Top layer at 2.19 hPa	No information available to us.
GFDL-CM3	48 layers	Interactive chemistry (*Donner et al.*, 2011).
GFDL-ESM2M	24 layers	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
GISS-E2-R	40 layers Top layer at 0.1 hPa	Prescribed; Observational analyses of *Randel and Wu* (1999).
HadCM3	19 layers	Prescribed; Observational analyses of *Randel and Wu* (1999).
HadGEM2-CC	60 layers Top layer at 0.006 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
HadGEM2-ES	38 layers Top layer at 4 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).

Table 2.2: Continued.

Model name	Atmosphere vertical resolution	Stratospheric ozone
INM-CM4	21 layers Top layer at 10 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
IPSL-CM5A-LR	39 layers Top layer at 0.04 hPa	Prescribed; Data from an offline simulation of the LMDz-REPROBUS model (*Szopa et al.*, 2012).
IPSL-CM5A-MR	39 layers Top layer at 0.04 hPa	Prescribed; Data from an offline simulation of the LMDz-REPROBUS model (*Szopa et al.*, 2012).
MIROC4h	56 layers Top layer at 40 km	Prescribed; Data from an offline simulation of *Kawase et al.* (2011).
MIROC5	40 layers Top layer at 3 hPa	Prescribed; Data from an offline simulation of *Kawase et al.* (2011).
MIROC-ESM	80 layers Top layer at 0.003 hPa	Prescribed; Data from an offline simulation of *Kawase et al.* (2011).
MIROC-ESM-CHEM	80 layers Top layer at 0.003 hPa	Interactive chemistry (*Watanabe et al.*, 2011).
MPI-ESM-LR	47 layers Top layer at 0.01 hPa	Prescribed; AC&C/SPARC ozone database (*Cionni et al.*, 2011).
MRI-CGCM3	48 layers Top layer at 0.01 hPa	Interactive chemistry (*Yukimoto et al.*, 2011).
NorESM1-M	26 layers Top layer at 2.9 hPa	No information available to us.

Table 2.3: Data assimilation methods used by the 10 models providing hindcast simulations.

Model name	Data assimilation method	References
BCC-CSM1.1	Nudging of 3-D ocean temperature (raw data).	*Gao et al.* (2012)
CCSM4	No information available to us.	
CNRM-CM5	Nudging of 3-D ocean temperature and salinity (raw data) as a function of depth and space, sea surface temperature and salinity nudging (raw data).	`ftp://ftp.cerfacs.fr/pub/globc/exchanges/cassou/Michael/Aspen_CMIP5_wrkshop_cassou_2.ppt`
FGOALS-g2	No information available to us.	
HadCM3	Nudging of 3-D ocean temperature and salinity (anomalies), nudging of 3-D atmosphere temperature and wind speed, nudging of surface pressure.	`http://www.met.reading.ac.uk/~swr06jir/presentations/JIR_dept_seminar.pptx`
IPSL-CM5A-LR	Nudging of sea surface temperature (anomalies).	*Swingedouw et al.* (2012)
MIROC4h	Incremental analysis update (IAU) of 3-D ocean temperature and salinity (anomalies).	*Chikamoto et al.* (2012)
MIROC5	Incremental analysis update (IAU) of 3-D ocean temperature and salinity (anomalies).	*Chikamoto et al.* (2012)
MPI-ESM-LR	Nudging of 3-D ocean temperature and salinity (anomalies), except in the area covered by sea ice.	*Matei et al.* (2012a)
MRI-CGCM3	No information available to us.	

2.3 Historical simulations

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

2.3.1 Mean state and variability

In a first step, we analyse the mean sea ice concentration over the period 1979–2005. Figure 2.1 shows the multi-model mean of sea ice concentration in the Southern Ocean and compares the simulated sea ice edge to the observed one. Results are given for February (September), the month during which the observed sea ice extent reaches its minimum (maximum). In February the multi-model mean underestimates the sea ice cover in the Bellingshausen and Amundsen Seas as well as in the eastern part of the Ross Sea. In the

Western Ross Sea and in small parts of the Weddell Sea and of the Indian Ocean sector, the multi-model mean overestimates the sea ice extent. In September the shape of the sea ice edge computed from multi-model mean roughly fits the observations. However, the multi-model mean overestimates the sea ice cover everywhere except in the Indian Ocean sector and in the eastern part of the Ross Sea sector.

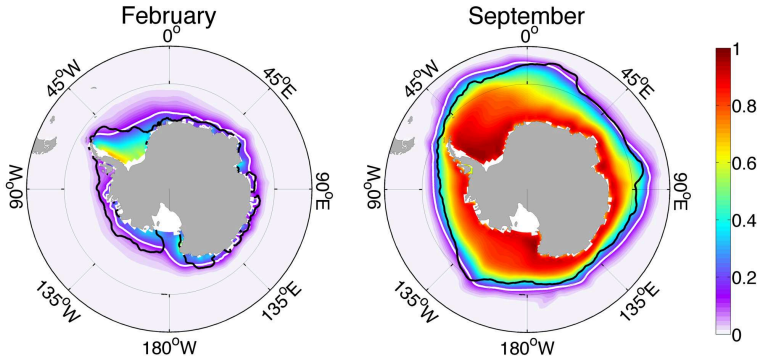

Figure 2.1: Multi-model mean of sea ice concentration, computed from historical simulations over the period 1979–2005. White (black) line refers to the sea ice edge, i.e. the 15 % concentration limit of the multi-model ensemble mean (observations, *Comiso*, 1999, updated daily).

This reasonable multi-model mean extent is the result of the average of a wide range of individual behaviours. To account for this variety of mean model states, we have plotted, for individual models, the mean of sea ice extent of each month of the year during the period 1979–2005. Figure 2.2a confirms that the multi-model mean fits quite well the observations, especially during winter months. However, the seasonal cycle of sea ice extent of the various models is largely spread around the observations and the timing of the minimum/maximum sea ice extent varies from one model to the other. In summer, 16 of the models underestimate the sea ice extent. In particular, CNRM-CM5 and MIROC5 are nearly sea ice free during summer. The latter strongly underestimates the ice extent all over the year, and its winter sea ice extent is smaller than some models' summer sea ice extent. On the contrary, CCSM4 and CSIRO-Mk3.6.0 overestimate the sea ice extent during the whole year, especially during summer. In winter, when the simulated sea ice cover reaches its maximum, the sea ice extent ranges from approximately 5×10^6 to 24×10^6 km^2, while the observations display a sea ice extent of about 17×10^6 km^2. 10 models underestimate the sea ice extent in September.

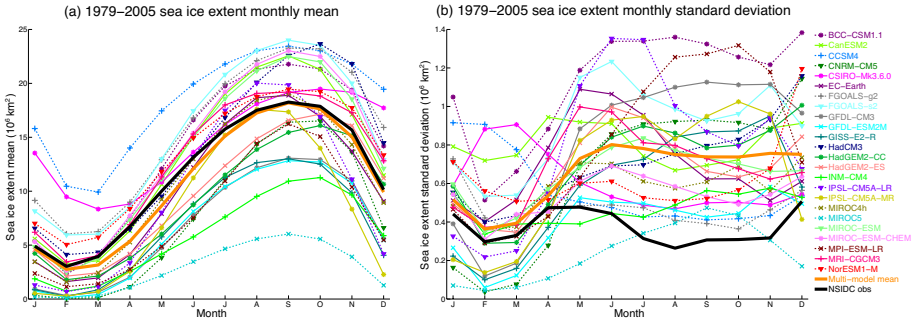

Figure 2.2: (a) Monthly mean of Antarctic sea ice extent, computed over the period 1979–2005. (b) Standard deviation of detrended Antarctic sea ice extent, computed over the period 1979–2005 for each month of the year. Colours correspond to the ensemble mean of historical simulations from 24 different models. Dotted lines refer to models that provide both historical and hindcast simulations, but here results are only from historical simulations. Orange bold line is the multi-model mean. Black bold line refers to observations (*Cavalieri and Parkinson*, 2008).

Since the internal variability of the climate system may also have played a role in the observed expansion of sea ice cover, we assess its representation in models by computing the standard deviation of the sea ice extent for each month of the year, over the period 1979–2005 (Fig. 2.2b). Here, to obtain both the ensemble mean of each model and the multi-model mean of standard deviations, an average of the individual standard deviations has been performed. We have chosen to detrend data before computing the standard deviation in order to suppress the direct impact of a trend on the standard deviation that could obscure our analysis of the potential links between those two variables discussed in Sect. 2.3.2. The monthly standard deviation indicates that the variability strongly differs between models. In February, 15 models have a standard deviation higher than the observed one, and all of the 24 models overestimate the standard deviation during September. Consequently, the multi-model mean of standard deviations does not fit very well the observations. It overestimates the standard deviation all over the year, particularly during winter. The interannual variability in some models is significantly larger during winter months than during summer months. As a result, these models have a pronounced seasonal cycle of their standard deviation, in contrast to the observations, which display a relatively constant value throughout the year. The causes of the overestimated winter variability of modelled sea ice have not been identified yet. We have performed some preliminary analyses that indicate that, for some models, changes in the oceanic convection, and thus in the entrainment modulated heat flux to the surface, could be associated to the higher winter variability (not shown). The heat supplied by the atmosphere may also contribute to

the variability in sea ice, but the relationship between the sea ice, the heat provided by the ocean and by the atmosphere is very disparate among CMIP5 models (*Close and Goosse*, 2013). However, this aspect is out of the scope of the present study and it will be addressed in future work.

The analysis of Fig. 2.2b tells us two important things. On the one hand, it points out the inability of the majority of models to reproduce the observed interannual variability. In particular, they all overestimate the winter interannual variability. On the other hand, it highlights the fact that some models are characterised by a very different magnitude of the interannual variability from one season to the other. In order to avoid a loss of information, we have thus chosen in the following analysis to work with seasonal mean rather than with annual mean and to treat the summer and winter separately.

2.3.2 Trend over the period 1979–2005

For the historical simulations, we have computed for each member of the ensemble the trend from 1979 to 2005 of summer (average of January, February and March) and winter (average of July, August and September) sea ice extent. Each trend has been computed through a linear regression of the yearly values (between 1979 and 2005) of the summer or winter sea ice extent. We have checked if the trends were significant at the 95 % level (see Table B.2 and B.3 of the Appendix B). The autocorrelation of the residuals has been taken into account in the computation of the standard deviation of each trend as well as in the number of degrees of freedom used to determine the threshold of significance, as proposed by *Santer et al.* (2000) and applied, for instance, by *Stroeve et al.* (2012). In addition to a direct evaluation of model performance, one of our goals is to analyse if a relationship can be established between the mean state, the interannual variability simulated by the model and the ability to reproduce the observed trend.

Observations show that the summer sea ice extent expanded between 1979 and 2005 at a rate of approximately 149 000 km^2 per decade. The seasonal trends of the observations are not statistically significant at the 95 % level, in contrast to the trend of the annual mean (not shown). In Fig. 2.3a it appears that almost all of the simulations performed with the 24 models fail in simulating the sign of this observed trend. Only three models (FGOALS-g2, GFDL-CM3 and GISS-E2-R) have an ensemble mean with a positive trend, while most of them simulate a relatively large negative trend. For four additional models (CCSM4, CSIRO-Mk3.6.0, HadCM3 and MRI-CGCM3), some ensemble members display a positive trend. Nevertheless, CCSM4, CSIRO-Mk3.6.0 and FGOALS-g2 have a mean summer sea ice extent much larger than what is observed, while GFDL-CM3 and GISS-E2-R are well below the observations. Moreover, CCSM4 and CSIRO-Mk3.6.0 have an interannual variability which is on average twice the one of the observations.

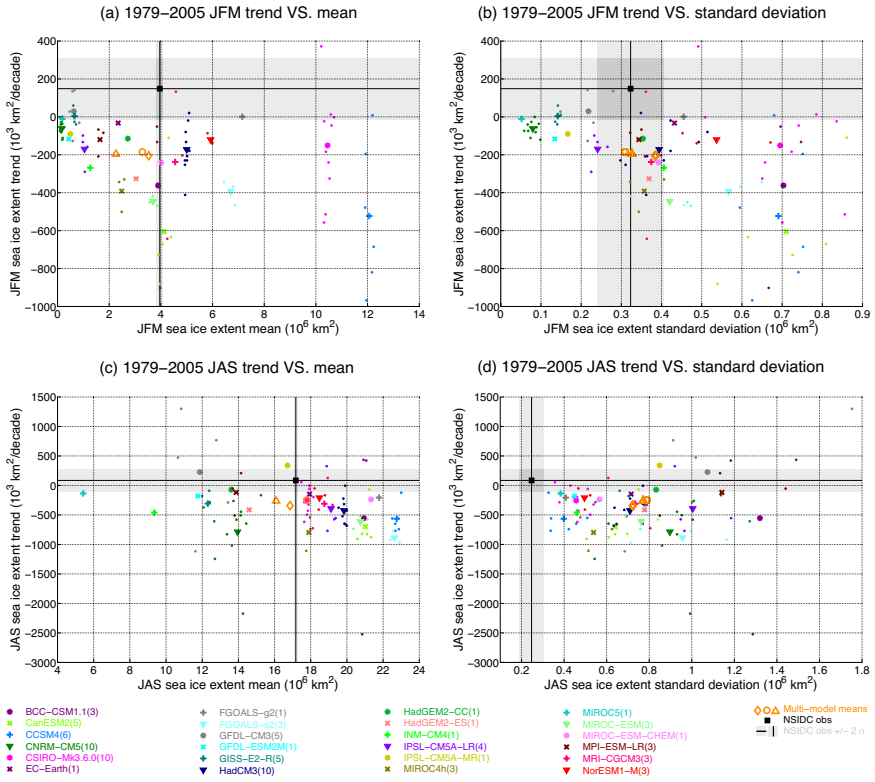

Figure 2.3: Sea ice extent trend for the period 1979–2005 over the whole Southern Ocean vs. mean (a, c) and standard deviation (b, d). The first row corresponds to summer (JFM), the second to winter (JAS). The different colours correspond to the historical simulations from 24 different models. For each colour, the small dots refer to model individual members and the symbol specified in the legend is for the model ensemble mean. The number of members in each model is indicated in brackets in the legend. Orange refers to multi-model means, for which the diamond sign is for the average over all the models, circle sign is for the mean of models with interactive chemistry (in bold in Table 2.2) and triangle sign is for the mean of models with 35 atmospheric levels or more on the vertical. Black square is for the observations (*Cavalieri and Parkinson*, 2008), surrounded by 2 standard deviations (dark-grey rectangle). Horizontal (vertical) solid black line with the light-grey shade refers to the trend (mean/standard deviation) of the observations along with 2 standard deviations. The computed standard deviation of the observed trend takes into account the autocorrelation of the residuals (see for instance *Santer et al.*, 2000; *Stroeve et al.*, 2012).

For summer sea ice extent, some given models display a standard deviation that could be quite different between members (Fig. 2.3b). Besides, the individual means of ensemble members performed with the same model are relatively similar (Fig. 2.3a). The range of values reached by the trends of the different members belonging to one model's simulation also differs strongly from one model to the other (Fig. 2.4a). We quantify the various ranges provided by the different models, thanks to the ensemble standard deviation of the trends, for models that have at least 3 members in their historical simulations. This ensemble standard deviation of the trends stands between 26 000 km^2 per decade for MIROC-ESM and 470 000 km^2 per decade for BCC-CSM1.1 (see Table B.2 of the Appendix B). On average, the ensemble standard deviation of the trend equals 166 000 km^2 per decade. If we consider this average as an estimate of the range of the trends that can be associated with internal variability, the observed positive trend of 149 000 km^2 per decade is well among the values that could be due to natural processes alone and compatible with the available ensemble of model results. Nevertheless, given that many models have an interannual variability that is much larger than the one of the observations, it is not sure whether the range of the trends they provide is representative of the reality.

The comparison between the trend, the mean extent, and standard deviation does not display any clear link in summer between those variables: some of the models that simulate an increase in the ice extent in at least one of their members overestimate the observed mean and variability, some underestimate it. Figure 2.3b also underlines the fact that models with little ice during summer often have a small interannual variability of summer sea ice extent, in agreement with results of *Goosse et al.* (2009b). Moreover, the spread of the sea ice extent trends and standard deviations of members belonging to one model ensemble grows with the mean summer sea ice extent.

Winter sea ice extent also increased between 1979 and 2005 by approximately 86 000 km^2 per decade. Two models have an ensemble mean whose trend is positive: GFDL-CM3 and IPSL-CM5A-MR (Fig. 2.3c). The ensemble mean of GFDL-CM3 (5 members) has a positive trend which is close to the observed one, but it strongly underestimates the mean winter sea ice extent. It is also an ensemble whose members are highly scattered along the trend axis, three having a positive trend (from approximately 470×10^3 to $1300 \times 10^3\,km^2\,decade^{-1}$) and two having a negative one (from approximately -290×10^3 to $-1120 \times 10^3\,km^2\,decade^{-1}$). The IPSL-CM5A-MR ensemble is made up of one member only. Its trend and its mean are both close to observations.

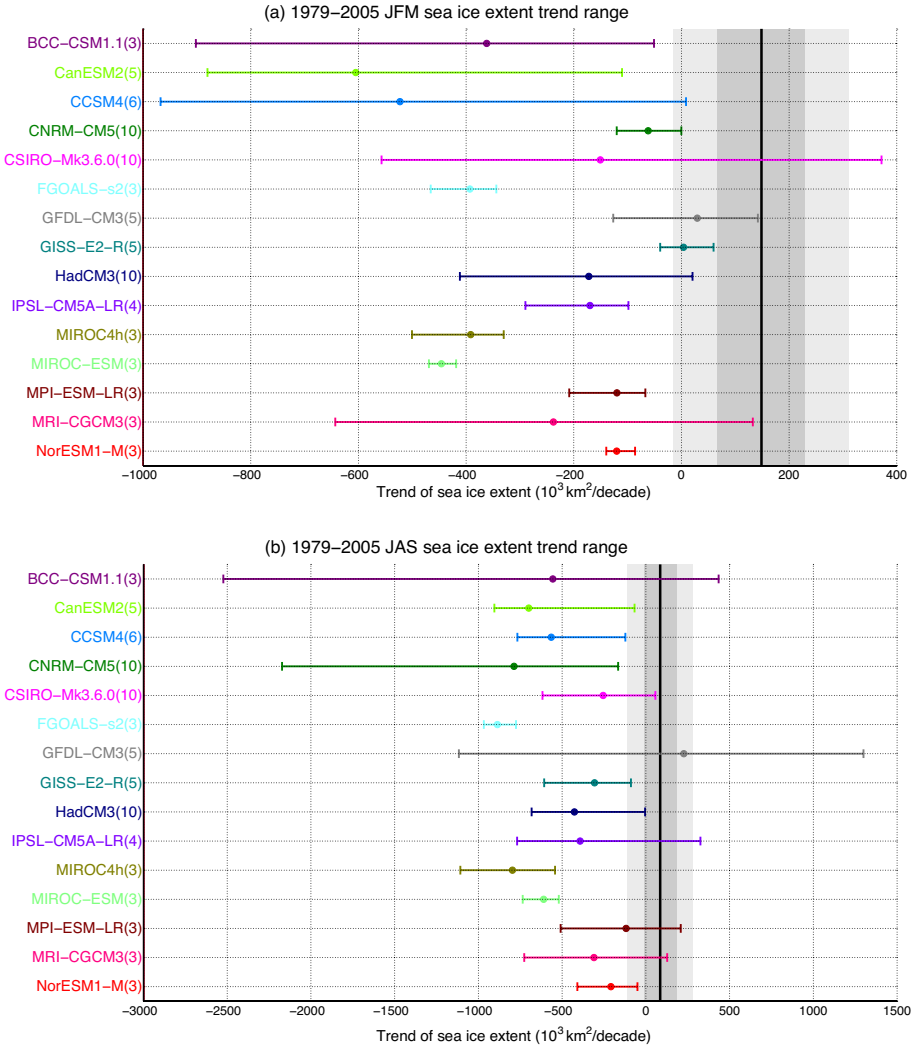

Figure 2.4: Ensemble mean, minimum and maximum values of the sea ice extent trend for the period 1979–2005 over the whole Southern Ocean for summer (a) and winter (b). The different colours correspond to the historical simulations from the 15 models that have at least 3 members in their ensemble. Dots refer to the ensemble means of the trends. Horizontal bars show the minimum and maximum values of the trend reached by the members of one model ensemble. Solid black line is for the trend of the observations (*Cavalieri and Parkinson*, 2008) surrounded by 1 standard deviation (dark grey shade) and 2 standard deviations (light grey shade). The computed standard deviation of the observed trend takes into account the autocorrelation of the residuals (see, for instance, *Santer et al.*, 2000; *Stroeve et al.*, 2012).

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

Figure 2.3d confirms that all the 24 models overestimate the interannual variability in winter. It also underlines the fact that simulations that have an ensemble mean of the trends close to the observed one have generally a standard deviation which is much larger than the one of the observations. IPSL-CM5A-MR single member, which has a trend and a mean state relatively close to the observations, has a standard deviation equals to $0.85{\times}10^6\,\mathrm{km^2}$, while the observed standard deviation stands around $0.25{\times}10^6\,\mathrm{km^2}$. GFDL-CM3 is a model that has a very high standard deviation (around 4 times the standard deviation of the observations). It is also a model with a large range of trends reached by its members (Fig. 2.4b).

For winter sea ice extent, considering again models that have at least 3 members in their historical simulations, the ensemble standard deviation of the trends varies between $100{\times}10^3\,\mathrm{km^2\,decade^{-1}}$ for FGOALS-s2 and $1\,704{\times}10^3\,\mathrm{km^2\,decade^{-1}}$ for BCC-CSM1.1 (see Table B.3 of the Appendix B). On average, this ensemble standard deviation of the trends equals $428\,000\,\mathrm{km^2\,decade^{-1}}$. As for summer, if this value is representative of the range of trends due to internal variability, the observed trend of $86\,000\,\mathrm{km^2}$ per decade appears compatible with natural processes and the model ensemble. However, the model biases in their representation of the variance in winter during the last 30 years is even larger than in summer, making this estimate of the uncertainty based on model results very questionable.

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

2.3.3 Stratospheric ozone

CMIP5 models all take into account the stratospheric ozone depletion that occurred during the last 30 years (see Table 2.2 for details). However, this improvement compared to CMIP3 brought to the stratospheric ozone does not lead to major changes in their representation of the trend in sea ice extent in the Southern Ocean.

To go a step further, we discuss if the way stratospheric ozone is treated has an influence on the results. The models with interactive chemistry (activated during the simulation or used in an offline simulation to compute the ozone dataset) and the ones with higher atmospheric vertical resolution (≥ 35 layers) have on average a slightly smaller extent of sea ice in summer (Fig. 2.3a, respectively circle and triangle orange symbols). In winter, the models with high atmospheric resolution underestimate the sea ice extent, while the ones with interactive chemistry overestimate it (Fig. 2.3c), but no specific reason for this difference has been identified. The influence on the trend is hardly detected. This shows that, on average, the inclusion of an interactive chemistry or an increased vertical resolution does not make major differences compared to other models.

Looking now at individual models, we have seen in Sect. 2.3.2 that CSIRO-Mk3.6.0, GFDL-CM3 and IPSL-CM5A-MR provide results for sea ice extent trends in winter in relatively good agreement with observations, but with much too high a standard deviation for GFDL-CM3 and IPSL-CM5A-MR. CSIRO-Mk3.6.0 has a quite coarse resolution in its atmosphere component (18 vertical layers) and prescribes the ozone from the AC&C/SPARC database. GFDL-CM3 and IPSL-CM5A-MR have a finer resolution (48 and 39 layers, respectively). They both have interactive chemistry, but IPSL-CM5A-MR treats the interaction between ozone and climate through a semi-offline approach. Again, from the available ensemble, the representation of ozone in models does not seem to be the dominant factor influencing the simulation of the trend in sea ice extent.

2.4 Hindcast simulations

We have shown in Sect. 2.3 that the lack of agreement between simulated and observed variance over the last 30 years does not allow us to confidently establish the link between the internal variability and the positive trend found in observations of the sea ice extent. Nevertheless, if this link exists and if the internal variability in the Southern Ocean is in some way predictable, an adequate initialisation of the system should improve the results of the simulated evolution of the sea ice extent. This hypothesis is tested in this section using the hindcast simulations performed in the framework of CMIP5. In contrast to the historical simulations, the hindcasts are initialised through data assimilation of

observations. The data assimilation method and the variables assimilated vary from one model to the other, as summarised in Table 2.3. Note that no bias nor drift corrections are applied on the hindcasts.

2.4.1 Impact of the initialisation on the simulated trends

The models used for the hindcast analysis have been chosen on the basis of the availability of their results. Fortunately, we see in Fig. 2.2 that these 10 models (dotted lines) constitute a subset which represents reasonably well the variety of general circulation models. In order to outline the effect of the initialisation on the simulated trend in sea ice extent for each model, we have computed the ensemble mean of the trends in hindcast simulations spanning the period 1981–2005, for winter and summer extent, and compared them to the ones from historical simulations (i.e. uninitialised) over the same time period. This period has been chosen as no hindcast was started in 1979. Here the hindcasts were initialised in January 1981 for all the models except HadCM3, whose hindcast members were started in November 1980. In Fig. 2.5, showing the trend in sea ice extent computed from hindcast simulations against the one computed from historical simulations, a dot located on the line $y(x) = x$ means that the trend in hindcast simulation equals the one of historical simulation. If the trend simulated by hindcast is greater (smaller) than the one computed from historical simulation, then the dot will be above (below) the line $y(x) = x$.

Regarding summer sea ice extent (Fig. 2.5a), the initialisation through a data assimilation procedure does not improve systematically the simulated trend. HadCM3, MIROC4h and MRI-CGCM3 hindcasts trends are closer to the observation than are their historical trends, but they remain negative. BCC-CSM1.1, CNRM-CM5, IPSL-CM5A-LR and MPI-ESM-LR simulate a more negative trend in their hindcasts than in their historical runs. FGOALS-g2 has a largely positive trend in its hindcast, while the trend in its historical simulation is slightly negative. CCSM4 hindcast displays a slightly positive trend, while the one of its historical simulation is negative.

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

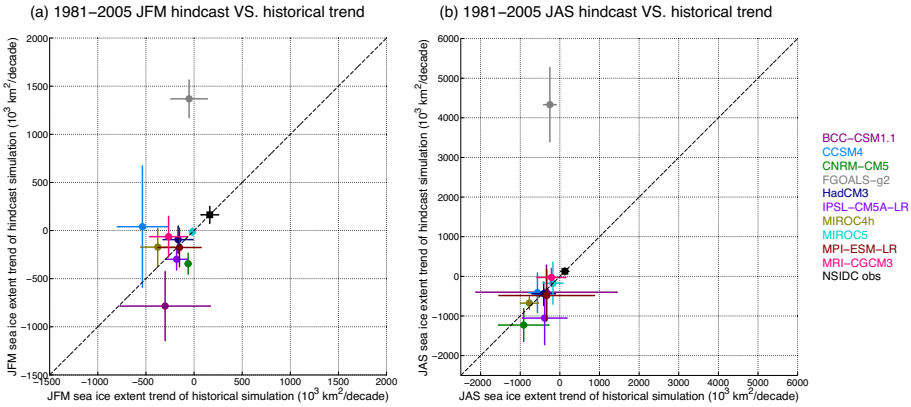

Figure 2.5: Hindcast vs. historical Antarctic sea ice extent trend for summer (a) and winter (b), computed over the period 1981–2005. The different colours refer to the different models. For each model the dot refers to the ensemble mean of the trends and the horizontal (vertical) bar shows the ensemble mean of the standard deviations of the trends in the historical (hindcast) simulations. Black square is for the trend of the observations (*Cavalieri and Parkinson*, 2008). The vertical and the horizontal black bars are for the standard deviation of the observed trend which are barely distinguishable due to their small values. Dashed line represents the line $y(x) = x$. The computed standard deviations of the trends takes into account the autocorrelation of the residuals (see, for instance, *Santer et al.*, 2000; *Stroeve et al.*, 2012).

Similarly, for winter sea ice extent, the initialisation with observations does not systematically lead to a simulated trend in better agreement with observations. Figure 2.5b shows that hindcast simulations of MIROC4h, MIROC5 and MRI-CGCM3 have trends that are slightly closer to the observation than are the historical trends. The 7 other models perform worse or do not offer any improvement when they are initialised with observations. As in the case of summer sea ice extent (Fig. 2.5a), FGOALS-g2 simulates a large positive trend in its winter sea ice extent when it is initialised with observations, and CNRM-CM5 has a more negative trend in its hindcast for the same reasons as the one proposed above. For BCC-CSM1.1, the hindcast trend in winter sea ice extent does not differ significantly from the historical trend.

Results presented in Fig. 2.5 show that the initialisation of models through data assimilation of observations does not bring significant improvement to the simulated trend. When raw data are used instead of anomalies, the initialisation apparently deteriorates the trend in sea ice extent simulated by models. As no bias nor drift corrections are applied on the hindcasts before computing the trend, their performance may be underestimated. Drift correction methods

have been proposed in several studies (e.g., *Troccoli and Palmer*, 2007; *Vannitsem and Nicolis*, 2008). Nevertheless, this is a complicated issue, in particular for sea ice, and a drift correction procedure requires a larger amount of initialised simulations spanning several decades. Proposing such a method for sea ice and assessing how it would impact the analysis of the trends is out of the scope of our study.

2.4.2 Correlation between models and observations

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

$$COR(t) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[x_{ij}(t)-\bar{x}\right]\left[o_i(t)-\bar{o}\right]}{\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{M}\left[x_{ij}(t)-\bar{x}\right]^2\sum_{i=1}^{N}M\left[o_i(t)-\bar{o}\right]^2}}, \qquad (2.1)$$

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

The correlation between hindcast simulations and observations is shown for summer (Fig. 2.6) and winter (Fig. 2.7) sea ice extents. This correlation has been computed from a series of 4 hindcast ensemble simulations, initialised every 5 years between January 1981 and January 1996 (every 5 years between November 1980 and November 1995 for HadCM3). The 95 % significance level is computed using a t-test. This significance level varies from one model to another because of the different number of members in each model ensemble (see Table 2.1).

In summer, none of the 10 models analysed here has a significant correlation for the first year after initialisation (Fig. 2.6). HadCM3, IPSL-CM5A-LR and MIROC4h never outstrip the 95 % significant level. The 7 remaining models present one or two peaks of significant correlation several years after the initialisation, and almost all the models have a negative correlation during most of the 10 years. The emergence of correlation later on in the simulation can occur randomly, or it might still be a consequence of the initialisation. Indeed, models might undergo an initial shock due to the initialisation procedure before getting stabilised and benefit from the initialisation. For winter sea ice extent (Fig. 2.7), the correlation is significantly positive during the first year for CCSM4, MIROC5 and MPI-ESM-LR models, indicating some predictive skill.

Then the correlation decreases and reaches negative values. A negative correlation is also found in the other models. The significant correlation after one year in three models in winter likely arises from the initialisation, but the memory of the system is apparently not sufficient to keep a significant correlation during the following years. Unlike in the Arctic, sea ice around Antarctica is relatively young. It disappears almost entirely during the melting season and recovers during winter months, preventing this sea ice to retain information from initialisation. The ocean can keep the information over longer periods, but in the available experiments its role appears weak during the first year after initialisation. This could be due, for instance, to a spurious variability at the beginning of the hindcast associated to the initial shock that would mask the signal provided by the ocean. Still, the ocean may be responsible for the emergence of correlation several years after initialisation, for both summer and winter sea ice extent, through local interactions or teleconnections with remote areas.

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

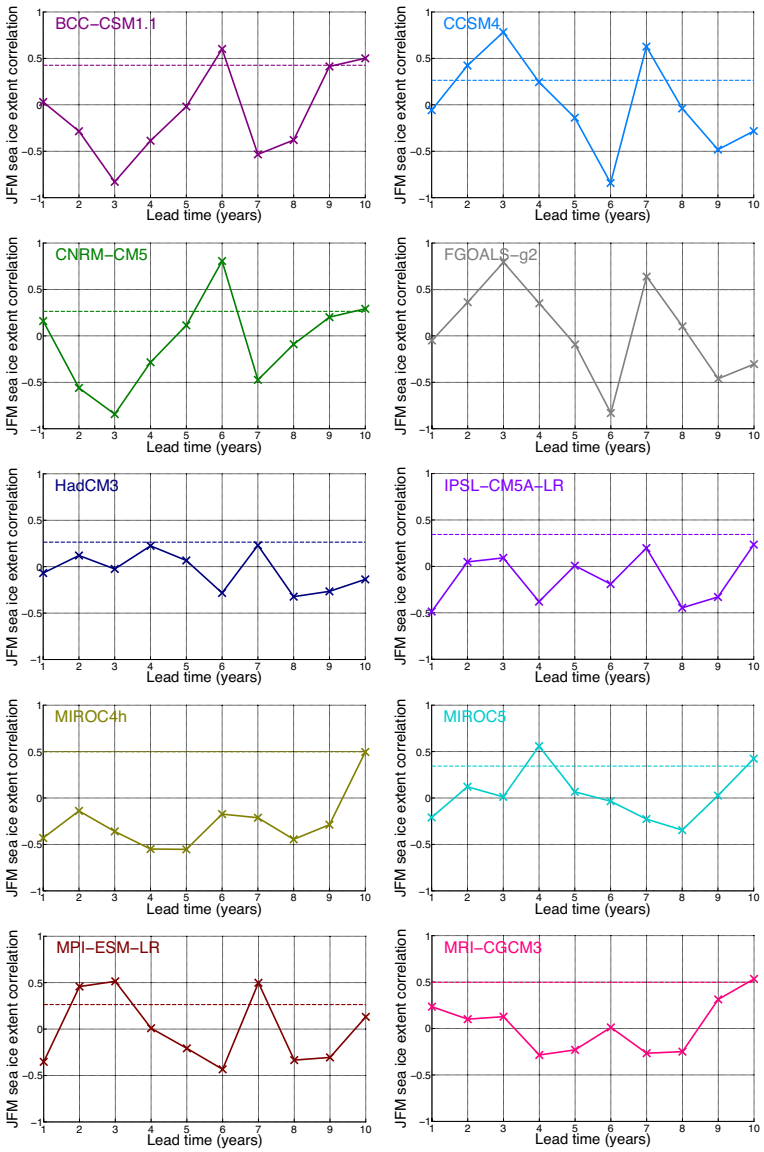

Figure 2.6: Correlation between Antarctic summer (JFM) sea ice extent in model results and observations. For each model, the correlation is computed from a series of 4 hindcast ensembles, initialised every 5 years between January 1981 and January 1996 (between November 1980 and November 1995 for HadCM3). In each plot the dashed line refers to the 95 % significance level.

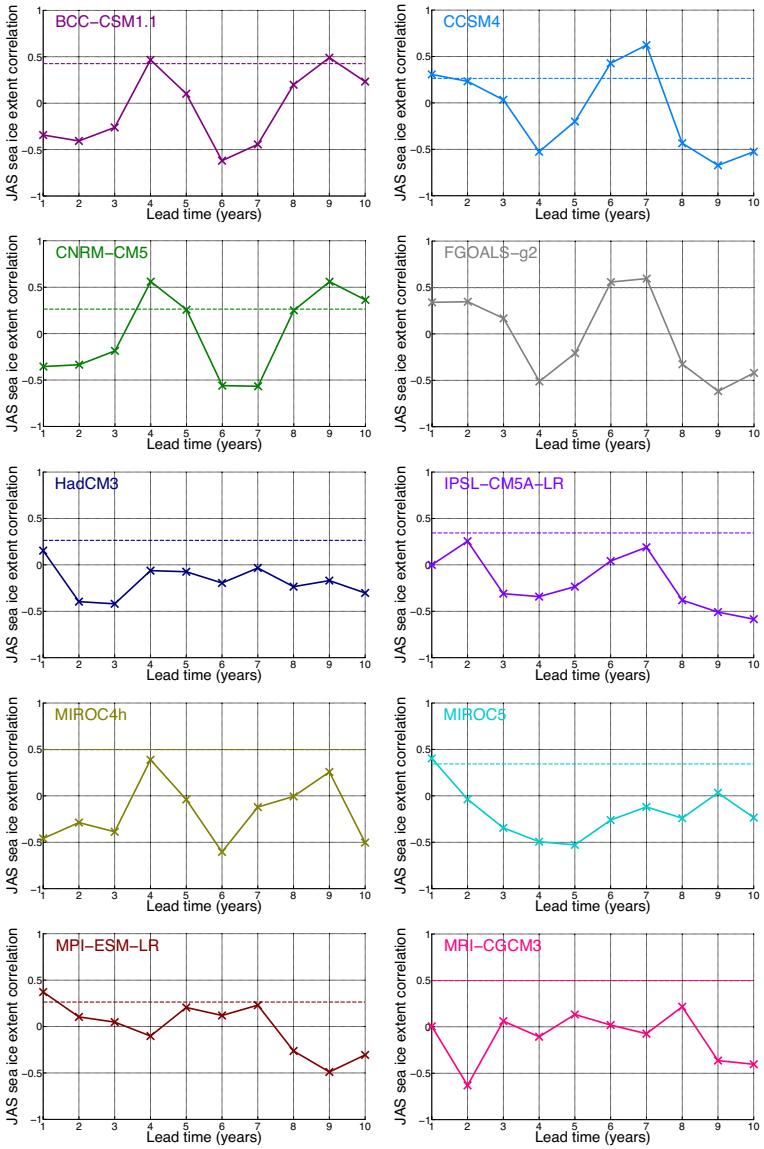

Figure 2.7: Correlation between winter (JAS) Antarctic sea ice extent in model results and observations. For each model, the correlation is computed from a series of 4 hindcast ensembles, initialised every 5 years between January 1981 and January 1996 (between November 1980 and November 1995 for HadCM3). In each plot the dashed line refers to the 95 % significance level.

2.5 Summary and conclusion

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

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

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

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

Secondly, does the models' internal variability agree with the one of the observations? From our model analysis, positive trends in sea ice extent, such as the observed one, can arise from internal variability. Nevertheless, to have confidence in this conclusion, the models' internal variability must fit the one of the observations. Unfortunately, we have shown that the models often have a climatological mean which is far from the observations, or too high an interannual variability, or even both. None of the CMIP5 models provides thus a reasonable estimate of all the main characteristics of the sea ice cover over

the last decades in the Southern Ocean, in contrast to the Arctic (e.g., *Stroeve et al.*, 2012; *Massonnet et al.*, 2012). Moreover, the few models that display an increase in sea ice extent have such a large variability that the sign of the trend is not robust. One may argue that the higher internal variability found in the models, compared to the one of the observations, is due to some transient, specific characteristics of the last decades. However, this hypothesis has not been confirmed since the mean state and the amplitude of the internal variability of the models are roughly constant over the past 150 years. Because of those models' biases, we cannot reasonably consider the results of these models as a good representation of the behaviour of the Southern Ocean sea ice. As a consequence, even if the positive observed trend in sea ice extent is compatible with the models internal variability, the biases of these models prevent us from firmly assessing the link between the internal variability in the Southern Ocean and the observed increase in sea ice extent.

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

To sum up, from an exclusive model approach, a positive trend in Antarctic sea ice extent spanning the last 30 years, though being a rare event, can be accounted for by the internal variability of the system. Nevertheless, we have shown that the models display a mean state or an interannual variability, or even both that disagree with what is observed. As a consequence, this raises the question whether we can consider these model results as reliable estimates of what happens in reality, and it affects the level of confidence one has in decadal predictions or projections of the evolution of the sea ice around Antarctica performed with those models.

CHAPTER 3

Impact of initialisation on the Antarctic sea ice predictability

This Chapter is based on the following paper: Zunz, V., H. Goosse, and S. Dubinkina (2014), Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multi-decadal timescales, *Climate Dynamics (submitted in revised version)*.

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 initialisation 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.

3.1 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*, 2008a; *Stammerjohn et al.*, 2008; *Goosse et al.*, 2009a; *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 hindcasts 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 incompatibility 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. 3.2.1 and the initialisation methods tested here are described in Sect. 3.2.2. Sect. 3.2.3 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. 3.3.1) and the multi-decadal (Sect. 3.3.2) predictions. Finally, the main results are summarised and conclusions are proposed in Sect. 3.4.

3.2 Methodology

3.2.1 Model description

The three-dimensional model LOVECLIM1.2 (*Goosse et al.*, 2010a) 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°×5.6°). The stratosphere dynamics is not represented in ECBilt2 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°×3°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. 3.1). 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).

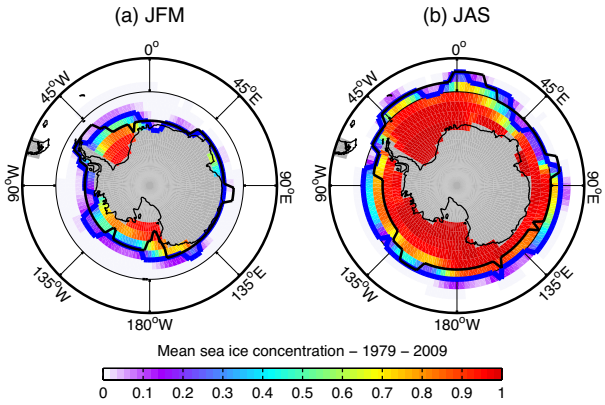

Figure 3.1: Mean sea ice concentration over the period 1979-2009 computed from a reference simulation performed with the model LOVECLIM1.2 driven by external forcing. Results are shown for (a) summer and (b) winter. The blue (black) line refers to the ice edge, i.e., the 15% concentration limit of the model simulation (observations interpolated on LOVECLIM1.2 ocean model grid, *Comiso*, 1999, updated daily).

3.2.2 Initialisation of the hindcasts

Eight initialisation methods are tested in this study. The methods are presented in this section and are summarised in Table 3.1. 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-year 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).

Table 3.1: Summary of the initialisation methods applied in the hindcasts analysed in this study. All the hindcasts are 96-member ensemble 30-year long simulations.

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_$\text{NUD}_{\text{dense}}$	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_$\text{NUD}_{\text{sparse}}$	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_$\text{SIR}_{\text{dense}}$	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_$\text{SIR}_{\text{sparse}}$	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_$\text{NPPF}_{\text{dense}}$	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_$\text{NPPF}_{\text{sparse}}$	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.

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.*, 2013a) 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 cor-

responds to an additional heat flux between the atmosphere and the ocean $Q = \gamma(T_{\mathrm{mod}} - T_{\mathrm{obs}})$. 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 3.1.

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$_{\text{dense}}$ (HIND_SIR$_{\text{sparse}}$).

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$_{\text{dense}}$ (HIND_NPPF$_{\text{sparse}}$).

3.2.3 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;

Germe et al., 2014). 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,

$$\mathrm{PPP}(t) = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{M-1}\sum_{j=1}^{M}\left[x_{ij}(t) - X_i(t)\right]^2}{\sigma^2_{\mathrm{clim}}}, \tag{3.1}$$

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^2_{clim} is the variance of a 1000-year 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. 3.3.1, the variance σ^2_{clim} 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. 3.3.1) or the pseudo-observed trends spanning several decades (Sect. 3.3.2). For the ACC, we follow the formulation of *Pohlmann et al.* (2009):

$$ACC(t) = \frac{\sum_{i=1}^{N}\left[X_i(t) - \bar{X}\right]\left[o_i(t) - \bar{o}\right]}{\sqrt{\sum_{i=1}^{N}\left[X_i(t) - \bar{X}\right]^2\sum_{i=1}^{N}\left[o_i(t) - \bar{o}\right]^2}}, \tag{3.2}$$

where t is the lead time (in years), X_i is the ensemble mean of the ith 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 across 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 100 m 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 more 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 100 m 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.

3.3 Results

3.3.1 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 predictability between the seasons.

First, we analyse the PPP and the ACC for the hindcasts HIND_perfect (Sect. 3.3.1.1). In these hindcasts, all the model variables are initialised with

values provided by the pseudo-observations (see Table 3.1). Second, the predictability provided through distinct initialisation methods is discussed for the sea ice extent (Sect. 3.3.1.2).

3.3.1.1 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. 3.2a). 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. 3.2b). 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 100 m below the surface) during the first two years of integration agrees well with this hypothesis (Fig. 3.3a). 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.

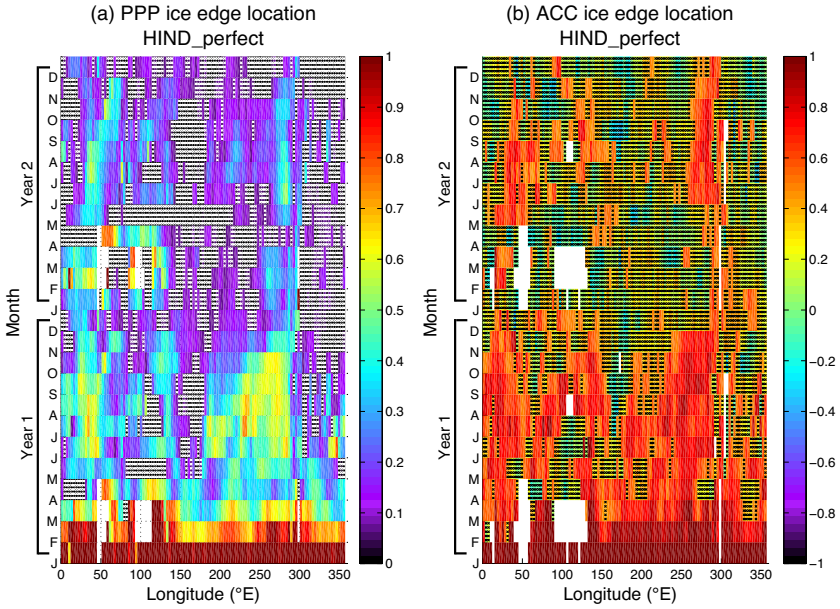

Figure 3.2: (a) Prognostic potential predictability and (b) anomaly correlation coefficient of the ice edge location computed from the hindcast HIND_perfect. The white (a) or black (b) crosses highlight the values that are not significant at the 95% level. The white areas correspond to undefined values coinciding with longitudes nearly free of sea ice during summer months.

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. 3.3b). 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. 3.3 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.

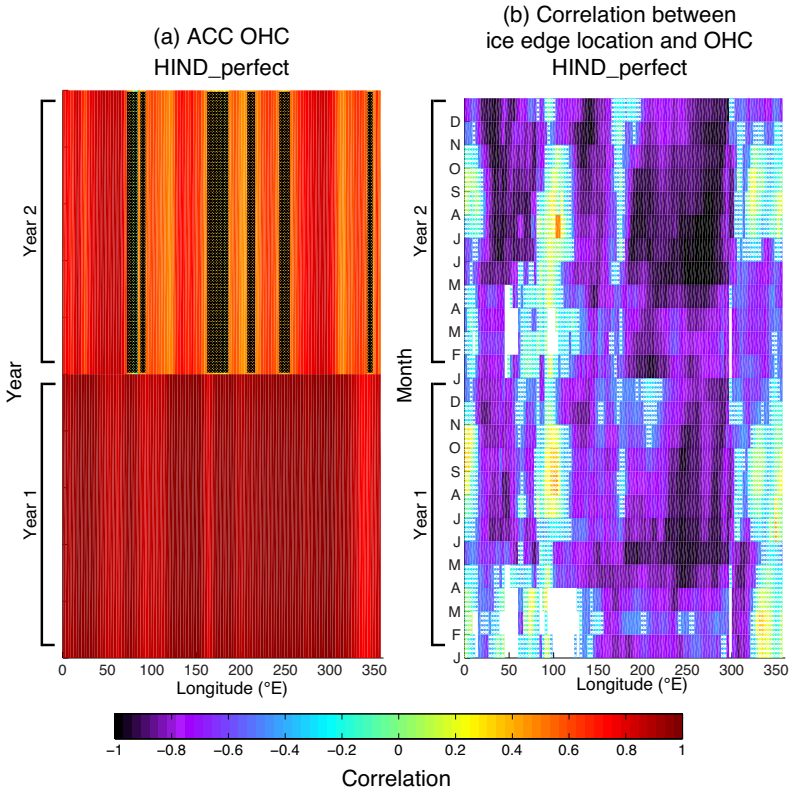

Figure 3.3: (a) ACC of the annual mean ocean heat content (OHC). (b) Correlation between the annual mean ocean heat content and the ice edge location. The ocean heat content is computed southward of 60°S and between 0 and 100 m depth. The black (a) or white (b) crosses highlight the values that are not significant at the 95% level. The white areas correspond to undefined values coinciding with longitudes nearly free of sea ice during summer months.

3.3.1.2 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$_{\text{sparse}}$ whose PPP never reaches statistically significant values. After three years of integration the PPP barely reaches significant values for any initialisation method (Fig. 3.4a). In the first year, whether the dataset used to initialise the hindcasts is dense or sparse strongly impacts the PPP, especially when the NPPF or the SIR are used to assimilate the data (blue and orange lines in Fig. 3.4a). Indeed, in the first year the values of the PPP of the hindcasts HIND_NPPF$_{\text{dense}}$ and HIND_SIR$_{\text{dense}}$ reach 0.70 and 0.65 respectively, while the PPP is much lower in the first year for the hindcasts HIND_NPPF$_{\text{sparse}}$ and HIND_SIR$_{\text{sparse}}$ (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$_{\text{dense}}$ and HIND_NUD$_{\text{sparse}}$ indicates a large spread of the ensemble, already present at the initialisation of the hindcasts, because the nudging is not applied over the sea ice covered area, which obviously reduces the constraint from the pseudo-observations on the sea ice extent. Even though the nudging term tends to maintain the members close to each other, the perturbation applied to each member every three months for the consistency of the experimental setup with the other simulations and the reduction of the assimilation domain to the sea ice free area have a large impact on the spread of the ensemble for the sea ice extent. This conclusion is confirmed by the higher values of the PPP for the hindcasts initialised with the SIR or the NPPF where sea ice covered area is included in the assimilation domain of the particle filters, which provides a stronger constraint from the pseudo-observations on the sea ice extent.

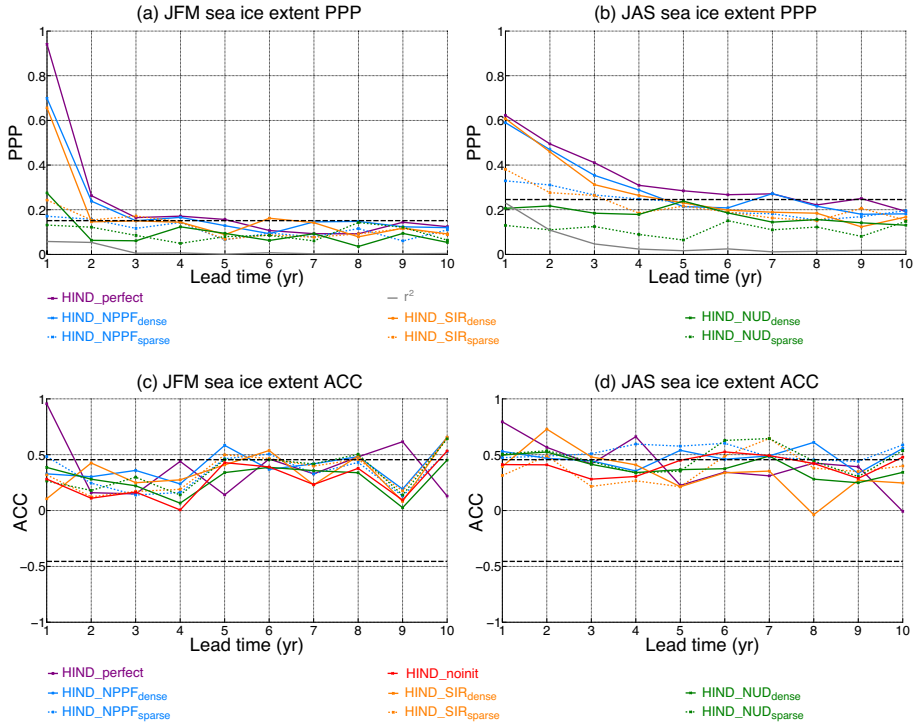

Figure 3.4: Prognostic potential predictability (a, b) and anomaly correlation coefficient (c, d) for summer (left column) and winter (right column) sea ice extent. The different colours correspond to different initialisation methods. Coloured solid lines correspond to an initialisation with dense data, while coloured 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.

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. 3.2.2, the reduction of the assimilation domain for the particle filters was required in order to avoid filter degeneracy in the simulation with assimilation. Fig. 3.4a 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. 3.4b). In the hindcasts HIND_NPPF$_{\text{dense}}$ and HIND_SIR$_{\text{dense}}$ the PPP remains significant during the first four years and in the hindcasts HIND_NPPF$_{\text{sparse}}$ and HIND_SIR$_{\text{sparse}}$ during the first three years. The hindcasts HIND_NUD$_{\text{dense}}$ and HIND_NUD$_{\text{sparse}}$ 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. 3.4a 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. 3.3.1.1.

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$_{\text{sparse}}$. 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. 3.4d), 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$_{\text{sparse}}$, outstrips the ACC of the hindcasts that were not initialised with pseudo-observations, though the improvement is weak and none of the tested initialisation 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$_{\text{dense}}$ and HIND_NUD$_{\text{sparse}}$ for which the salt content in the first 100 m 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).

3.3.2 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 initialisation. 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 initialisation 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. 3.3.1), we use the prognostic potential predictability (Eq. (3.1)) as a measure of the uncertainty of the simulated trends in sea ice. Here, the variable x in Eq. 3.1 stands for the trend in ice edge location (sea ice extent) for each ensemble member and the time t in Eq. (3.1) represents the length of the time period over which the trend is computed. The climatological variance (σ^2_{clim} in Eq. (3.1)) is the variance of the trends computed over successive time periods, spaced by 5 years, of a 1000-year 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-year trends is smaller than the number of hindcasts used in Sect. 3.3.1 (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.

3.3.2.1 Predictability of the 10- to 30-year 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. 3.5). 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. 3.5a). 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. 3.5b). 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.

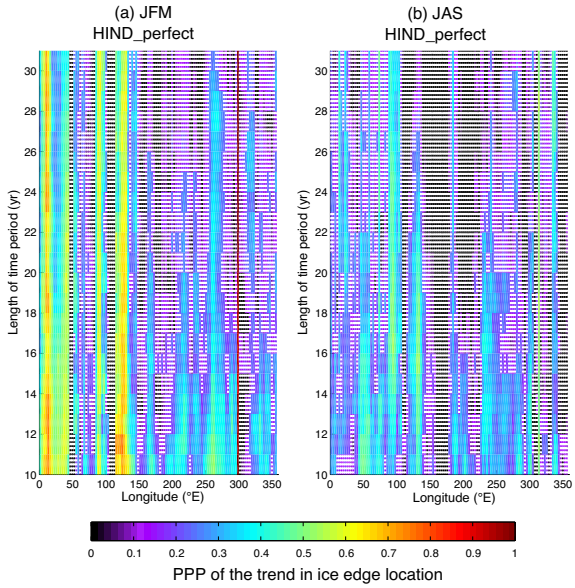

Figure 3.5: PPP of the trend for the hindcasts HIND_perfect in (a) summer and (b) winter ice edge location for increasing length of the time period over which the trends are computed. The white crosses highlight the values that are not significant at the 95% level.

The correlation between the trends in ice edge location from the hindcasts HIND_perfect and 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. 3.6). 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. 3.6a). 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. 3.6b). 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. 3.5 and 3.6. 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 uninitialised 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 $\sigma^2_{\text{clim}}(t)$ 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. 3.6c, d). In summer (winter) 8% (4%) of the values of the correlation shown on Fig. 3.6c (Fig. 3.6d) 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. 3.6a,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.

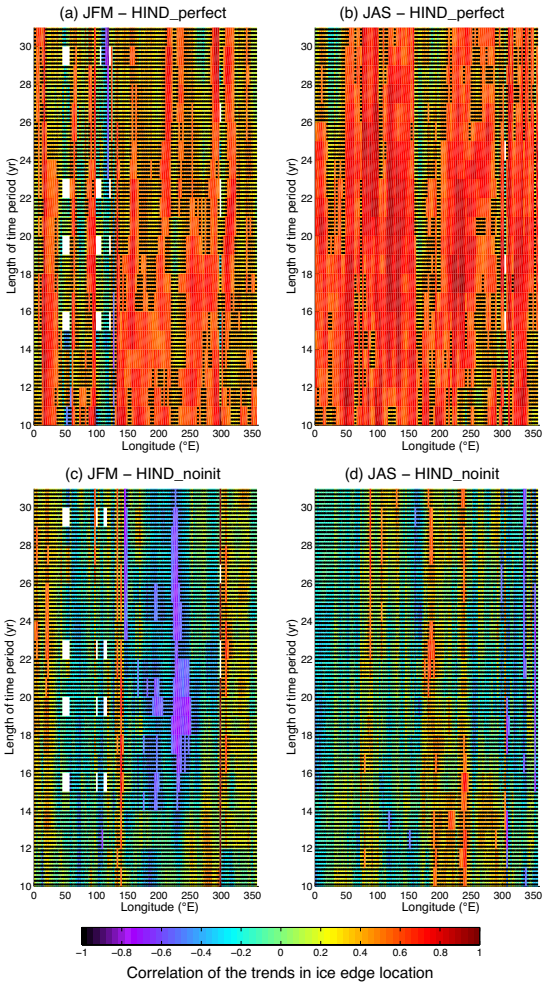

Figure 3.6: Correlation between the trend of the hindcasts ensemble mean and the trend of the pseudo-observations of the ice edge location in summer (left column) and in winter (right column) for the hindcasts HIND_perfect (a, b) and HIND_noinit (c, d). The vertical axis refers to increasing length of 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.

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. 3.3.1) the 10- to 30-year trend in ice edge location is highly correlated between the hindcasts and the pseudo-observations.

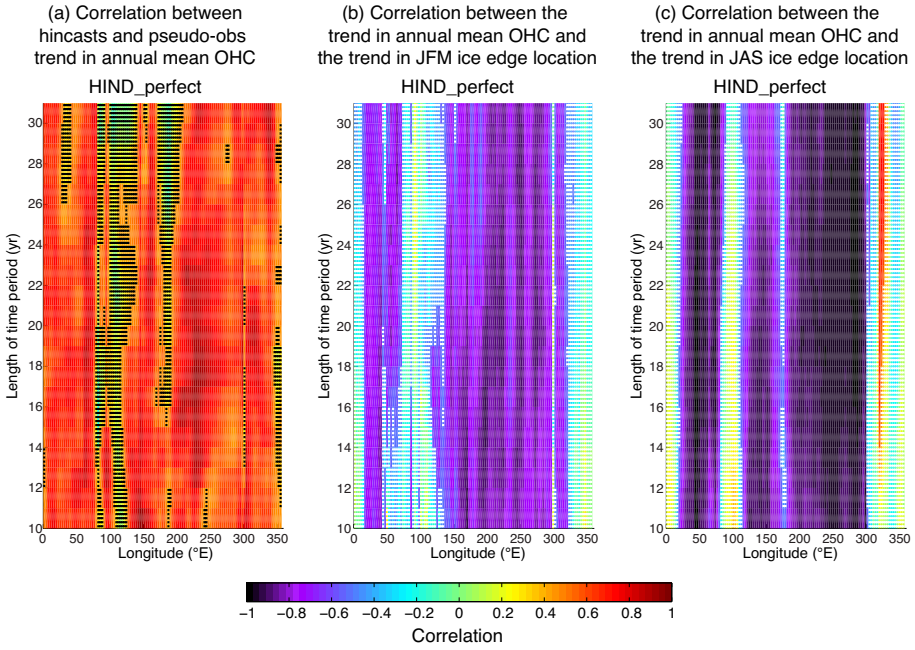

Figure 3.7: (a) Correlation between the trend in annual mean ocean heat content from the hindcasts and the one from the pseudo-observations. Correlation between the trend in summer (b) and winter (c) ice edge and the trend in annual mean ocean heat content. The vertical axis refers to increasing length 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.

As already proposed in Sect. 3.3.1 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 100 m 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 340°E (Fig. 3.7a). 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. 3.7c) with less negative values in summer (Fig. 3.7b). The same analysis performed for the ocean heat content in the upper 300 m provides results similar to the one shown here for the ocean heat content in the upper 100 m. 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. 3.6. 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. 3.6b), 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.

3.3.2.2 Predictability of the 10- to 30-year 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-year trend in ice edge location (Sect. 3.3.2.1), we have computed the prognostic potential predictability of the trend in sea ice extent (Fig. 3.8a, 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. 3.8c, d). These measures of the predictability are computed for the trends over 10- to 30-year long time period (starting on the first year after the initialisation) shown on the x-axis of Fig. 3.8 and for different initialisation methods.

As discussed in Sect. 3.3.1.2 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$_{\mathrm{dense}}$. The skill of HIND_SIR$_{\mathrm{dense}}$ is among the highest of all the experiments though HIND_SIR$_{\mathrm{dense}}$ 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-year long time period for the hindcasts HIND_perfect (purple solid lines in Fig. 3.8a, b). The hindcasts HIND_SIR$_{\text{dense}}$ (orange solid line in Fig. 3.8a, b) is above the 95% significance level, for 10- to 13-year 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$_{\text{dense}}$. 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. 3.3.1.2), 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. 3.8c, 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. 3.8c) 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$_{\text{dense}}$ 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.

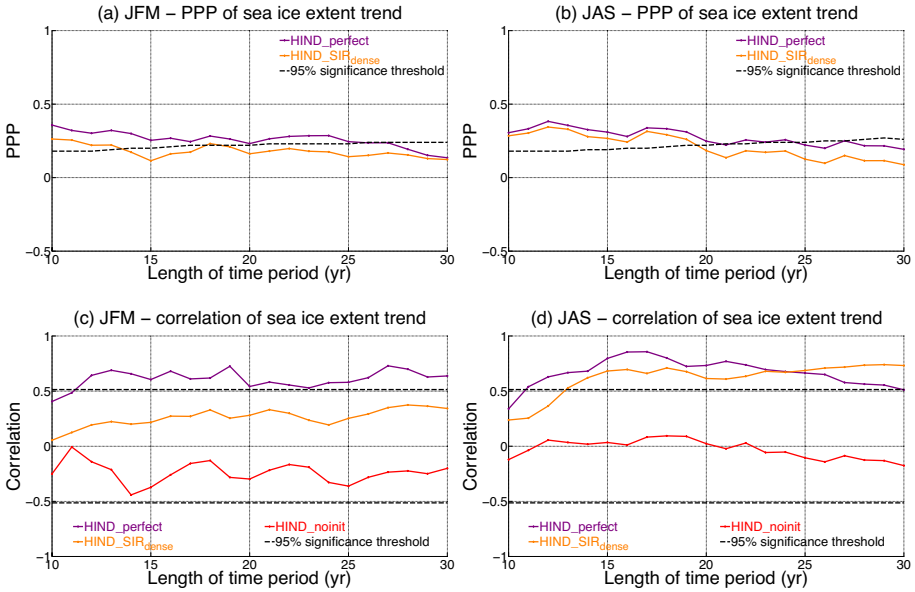

Figure 3.8: Prognostic potential predictability (a, b) and correlation with the pseudo-observations (c, d) of the trends in summer (left column) and winter (right column) sea ice extent, for increasing length of the time period over which the trends are computed. The different colours correspond to different 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. (3.1)). This autocorrelation depends on the length of the time period used to compute the trends.

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. 3.9): 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 sea 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. 3.9). 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 stationary but slightly decreases in response to the external forcing.

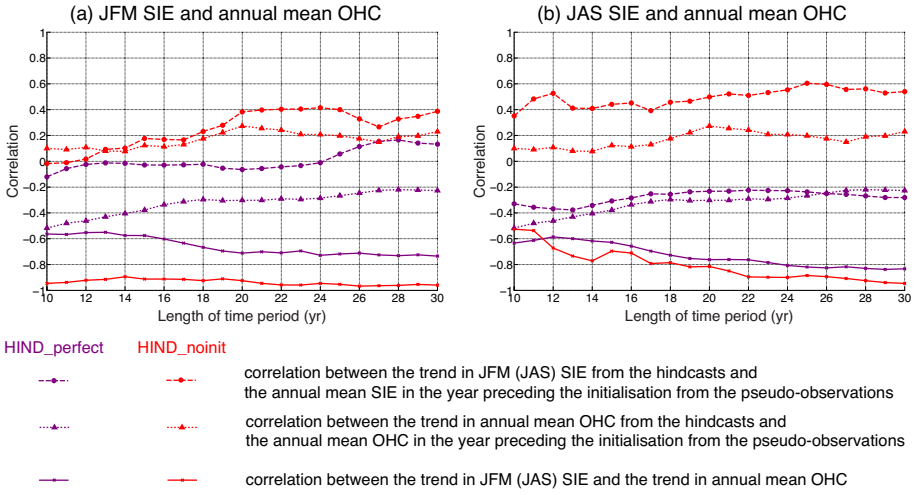

Figure 3.9: Correlation between the trends in (a) summer and (b) winter sea ice extents from the hindcasts and the annual mean sea ice extent from the pseudo-observations the year preceding the initialisation of the hindcasts (coloured dashed lines with circle markers). Correlation between the trend 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 (coloured 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 to the increasing length of the time period over which the trends are computed.

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. 3.3.2.1. 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 100 m below the surface (solid lines in Fig. 3.9), similar results being obtained with the ocean heat content in the upper 300 m (not shown). Consequently, it is essential to reproduce well this trend (Fig. 3.10) and thus the initial oceanic heat content as both are well anti-correlated (purple dotted lines with triangle markers in Fig. 3.9). 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. 3.10).

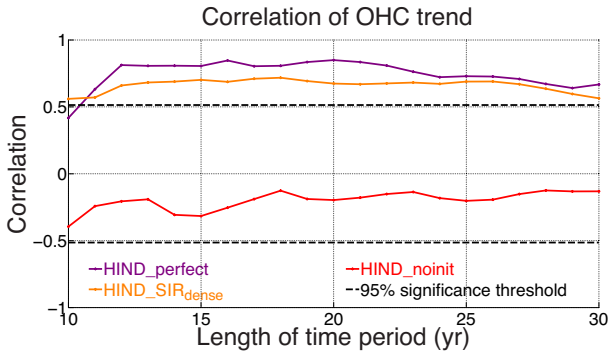

Figure 3.10: Correlation of the trends in annual mean ocean heat content between the hindcasts and the pseudo-observations, for increasing length of the time period over which the trends are computed. The ocean heat content is computed around Antarctica, southward of 60°S and between 0 and -100 m in the ocean. The dashed black lines show the 95% significance level.

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. 3.3.1. 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.

3.4 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 multi-decadal 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 variability 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 multi-decadal 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 the 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 constraints 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 initialisation 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.

CHAPTER 4

Influence of a meltwater input on the Antarctic sea ice forecast skill

This Chapter is based on the following paper: Zunz, V., and H. Goosse (2014), Influence of meltwater input on the skill of decadal forecast of sea ice in the Southern Ocean, *The Cryosphere Discussions*, *8*(4), 3563–3602, doi: 10.5194/tcd-8-3563-2014

Abstract

Recent studies have investigated the potential link between the freshwater input derived from the melting of the Antarctic ice sheet and the observed recent increase in sea ice extent in the Southern Ocean. In this study, we assess the impact of an additional freshwater flux on the trend in sea ice extent and concentration in a simulation with data assimilation, spanning the period 1850–2009, as well as in retrospective forecasts (hindcasts) initialised in 1980. The data assimilation procedure used in this study relies on particle filtering. A random freshwater flux, driven by an autoregressive process is added to the members of an ensemble of simulation. The particle filter selects the members

that are closer to the assimilated observations and the additional freshwater flux associated to these selected members provides a reconstruction of the additional freshwater input in the Southern Ocean. In a simulation with data assimilation, including the random additional freshwater flux improves the reconstruction of the trend in ice extent and concentration between 1980 and 2009 compared to a control simulation with data assimilation not including this flux. This is partly due to a better representation of the freshwater cycle in the Southern Ocean, but the additional flux could also compensate for some model deficiencies. The additional freshwater flux diagnosed from the simulation with data assimilation is mostly negative, i.e. it increases the sea surface salinity, and it involves a downward shift in the simulated mean state of the sea ice. A hindcast initialised from this shifted state has to be forced by an additional freshwater flux with a magnitude similar to the one included in the simulation with data assimilation, i.e., a negative one, in order to avoid a model drift. This points out the importance of the experimental design that has to be consistent between the simulation used to compute the initial state and the hindcast initialised from this initial state. The hindcast including this negative additional freshwater flux, constant in time, provides trends in sea ice extent and concentration that are in satisfying agreement with satellite observations. This constitutes encouraging results for sea ice predictions in the Southern Ocean. In our simulation, no increase in meltwater flux from Antarctica with time is required and the positive trend in ice extent over the last 30 years is largely determined by the state of the system in the late 1970's.

4.1 Introduction

The sea ice extent in the Southern Ocean has been increasing at a rate of 0.13 to 0.2 million km^2 per decade between November 1978 and December 2012 (*Kirtman et al.*, 2013). However, the processes that drive this evolution and the causes of sea ice expansion are still debated. The hypothesis that the stratospheric ozone depletion (*Solomon*, 1999) could have been responsible for the increase in sea ice extent is not compatible with the results of recent analyses (e.g., *Sigmond and Fyfe*, 2010; *Bitz and Polvani*, 2012; *Smith et al.*, 2012; *Sigmond and Fyfe*, 2013). Besides, other studies have underlined the fact that the positive trend in sea ice extent could be attributed to the internal variability of the system (e.g., *Mahlstein et al.*, 2013; *Zunz et al.*, 2013; *Polvani and Smith*, 2013; *Swart and Fyfe*, 2013). Nevertheless, this explanation cannot be confirmed by present-day general circulation models (GCMs) involved in the 5th Coupled Model Intercomparison Project (CMIP5, *Taylor et al.*, 2011). Indeed, because of the biases present in those models, they often simulate a seasonal cycle or an internal variability (or both) of the Antarctic sea ice

that disagrees with what is observed (e.g., *Turner et al.*, 2013; *Zunz et al.*, 2013).

Hypotheses related to changes in the atmospheric circulation or in the ocean stratification (e.g., *Bitz et al.*, 2006; *Zhang*, 2007; *Lefebvre and Goosse*, 2008a; *Stammerjohn et al.*, 2008; *Goosse et al.*, 2009a; *Kirkman and Bitz*, 2010; *Landrum et al.*, 2012; *Holland and Kwok*, 2012; *Goosse and Zunz*, 2014; *de Lavergne et al.*, 2014) have also been proposed. In particular, a potential link between the melting of the Antarctic ice sheet, especially the ice shelves, and the formation of sea ice has been recently underlined (e.g., *Hellmer*, 2004; *Swingedouw et al.*, 2008; *Bintanja et al.*, 2013). The meltwater input from the ice sheet leads to a fresher and colder surface layer in the ocean surrounding Antarctica. As a consequence, the ocean gets more stratified and there is less interaction between the surface and the warmer and saltier interior ocean, leading to an enhanced cooling of the surface. This negative feedback could counteract the greenhouse warming and could thus contribute to the expansion of the sea ice. Estimates of the Antarctic ice sheet mass imbalance are available thanks to satellite observations and climate modelling. These estimates report an increase in the melting of the Antarctic ice sheet over the past decade, mainly coming from West Antarctica (e.g., *Rignot et al.*, 2008; *Velicogna*, 2009; *Pritchard et al.*, 2012; *Shepherd et al.*, 2012). According to *Bintanja et al.* (2013), incorporating realistic changes in the Antarctic ice sheet mass in a coupled climate model could lead to a better simulation of the evolution of the sea ice in the Southern Ocean. For past periods, this may be achieved using estimates of changes in mass balance but, for the future, this requires a comprehensive representation of the polar ice sheets in models. Besides, *Swart and Fyfe* (2013) have shown that the freshwater derived from the Antarctic ice sheet is unlikely to affect significantly the recent trend in Antarctic sea ice extent simulated by CMIP5 models, when imposing a flux whose magnitude is constrained by the observations.

In addition to the studies devoted to a better understanding of the causes of the recent variations, models are also employed to perform projections of the changes for the end of the 21st century and predictions for the next months to decades. Such predictions are generally performed using GCMs. Unfortunately, as mentioned above, current GCMs have biases that reduce the accuracy of the simulated sea ice in the Southern Ocean. In addition, taking into account observations to initialise these models, generally through simple data assimilation (DA) methods, does not improve the quality of the predictions in the Southern Ocean (*Zunz et al.*, 2013). However, two recent studies performed in a perfect model framework, i.e. using pseudo-observations provided by a reference simulation of the model instead of actual observations, underlined some predictability of the Antarctic sea ice (e.g., *Holland et al.*, 2013; *Zunz et al.*, 2014). According to these studies, at interannual timescales, the predictability is limited to a few years ahead. Besides, significant predictability is found for

the trends spanning several decades. Both studies have pointed out that the heat anomalies stored in the interior ocean could play a key role in the sea ice predictability. In particular, *Zunz et al.* (2014) have underlined a link between the skill of the prediction of the sea ice cover and the quality of the initialisation of the ocean below it.

On the basis of those results, the present study aims at identifying a procedure that could improve the quality of the predictions of the sea ice in the Southern Ocean at multi-decadal timescales. Unlike *Holland et al.* (2013) and *Zunz et al.* (2014), the results discussed here have been obtained in a realistic framework. It means that actual observations are used to initialise the model simulations as well as to assess the skill of the model. The results of *Holland et al.* (2013) and *Zunz et al.* (2013, 2014) encouraged us to focus on the prediction of the multi-decadal trends in sea ice concentration or extent rather than on its evolution at interannual timescales. Our study deals with two aspects that could influence the quality of the predicted trend in sea ice in the Southern Ocean: the initial state of the simulation and the magnitude of the freshwater input associated to the Antarctic ice sheet mass imbalance. The initialisation procedure is based on the nudging proposal particle filter (NPPF, *Dubinkina and Goosse*, 2013), a data assimilation method that requires large ensemble of simulations. Such a large amount of simulations cannot be afforded with GCMs because of their requirements in CPU time. We have thus chosen to work with an Earth system model of intermediate complexity, LOVECLIM1.3. It has a coarser resolution and a lower level of complexity than a GCM, resulting in a lower computational cost. However, it behaves similarly to the GCMs in the Southern Ocean (*Goosse and Zunz*, 2014). It thus seems relevant to use this model to study the evolution of the Antarctic sea ice.

The climate model LOVECLIM1.3 is briefly described in Sect. 4.2.1, along with a summary of the simulations performed in this study. The data assimilation method used to compute the initial conditions of the hindcast simulations is presented in Sect. 4.2.2. Section 4.2.3 explains how the additional freshwater flux is taken into account in the simulations. Details about the measurement of the model skill are given in Sect. 4.2.4. The discussion of the results is divided into three parts: the simulations with data assimilation that provide the initial states (Sect. 4.3.1), the link between the freshwater flux applied during data assimilation and the model biases (Sect. 4.3.2) and the hindcast simulations (Sect. 4.3.3). Finally, Sect. 4.4 summarises the main results and proposes conclusions.

4.2 Methodology

4.2.1 Model and simulations

The three dimensional Earth-system model of intermediate complexity LOVECLIM1.3 (*Goosse et al.*, 2010a) used here includes representations of the atmosphere (ECBilt2, *Opsteegh et al.*, 1998), the ocean and the sea ice (CLIO3, *Goosse and Fichefet*, 1999) and the vegetation (VECODE, *Brovkin et al.*, 2002). The atmospheric component is a T21 (corresponding to an horizontal resolution of about $5.6° \times 5.6°$), three-level quasi geostrophic model. ECBilt2 does not include a representation of the stratospheric dynamics and the influence of stratospheric ozone depletion thus cannot be adequately taken into account in the simulation. The oceanic component consists of an ocean general circulation model coupled to a sea ice model with horizontal resolution of $3° \times 3°$ and 20 unevenly spaced vertical levels in the ocean. The vegetation component simulates the evolution of trees, grasses and desert, with the same horizontal resolution as ECBilt2.

In a control simulation with LOVECLIM driven by constant pre-industrial greenhouse gas concentrations, *Goosse and Zunz* (2014) have identified eleven 30-year independent periods, scattered over 1000 years, that display an increase in sea ice extent similar to the observations since 1979. Trend in sea ice extent such as the observed one can thus be simulated thanks to the internal variability of the LOVECLIM model in the absence of varying external forcings. The simulations performed in the present study span the period 1850–2009 and are driven by the same natural and anthropogenic forcings (greenhouse gases increase, variations in volcanic activity, solar irradiance, orbital parameters and land use) as the ones adopted in the historical simulations performed in the framework of CMIP5 (*Taylor et al.*, 2011). Three kinds of simulations are performed and all of them consist of 96-member ensembles. First, a simulation driven by external forcing only provides a reference to measure the predictive skill of the model that can be accounted for by the external forcing alone (NODA in Table 4.1). This numerical experiment does not take into account any observation, neither in its initialisation nor during the integration. At the initialisation and every three months of simulation, the surface air temperature of each members of NODA is slightly perturbed to have an experimental design as close as possible to the simulations with data assimilation (see below). Second, simulations that assimilate observations of surface air temperature anomalies (see Sect. 4.2.2 for details) are used to reconstruct the past evolution of the system, from January 1850 to December 2009, and to provide initial conditions for hindcast simulations. Assimilating surface air temperature data rather than sea ice data allows applying the assimilation procedure over a longer time period since reliable observations of the sea ice concentration are available from 1979 onwards only. The information provided at the surface

is then propagated towards the ocean interior over a longer period in order to improve the reconstruction of ocean variables. This approach has provided satisfying reconstructions of the changes in sea ice, atmospheric and oceanic variables over the past half century (*Goosse et al.*, 2009a) in a configuration similar to the one used in the present study. Third, the hindcast simulations are initialised from a state the 1 January 1980 extracted from a simulation with data assimilation and are not constrained by the observations during the model integration.

Table 4.1: Summary of the simulations analysed in this study. All the simulations are 96-member ensembles.

Simulation	Time period	Initialisation	Data assimilation	Additional freshwater flux during the simulation
NODA	Jan 1850–Dec 2009	on 1 Jan 1850 from a transient simulation starting from an equilibrium state in 850	NO	NO
DA_NOFWF	Jan 1850–Dec 2009	on 1 Jan 1850 from a transient simulation starting from an equilibrium state in 850	YES	NO
DA_FWF	Jan 1850–Dec 2009	on 1 Jan 1850 from a transient simulation starting from an equilibrium state in 850	YES	Autoregressive FWF
HINDCAST_1	Jan 1980–Dec 2009	on 1 Jan 1980 from DA_NOFWF	NO	NO
HINDCAST_2	Jan 1980–Dec 2009	on 1 Jan 1980 from DA_FWF	NO	NO
HINDCAST_3	Jan 1980–Dec 2009	on 1 Jan 1980 from DA_FWF	NO	Ensemble mean of the FWF computed in DA_FWF between 1980 and 2009 (see Fig. 4.6).
HINDCAST_4	Jan 1980–Dec 2009	on 1 Jan 1980 from DA_FWF	NO	Ensemble mean of the FWF computed in DA_FWF, averaged over the period 1980–2009 (= −0.03 Sv).

Two simulations with data assimilation, from 1850 to 2009, are analysed here: one without additional freshwater flux (DA_NOFWF in Table 4.1) and one that is forced by an autoregressive freshwater flux described in Sect. 4.2.3 (DA_FWF in Table 4.1), representing crudely an additional freshwater input to the Southern Ocean, associated for instance to the Antarctic ice sheet mass

imbalance or to the changes in the hydrological cycle, that may not be adequately represented in the original model. The simulation DA_NOFWF provides the initial state of the first hindcast (HINDCAST_1 in Table 4.1) while the three other hindcasts (HINDCAST_2, HINDCAST_3 and HINDCAST_4 in Table 4.1) are initialised from a state extracted from DA_FWF. These three hindcasts differ from each other in the additional freshwater flux they receive during the model integration. No additional freshwater flux is applied for HINDCAST_2. HINDCAST_3 is forced by a time series resulting from the ensemble mean of the additional freshwater flux diagnosed in DA_FWF. The average over the period 1980–2009 of the ensemble mean diagnosed from DA_FWF is applied in HINDCAST_4 as a constant additional flux.

4.2.2 Data assimilation: the nudging proposal particle filter

Data assimilation consists in a combination of the model equations and the available observations, in order to provide an estimate of the state of the system as accurate as possible (*Talagrand*, 1997). The data assimilation simulations performed here provide a reconstruction of the past evolution of the climate system over the period 1850–2009. Such a long period appears necessary because of the long memory of the Southern Ocean. It allows the ocean to be dynamically consistent with the surface variables, constrained by the observations, over a wide depth range. The state of the system on 1 January 1980 is then extracted and used to initialise the hindcast. After the initialisation, the hindcast is driven by external forcing only and no observations are taken into account anymore.

In this study, observed anomalies of surface air temperature are assimilated in LOVECLIM1.3 thanks to a nudging proposal particle filter (*Dubinkina and Goosse*, 2013). The assimilated observations are from the HadCRUT3 dataset (*Brohan et al.*, 2006). This dataset has been derived from in situ land and ocean observations and provides monthly values of surface air temperature anomalies (with regard to 1961–1990) since January 1850. Model anomalies of surface air temperature are computed with regard to a reference computed over 1961–1990 as well, from a simulation driven by the external forcing only, without data assimilation and additional freshwater flux.

The NPPF is based on the particle filter with sequential importance resampling (e.g., *van Leeuwen*, 2009; *Dubinkina et al.*, 2011) that consists of three steps. First, an ensemble of simulations, the *particles*, is integrated forward in time with the model. These particles are initialised from a set of different initial conditions. Therefore, each particle represents a different solution of the model. Second, after three months of simulation, a weight is attributed to each particle of the ensemble based on the likelihood of the solution of the particle, i.e., the conditional probability of the observations given the model solution (the closer to the observations, the larger the weight). To compute this weight,

only anomalies of surface air temperature southward of 30° S are taken into account. This choice of assimilation domain prevents the degeneracy of the particle filter that would occur if the particle filter was applied globally (e.g., *Dubinkina and Goosse*, 2013). Third, the particles are resampled: the ones with small weight are eliminated, while the ones with large weight are retained and duplicated, in proportion to their weight. This way, a constant number of particles is maintained throughout the procedure. A small perturbation is applied on the duplicated particles to generate different solutions of the model and the three steps are repeated until the end of the period of interest.

In the NPPF, a nudging is applied on each particle during the model integration. It consists in adding to the model equations a term that pulls the solution towards the observations (e.g., *Kalnay*, 2007). The nudging alone, i.e. not in combination with another DA method, has been used in many recent studies on decadal predictions (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). In LOVECLIM1.3, the nudging has been implemented as an additional heat flux between the atmosphere and the ocean $Q = \gamma(T_{\mathrm{mod}} - T_{\mathrm{obs}})$. T_{mod} and T_{obs} are the monthly mean surface air temperature simulated by the model and from the observations respectively. γ determines the relaxation time and equals $120\,\mathrm{W\,m^{-2}\,K^{-1}}$, a value similar to the ones 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). The nudging is applied on every ocean grid cell, except the ones covered by sea ice and the amplitude of the nudging applied on a particle is taken into account in the computation of its weight (*Dubinkina and Goosse*, 2013). Excluding the ice covered area from the nudging procedure is a common practice (e.g., *Keenlyside et al.*, 2008; *Pohlmann et al.*, 2009; *Matei et al.*, 2012a; *Servonnat et al.*, 2014) that prevents the inclusion of spurious forcing in the case when sea ice is present in the observed state but not in the simulation to be nudged.

4.2.3 Autoregressive additional freshwater flux

As the freshwater related to the melting of the Antarctic ice sheet may contribute to the variability of the sea ice extent (e.g., *Hellmer*, 2004; *Swingedouw et al.*, 2008; *Bintanja et al.*, 2013), it appears relevant to check its impact on the data assimilation simulations as well as on the hindcasts. However, deriving the distribution of the freshwater flux from estimate of the observed Antarctic ice sheet mass imbalance is not possible for the whole period covered by our simulations, because of the lack of data. Furthermore, the configuration of the model used in our study does not allow simulating this freshwater flux in an interactive way. We have thus chosen to apply a random freshwater flux, described in term of an autoregressive process as in *Mathiot et al.* (2013), on each particle during

the data assimilation simulation DA_FWF (see Table 4.1 for details). This allows determining the most adequate value of the additional freshwater flux for the model using the NPPF. Because of this additional freshwater flux, the parameters selected to define the error covariance matrix, required to compute the weight of each particle (see *Dubinkina et al.*, 2011), are slightly modified in comparison to the values applied for these parameters in the data assimilation without additional freshwater flux (DA_NOFWF).

The freshwater flux is computed every three months, i.e. with the same frequency as the particle filtering, as follow:

$$\mathrm{FWF}(t) = \mathrm{FWF}(t-1) + 0.25\epsilon_{\mathrm{FWF}}(t-1) + \epsilon_{\mathrm{FWF}}(t) \tag{4.1}$$

where ϵ_{FWF} is a random noise following a Gaussian distribution $N(0, \sigma_{\mathrm{FWF}})$, with σ_{FWF} equal to 10 mSv. The parameters of the autoregressive process have been chosen in order to obtain a freshwater flux compatible with the estimates of the current Antarctic ice sheet mass loss.

The melting of the Antarctic ice sheet being particularly strong over West Antarctica (e.g., *Rignot et al.*, 2008; *Velicogna*, 2009; *Pritchard et al.*, 2012; *Shepherd et al.*, 2012), we have chosen to distribute uniformly the freshwater flux in the ocean between 0° and 170° W, southern of 70° S (area in blue on Fig. 4.1). Here, the distribution of the freshwater flux is thus not limited to the cells adjacent to the Antarctic shelf, unlike *Bintanja et al.* (2013); *Swart and Fyfe* (2013). This is based on the assumption that a part of the freshwater might be redistributed offshore by icebergs (e.g., *Silva et al.*, 2006) or coastal currents not well represented in a coarse-resolution model. Alternatively, we can also consider that the ice sheet mass imbalance is not the only contributor to the additional freshwater flux required by the model. For instance, variations in precipitation are also expected to impact the freshwater balance in the Southern Ocean and might not be simulated adequately by the model. In that case, including an additional freshwater flux all around Antarctica might be more appropriate, but testing the sensitivity of the model to the area over which the freshwater is included is out of the scope of the present study.

This freshwater flux increases the range of solutions reached by the particles and could randomly bring some of them closer to the observations. When a particle is picked up because of its large weight, it is duplicated and the copied particles inherit the value of the freshwater flux that possibly brought the particle close to the observations. This value keeps influencing the copied particles because the freshwater flux is autoregressive. It could thus improve the efficiency of the particle filter. Furthermore, by selecting the solutions that best fit the observations, the particle filter allows estimating the freshwater flux that is more likely to provide a state compatible with the observations.

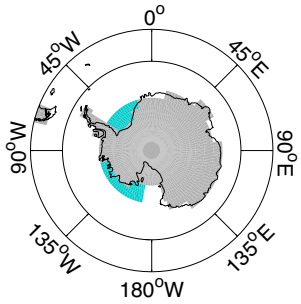

Figure 4.1: Spatial distribution of the additional freshwater flux included in model simulations.

4.2.4 Skill assessment

In order to measure the skill of the model combined with the assimilation of observations, the results of the data assimilation simulations and of the hindcasts are compared to observations of the annual mean sea ice concentration (the fraction of grid cell covered by sea ice) and sea ice extent (the sum of the areas of all grid cells having a sea ice concentration above 15 %), between 1980 and 2009. This corresponds to the period for which reliable observations of the whole ice covered area are available. The sea ice concentration data used here are derived from the Nimbus-7 SMMR and DMSP SSM/I-SSMIS satellite observations through the bootstrap algorithm (*Comiso*, 1999, updated daily). The sea ice extent is provided by the Sea Ice Index from the National Snow and Ice Data Center (*Fetterer et al.*, 2002, updated daily).

Particular attention is paid to the trend in sea ice concentration and extent. Significance levels for the trends are computed on the basis of a two-tailed t test. The autocorrelation of the residuals is taken into account in both the standard deviation of the trend and in the number of degrees of freedom used to to determine the significance threshold (e.g., *Santer et al.*, 2000; *Stroeve et al.*, 2012). This statistical test provides an estimate of the relative significance of the trend, but we have to keep in mind that the assumptions inherent to this kind of test are rarely totally satisfied in the real world (e.g., *Santer et al.*, 2000).

The ensemble means computed for the results of the data assimilation simulations consists of weighted averages. The ensemble mean $X(y,m)$ of the variable x, for the month m in the year y is thus defined as

$$X(y,m) = \frac{1}{K}\sum_{k=1}^{K} x_k(y,m).w_k(y,m), \tag{4.2}$$

where k is the member index, K is the number of members within the ensemble and $w_k(y,m)$ is the weight attributed to the member k during the data assimilation procedure. The ensemble means of each month of the year are then averaged over a year to obtain the annual mean.

The standard deviation of the annual mean of the ensemble cannot be computed explicitly because of the possible time discontinuity in the results of individual members, arising from the resampling occurring every three months. An estimate of this standard deviation is however assessed by multiplying the weighted standard deviation for each month of a year by a coefficient and averaging it over the year. These coefficients are introduced to take into account the fact that the standard deviation of the annual mean is not the mean of the standard deviation from every month. They are obtained here by computing the mean ratio between the ensemble standard deviation of the annual mean and the ensemble standard deviation of each month in the simulation NODA.

The ensemble means and standard deviations calculated for NODA and for the hindcast simulations correspond to classical values that do not include any weight as this procedure is only required when data assimilation is applied.

4.3 Results

In this section, the results of the various simulations (see Table 4.1 for details) are discussed. First, the reconstruction of the evolution of the sea ice between 1850 and 2009, provided by the simulations NODA, DA_NOFWF and DA_FWF, is presented in Sect. 4.3.1 and compared to observations. Second, the link between the model biases and the freshwater flux diagnosed from the data assimilation is presented in Sect. 4.3.2. Third, the hindcasts initialised with a state extracted from a data assimilation simulation are analysed to measure the skill of the prediction system tested in this study (Sect. 4.3.3).

4.3.1 Data assimilation simulations

The observations of yearly mean sea ice extent display a positive trend between 1980 and 2009 (Fig. 4.2) equal to $17.4 \times 10^3\,\mathrm{km^2\,yr^{-1}}$, significant at the 99 % level. This trend in sea ice extent is the result of an increase in sea ice concentration in most part of the Southern Ocean, particularly in the Ross Sea (Fig. 4.3a).

When no data assimilation is included in the model simulation (NODA), the ensemble mean displays a decreasing trend in sea ice extent in response to the external forcing (Fig. 4.2a and b), similar to the one found in other climate models (e.g., *Zunz et al.*, 2013). Consequently, for the ensemble mean, 30-year trends are negative during the whole period of the simulation without data assimilation (Fig. 4.2b). Over the period 1980–2009, the ensemble mean of the trend in sea ice extent equals $-15.5 \times 10^3\,\mathrm{km^2\,yr^{-1}}$, with an ensemble standard deviation of $14.5 \times 10^3\,\mathrm{km^2\,yr^{-1}}$, and the melting of sea ice occurs everywhere in the Southern Ocean (Fig. 4.3b), except in the Ross Sea and in the western Pacific sector. This negative trend obtained for the ensemble mean is the result of a wide range of behaviours simulated by the different members belonging to the ensemble (light green shade in Fig. 4.2a and b) and, considered individually, the members can thus provide positive or negative values for the trend. This indicates thus that, for some members, the natural variability could compensate for the negative trend in sea ice extent simulated in response to the external forcing. Positive trends similar to the one observed over the last 30 years are however rare in NODA. For instance, only 14 of the 96 members have a positive trend over the period 1980–2009 and none of them have a trend larger than the observed one.

In NODA, the ensemble mean displays an increase in the heat contained in both the upper ocean, defined here as the first 100 m below the surface, and the interior ocean, considered to lie between −100 and −500 m (green solid line in Fig. 4.4a). The correlation between the heat contained in the upper ocean and the one in the interior ocean equals 0.89 over the period 1980–2009 (Table 4.3). This warming of the ocean results directly from the increase in the external forcing and is consistent with the decrease in sea ice extent (Fig. 4.2a). Besides, the ocean salt content in the first 100 m decreases (Fig. 4.4c). This decrease in salt content is due to an increase in precipitation over the Southern Ocean (not shown), in agreement with the enhanced hydrological cycle in a global warming context that tends to freshen the ocean surface at high southern latitudes (e.g., *Liu and Curry*, 2010; *Fyfe et al.*, 2012). In the simulation NODA, the negative correlation of −0.94 between the ocean heat and salt content in the first 100 m below the surface over the period 1980–2009 (see Table 4.3) is linked to the response of these two variables to the external forcing.

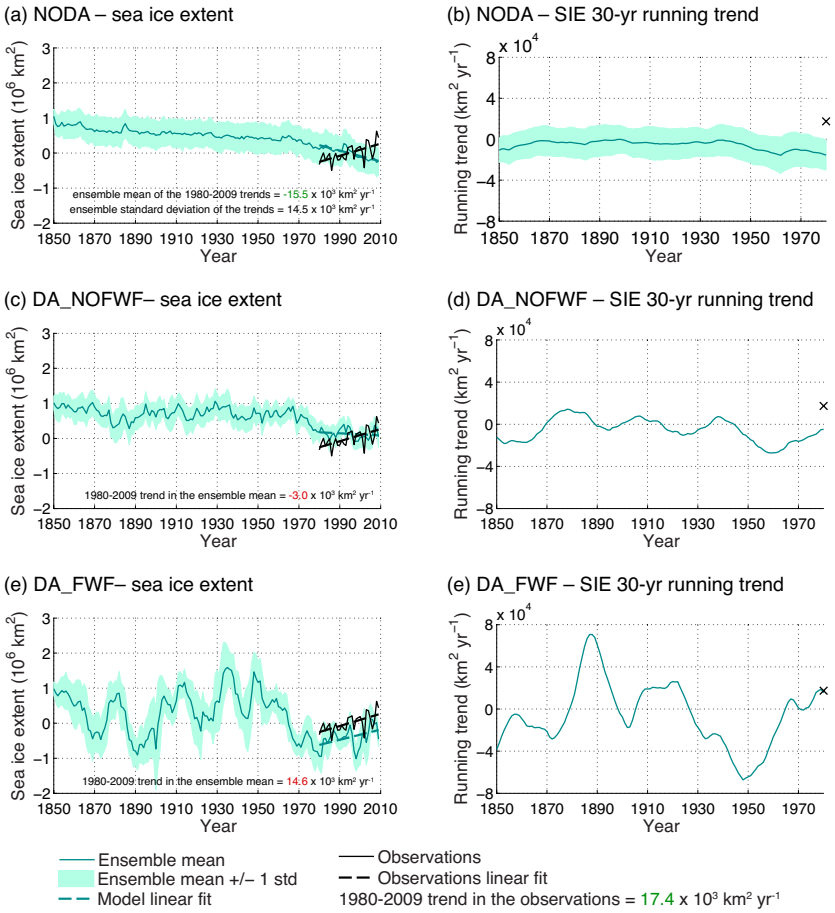

Figure 4.2: (a, c, e) Yearly mean sea ice extent anomalies with regard to 1980–2009 and (b, d, f) 30-year running trend in sea ice extent. Results are from (a, b) the simulation without data assimilation (NODA), (c, d) the model simulation that assimilates anomalies of surface air temperature (DA_NOFWF) and (e, f) the model simulation that assimilates anomalies of surface air temperature and that is forced by an additional autoregressive freshwater flux (DA_FWF). The model ensemble mean is shown as the dark green line surrounded by one standard deviation shown as the light green shade. Observations (*Fetterer et al.*, 2002, updated daily) are shown as the black line (cross) in (a, c, e) (in (b, d, f)). The green (black) dashed line shows the linear fit of the model simulation (observations) in (a, c, e). The values of the trend indicated in the (a, c and e) correspond to the ensemble mean of the trends along with the ensemble standard deviation for NODA. Trends that are (non-)significant at the 99 % level are shown in green (red).

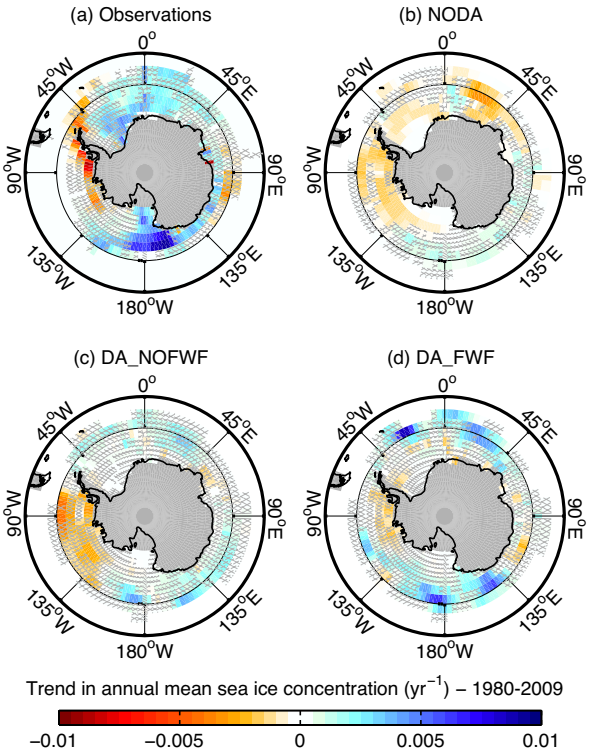

Figure 4.3: Trend in yearly mean sea ice concentration between 1980 and 2009, shown for (a) the observations (*Comiso*, 1999, updated daily), (b) the model simulation without data assimilation (NODA), (c) the model simulation that assimilates anomalies of surface air temperature (DA_NOFWF) and (d) the model simulation that assimilates anomalies of surface air temperature and that is forced by an additional autoregressive freshwater flux (DA_FWF). Hatched areas highlight the grid cells where the trend is not significant at the 99 % level. The shaded grey areas correspond to the land mask of the ocean model.

If observations of the anomalies of the surface air temperature are assimilated during the simulation, without additional freshwater flux (DA_NOFWF), the ensemble is able to capture the observed interannual and multi-decadal variability of this variable, as expected (Fig. 4.5b). Consequently, the trend in the ensemble mean sea ice extent is more variable than in NODA (Table 4.2). Over the period 1850–2009, the values of the 30-year trend in sea ice extent, computed from the ensemble mean, stand between $-29.1 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ and $13.6 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ and the standard deviation of these running trends equals $10.80 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ (Fig. 4.2d and Table 4.2). This standard deviation is

still in the range of the internal variability generated by the model since the last 1000 years of a 5000-year control simulation provides a standard deviation of the 30-year running trend equals to $13.85 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. Between 1980 and 2009, the trend in sea ice extent equals $-3.0 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. On average over the ensemble, the trend is thus less negative than in the case where no observation are taken into account during the simulation but it still has a sign opposite to the observed one. Nevertheless, the trends in sea ice concentration display a pattern roughly similar to the observed one (Fig. 4.3a and c), with an increase in the Weddell Sea, in the eastern Indian sector, in the western Pacific sector and in the Ross Sea, and a decrease elsewhere. The decrease in sea ice concentration occurring in the Bellingshausen and Amundsen Seas is, however, overestimated by the model, leading to the decrease of the overall extent.

Table 4.2: Standard deviation of the sea ice extent and of the 30-year running trend in sea ice extent computed from the ensemble mean over the time period spanned by the simulation (see Table. 4.1). The standard deviations computed from the last 1000 years of a control simulation (CNTRL) are given as reference.

Simulation	Standard deviation of the sea ice extent (10^6 km^2)	Standard deviation of the 30-year running trend (10^3 km^2 yr^{-1})
NODA	0.26	4.16
DA_NOFWF	0.29	10.80
DA_FWF	0.63	30.42
CNTRL	0.49	13.85

In the simulation DA_NOFWF, the ocean heat content in both the upper and interior ocean is lower than the ones obtained in the simulation NODA until about 1980 (Fig. 4.4a and b). This arises from the lower surface air temperature in DA_NOFWF compared to NODA (Fig. 4.5a and b) that cools down the whole system. The correlation between the upper and interior ocean heat contents equals 0.34 over the period 1980–2009 (Table 4.3) and is thus lower than in NODA. This could be due to the interannual variability captured thanks to the data assimilation that mitigates the global warming signal (see below). The ocean salt content is larger in DA_NOFWF than in NODA until 1980, likely because of the weakening of the hydrological cycle associated to the lower simulated temperature. From 1980 ahead, the ocean heat content, in both the upper and middle layers, increases and the salt content decreases in response to the external forcing, as in NODA. Nevertheless, as the ocean heat content is still slightly lower in the simulation DA_NOFWF than in the simulation NODA, the quantity of energy available to melt the sea ice is also lower. This can explain why the absolute value of the trend in sea extent between 1980 and 2009 is smaller in DA_NOFWF than in NODA.

Including a random freshwater flux following an autoregressive process in the data assimilation simulation (DA_FWF), as described in Sect. 4.2.3, increases the number of degrees of freedom of the modelled system and stimulates the variability of the ensemble mean sea ice extent, at both interannual and multi-decadal timescales (Fig. 4.2e and f). This is particularly clear before 1950, i.e. during the time period when fewer observations are available to constrain the model (*Dubinkina and Goosse*, 2013). Over the period 1850–2009, the ensemble mean of the 30-year trend in sea ice extent varies between $-68.3 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ and $70.9 \times 10^3\,\mathrm{km^2\,yr^{-1}}$, and the standard deviation of the 30-year running trend equals $30.42 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. This standard deviation is larger than the one provided by a control simulation (Table 4.2), indicating that the additional freshwater flux applied in the simulation DA_FWF excites significantly the variability of the model. Between 1980 and 2009, the average simulated trend equals $14.7 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ (not significant at the 99 % level), which is very close to the observed value of $17.4 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. In addition, the distribution of the trend in sea ice concentration, between 1980 and 2009, fits well the observations (Fig. 4.3a and d). In particular, the decrease in sea ice concentration occurring in the Bellingshausen and Amundsen seas is weaker than in DA_NOFWF and it is thus in better agreement with satellite data. We should stress here that this good match with observed trends is obtained from the constraints provided by (scarce) surface temperature observations, as no sea ice data is used in the assimilation process. Besides, the anomalies of the sea ice extent, with regard to the simulation NODA, have a mean of $-0.42 \times 10^6\,\mathrm{km^2}$ over the period 1980–2009. This shift in the mean state of the sea ice is discussed in Sect. 4.3.2.

Table 4.3: Correlation between the ocean heat content in the first 100 m below the surface and the ocean heat content between −500 m and −100 m (2nd column) and correlation between the ocean heat content and the ocean salt content in the first 100 m below the surface (3rd column), for the different simulations summarised in Table 4.1. The correlation is computed over the period 1980 and 2009, from the ensemble mean of the variables. For the simulation NODA, the correlation computed for each member of the simulation and averaged over the ensemble is given in brackets.

Simulation	Correlation between the upper and interior ocean heat contents	Correlation between the upper ocean heat and salt contents
NODA	0.89 (0.03)	−0.94 (−0.02)
DA_NOFWF	0.34	−0.28
DA_FWF	−0.84	0.78
HINDCAST_1	0.86	−0.94
HINDCAST_2	−0.92	0.89
HINDCAST_3	−0.40	0.37
HINDCAST_4	−0.23	0.29

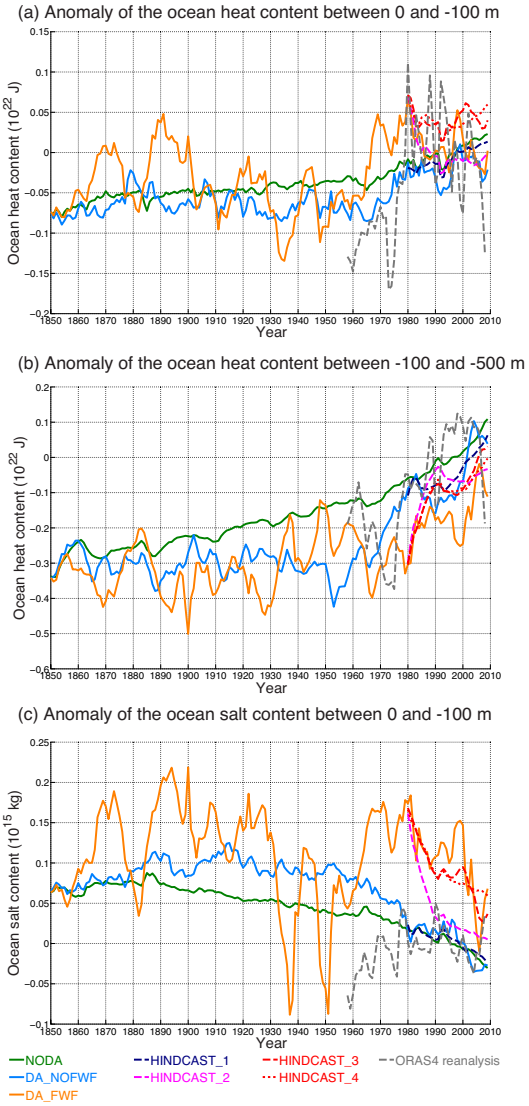

Figure 4.4: Ensemble mean of yearly mean (a) ocean heat content in the first 100 m below the surface, (b) ocean heat content between −100 and −500 m and (c) ocean salt content in the first 100 m below the surface, for the simulations summarised in Table 4.1. The ORAS4 ocean reanalyses (*Balmaseda et al.*, 2013) are shown in dashed grey for comparison. The displayed values are anomalies computed with regard to 1980-2009. The ocean heat and salt contents are computed southward of 60° S.

In DA_FWF, the correlation between the heat content in the upper ocean (Fig. 4.4a) and the one in the interior ocean (Fig. 4.4b) has a negative value of −0.84 over the period 1980–2009 (see Table 4.3). This means that when the ocean surface is colder, the middle layer is warmer and vice-versa. This suggests that the evolution of the distribution of the heat content in the water column critically depends on the exchanges between the surface and the interior of the ocean, the impact of the external forcing being apparently weaker than in DA_NOFWF. The strength of the mixing depends on the difference in density between the surface and the deeper layers, which is in turn determined by the difference in temperature and salinity. In the simulation DA_FWF, the correlation between the ocean salt and heat contents in the first 100 m reaches a value of 0.78, while it is negative in NODA and in DA_NOFWF (see Table 4.3). This confirms that, during periods of increase in salt content in the upper layer, the vertical mixing in the ocean is enhanced, allowing positive heat anomalies to be transported from the interior to the upper ocean. The heat content in the first 100 m increases while the one between −100 m and −500 m decreases. On the contrary, when the salt content in the upper layer decreases, the ocean becomes more stratified, preventing the heat exchange between the surface and the interior ocean. The heat is trapped in the interior ocean that gets warmer, and the upper ocean cools down. The ocean heat and salt content in the surface layer display stronger interannual variability than in NODA and DA_FWF. A comparison with the ocean heat and salt content obtained in various ocean reanalyses dataset suggests that such strong variability also occurs in these reanalyses, but they sometimes display discontinuities in the Southern Ocean. The results provided by the ORAS4 ocean reanalysis (*Balmaseda et al.*, 2013) are shown in Fig. 4.4) as an indication, but they must be considered cautiously since their reliability southward of 60°S has not been completely assessed. It is thus not possible from this comparison to estimate which simulation appears as the most realistic.

Because of the additional freshwater flux that tends to stabilise the water column during some periods and to destabilise it in others (Fig. 4.6), the general behaviour of the ocean in the simulation DA_FWF differs from the simulation NODA and DA_NOFWF. While the latter simulations appear mainly driven by the external forcing, the interaction between the different layers in the ocean seems to be dominant in DA_FWF. In the simulation DA_FWF, the ocean heat and salt contents of the surface layer are particularly large in 1980 while the heat content between 100 and 500 m is low. This implies that the heat storage at depth is much lower in DA_FWF than in NODA. Note that the heat content of the top 500 m in DA_FWF is also lower than in NODA. After 1980, the salt content decreases until 2009 (Fig. 4.4c). This is associated with a decrease in the upper ocean heat content and a strong increase in the ocean temperature between 100 and 500 m, suggesting a reduction of the vertical ocean heat flux. This is likely responsible for the increase in sea ice extent

between 1980 and 2009 (Fig. 4.2e). In DA_FWF, the additional freshwater flux is the main cause of the variability of the stratification. Additionally, internal processes can be responsible for such changes in vertical exchanges, as discussed in details in *Goosse and Zunz* (2014), also leading to a negative correlation between the heat content in surface and intermediate layers. This explains why the correlation between those two variables is lower for the ensemble mean of DA_NOFWF than in NODA. It is also much lower in individual simulation of NODA (0.03 on average, Table 4.3) than in the ensemble mean (0.89, Table 4.3), the ensemble mean amplifying the contribution of the response to the forcing associated with high positive value.

4.3.2 Model biases and freshwater flux

Over the years 1980–2009, the model, without data assimilation, simulates too cold a surface air temperature on average southward of 30° S compared to the reference period 1961–1990, i.e. a mean anomaly over 1980–2009 of 0.06 °C in NODA against 0.13 °C in the observations. Besides, the model is much too warm before 1960. This bias is clearly reduced in the two data assimilation simulations DA_NOFWF and DA_FWF, that furthermore provide a better synchronisation between the model solutions and the observations (Fig. 4.5). Nevertheless, this bias reduction is likely achieved differently whether an additional freshwater flux is included or not during the data assimilation. If no additional freshwater flux is taken into account, the shift in the model state induced by the data assimilation procedure is partly due to the nudging and partly to the selection of the particles whose simulated temperature is closer to the observations. The sea ice simulated by a particle is then linked to the surface air temperature through the model dynamics. Adding a freshwater flux during the data assimilation process modifies the structure of the ocean that in turn impacts the sea ice formation and the temperature at the ocean surface. The covariance between the sea ice and the surface air temperature is thus modified. However, the particles are still selected on the basis of the agreement between the surface air temperature they simulate and the observed one. As a consequence, the state of the mean surface air temperature simulated in DA_NOFWF is very similar to the one in DA_FWF but the state of the sea ice differs between these two data assimilation simulation, the simulation DA_FWF displaying a lower sea ice extent over the period 1980–2009.

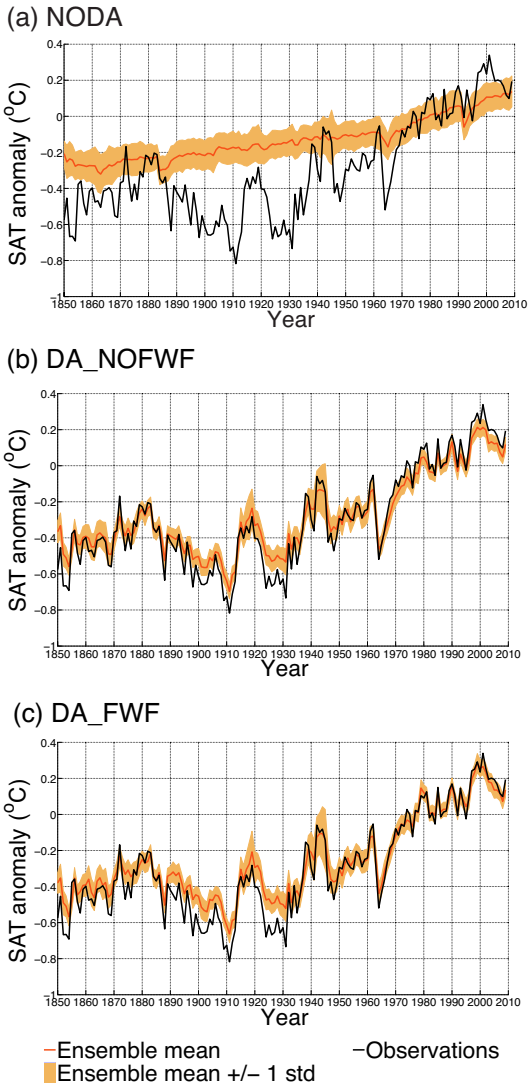

Figure 4.5: Yearly mean surface air temperature anomalies with regard to 1961–1990, averaged over the area southward of 30° S, from (a) the model simulation without data assimilation (NODA), (b) the model simulation that assimilates anomalies of surface air temperature (DA_NOFWF) and (c) the model simulation that assimilates anomalies of surface air temperature and that is forced by an additional autoregressive freshwater flux (DA_FWF). The model ensemble mean is shown as the orange line, surrounded by one standard deviation shown as the light orange shade. Observations (HadCRUT3, *Brohan et al.*, 2006) are shown as the black line. For both model results and observations, the model grid cells where no observation is available are not included in the spatial average.

On average over the period 1850–2009, the weighted ensemble mean of the additional freshwater flux diagnosed from the simulation DA_FWF is negative (−0.02 Sv, Fig. 4.6). The freshwater flux also displays a large amplitude of variation, with a standard deviation over the period 1850–2009 equal to 0.03 Sv. This standard deviation is larger than the standard deviation of the total freshwater flux southward of 30°S computed from the last 1000 years of a 5000-year control simulation, which equals 0.02 Sv. The amplitude of the variations in the additional freshwater flux in DA_FWF decreases after 1980 likely thanks to the larger amount of observations available to constrain the ensemble and thus the random additional freshwater flux associated to each member selected by the particle filter. Indeed, until 1950, the assimilated observations are scarce (*Dubinkina and Goosse*, 2013) and the freshwater flux obtained before 1950 should thus be interpreted cautiously. Over the period 1980–2009, the freshwater flux equals −0.03 Sv on average. This mean negative flux is relatively small compared to the standard deviation over the whole period but likely participates to the model bias reduction over the period 1980–2009. Indeed, a negative freshwater flux makes the ocean surface saltier and destabilises the water column. This enhances the vertical mixing and warmer water from the interior ocean reaches the surface that consequently warms up. Therefore, particles receiving a negative freshwater flux are more likely to get closer to the observations compared to the mean of NODA that is too cold over this period. They have thus a higher probability to be selected by the particle filter, reducing the model bias. We should stress that the magnitude of the diagnosed additional freshwater flux is specific to the model LOVECLIM and an other model with different biases would probably require a different correction.

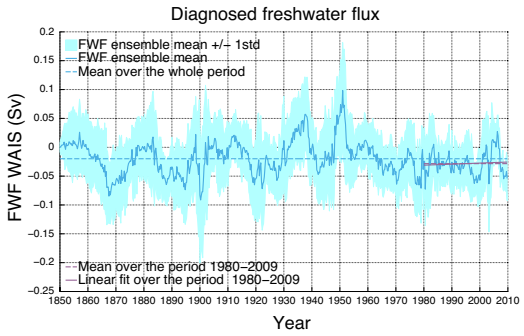

Figure 4.6: Freshwater flux from the model simulation with data assimilation and additional autoregressive freshwater flux (DA_FWF). The ensemble mean is shown as the blue solid line, surrounded by one standard deviation shown as the light blue shade. The dashed blue (purple) line shows the mean over the period 1850–2009 (1980–2009). The linear fit between 1980 and 2009 is shown as the solid purple line.

The negative value obtained for the ensemble mean of the freshwater flux between 1980 and 2009 may appear in contradiction with the estimates of the Antarctic ice sheet mass imbalance. Indeed, these clearly indicate a melting of the ice sheet that results in a freshwater input in the Southern Ocean. Nevertheless, the freshwater flux applied in our simulation allows first to compensate for model biases thanks to this negative mean value. Starting from a negative value in 1980, the ensemble mean of the freshwater flux slightly increases until 2009 at a rate of $4.53 \times 10^{-5}\,\mathrm{Sv}\,yr^{-1}$, equivalent to a change in melting rate of $42\,\mathrm{Gt\,yr^{-1}}$ between 1980 and 2009 (Fig. 4.6). This value is much smaller than the increase in freshwater flux derived from the recent estimates of the ice sheet mass imbalance but the values are only available on shorter timescales. For instance, in their reconciled estimates, *Shepherd et al.* (2012) reported a freshwater input from the West Antarctic ice sheet melting of $38 \pm 32\,\mathrm{Gt\,yr^{-1}}$ ($\simeq 10^{-3}\,\mathrm{Sv}$) over 1992–2000 and of $102 \pm 18\,\mathrm{Gt\,yr^{-1}}$ ($\simeq 3 \times 10^{-3}\,\mathrm{Sv}$) over 2005–2010. To sum up, our results show that the mean value of the additional freshwater flux does impact the simulation results by compensating for model biases but the increase in this flux may not be a determinant feature.

4.3.3 Hindcast simulations

In this section, we focus on simulations that are initialised on 1 January 1980 with a state that has been extracted from the data assimilation simulations discussed in Sect. 4.3.1. After the initialisation, the hindcast simulation is driven by external forcing but no observation is taken into account anymore. The analyses discussed here aims at answering two questions. (1) Can the information contained in the initial state persist long enough to impact the simulated trend in sea ice extent? (2) How does an additional freshwater flux impact the sea ice in hindcast simulations? Including an additional freshwater flux appears indeed to be relevant to improve the efficiency of data assimilation (see Sect. 4.3.1). The results of HINDCAST_1, initialised from DA_NOFWF, and of HINDCAST_2, initialised from DA_FWF, bring answers to the first question, these hindcasts including no additional freshwater flux. The second question is specifically addressed in the analyses of HINDCAST_3 and HINDCAST_4. Both are initialised from a state provided by the simulation DA_FWF and include an additional freshwater flux. Given that it is still not clear whether it is the mean value of the additional freshwater flux or its variations that matters, two configurations for the additional freshwater flux have been tested. In HINDCAST_3, the additional freshwater flux corresponds to the one that has been diagnosed from DA_FWF, shown on Fig. 4.6, and evolves in time. On the contrary, in HINDCAST_4, the freshwater flux is constant in time and equals $-0.03\,\mathrm{Sv}$, the average freshwater flux diagnosed in DA_FWF between 1980 and 2009.

In HINDCAST_1, the sea ice extent is high at the beginning of the simulation and decreases between 1980 and 2009 (Fig. 4.7a). The ensemble mean of the trends equals $-14.2 \times 10^3\,\mathrm{km^2\,yr^{-1}}$, with an ensemble standard deviation of $13.2 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. This provides a 95 % range that does not encompass the observed trend of $17.4 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. In this hindcast, the trend in sea ice concentration is negative over a large area in the Bellingshausen and Amundsen Seas and slightly positive elsewhere (Fig. 4.8a). This pattern thus roughly fits the observed one but the decrease obtained in the western part of the Southern Ocean covers too large an area. The regional distribution of the trend in sea ice concentration in HINDCAST_1 (Fig. 4.8a) is very similar to the one in DA_NOFWF, i.e. the simulation that provided the initial state for HINDCAST_1. This suggests that the information provided at the initialisation can impact the solution of the hindcast over multi-decadal timescales, in agreement with the results discussed in *Zunz et al.* (2014). The too large decrease in sea ice concentration occurring in the Bellingshausen and Amundsen Seas already noticed in DA_NOFWF is however amplified in HINDCAST_1, leading to an overall decrease similar to the mean of NODA.

The ocean heat and salt contents in HINDCAST_1 follow roughly the evolution of these variables for the ensemble mean in NODA (Fig. 4.4). The correlation between the upper and interior ocean heat content equals 0.86 and the correlation between the upper ocean heat and salt content equals -0.94 (see Table 4.3). This points out the role played by the external forcing in this hindcast, as discussed in Sect. 4.3.1.

By contrast, the results of HINDCAST_2 display a low sea ice extent at the beginning of the simulation (Fig. 4.7b). During the first 5 years following the initialisation, the sea ice extent rapidly increases until the solution reaches the model climatology and then remains more or less stable. Overall, the trend in sea ice extent between 1980 and 2009 computed from this hindcast has an ensemble mean equal to $19.1 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ and a standard deviation of $15.7 \times 10^3\,\mathrm{km^2\,yr^{-1}}$. The ensemble is thus shifted towards positive values of the trend in sea ice extent compared to HINDCAST_1, with an ensemble mean that is very close to the observed one. This positive value is, however, mainly due to the model drift caused by a change in the conditions of the experiment compared to DA_FWF that provided the initial state. As HINDCAST_2 is not driven by any additional freshwater flux, the sea ice extent rapidly tends to its mean climatological state in this configuration: this is the one obtained in NODA which is characterised by a higher ice extent than in DA_FWF. The model drift is also clearly seen in the ocean heat and salt contents (Fig. 4.4a and b). The regional distribution of the trend in sea ice concentration is in good agreement with the observations (Fig. 4.8b). Nevertheless, this apparent satisfying results provided by HINDCAST_2 has to be moderated given the drift that produces unrealistic trends at the beginning of the simulation.

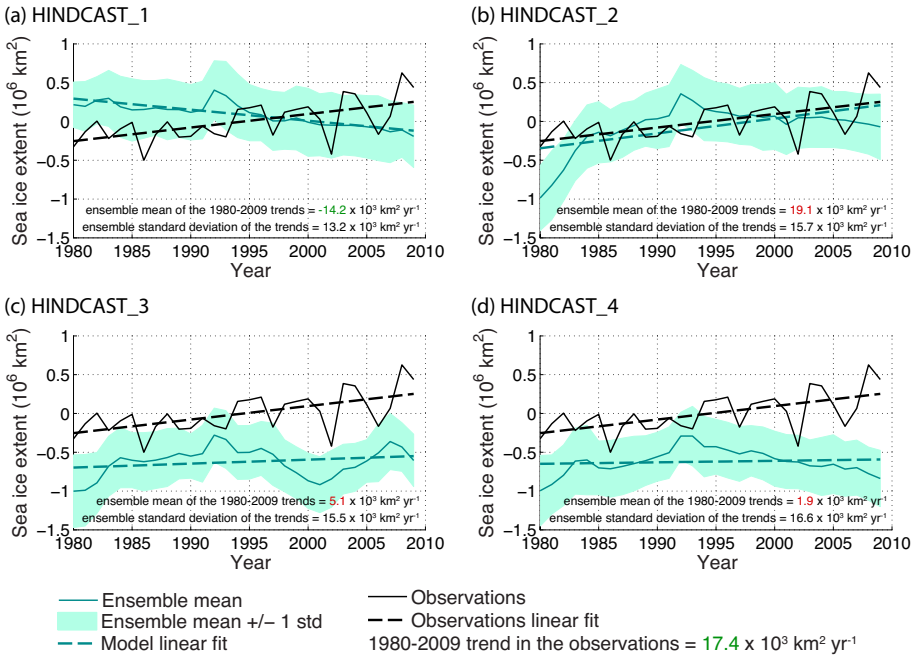

Figure 4.7: Yearly mean sea ice extent anomalies with regard to 1980–2009, for the four hindcast simulations initialised on 1 January 1980 through data assimilation (see Table 4.1 for details). The model ensemble mean is shown as the dark green line, surrounded by one standard deviation shown as the light green shade. Observations (*Fetterer et al.*, 2002, updated daily) are shown as the black line. The green (black) dashed line shows the linear fit of the model simulation (observations). The values of the trend indicated in each panel correspond to the ensemble mean of the trends along with the ensemble standard deviation. Trends that are (non-)significant at the 99 % level are shown in green (red).

In HINDCAST_3, the additional freshwater flux applied during the simulations slows down the increase in sea ice extent at the beginning of the simulation (Fig. 4.7c), resulting in a weaker trend compared to HINDCAST_2 (Fig. 4.7b). The ensemble mean (standard deviation) of the trends equals $5.1 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ ($15.5 \times 10^3\,\mathrm{km^2\,yr^{-1}}$), the observed value of $17.4 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ is thus well within the ensemble range. The trend is relatively stable over the whole 30-year period and not concentrated on the first years of simulation, as in HINDCAST_2. Furthermore, the experimental conditions are much closer to DA_FWF. There is thus no reason to suspect that the increase in sea ice extent in HINDCAST_3 is due to a spurious drift. Such a weak or even non existent

drift is ensured by the experimental design, consistent with the behaviour of the ocean heat and salt contents that remain relatively far from the results of NODA (Fig. 4.4). The pattern of the trend in sea ice concentration reasonably fits the observations (Fig. 4.8c). Including an additional freshwater flux during the hindcast simulation ensures thus a compensation of the model biases, as in the simulation DA_FWF, and avoids the generation of unrealistic trend just after the initialisation. This improves the results of the hindcast but the method applied in HINDCAST_3 requires using information spanning the period being predicted to determine the time evolution of the additional freshwater flux.

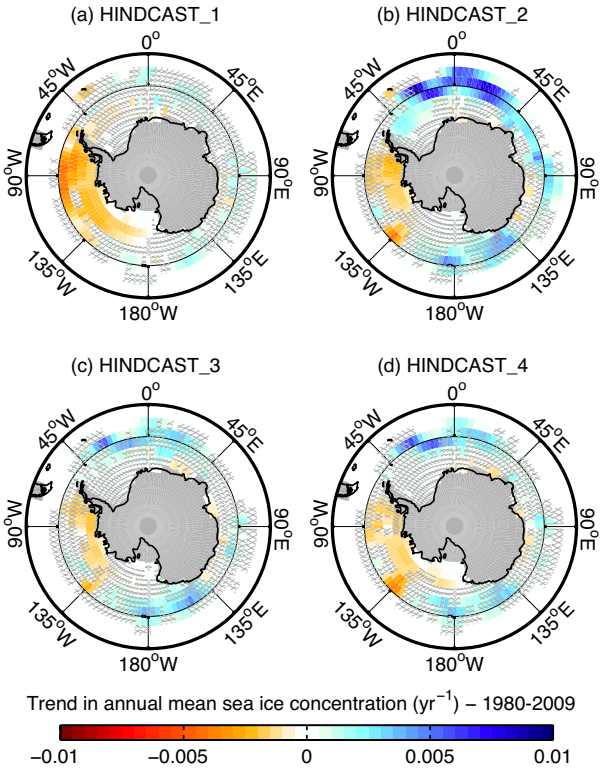

Figure 4.8: Trend in yearly mean sea ice concentration between 1980 and 2009, for the four hindcast simulations initialised on 1 January 1980 through data assimilation (see Table 4.1 for details). Hatched areas highlight the grid cells where the trend is not significant at the 99 % level. The shaded grey areas correspond to the land mask of the ocean model.

The additional freshwater flux applied during the simulation HINDCAST_4, equal to $-0.03\,\mathrm{Sv}$, corresponds to the mean of the diagnosed freshwater flux over the period 1980–2009 in DA_FWF and thus does not require a detailed knowledge of its variation in time. Note that this value is very close to the one of the period preceding the hindcast. The trend in sea ice extent in HINDCAST_4 has an ensemble mean equal to $1.9 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ and a standard deviation of $16.6 \times 10^3\,\mathrm{km^2\,yr^{-1}}$ (Fig. 4.7d). The ensemble mean of the trend is thus slightly smaller than the one of HINDCAST_3 but the ensemble still contains the observed trend. Furthermore, the sea ice extent does not display a rapid change during the first years of simulation. This suggests that the model drift is also prevented by the addition of a constant freshwater flux during the hindcast simulation. The ocean heat and salt contents stay relatively far from the model climatology (Fig. 4.4), confirming the absence of a significant model drift in HINDCAST_4. The regional distribution of the trend in sea ice concentration is in a satisfying agreement with the observed one (Fig. 4.8d). This last hindcast thus provides trends in sea ice extent and concentration that fit the observations. Forcing the hindcast with a mean value for the additional freshwater flux also allows compensating for model biases and avoids model drift. Therefore, while adding a freshwater flux is needed to maintain the sea ice of the hindcast around a mean state compatible with the initial state, a detailed knowledge of the time evolution of the freshwater flux does not seem to be crucial.

4.3.4 Link with the Southern Annular Mode

The Southern Annular Mode (SAM) is one of the dominant mode of atmospheric variability in the Southern Hemisphere. The SAM is characterised by a zonally symmetric structure of the surface pressure that implies exchanges of mass between the mid- and high latitudes (e.g., *Thompson and Wallace*, 2000). A positive (negative) phase of the SAM implies low (high) pressure over the continent and high (low) pressure in the extra-tropics and stronger (weaker) westerly winds. The SAM index is generally computed as the difference of the zonally averaged pressure between low and high southern latitudes (e.g., *Gong and Wang*, 1999). Alternatively, the time series of the principal component (PC) corresponding to the first Empirical Orthogonal Function (EOF) of the monthly anomalies of the surface pressure or of the geopotential height southward of 20°S is also used as a SAM index (e.g. *Thompson and Wallace*, 2000; *Lefebvre and Goosse*, 2008a).

In all our simulations, the first EOF mode of the monthly mean anomalies of the geopotential height at 800 hPa has an annular structure similar to the SAM pattern (see Chapter 1). Consequently, we have computed for each simulation the SAM index as the time series of the first EOF coefficients of the monthly anomalies of the geopotential height at 800 hPa southward of 20°S. The SAM

index in the simulation is compared to the SAM index computed from the NCEP-NCAR reanalyses (*Kalnay et al.*, 1996). Over the period 1980-2009, all the simulations analysed here display a positive trend in the annual mean SAM index similar to the one derived from the atmospheric reanalyses and all the trends in SAM index are statistically significant at the 95% level (Fig. 4.9 and Table 4.4).

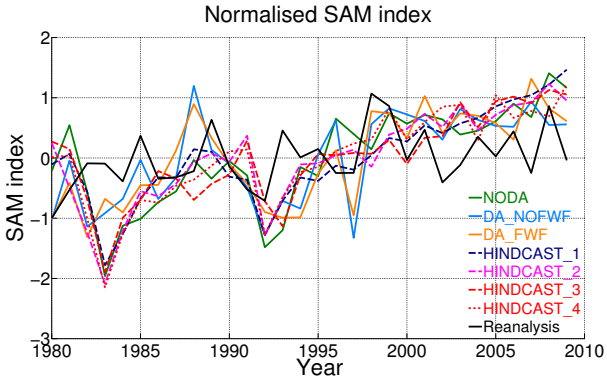

Figure 4.9: SAM index normalised over the period 1980-2009 for the simulations summarised in Table 4.1. Black line refers to the SAM index derived from NCEP-NCAR reanalyses (*Kalnay et al.*, 1996), available at `http://www.jisao.washington.edu/aao/slp`.

The correlation between the annual mean SAM index in the simulation NODA and the one in the reanalyses equals 0.28 (Table 4.4). The correlation provided by the detrended time series is much lower ($=-0.06$), suggesting that the positive trend in SAM index in both the simulation and the reanalyses account for a large part of the correlation between non-detrended data. The correlations of both non-detrended and detrended time series reach higher values for the simulations with data assimilation DA_NOFWF and DA_FWF, and are significant at the 95% level for the non-detrended time series. This suggests that the assimilation of surface temperature observations might lead to a better reconstruction of the variations in SAM. Nevertheless, as discussed in Chapter 1, the link between the SAM and the variations in sea ice concentration is not well represented in the model LOVECLIM. Therefore, it cannot be reasonably concluded that the rather good reconstruction of the trends in sea ice concentration and extent provided by the simulation DA_FWF is achieved thanks to a better representation of the SAM variations in this simulation. However, our results do not exclude a link between the behaviour of the Antarctic sea ice and the one of the atmosphere since some modes of variability not analysed here might play a role in the evolution of the sea ice cover.

In the hindcast simulations, the correlations of the non-detrended SAM index between the model results and the reanalyses are also positive but they are not statistically significant at the 95% level. The correlations for detrended time series reach negative values that are relatively large in absolute value in the hindcasts HINDCAST_2, HINDCAST_3 and HINDCAST_4. Identifying the causes responsible for these large negative values requires a deeper investigation that is out of the scope of the present study.

Table 4.4: SAM index in the various simulations compared to atmospheric reanalyses. Trend in the annual mean SAM index (2nd column). Correlation between the SAM index from the simulations and the one from the NCEP-NCAR reanalyses (*Kalnay et al.*, 1996), available at `http://www.jisao.washington.edu/aao/slp` (3rd column). The correlations are computed over the period 1980-2009 and the model results correspond to the ensemble mean of each simulation. The values in bracket correspond to the correlation obtained after detrending the time series of the SAM index of both the simulations and the reanalyses. The last line of the table displays the trend in SAM index computed from NCEP-NCAR reanalyses. Values in bold are significant at the 95% level and the autocorrelation is taken into account in the computation of the significance threshold.

Simulation	Trend in annual mean SAM index (yr^{-1})	Correlation between SAM index from the simulations and from the reanalyses	
NODA	**0.02**	0.28	(−0.06)
DA_NOFWF	**0.01**	**0.45**	(0.26)
DA_FWF	**0.02**	**0.45**	(0.23)
HINDCAST_1	**0.02**	0.32	(−0.03)
HINDCAST_2	**0.01**	0.27	(−0.13)
HINDCAST_3	**0.01**	0.15	(−0.30)
HINDCAST_4	**0.01**	0.27	(−0.12)
NCEP-NCAR reanalyses	**0.01**	–	

4.4 Summary and conclusions

In our simulations, assimilating anomalies of the surface air temperature through the nudging proposal particle filter induces an increase in the trend in simulated sea ice extent over recent decades in the Southern Ocean, compared to the case where no observation is taken into account. This leads to a better agreement with satellite data than in the simulation without data assimilation, the latter displaying a reduction of the extent in response to the external forcings. Further improvement is achieved if an additional autoregressive freshwater flux is included during the data assimilation. This freshwater flux allows compensating for model deficiencies that affect the representation of

the hydrological cycle, the meltwater input from the ice sheet mass imbalance, the ocean dynamics, the internal variability, etc.

In combination with the data assimilation, the additional freshwater input leads to simulated trends in Antarctic sea ice extent and concentration between 1980 and 2009 that agree reasonably well with the observations. The magnitude of the total freshwater flux that is injected in the Southern Ocean thus appears to play an important role on the simulated evolution of the sea ice, as already pointed out in previous studies (e.g., *Hellmer*, 2004; *Swingedouw et al.*, 2008; *Bintanja et al.*, 2013). In our simulation, an overall weaker freshwater input, compared to the one simulated by the original configuration of the model, seems to be required to reproduce the observed sea ice state over the last decades and compensates for some model biases.

While the model LOVECLIM reproduces reasonably well the increase in SAM index characterising the last decades, the interannual fluctuations of this index are generally not well simulated, although the assimilation of surface temperature slightly improves the simulation of the SAM index. Nevertheless, no robust link has been obtained between this trend in SAM index and the changes in ice concentration in our experiments.

Hindcasts initialised from the simulations with data assimilation have allowed illustrating factors that can potentially increase the model skill to predict the trend in Antarctic sea ice over the next decades. This is summarised into three points below.

1. Initialising a hindcast simulation with a state extracted from a simulation that has assimilated observations through a nudging proposal particle filter has a significant impact on the simulated trends in sea ice extent and concentration over the period 1980–2009. This indicates that the information contained in the initial state influences the results of the simulation over multi-decadal timescales, confirming the results of *Zunz et al.* (2014). As a consequence, an initial condition that adequately represents the observed state is required in order to perform skillful predictions for the trend in sea ice extent over the next decades. Nevertheless, the conclusions drawn from our hindcast simulations have to be considered cautiously since they are based on the analyses of the only 30-year period for which we have relevant observations. Similar analyses could be performed for periods starting before 1980, using the reconstruction of the sea ice provided by simulations with data assimilation as target for the hindcast instead of actual observations. However, this approach would be nearly equivalent to a perfect model study, as proposed in *Zunz et al.* (2014).

2. It has been shown that the experimental design used to perform a prediction has to be consistent with the one applied in the simulation providing the initial state for the forecast simulation. In particular, a shift in the mean state of the sea ice has been pointed out in the simulation with data assimilation and additional freshwater flux, that could lead to a model drift in the hindcast initialised from this simulation. Such a drift could be prevented if a freshwater flux of amplitude similar to the one applied during the simulation with data assimilation is included in the hindcast simulation.

3. In hindcast simulations, the variations in the freshwater flux diagnosed through the data assimilation are not needed to simulate a positive trend in ice extent over the last decades. An additional freshwater flux constant in time of adequate magnitude (negative in our case) leads to results similar to the ones obtained when the variations are taken into account. The time evolution of the additional freshwater flux, in particular its trend, has thus little impact on the hindcast simulations, in agreement with the results of *Swart and Fyfe* (2013). The strong link between the increase in freshwater input derived from the melting of the Antarctic ice sheet and the growth in sea ice extent between 1980 and 2009, suggested by *Bintanja et al.* (2013), is thus not confirmed in the present study.

Those results indicate that the increase in ice extent, the surface cooling and the freshening simulated between 1980 and 2009, in both simulations with data assimilation and hindcasts using additional freshwater flux, could not be attributed to the anthropogenical forcing or to a particular large melting of the ice sheet during this period. The evolution of the variables at the surface of the ocean after 1980 seems rather influenced by the state of the ocean in the 1970's, characterised by a warm and salty surface layer, a cold intermediate layer and a strong convection. This state is consistent with the results of *de Lavergne et al.* (2014) and evolves towards a fresher and cooler upper ocean that allows a greater production of sea ice. In our experiments, this state in the late 1970's is reached thanks to variations in the freshwater input to the Southern Ocean. This flux is very likely playing a role but the retroaction mechanisms that could involve this additional freshwater flux cannot be taken into account since our experimental design allows variations of this flux only and not of other forcings or model parameters. Whether the addition of a freshwater flux could compensate for biases in the simulated sea ice in other climate models still needs to be investigated, a reduction of the model biases being also possible through other approaches. Overall, the results that have been discussed here are rather encouraging and open perspectives to perform predictions of the sea ice in the Southern Ocean over the next decades.

CHAPTER 5

Conclusions

5.1 Main achievements

The present thesis was devoted to the assessment of the variability and the predictability of the sea ice in the Southern Ocean at decadal timescales. Possible improvements of the prediction systems for the Antarctic sea ice were investigated as well. The major difficulty we had to face was related to the very few observations available in this area. On the one hand, this lack of observation leads to a rather poor knowledge of the evolution of the climate at multi-decadal timescales in the Antarctic region, complicating the task of model validation. On the other hand, the initialisation of prediction simulations with observations is not as efficient as it might be if more observations of the state variables were available in the Southern Ocean. In the framework of the present thesis, we tried to combine at best the information provided by both the observations and the model simulations, taking into account both measurement uncertainties and model biases, in order to improve our knowledge of the evolution and the predictability of the Antarctic sea ice. Step by step, the research carried out here brought new insights, opening perspectives that determined progressively the focus of our work. The main achievements of this thesis are summarised in the following paragraphs.

Given the important role played by the internal variability of the system at these timescales, it is essential to understand the impact of the internal variability on the observed changes in the Antarctic sea ice cover. In particular,

the recent increase in sea ice extent could arise from the large multi-decadal variability that characterises the sea ice in the Southern Hemisphere. Our analyses of the results provided by the CMIP5 models (Chapter 2) have shown that the positive trend in sea ice extent observed over the last 30 years, although being rare among the CMIP5 historical ensemble, is compatible with the internal variability simulated by these models, in agreement with other recent studies (e.g., *Turner et al.*, 2013; *Polvani and Smith*, 2013). Nevertheless, the mean state and the internal variability of the Southern Ocean sea ice represented by these models are strongly biased and the conclusions drawn from these models have thus to be considered cautiously. Besides, we have shown that an adequate initialisation of a hindcast simulation, in combination with a suitable experimental design, can lead to the simulation of a positive trend in sea ice extent by an Earth system model of intermediate complexity between 1980 and 2009, while the simulated trend is negative over this period in the absence of an initialisation with observations (Chapter 4). The trend in sea ice extent simulated by this adequately initialised hindcast is positive even if the additional freshwater flux applied in the western Southern Ocean is constant in time. This hindcast simulation thus suggests that the growth in sea ice is not driven by an increase in the freshwater input. Besides, the synchronisation between the internal variability of the hindcast and the one of the observations, achieved thanks to the initialisation, seems to contribute to the simulation of a positive trend in sea ice extent. This result supports the hypothesis that the observed trend in sea ice extent over the last 30 years is due to the internal variability of the system.

In an idealised framework, the potential predictability and the forecast quality of hindcast simulations indicate that the state of the Antarctic sea ice during a particular year cannot be predicted more than a few years in advance (Chapter 3). Nevertheless, the trend in sea ice extent spanning up to three decades is potentially predictable. The predictive skill at multi-decadal timescales is not entirely accounted for by the external forcings and the initialisation of the hindcasts through the assimilation of pseudo-observations clearly improves the quality of the simulated trends. Our analyses have also allowed identifying mechanisms that potentially account for the rather good predictability of Antarctic sea ice at multi-decadal timescales. However, we have to keep in mind that this idealised study was performed with an intermediate complexity model and some important mechanisms may thus be represented rather crudely. In realistic conditions, the simulations initialised with observations, generally through a nudging, performed with the GCMs involved in CMIP5 do not provide clearly better trends compared to the simulations that are not initialised with observations (Chapter 2). However, the initialisation of a hindcast simulation through a more sophisticated data assimilation method (a nudging proposal particle filter) appears to increase the quality of the simulated trends in sea ice extent simulated between 1980 and 2009 (Chapter 4). Our results

thus suggest that, in realistic conditions, the trend in Antarctic sea ice extent at multi-decadal timescales could properly be simulated if the prediction is adequately initialised, in particular through data assimilation methods that are more sophisticated than a nudging.

5.2 Predicting sea ice in the Southern Ocean

The initialisation strategy implemented for the hindcasts discussed in Chapter 4 have allowed a clear improvement of the simulated trends in Antarctic sea ice extent and concentration over the past 30 years. This method is here applied to perform a forecast of the evolution of the sea ice in the Southern Hemisphere over the next 30 years with the model LOVECLIM1.3. This forecast is initialised on January 1, 2010, from the simulation with data assimilation and an additional freshwater flux (DA_FWF in Chapter 4). A constant additional freshwater flux of -0.03 Sv (the averaged additional freshwater flux between 1980 and 2009 in DA_FWF), evenly distributed over the area southern of 70°S and between 0°and 170°W, is applied during the forecast simulation. This forecast simulation is referred to as the *initialised forecast*. It is compared to a simulation also initialised on January 1, 2010, but from the simulation without data assimilation (NODA in Chapter 4), referred to as the *non-initialised forecast*. Both forecast simulations are driven by the external radiative forcing corresponding to the scenario RCP4.5, i.e. the intermediate emission scenario used in the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC, *Kirtman et al.*, 2013).

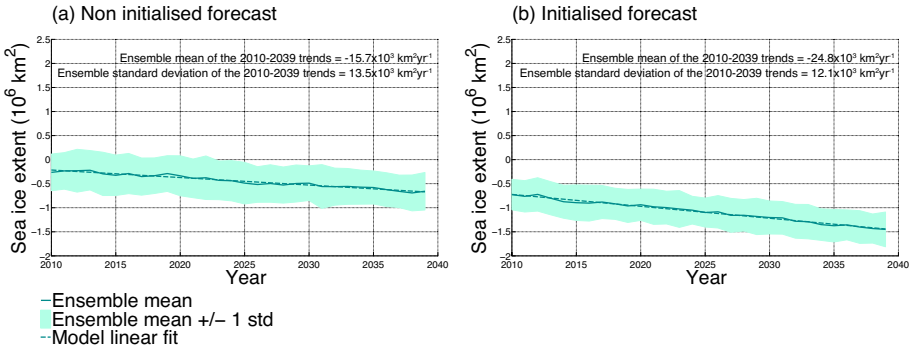

Figure 5.1: Antarctic sea ice extent anomalies with regard to 1980-2009, from (a) the non-initialised forecast and (b) the initialised forecast. The model ensemble mean is shown as the dark green line, surrounded by one standard deviation shown as the light green shade. The trends indicated in the figure are both statistically significant at the 99% level.

The non-initialised and the initialised forecasts both display a decrease in sea ice extent between 2010 and 2039 (Fig. 5.1). This decline is slightly stronger in the ensemble mean of the initialised forecast (-24.8$\times 10^3$km^2) than in the non-initialised one (-15.7$\times 10^3$km^2). The stronger decreasing trend in the initialised prediction might be accounted for by the superposition of an internally generated and a forced signals, that would both result in a decrease in sea ice extent. Indeed, the internal variability may generate a decrease in sea ice after an increase such as the one the occurring since 1979. The trends displayed on Fig. 5.1 are in the range of the changes in sea ice extent over the period 2010-2039 simulated by the models involved in CMIP5, driven by the radiative forcing corresponding to the RCP4.5 scenario (*Kirtman et al.*, 2013). Furthermore, as in CMIP5 models, the absolute rate of decline is larger for the September sea ice extent than for the February sea ice extent, for both the non-initialised and the initialised forecast (not shown). The CMIP5 database also provides initialised prediction simulations for the next decades whose results for Antarctic sea ice certainly deserve to be analysed in future work.

The trend in sea ice extent provided by the non-initialised forecast almost equals the trend over the period 1980-2009 computed from the simulation NODA (-15.5$\times 10^3$km^2), discussed in Chapter 4. Therefore, the decrease in sea ice extent in the Southern Hemisphere in response to the external forcings is not expected to accelerate during the next 30 years. In the initialised forecast, the absolute value of the trend in sea ice extent over the period 2010-2039 equals approximately 1.5 times the observed trend between 1980 and 2009 (17.4$\times 10^3$km^2). Only 5% of the members of the initialised forecast simulate a positive trend between 2010 and 2039, suggesting that the extent of sea ice that has been gained during the last 30 years is expected to be lost during the next 30 years. Nevertheless, the recent observations of the sea ice extent do not indicate any decrease in sea ice extent since 2010 and the Antarctic sea ice even settled maximum extent records in July 2013 and 2014 compared to the 1981-2010 average (`http://nsidc.org`). The beginning of the decline period may be delayed compared to the results provided by our prediction simulation.

The stronger decline in sea ice extent obtained from the initialised forecast is essentially due to a larger melting taking place in the Amundsen and Ross Seas, compared to the non-initialised forecast (Fig. 5.2). This region coincides with the one over which the additional (negative) freshwater flux is distributed. The melting of the sea ice occurring there might thus be associated to a weaker stratification of the ocean subsequent to the salinity increase at the ocean surface imposed by the additional freshwater flux. Besides, the Ross Sea is the sector of the Southern Ocean that underwent the strongest increase in sea ice concentration during the last 30 years. The decrease in sea ice concentration forecasted in this sector might thus belong to the potentially strong multi-decadal variability of the sea ice occurring there.

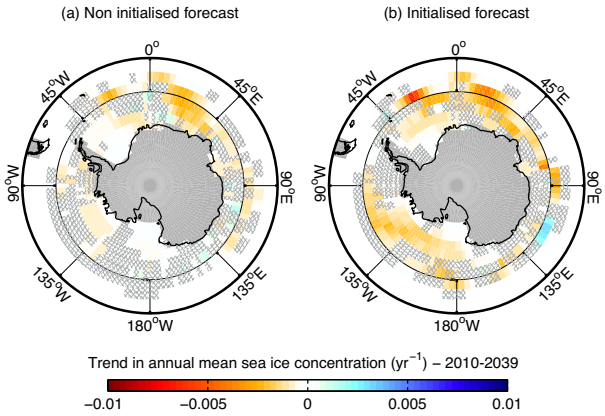

Figure 5.2: Trends in annual mean sea ice concentration between 2010 and 2039, from (a) the non-initialised forecast and (b) the initialised forecast. Hatched areas highlight the grid cells where the trend is not significant at the 99% level.

The magnitude of the additional freshwater flux applied in the initialised forecast has been determined on the basis of the results of a simulation spanning the last 160 years. This magnitude may be specific for the conditions, in particular the external forcings, characterising the period used to estimate the additional freshwater flux. It is thus not sure whether this magnitude is still adequate for the next decades. Consequently, the uncertainty on the value of the additional freshwater flux has to be kept in mind and the results of the initialised forecast have to be interpreted cautiously.

5.3 Perspectives

Any climate model is subject to imperfections due to its limited spatial and temporal resolution as well as to the parameterisations used to describe the processes that are not explicitly resolved by the model equations. The fact that all of the 24 GCMs analysed in Chapter 2 systematically overestimate the standard deviation of the winter Antarctic sea ice extent suggests that some characteristics of the sea ice in the Southern Ocean or (and) of its response to external processes are systematically misrepresented by these models (e.g., *Massonnet et al.*, 2011; *Turner et al.*, 2013). Efforts should thus focus on possible improvements that could help current climate models to simulate properly the mean state and the internal variability of the sea ice in the Southern Hemi-

sphere. In particular, given the apparently strong links between the Antarctic sea ice variability and the state of the underlying ocean (e.g., *de Lavergne et al.*, 2014), it may be useful to ensure that the processes that impact the ocean stratification are correctly represented in the models.

Despite these model biases, it has been shown that a suitable initialisation of the simulations combined with an adequate constant additional freshwater flux distributed in the Southern Ocean can improve the trends in Antarctic sea ice extent and concentration simulated by a climate model. However, this conclusion is based on the analysis of the results provided by a climate model of intermediate complexity only. In addition, the investigation of the mechanisms accounting for the predictability of the Antarctic sea ice was here essentially based on ocean processes and the role of the atmosphere, in particular its internal modes of variability having a significance variance at decadal timescales, may have been underestimated. The impact of different initialisation strategies as well as the possible approaches to compensate for model deficiencies still need to be investigated with GCMs. Furthermore, it would be relevant to test the influence of the initialisation with observations in a coupled model that includes an ice sheet component. Such a model allows simulating in a dynamic way the freshwater input derived from the Antarctic ice sheet mass imbalance and would certainly provide insightful results about the role played by the ice sheet melting in the ocean stratification.

The research carried out during this thesis has essentially focused on the quality of decadal climate predictions achieved thanks to the initialisation of the model simulations. Improvements of the forecast results might also be provided by post-processing tools that can reduce systematic biases or compensate for the drift of the model (e.g., *Meehl et al.*, 2013). Such corrections have been applied in several prediction systems (e.g., *Troccoli and Palmer*, 2007; *Doblas-Reyes et al.*, 2011; *Magnusson et al.*, 2013b; *Smith et al.*, 2013b). However, the impact of an adjustment of the (nonlinear) model drift on the prediction of the multi-decadal trends in sea ice in the Southern Ocean has not been specifically assessed yet. The methods that can correct a model drift may require a large amount of simulations with different start dates to properly estimate the correction that has to be applied. A model of intermediate complexity may thus again constitutes a useful tool to identify the most suitable post-processing strategy.

Appendices

APPENDIX A

Acronyms

ACC	Anomaly Correlation Coefficient
CMIP5	Coupled Model Intercomparison Project Phase 5
DA	Data Assimilation
EMIC	Earth-system Model of Intermediate Complexity
FWF	Freshwater Flux
GCM	General Circulation Model
NPPF	Nudging Proposal Particle Filter
NSIDC	National Snow and Ice Data Center
PF	Particle Filter
PPP	Prognostic Potential Predictability
SIC	Sea Ice Concentration
SIE	Sea Ice Extent

APPENDIX B

Supplementary material for Chapter 2

Table B.1: Summary of the 24 models used in the analysis.

Model Name	Atmospheric component	Oceanic component	Sea ice component	References
BCC-CSM1.1	BCC_AGCM2.1; 26 vertical layers, T42	MOM4_L40; 40 vertical layers, tripolar grid 1° × (1-1/3)°	SIS; tripolar grid 1° × (1-1/3)°	http://www.lasg.ac.cn/C20C/UserFiles/File/C20C-xin.pdf
CanESM2	CanAM4; 35 vertical layers, T63	CanOM4; 40 vertical layers, ∼ 1.4° ×0.9°	CanSIM1; T63 Gaussian Grid	http://www.cccma.ec.gc.ca/models
CCSM4	CAM4; 26 vertical layers, 1.25° ×0.9°	POP2; 60 vertical layers, 1.11° ×(0.27-0.54)°	CICE4; 1.11° ×(0.27-0.54)°	*Gent et al.* (2011)
CNRM-CM5	ARPEGE-CLIMAT v5.2; 31 vertical layers, T127	NEMO v3.2; 42 vertical layers, ORCA-1°	GELATO v5; ORCA-1°	*Voldoire et al.* (2012)
CSIRO-Mk3.6.0	Mk3.6 atmospheric component; 18 vertical layers, T63	Mk3.6 ocean component; 31 vertical layers, ∼ 1.875° × 0.9375 °	Sea ice sub-component of Mk3.6 (part of the AGCM); T63	*Rotstayn et al.* (2010)
EC-EARTH	IFS; 62 vertical layers, T159	NEMO v2; 42 vertical layers, ORCA-1°	LIM2; ORCA-1°	*Hazeleger et al.* (2012)
FGOALS-g2	GAMIL2; 26 vertical layers, ∼ 2.8° ×3°	LICOM2; 30 vertical layers, ∼1° × 1°	CICE; ∼1° × 1°	*Zhang and Yu* (2011)
FGOALS-s2	SAMIL2; 26 vertical layers, R42	LICOM2; 30 vertical layers, (0.5-1)° ×(0.5-1)°	CSIM5; (0.5-1)° ×(0.5-1)°	*Bao et al.* (submitted)
GFDL-CM3	AM3p9; 48 vertical layers, C48	MOM4p1; 50 vertical layers, tripolar grid ∼1° × 1°	SISp2; tripolar grid, ∼1° × 1°	*Griffies et al.* (2011)
GFDL-ESM2M	AM2p14; 24 vertical layers, M45	MOM4p1; 50 vertical layers; tripolar grid ∼ 1° × 1°	SISp2; tripolar grid ∼ 1° × 1°	*Dunne et al.* (2012)
GISS-E2-R	ModelE; 40 vertical layers, 2° ×2.5°	Russell; 32 vertical layers, 1° × 1.25°	Russel; 1° × 1.25°	http://data.giss.nasa.gov/modelE/ar5/
HadCM3	HadAM3; 19 vertical layers, 3.75° ×2.5°	HadOM3; 20 vertical layers, 1.25° ×1.25°	Sea ice component of HadOM3; 1.25° ×1.25°	*Collins et al.* (2001)
HadGEM2-CC	HadGAM2; 60 vertical layers, N96	HadGOM2; 40 vertical layers, (1-0.3)° ×1°	Inspired from CICE	*Martin et al.* (2011)
HadGEM2-ES	HadGAM2; 38 vertical layers, N96	HadGOM2; 40 vertical layers, (1-0.3)° ×1°	Inspired from CICE	*Martin et al.* (2011)
INM-CM4	INM-CM4 atmospheric component, 21 vertical layers, 2° ×1.5°	INM-CM4 ocean component; 40 vertical layers, 1° ×0.5°	INM-CM4 ocean component; 1° ×0.5°	*Volodin et al.* (2010)

Table B.1: Continued.

Model Name	Atmospheric component	Oceanic component	Sea ice component	References
IPSL-CM5A-LR	LMDZ4 v5; 39 vertical layers, $\sim 2^\circ \times 4^\circ$	NEMO v2.3; 31 vertical layers, ORCA-2°	LIM2; ORCA-2°	http://icmc.ipsl.fr/
IPSL-CM5A-MR	LMDZ4 v5; 39 vertical layers, $\sim 1.25^\circ \times 2.5^\circ$	NEMO v2.3; 31 vertical layers, ORCA-2°	LIM2; ORCA-2°	http://icmc.ipsl.fr/
MIROC4h	CCSR/NIES/FRCGC AGCM v5.7; 56 vertical layers, T213	COCO v3.4; 48 vertical layers, rotated pole, $0.28^\circ \times 0.19^\circ$	COCOv3.4; rotated pole, $0.28^\circ \times 0.19^\circ$	*Sakamoto et al.* (2012)
MROC5	CCSR/NIES/FRCGC AGCM; 40 vertical layers, T85	COCO v4.5; 49 vertical layers, $1.4^\circ \times (0.5\text{-}1.4)^\circ$	COCO v4.5; $1.4^\circ \times (0.5\text{-}1.4)^\circ$	*Watanabe et al.* (2010)
MIROC-ESM	CCSR/NIES/FRCGC AGCM; 80 vertical layers, T42	COCO v3.4; 44 vertical layers, $\sim 1.4^\circ \times 1^\circ$	COCO v3.4; $\sim 1.4^\circ \times 1^\circ$	*Watanabe et al.* (2011)
MIROC-ESM-CHEM	CCSR/NIES/FRCGC AGCM; 80 vertical layers, T42	COCO v3.4; 44 vertical layers; $\sim 1.4^\circ \times 1^\circ$	COCO v3.4; $\sim 1.4^\circ \times 1^\circ$	*Watanabe et al.* (2011)
MPI-ESM-LR	ECHAM6; 47 vertical layers, T63	MPI-OM; 40 vertical layers, $\sim 1.5^\circ \times 1.5^\circ$	Sea ice component of MPI-OM; $\sim 1.5^\circ \times 1.5^\circ$	*Raddatz et al.* (2007)
MRI-CGCM3	GSMUV; 48 vertical layers, TL159	MRI.COM3; 51 vertical layers, $1^\circ \times 0.5^\circ$	MRI.COM3; $1^\circ \times 0.5^\circ$	*Yukimoto et al.* (2011)
NorESM1-M	26 vertical layers, F19	53 vertical layers	No information available to us.	No information available to us.

Table B.2: Summer (JFM) sea ice extent: 1979-2005 seasonal mean and trend, computed from the historical simulations. The ensemble mean of seasonal means is the average over all the JFM extents of the individual members of one model historical simulation. The ensemble mean of seasonal standard deviations is the mean of all the seasonal standard deviations of the individual members. The ensemble mean of the trends is the mean of all the trends of the individual members and the ensemble standard deviation of the trend is the standard deviation of the trend between members. Trends that are significant at the 95% level are in bold. The autocorrelation of the residuals has been taken into account in both the standard deviation of each trend and in the number of degrees of freedom used to to determine the significance threshold (see for instance *Santer et al.*, 2000; *Stroeve et al.*, 2012). Details about the observations are given in *Cavalieri and Parkinson* (2008).

	1979-2005 sea ice extent (10^6km^2)		**1979-2005 trend in sea ice extent (10^3km^2/decade)**		
	Ensemble mean of seasonal means	Ensemble mean of seasonal standard deviations	Individual members	Ensemble mean	Ensemble standard deviation
BCC-CSM1-1	3.89	0.70	**-902.03** -132.44 -50.97	-361.81	469.61
CanESM2	4.13	0.71	**-880.51** **-728.81** -671.28 **-634.06** -110.28	-604.99	292.07
CCSM4	12.06	0.69	**-967.65** **-819.56** **-685.12** **-478.24** -195.45 8.56	-522.91	375.18
CNRM-CM5	0.16	0.08	-120.24 **-111.03** **-80.98** **-73.90** **-73.72** -54.79 -40.41 -36.38 -26.56 -0.19	-61.82	37.54
CSIRO-Mk3.6.0	10.45	0.70	-557.15 **-514.10** -325.14 -240.38 -183.97 -45.54 -23.27 -2.12 13.01 **371.72**	-150.69	276.07
EC-Earth	2.35	0.43	-32.41	-32.41	–
FGOALS-g2	7.15	0.46	0.83	0.83	–
FGOALS-s2	6.71	0.57	-465.78 -369.16 -343.86	-392.93	64.34

Table B.2: Continued.

	1979-2005 sea ice extent (10^6km^2)		**1979-2005 trend in sea ice extent (10^3km^2/decade)**		
	Ensemble mean of seasonal means	Ensemble mean of seasonal standard deviations	Individual members	Ensemble mean	Ensemble standard deviation
GFDL-CM3	0.63	0.22	-126.66 -31.00 27.83 134.95 142.06	29.44	113.84
GFDL-ESM2M	0.44	0.13	**-116.49**	-116.49	–
GISS-E2-R	0.66	0.14	-39.52 -25.73 10.50 14.65 59.69	3.92	38.84
HadCM3	5.00	0.39	**-411.58** **-252.60** **-229.59** **-229.57** -207.29 -179.35 -132.36 -79.55 -19.43 20.64	-172.07	125.76
HadGEM2-CC	2.72	0.35	-114.61	-114.61	–
HadGEM2-ES	3.04	0.37	**-326.27**	-326.27	–
INM-CM4	1.27	0.41	**-268.62**	-268.62	–
IPSL-CM5A-LR	1.04	0.24	**-289.85** **-158.40** -132.87 -98.51	-169.91	83.64
IPSL-CM5A-MR	0.50	0.17	-89.76	-89.76	–
MIROC4h	2.48	0.36	**-500.60** **-343.58** -330.13	-391.43	94.78
MIROC5	0.19	0.05	-10.94	-10.94	–
MIROC-ESM	3.70	0.42	-469.10 **-450.42** **-418.50**	-446.01	25.59
MIROC-ESM-CHEM	4.02	0.39	**-240.84**	-240.84	–
MPI-ESM-LR	1.64	0.34	-208.42 -83.99 -67.01	-119.81	77.21
MRI-CGCM3	4.55	0.37	**-643.00** -203.22 132.63	-237.86	388.98
NorESM1-M	5.93	0.54	-139.58 -135.12 -86.09	-120.27	29.68
Observations	3.96	0.32	148.69	–	–

Table B.3: Winter (JAS) sea ice extent: 1979-2005 seasonal mean and trend, computed from the historical simulations. The ensemble mean of seasonal means is the average over all the JAS extents of the individual members of one model historical simulation. The ensemble mean of seasonal standard deviations is the mean of all the seasonal standard deviations of the individual members. The ensemble mean of the trends is the mean of all the trends of the individual members and the ensemble standard deviation of the trend is the standard deviation of the trend between members. Trends that are significant at the 95% level are in bold. The autocorrelation of the residuals has been taken into account in both the standard deviation of each trend and in the number of degrees of freedom used to determine the significance threshold (see for instance *Santer et al.*, 2000; *Stroeve et al.*, 2012). Details about the observations are given in *Cavalieri and Parkinson* (2008).

	1979-2005 sea ice extent (10^6km^2)		**1979-2005 trend in sea ice extent (10^3km^2/decade)**		
	Ensemble mean of seasonal means	Ensemble mean of seasonal standard deviations	Individual members	Ensemble mean	Ensemble standard deviation
BCC-CSM1-1	20.94	1.32	-2522.87 422.24 434.57	-555.35	1703.93
CanESM2	21.02	0.64	**-904.52** **-878.38** **-826.56** -819.50 -67.45	-699.28	354.99
CCSM4	22.76	0.40	**-767.07** **-741.68** **-649.03** **-559.13** -551.02 -122.50	-565.07	234.58
CNRM-CM5	13.95	0.90	**-2172.40** **-1245.13** **-1019.92** -827.53 **-646.85** -580.44 -506.43 -445.84 -262.46 -165.47	-787.25	587.27
CSIRO-Mk3.6.0	17.81	0.46	**-617.24** **-494.90** **-427.46** -323.45 **-285.16** -201.11 -196.83 -58.77 -3.14 56.81	-255.13	218.47
EC-Earth	17.93	0.72	-147.14	-147.14	–
FGOALS-g2	21.78	0.41	-205.75	-205.75	–
FGOALS-s2	22.62	0.96	-967.45 -917.19 -775.29	-886.64	99.66

Table B.3: Continued.

	1979-2005 sea ice extent (10^6km^2)		**1979-2005 trend in sea ice extent (10^3km^2/decade)**		
	Ensemble mean of seasonal means	Ensemble mean of seasonal standard deviations	Individual members	Ensemble mean	Ensemble standard deviation
GFDL-CM3	11.86	1.07	**-1116.57** -288.07 472.70 766.19 1299.64	226.78	945.00
GFDL-ESM2M	11.76	0.45	-178.78	-178.78	–
GISS-E2-R	12.31	0.78	-607.23 -373.34 -282.37 -179.70 -88.96	-306.32	199.30
HadCM3	19.84	0.71	**-682.10** -654.18 **-647.19** **-521.33** **-424.63** **-414.32** **-377.90** -317.35 -222.93 -4.44	-426.64	213.14
HadGEM2-CC	13.61	0.83	-72.26	-72.26	–
HadGEM2-ES	14.60	0.78	-412.92	-412.92	–
INM-CM4	9.35	0.46	**-459.18**	-459.18	–
IPSL-CM5A-LR	19.12	1.00	-768.83 -573.81 -553.79 325.71	-392.68	488.65
IPSL-CM5A-MR	16.72	0.85	338.90	338.90	–
MIROC4h	17.89	0.54	**-1107.68** **-740.15** **-542.24**	-796.69	286.93
MIROC5	5.42	0.38	-135.04	-135.04	–
MIROC-ESM	20.75	0.76	**-735.34** -575.80 **-519.86**	-610.33	111.82
MIROC-ESM-CHEM	21.33	0.57	-237.01	-237.01	–
MPI-ESM-LR	13.87	1.14	-509.02 -53.14 208.48	-117.89	363.11
MRI-CGCM3	18.75	0.73	**-726.16** -330.31 127.28	-309.73	427.09
NorESM1-M	18.48	0.50	-409.14 -166.62 -50.12	-208.63	183.16
Observations	17.17	0.25	85.57	–	–

References

Arzel, O., T. Fichefet, and H. Goosse (2006), Sea ice evolution over the 20th and 21st centuries as simulated by current AOGCMs, *Ocean Modelling*, *12*(3–4), 401–415. *Cited on page 34.*

Balmaseda, M. A., K. Mogensen, and A. T. Weaver (2013), Evaluation of the ECMWF ocean reanalysis system ORAS4, *Quarterly Journal of the Royal Meteorological Society*, *139*(674), 1132–1161, doi: 10.1002/qj.2063. *Cited on pages 109 and 110.*

Bao, Q., P. Lin, T. Zhou, Y. Liu, Y. Yu, and G. Wu (submitted), The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2, *Advances in Atmospheric Sciences*. *Cited on page 134.*

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. *Cited on pages 22, 62, 95, 100, 101, 121, and 122.*

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. *Cited on pages 21, 35, 62, and 94.*

Bitz, C. M., and G. H. Roe (2004), A mechanism for the high rate of sea ice thinning in the Arctic Ocean, *Journal of Climate*, *17*(18), 3623–3632, doi: 10.1175/1520-0442(2004)017<3623:AMFTHR>2.0.CO;2. *Cited on page 17.*

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. *Cited on pages 35, 62, and 95.*

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). *Cited on pages 65 and 70.*

Bloom, S. C., L. L. Takacs, A. M. da Silva, and D. Ledvina (1996), Data assimilation using incremental analysis updates, *Monthly Weather Review*, *124*(6), 1256–1271. *Cited on page 38.*

Boer, G. (2011), Decadal potential predictability of twenty-first century climate, *Climate Dynamics*, *36*(5-6), 1119–1133, doi: 10.1007/s00382-010-0747-9. *Cited on page 13.*

Boer, G. J., and S. J. Lambert (2008), Multi-model decadal potential predictability of precipitation and temperature, *Geophysical Research Letters*, *35*(5), L05,706. *Cited on page 15.*

Boer, G. J., V. V. Kharin, and W. J. Merryfield (2013), Decadal predictability and forecast skill, *Climate Dynamics*, *41*(7-8), 1817–1833, doi: 10.1007/s00382-013-1705-0. *Cited on page 13.*

Bouttier, F., and F. Rabier (1997), The operational implementation of 4D-Var, *ECMWF Newsletter*, *78*, 2–5. *Cited on page 9.*

Brandt, R. E., S. G. Warren, A. P. Worby, and T. C. Grenfell (2005), Surface albedo of the Antarctic sea ice zone, *Journal of Climate*, *18*(17), 3606–3622, doi: 10.1175/JCLI3489.1. *Cited on page 16.*

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. *Cited on pages 5, 30, 68, 99, and 112.*

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. *Cited on pages 26, 64, and 97.*

Cavalieri, D. J., and C. L. Parkinson (2008), Antarctic sea ice variability and trends, 1979–2006, *J. Geophys. Res.*, *113*(C7), doi: 10.1029/2007JC004564. *Cited on pages 34, 35, 39, 46, 48, 50, 54, 136, and 138.*

Cavalieri, D. J., C. L. Parkinson, P. Gloersen, J. C. Comiso, and H. J. Zwally (1999), Deriving long-term time series of sea ice cover from satellite passive-microwave multisensor data sets, *J. Geophys. Res.*, *104*(C7), 15,803–15,814. *Cited on page 39.*

Cavalieri, D. J., C. L. Parkinson, and K. Y. Vinnikov (2003), 30-Year satellite record reveals contrasting Arctic and Antarctic decadal sea ice variability, *Geophys. Res. Lett.*, *30*(18), doi: 10.1029/2003GL018031. *Cited on pages 20, 23, and 36.*

Chikamoto, Y., M. Kimoto, M. Ishii, T. Mochizuki, T. Sakamoto, H. Tatebe, Y. Komuro, M. Watanabe, T. Nozawa, H. Shiogama, M. Mori, S. Yasunaka, and Y. Imada (2012), An overview of decadal climate predictability in a multi-model ensemble by climate model MIROC, *Climate Dynamics*, pp. 1–22, doi: 10.1007/s00382-012-1351-y. *Cited on page 44.*

Cionni, I., V. Eyring, J. F. Lamarque, W. J. Randel, D. S. Stevenson, F. Wu, G. E. Bodeker, T. G. Shepherd, D. T. Shindell, and D. W. Waugh (2011), Ozone database in support of CMIP5 simulations: results and corresponding radiative forcing, *Atmos. Chem. Phys.*, *11*(21), 11,267–11,292, doi: 10.5194/acp-11-11267-2011. *Cited on pages 38, 42, and 43.*

Close, S. E., and H. Goosse (2013), Entrainment-driven modulation of Southern Ocean mixed layer properties and sea ice variability in CMIP5 models, *Journal of Geophysical Research: Oceans*, pp. n/a–n/a, doi: 10.1002/jgrc.20226. *Cited on page 47.*

Collins, M., and M. R. Allen (2002), Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability, *Journal of Climate*, *15*(21), 3104–3109, doi: 10.1175/1520-0442(2002)015<3104:ATRROI>2.0.CO;2. *Cited on page 4.*

Collins, M., S. F. B. Tett, and C. Cooper (2001), The internal climate variability of HadCM3, a version of the Hadley Centre coupled model without flux adjustments, *Climate Dynamics*, *17*(1), 61–81. *Cited on page 134.*

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. *Cited on pages 17, 20, 21, 29, 30, 39, 45, 65, 102, and 106.*

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. *Cited on pages 34 and 62.*

Corti, S., A. Weisheimer, T. N. Palmer, F. J. Doblas-Reyes, and L. Magnusson (2012), Reliability of decadal predictions, *Geophysical Research Letters*, *39*(21), n/a–n/a, doi: 10.1029/2012GL053354. *Cited on page 13.*

Cotté, C., and C. Guinet (2007), Historical whaling records reveal major regional retreat of Antarctic sea ice, *Deep Sea Research Part I: Oceanographic Research Papers*, *54*(2), 243–252, doi: 10.1016/j.dsr.2006.11.001. *Cited on pages 23 and 36.*

Crespin, E., H. Goosse, T. Fichefet, and M. E. Mann (2009), The 15th century Arctic warming in coupled model simulations with data assimilation, *Climate of the Past*, *5*(3), 389–401, doi: 10.5194/cp-5-389-2009. *Cited on page 10.*

Curran, M. A. J., T. D. van Ommen, V. I. Morgan, K. L. Phillips, and A. S. Palmer (2003), Ice core evidence for Antarctic sea ice decline since the 1950s, *Science*, *302*(5648), 1203–1206, doi: 10.1126/science.1087888. *Cited on pages 20, 23, and 36.*

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, doi: 10.1175/JCLI-D-13-00614.1. *Cited on page 65.*

de la Mare, W. K. (1997), Abrupt mid-twentieth-century decline in Antarctic sea-ice extent from whaling records, *Nature*, *389*(6646), 57–60. *Cited on pages 20, 23, and 36.*

de la Mare, W. K. (2009), Changes in Antarctic sea-ice extent from direct historical observations and whaling records, *Climatic Change*, *92*(3), 461–493. *Cited on pages 20, 23, and 36.*

de Lavergne, C., J. B. Palter, E. D. Galbraith, R. Bernardello, and I. Marinov (2014), Cessation of deep convection in the open Southern Ocean under anthropogenic climate change, *Nature Clim. Change*, *4*(4), 278–282. *Cited on pages 95, 122, and 128.*

Deser, C., R. Knutti, S. Solomon, and A. S. Phillips (2012), Communication of the role of natural variability in future North American climate, *Nature Clim. Change*, *2*(11), 775–779, doi: 10.1038/nclimate1562. *Cited on page 35.*

Doblas-Reyes, F. J., M. A. Balmaseda, A. Weisheimer, and T. N. Palmer (2011), Decadal climate prediction with the European Centre for Medium-Range Weather Forecasts coupled forecast system: Impact of ocean observations, *J. Geophys. Res.*, *116*(D19), doi: 10.1029/2010JD015394. *Cited on pages 2 and 128.*

Donner, L. J., et al. (2011), The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component AM3 of the GFDL global coupled model CM3, *Journal of Climate*, *24*(13), 3484–3519, doi: 10.1175/2011JCLI3955.1. *Cited on page 42.*

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. *Cited on pages 13 and 70.*

Du, H., F. Doblas-Reyes, J. García-Serrano, V. Guemas, Y. Soufflet, and B. Wouters (2012), Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations, *Climate Dynamics*, pp. 1–11, doi: 10.1007/s00382-011-1285-9. *Cited on page 2.*

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. *Cited on pages 10, 28, 66, 68, 70, 79, 86, 92, 96, 99, 100, 108, and 113.*

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. *Cited on pages 10, 69, 99, and 101.*

Dunne, J. P., J. G. John, A. J. Adcroft, S. M. Griffies, R. W. Hallberg, E. Shevliakova, R. J. Stouffer, W. Cooke, K. A. Dunne, M. J. Harrison, J. P. Krasting, S. L. Malyshev, P. C. D. Milly, P. J. Phillipps, L. A. Sentman, B. L. Samuels, M. J. Spelman, M. Winton, A. T. Wittenberg, and N. Zadeh (2012), GFDL's ESM2 global coupled climate-carbon Earth System Models Part I: Physical formulation and baseline simulation characteristics, *Journal of Climate*, doi: 10.1175/JCLI-D-11-00560.1. *Cited on page 134.*

Dunstone, N. J., and D. M. Smith (2010), Impact of atmosphere and subsurface ocean data on decadal climate prediction, *Geophys. Res. Lett.*, *37*(2), doi: 10.1029/2009GL041609. *Cited on pages 7, 11, 68, and 100.*

Ebert, E. E., and J. A. Curry (1993), An intermediate one-dimensional thermodynamic sea ice model for investigating ice-atmosphere interactions, *Journal of Geophysical Research: Oceans*, *98*(C6), 10,085–10,109, doi: 10.1029/93JC00656. *Cited on page 17.*

Evensen, G. (2003), The Ensemble Kalman Filter: theoretical formulation and practical implementation, *Ocean Dynamics*, *53*(4), 343–367, doi: 10.1007/s10236-003-0036-9. *Cited on page 7.*

Fan, T., C. Deser, and D. P. Schneider (2014), Recent Antarctic sea ice trends in the context of Southern Ocean surface climate variations since 1950, *Geophysical Research Letters*, *41*(7), 2014GL059,239. *Cited on pages 20 and 23.*

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. *Cited on pages 30, 65, 102, 105, and 116.*

Fichefet, T., and M. A. M. Maqueda (1999), Modelling the influence of snow accumulation and snow-ice formation on the seasonal cycle of the Antarctic sea-ice cover, *Climate Dynamics*, *15*(4), 251–268. *Cited on page 22.*

Fichefet, T., B. Tartinville, and H. Goosse (2003), Antarctic sea ice variability during 1958–1999: A simulation with a global ice-ocean model, *Journal of Geophysical Research: Oceans*, *108*(C3), 3102. *Cited on pages 23 and 25.*

Flanner, M. G., K. M. Shell, M. Barlage, D. K. Perovich, and M. A. Tschudi (2011), Radiative forcing and albedo feedback from the Northern Hemisphere cryosphere between 1979 and 2008, *Nature Geosci*, *4*(3), 151–155. *Cited on page 17.*

Flato, G. M. (2004), Sea-ice and its response to CO_2 forcing as simulated by global climate models, *Climate Dynamics*, *23*(3), 229–241, doi: 10.1007/s00382-004-0436-7. *Cited on page 34.*

Fyfe, J. C., G. J. Boer, and G. M. Flato (1999), The Arctic and Antarctic oscillations and their projected changes under global warming, *Geophysical Research Letters*, *26*(11), 1601–1604, doi: 10.1029/1999GL900317. *Cited on page 21.*

Fyfe, J. C., W. J. Merryfield, V. Kharin, G. J. Boer, W. S. Lee, and K. von Salzen (2011), Skillful predictions of decadal trends in global mean surface temperature, *Geophysical Research Letters*, *38*(22), L22,801, doi: 10.1029/2011GL049508. *Cited on page 15.*

Fyfe, J. C., N. P. Gillett, and G. J. Marshall (2012), Human influence on extratropical Southern Hemisphere summer precipitation, *Geophysical Research Letters*, *39*(23), doi: 10.1029/2012GL054199. *Cited on pages 22 and 104.*

Gao, F., X. Xin, and T. Wu (2012), Study on the prediction of regional and global temperature in decadal time scale with BCC_CSM1.1, *Chinese Journal of Atmospheric Sciences (in Chinese)*. *Cited on page 44.*

Gent, P. R., G. Danabasoglu, L. J. Donner, M. M. Holland, E. C. Hunke, S. R. Jayne, D. M. Lawrence, R. B. Neale, P. J. Rasch, M. Vertenstein, P. H. Worley, Z.-L. Yang, and M. Zhang (2011), The Community Climate System Model Version 4, *Journal of Climate*, *24*(19), 4973–4991, doi: 10.1175/2011JCLI4083.1. *Cited on page 134.*

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. *Cited on pages 13, 15, 71, and 72.*

Giles, K. A., S. W. Laxon, and A. P. Worby (2008), Antarctic sea ice elevation from satellite radar altimetry, *Geophysical Research Letters*, *35*(3), L03,503. *Cited on page 20.*

Goddard, L., A. Kumar, A. Solomon, D. Smith, G. Boer, P. Gonzalez, V. Kharin, W. Merryfield, C. Deser, S. J. Mason, B. P. Kirtman, R. Msadek, R. Sutton, E. Hawkins, T. Fricker, G. Hegerl, C. A. T. Ferro, D. B. Stephenson, G. A. Meehl, T. Stockdale, R. Burgman, A. M. Greene, Y. Kushnir, M. Newman, J. Carton, I. Fukumori, and T. Delworth (2012), A verification framework for interannual-to-decadal predictions experiments, *Climate Dynamics*, doi: 10.1007/s00382-012-1481-2. *Cited on pages 13, 14, and 70.*

Gong, D., and S. Wang (1999), Definition of Antarctic Oscillation index, *Geophysical Research Letters*, *26*(4), 459–462, doi: 10.1029/1999GL900003. *Cited on pages 19 and 118.*

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. *Cited on pages 26, 64, and 97.*

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. *Cited on pages 10, 18, 22, 28, 29, 62, 64, 74, 88, 95, 96, 97, and 111.*

Goosse, H., F. M. Selten, R. J. Haarsma, and J. D. Opsteegh (2001), Decadal variability in high northern latitudes as simulated by an intermediate-complexity climate model, *Annals of Glaciology*, *33*(1), 525–532. *Cited on page 27.*

Goosse, H., H. Renssen, A. Timmermann, and R. S. Bradley (2005), Internal and forced climate variability during the last millennium: a model-data comparison using ensemble simulations, *Quaternary Science Reviews*, *24*(12–13), 1345–1360. *Cited on page 27.*

Goosse, H., W. Lefebvre, A. de Montety, E. Crespin, and A. Orsi (2009a), 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. *Cited on pages 10, 19, 23, 35, 36, 62, 95, and 98.*

Goosse, H., O. Arzel, C. M. Bitz, A. de Montety, and M. Vancoppenolle (2009b), Increased variability of the Arctic summer ice extent in a warmer climate, *Geophys. Res. Lett.*, *36*(23). *Cited on pages 37 and 49.*

Goosse, H., V. Brovkin, T. Fichefet, R. Haarsma, P. Huybrechts, J. Jongma, A. Mouchet, F. Selten, P.-Y. Barriat, J.-M. Campin, E. Deleersnijder, E. Driesschaert, H. Goelzer, I. Janssens, M.-F. Loutre, M. A.

Morales Maqueda, T. Opsteegh, P.-P. Mathieu, G. Munhoven, E. J. Pettersson, H. Renssen, D. M. Roche, M. Schaeffer, B. Tartinville, A. Timmermann, and S. L. Weber (2010a), 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. *Cited on pages 5, 10, 26, 27, 64, and 97.*

Goosse, H., E. Crespin, A. de Montety, M. E. Mann, H. Renssen, and A. Timmermann (2010b), Reconstructing surface temperature changes over the past 600 years using climate model simulations with data assimilation, *Journal of Geophysical Research: Atmospheres*, *115*(D9), doi: 10.1029/2009JD012737. *Cited on pages 9 and 10.*

Gordon, A. L., P. C. Chu, and J. C. Gascard (1991), *Two Stable Modes of Southern Ocean Winter Stratification*, vol. Volume 57, pp. 17–35, Elsevier, doi: http://dx.doi.org/10.1016/S0422-9894(08)70058-8. *Cited on page 18.*

Griffies, S. M., and K. Bryan (1997), A predictability study of simulated North Atlantic multidecadal variability, *Climate Dynamics*, *13*(7-8), 459–487. *Cited on page 13.*

Griffies, S. M., M. Winton, L. J. Donner, L. W. Horowitz, S. M. Downes, R. Farneti, A. Gnanadesikan, W. J. Hurlin, H.-C. Lee, Z. Liang, J. B. Palter, B. L. Samuels, A. T. Wittenberg, B. L. Wyman, J. Yin, and N. Zadeh (2011), The GFDL CM3 coupled climate model: Characteristics of the ocean and sea ice simulations, *Journal of Climate*, *24*(13), 3520–3544. *Cited on page 134.*

Hall, A., and M. Visbeck (2002), Synchronous variability in the southern hemisphere atmosphere, sea ice, and ocean resulting from the annular mode*, *Journal of Climate*, *15*(21), 3043–3057. *Cited on page 19.*

Ham, Y.-G., M. Rienecker, M. Suarez, Y. Vikhliaev, B. Zhao, J. Marshak, G. Vernieres, and S. Schubert (2014), Decadal prediction skill in the GEOS-5 forecast system, *Climate Dynamics*, *42*(1-2), 1–20. *Cited on page 2.*

Hazeleger, W., X. Wang, C. Severijns, S. Stefanescu, R. Bintanja, A. Sterl, K. Wyser, T. Semmler, S. Yang, B. van den Hurk, T. van Noije, E. van der Linden, and K. van der Wiel (2012), EC-Earth V2.2: description and validation of a new seamless earth system prediction model, *Climate Dynamics*, *39*(11), 2611–2629. *Cited on page 134.*

Hellmer, H. H. (2004), Impact of Antarctic ice shelf basal melting on sea ice and deep ocean properties, *Geophysical Research Letters*, *31*(10), doi: 10.1029/2004GL019506. *Cited on pages 95, 100, and 121.*

Holland, M., and M. Raphael (2006), Twentieth century simulation of the southern hemisphere climate in coupled models. Part II: sea ice conditions and variability, *Climate Dynamics*, *26*(2), 229–245. *Cited on pages 19 and 35.*

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. *Cited on pages 13, 15, 63, 66, 70, 72, 73, 90, 95, and 96.*

Holland, P. R., and R. Kwok (2012), Wind-driven trends in Antarctic sea-ice drift, *Nature Geosci*, *5*(12), 872–875. *Cited on pages 21, 62, and 95.*

Hurrell, J. W., T. Delworth, G. Danabasoglu, H. Drange, S. Griffies, N. Holbrook, B. Kirtman, N. Keenlyside, M. Latif, J. Marotzke, G. A. Meehl, T. Palmer, H. Pohlmann, T. Rosati, R. Seager, D. Smith, R. Sutton, A. Timmermann, K. E. Trenberth, and J. Tribbia (2010), Decadal Climate Variability, Predictability and Prediction: Opportunities and Challenges, in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, vol. 1, Hall, J., Harrison, D.E., Stammer, D., Eds., ESA Publication WPP-306, Venice, Italy, 21-25 September 2009. *Cited on page 4.*

Huybrechts, P. (2002), Sea-level changes at the LGM from ice-dynamic reconstructions of the Greenland and Antarctic ice sheets during the glacial cycles, *Quaternary Science Reviews*, *21*(1–3), 203–231, doi: http://dx.doi.org/10.1016/S0277-3791(01)00082-8. *Cited on page 26.*

Jacobs, S. S., C. F. Giulivi, and P. A. Mele (2002), Freshening of the Ross Sea during the late 20th century, *Science*, *297*(5580), 386–389. *Cited on page 22.*

Kalnay, E. (2007), *Atmospheric Modeling, Data Assimilation and Predictability*, 4 ed., Cambridge University Press, Cambridge. *Cited on pages 6, 7, 8, 9, 38, 68, and 100.*

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, A. Leetmaa, R. Reynolds, M. Chelliah, W. Ebisuzaki, W. Higgins, J. Janowiak, K. C. Mo, C. Ropelewski, J. Wang, R. Jenne, and D. Joseph (1996), The NCEP/NCAR 40-year reanalysis project, *Bulletin of the American Meteorological Society*, *77*(3), 437–471. *Cited on pages 29, 119, and 120.*

Kawase, H., T. Nagashima, K. Sudo, and T. Nozawa (2011), Future changes in tropospheric ozone under Representative Concentration Pathways (RCPs), *Geophys. Res. Lett.*, *38*(5). *Cited on page 43.*

Kay, J. E., M. M. Holland, and A. Jahn (2011), Inter-annual to multi-decadal Arctic sea ice extent trends in a warming world, *Geophys. Res. Lett.*, *38*(15). *Cited on page 34.*

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. *Cited on pages 7, 15, 38, 68, 69, and 100.*

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), doi: 10.1029/2012GL051644. *Cited on pages 15, 56, and 90.*

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. *Cited on pages 35, 62, and 95.*

Kirtman, B., S. B. Power, J. A. Adedoyin, G. J. Boer, R. Bojariu, I. Camilloni, F. J. Doblas-Reyes, A. M. Fiore, M. Kimoto, G. A. Meehl, M. Prather, A. Sarr, C. Schär, R. Sutton, G. J. van Oldenborgh, G. Vecchi, and H. J. Wang (2013), Near-term Climate Change: Projections and Predictability, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. *Cited on pages 3, 12, 13, 20, 24, 94, 125, and 126.*

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. *Cited on pages 13, 15, and 70.*

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. *Cited on pages 7, 14, 15, 68, and 100.*

Kushner, P. J., I. M. Held, and T. L. Delworth (2001), Southern hemisphere atmospheric circulation response to global warming, *Journal of Climate*, *14*(10), 2238–2249, doi: 10.1175/1520-0442(2001)014<0001:SHACRT>2.0.CO;2. *Cited on page 21.*

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. *Cited on pages 35, 42, 62, 79, and 95.*

Latif, M., T. Delworrth, D. Dommenget, H. Drange, W. Hazeleger, J. Hurrell, N. Keenlyside, G. A. Meehl, and R. Sutton (2010), Dynamics of decadal climate variability and implications for its prediction, in *Proceedings of OceanObs'09: Sustained Ocean Observations and Information for Society*, vol. 2, Hall, J., Harrison D.E. & Stammer, D., Eds., ESA Publication WPP-306, Venice, Italy, 21-25 September 2009. *Cited on pages 4, 13, 37, and 60.*

Lecomte, O., T. Fichefet, M. Vancoppenolle, F. Domine, F. Massonnet, P. Mathiot, S. Morin, and P. Y. Barriat (2013), On the formulation of snow thermal conductivity in large-scale sea ice models, *Journal of Advances in Modeling Earth Systems*, *5*(3), 542–557, doi: 10.1002/jame.20039. *Cited on page 22.*

Lefebvre, W., and H. Goosse (2005), Influence of the Southern Annular Mode on the sea ice-ocean system: the role of the thermal and mechanical forcing, *Ocean Science*, *1*(3), 145–157, doi: 10.5194/os-1-145-2005. *Cited on pages 19 and 21.*

Lefebvre, W., and H. Goosse (2008a), 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. *Cited on pages 19, 21, 35, 62, 95, and 118.*

Lefebvre, W., and H. Goosse (2008b), Analysis of the projected regional sea-ice changes in the Southern Ocean during the twenty-first century, *Climate Dynamics*, *30*(1), 59–76–76. *Cited on pages 34 and 37.*

Lefebvre, W., H. Goosse, R. Timmermann, and T. Fichefet (2004), Influence of the Southern Annular Mode on the sea ice–ocean system, *Journal of Geophysical Research: Oceans*, *109*(C9), C09,005. *Cited on pages 19, 21, and 27.*

Liu, J., and J. A. Curry (2010), Accelerated warming of the Southern Ocean and its impacts on the hydrological cycle and sea ice, *Proceedings of the National Academy of Sciences*, *107*(34), 14,987–14,992. *Cited on pages 22, 35, and 104.*

Liu, J., D. G. Martinson, X. Yuan, and D. Rind (2002), Evaluating Antarctic sea ice variability and its teleconnections in global climate models, *International Journal of Climatology*, *22*(8), 885–900. *Cited on page 19.*

Liu, J., J. A. Curry, and D. G. Martinson (2004), Interpretation of recent Antarctic sea ice variability, *Geophys. Res. Lett.*, *31*(2). *Cited on page 21.*

Lorenz, E. N. (1963), Deterministic nonperiodic flow, *Journal of the Atmospheric Sciences*, *20*(2), 130–141. *Cited on page 3.*

Lorenz, E. N. (1975), Climate Predictability, in *The physical bases of climate and climate modelling*, no. 16 in WMO GARP Ser., pp. 132–136, World Meteorol Organ, Geneva, Switzerland. *Cited on page 3.*

Magnusson, L., M. Alonso-Balmaseda, and F. Molteni (2013a), On the dependence of ENSO simulation on the coupled model mean state, *Climate Dynamics*, *41*(5-6), 1509–1525. *Cited on pages 11 and 12.*

Magnusson, L., M. Alonso-Balmaseda, S. Corti, F. Molteni, and T. Stockdale (2013b), Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors, *Climate Dynamics*, *41*(9-10), 2393–2409. *Cited on pages 11, 12, and 128.*

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. *Cited on pages 23, 62, and 94.*

Mairesse, A., H. Goosse, P. Mathiot, H. Wanner, and S. Dubinkina (2013), Investigating the consistency between proxy-based reconstructions and climate models using data assimilation: a mid-Holocene case study, *Climate of the Past*, *9*(6), 2741–2757, doi: 10.5194/cp-9-2741-2013. *Cited on page 10.*

Marshall, G. J. (2003), Trends in the Southern Annular Mode from Observations and Reanalyses, *Journal of Climate*, *16*(24), 4134–4143, doi: 10.1175/1520-0442(2003)016<4134:TITSAM>2.0.CO;2. *Cited on page 21.*

Martin, G. M., et al. (2011), The HadGEM2 family of Met Office Unified Model Climate configurations, *Geosci. Model Dev. Discuss.*, *4*(2), 765–841. *Cited on page 134.*

Martinson, D. G. (1990), Evolution of the Southern Ocean winter mixed layer and sea ice: Open ocean deepwater formation and ventilation, *Journal of Geophysical Research: Oceans*, *95*(C7), 11,641–11,654, doi: 10.1029/JC095iC07p11641. *Cited on page 18.*

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. *Cited on pages 18 and 74.*

Massonnet, F., T. Fichefet, H. Goosse, M. Vancoppenolle, P. Mathiot, and C. König Beatty (2011), On the influence of model physics on simulations of Arctic and Antarctic sea ice, *The Cryosphere*, *5*(3), 687–699, doi: 10.5194/tc-5-687-2011. *Cited on page 127.*

Massonnet, F., T. Fichefet, H. Goosse, C. M. Bitz, G. Philippon-Berthier, M. M. Holland, and P. Y. Barriat (2012), Constraining projections of summer Arctic sea ice, *The Cryosphere*, *6*(6), 1383–1394. *Cited on page 60.*

Massonnet, F., P. Mathiot, T. Fichefet, H. Goosse, C. König Beatty, M. Vancoppenolle, and T. Lavergne (2013), A model reconstruction of the Antarctic sea ice thickness and volume changes over 1980–2008 using data assimilation, *Ocean Modelling*, *64*(0), 67–75, doi: 10.1016/j.ocemod.2013.01.003. *Cited on page 8.*

Massonnet, F., H. Goosse, and T. Fichefet (2014), Prospects for improved seasonal Arctic sea ice predictions from multivariate data assimilation, *Ocean Modelling (submitted)*. *Cited on page 8.*

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. *Cited on pages 7, 14, 15, 44, 68, 69, and 100.*

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. *Cited on pages 14, 15, 56, and 90.*

Mathiot, P., H. Goosse, X. Crosta, B. Stenni, M. Braida, H. Renssen, C. J. Van Meerbeeck, V. Masson-Delmotte, A. Mairesse, and S. Dubinkina (2013), Using data assimilation to investigate the causes of Southern Hemisphere high latitude cooling from 10 to 8 ka BP, *Climate of the Past*, *9*(2), 887–901. *Cited on pages 10 and 100.*

Meehl, G. A., L. Goddard, J. Murphy, R. J. Stouffer, G. Boer, G. Danabasoglu, K. Dixon, M. A. Giorgetta, A. M. Greene, E. Hawkins, G. Hegerl, D. Karoly, N. Keenlyside, M. Kimoto, B. Kirtman, A. Navarra, R. Pulwarty, D. Smith, D. Stammer, and T. Stockdale (2009), Decadal prediction: Can it be skillful?, *Bulletin of American Meteorological Society*, *90*(10), 1467–1485. *Cited on pages 3 and 4.*

Meehl, G. A., L. Goddard, G. Boer, R. Burgman, G. Branstator, C. Cassou, S. Corti, G. Danabasoglu, F. Doblas-Reyes, E. Hawkins, A. Karspeck, M. Kimoto, A. Kumar, D. Matei, J. Mignot, R. Msadek, A. Navarra, H. Pohlmann, M. Rienecker, T. Rosati, E. Schneider, D. Smith, R. Sutton, H. Teng, G. J. van Oldenborgh, G. Vecchi, and S. Yeager (2013), Decadal Climate Prediction: An Update from the Trenches, *Bulletin of the American Meteorological Society*, *95*(2), 243–267. *Cited on page 128.*

Meier, W. N., D. Gallaher, and G. G. Campbell (2013), New estimates of Arctic and Antarctic sea ice extent during September 1964 from recovered Nimbus I satellite imagery, *The Cryosphere*, *7*(2), 699–705, doi: 10.5194/tc-7-699-2013. *Cited on pages 20, 23, and 36.*

Mo, K. C., and G. H. White (1985), Teleconnections in the southern hemisphere, *Monthly Weather Review*, *113*(1), 22–37, doi: 10.1175/1520-0493(1985)113<0022:TITSH>2.0.CO;2. *Cited on page 27.*

Mouchet, A., and L. M. François (1996), Sensitivity of a global oceanic carbon cycle model to the circulation and the fate of organic matter: preliminary results, *Phys. Chem. Earth*, *21*, 511–516. *Cited on page 26.*

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), doi: 10.1029/2010GL044517. *Cited on pages 13 and 70.*

Murphy, A. H. (1988), Skill scores based on the mean square error and their relationships to the correlation coefficient, *Monthly Weather Review*, *16*, 2417–2424. *Cited on page 14.*

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. *Cited on pages 11 and 68.*

Notz, D. (2014), Sea-ice extent and its trend provide limited metrics of model performance, *The Cryosphere*, *8*(1), 229–243, doi: 10.5194/tc-8-229-2014. *Cited on page 72.*

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. *Cited on pages 26, 64, and 97.*

Parkinson, C. L., and D. J. Cavalieri (2012), Antarctic sea ice variability and trends, 1979–2010, *The Cryosphere*, *6*(4), 871–880. *Cited on pages 16, 20, and 62.*

Parkinson, C. L., K. Y. Vinnikov, and D. J. Cavalieri (2006), Evaluation of the simulation of the annual cycle of Arctic and Antarctic sea ice coverages by 11 major global climate models, *J. Geophys. Res.*, *111*(C7). *Cited on page 34.*

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. *Cited on pages 13 and 70.*

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. *Cited on pages 38 and 68.*

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. *Cited on pages 2, 13, 15, 37, 70, and 71.*

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. *Cited on pages 7, 14, 15, 38, 55, 63, 68, 69, 71, and 100.*

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. *Cited on pages 14, 15, and 63.*

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. *Cited on pages 23, 62, 72, 94, and 124.*

Pringle, D. J., H. Eicken, H. J. Trodahl, and L. G. E. Backstrom (2007), Thermal conductivity of landfast Antarctic and Arctic sea ice, *Journal of Geophysical Research: Oceans*, *112*(C4), n/a–n/a, doi: 10.1029/2006JC003641. *Cited on page 16.*

Pritchard, H. D., S. R. M. Ligtenberg, H. A. Fricker, D. G. Vaughan, M. R. van den Broeke, and L. Padman (2012), Antarctic ice-sheet loss driven by basal melting of ice shelves, *Nature*, *484*(7395), 502–505. *Cited on pages 22, 95, and 101.*

Raddatz, T., C. Reick, W. Knorr, J. Kattge, E. Roeckner, R. Schnur, K. G. Schnitzler, P. Wetzel, and J. Jungclaus (2007), Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, *Climate Dynamics*, *29*(6), 565–574. *Cited on page 135.*

Randel, W. J., and F. Wu (1999), A stratospheric ozone trends data set for global modeling studies, *Geophys. Res. Lett.*, *26*(20), 3089–3092. *Cited on page 42.*

Raphael, M. N. (2004), A zonal wave 3 index for the southern hemisphere, *Geophysical Research Letters*, *31*(23), L23,212. *Cited on page 27.*

Rignot, E., J. L. Bamber, M. R. van den Broeke, C. Davis, Y. Li, W. J. van de Berg, and E. van Meijgaard (2008), Recent Antarctic ice mass loss from radar interferometry and regional climate modelling, *Nature Geosci*, *1*(2), 106–110. *Cited on pages 22, 95, and 101.*

Rotstayn, L. D., M. A. Collier, M. R. Dix, Y. Feng, H. B. Gordon, S. P. O'Farrell, I. N. Smith, and J. Syktus (2010), Improved simulation of Australian climate and ENSO-related rainfall variability in a global climate model with an interactive aerosol treatment, *International Journal of Climatology*, *30*(7), 1067–1088. *Cited on page 134.*

Saha, S., et al. (2010), The NCEP climate forecast system reanalysis, *Bulletin of the American Meteorological Society*, *91*(8), 1015–1057. *Cited on page 9.*

Sakamoto, T. K., Y. Komuro, T. Nishimura, M. Ishii, H. Tatebe, H. Shiogama, A. Hasegawa, T. Toyoda, M. Mori, T. Suzuki, Y. Imada, T. Nozawa, K. Takata, K. Mochizuki, K. Ogochi, S. Emori, H. Hasumi, and M. Kimoto (2012), MIROC4h-a new high-resolution atmosphere-ocean coupled general circulation model, *Journal of the Meteorological Society of Japan*, *In press*. *Cited on page 135.*

Santer, B. D., T. M. L. Wigley, J. S. Boyle, D. J. Gaffen, J. J. Hnilo, D. Nychka, D. E. Parker, and K. E. Taylor (2000), Statistical significance of trends and trend differences in layer-average atmospheric temperature time series, *J. Geophys. Res.*, *105*(D6), 7337–7356, doi: 10.1029/1999JD901105. *Cited on pages 47, 48, 50, 54, 102, 136, and 138.*

Sen Gupta, A., A. Santoso, A. S. Taschetto, C. C. Ummenhofer, J. Trevena, and M. H. England (2009), Projected changes to the southern hemisphere ocean and sea ice in the IPCC AR4 climate models, *Journal of Climate*, *22*(11), 3047–3078. *Cited on page 34.*

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. *Cited on pages 63, 68, 69, and 100.*

Shepherd, A., et al. (2012), A reconciled estimate of ice-sheet mass balance, *Science*, *338*(6111), 1183–1189, doi: 10.1126/science.1228102. *Cited on pages 22, 95, 101, and 114.*

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. *Cited on pages 21, 35, 62, and 94.*

Sigmond, M., and J. C. Fyfe (2013), The Antarctic sea ice response to the ozone hole in climate models, *Journal of Climate*, *27*(3), 1336–1342, doi: 10.1175/JCLI-D-13-00590.1. *Cited on pages 21 and 94.*

Silva, T. A. M., G. R. Bigg, and K. W. Nicholls (2006), Contribution of giant icebergs to the Southern Ocean freshwater flux, *Journal of Geophysical Research: Oceans*, *111*(C3), doi: 10.1029/2004JC002843. *Cited on pages 22 and 101.*

Simpkins, G. R., L. M. Ciasto, D. W. J. Thompson, and M. H. England (2012), Seasonal relationships between large-scale climate variability and Antarctic sea ice concentration, *Journal of Climate*, *25*(16), 5451–5469. *Cited on page 21.*

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. 3643–3648, doi: 10.1002/grl.50715. *Cited on pages 23 and 62.*

Smith, D., R. Eade, and H. Pohlmann (2013a), A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction, *Climate Dynamics*, *41*(11-12), 3325–3338. *Cited on pages 11, 12, and 68.*

Smith, D., A. Scaife, G. Boer, M. Caian, F. Doblas-Reyes, V. Guemas, E. Hawkins, W. Hazeleger, L. Hermanson, C. Ho, M. Ishii, V. Kharin, M. Kimoto, B. Kirtman, J. Lean, D. Matei, W. Merryfield, W. Müller, H. Pohlmann, A. Rosati, B. Wouters, and K. Wyser (2013b), Real-time multi-model decadal climate predictions, *Climate Dynamics*, *41*(11-12), 2875–2888. *Cited on pages 15 and 128.*

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy (2007), Improved surface temperature prediction for the coming decade from a global climate model, *Science*, *317*(5839), 796–799. *Cited on page 38.*

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. *Cited on pages 7, 68, 69, and 100.*

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. *Cited on pages 21, 35, 62, and 94.*

Solomon, S. (1999), Stratospheric ozone depletion: A review of concepts and history, *Rev. Geophys.*, *37*(3), 275–316. *Cited on pages 35, 62, and 94.*

Son, S. W., L. M. Polvani, D. W. Waugh, H. Akiyoshi, R. Garcia, D. Kinnison, S. Pawson, E. Rozanov, T. G. Shepherd, and K. Shibata (2008), The impact of stratospheric ozone recovery on the southern hemisphere westerly jet, *Science*, *320*(5882), 1486–1489. *Cited on page 38.*

Sriver, R. L., A. Timmermann, M. E. Mann, K. Keller, and H. Goosse (2014), Improved representation of tropical Pacific Ocean–atmosphere dynamics in an intermediate complexity climate model, *Journal of Climate*, *27*(1), 168–185, doi: doi: 10.1175/JCLI-D-12-00849.1. *Cited on page 27.*

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. *Cited on pages 19, 21, 35, 62, and 95.*

Stroeve, J. C., V. Kattsov, A. Barrett, M. Serreze, T. Pavlova, M. Holland, and W. N. Meier (2012), Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, *Geophys. Res. Lett.*, *39*(16), doi: 10.1029/2012GL052676. *Cited on pages 47, 48, 50, 54, 60, 102, 136, and 138.*

Swart, N. C., and J. C. Fyfe (2012), Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress, *Geophysical Research Letters*, *39*(16), L16,711. *Cited on page 21.*

Swart, N. C., and J. C. Fyfe (2013), The influence of recent Antarctic ice sheet retreat on simulated sea ice area trends, *Geophysical Research Letters*, *40*(16), 4328–4332. *Cited on pages 22, 23, 94, 95, 101, and 122.*

Swingedouw, D., T. Fichefet, P. Huybrechts, H. Goosse, E. Driesschaert, and M.-F. Loutre (2008), Antarctic ice-sheet melting provides negative feedbacks on future climate warming, *Geophysical Research Letters*, *35*(17), doi: 10.1029/2008GL034410. *Cited on pages 95, 100, and 121.*

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. *Cited on pages 7, 15, 44, 68, 69, and 100.*

Szopa, S., Y. Balkanski, M. Schulz, S. Bekki, D. Cugnet, A. Fortems-Cheiney, S. Turquety, A. Cozic, C. Déandreis, D. Hauglustaine, A. Idelkadi, J. Lathière, F. Lefevre, M. Marchand, R. Vuolo, N. Yan, and J. L. Dufresne (2012), Aerosol and ozone changes as forcing for climate evolution between 1850 and 2100, *Climate Dynamics*, pp. 1–28. *Cited on page 43.*

Talagrand, O. (1997), Assimilation of Observations, an Introduction, *Journal of the Meteorological Society of Japan. Ser. II*, *75*(1B), 191–209. *Cited on pages 4, 7, 8, 66, and 99.*

Talagrand, O., and R. Vautard (1997), Evaluation of probabilistic prediction systems, in *ECMWF Workshop on Predictability*, pp. 1–25, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom. *Cited on page 14.*

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. *Cited on pages 2, 15, 37, 62, 64, 94, and 97.*

Tebaldi, C., J. M. Arblaster, and R. Knutti (2011), Mapping model agreement on future climate projections, *Geophys. Res. Lett.*, *38*(23). *Cited on page 35.*

Thépaut, J.-N., P. Caille, V. Cassé, J.-F. Geleyn, P. Haution, M. Janiskova, P. Moll, J. Pailleux, F. Taillefer, and F. Veersé (1996), 4D-Var developments at Météo-France, in *ECMWF Workshop on Non-linear aspects of data assimilation*, pp. 469–492, ECMWF, Shinfield Park, Reading, Berkshire RG2 9AX, United Kingdom. *Cited on page 9.*

Thompson, D. W. J., and S. Solomon (2002), Interpretation of recent southern hemisphere climate change, *Science*, *296*(5569), 895–899, doi: 10.1126/science.1069270. *Cited on page 21.*

Thompson, D. W. J., and J. M. Wallace (2000), Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability, *Journal of Climate*, *13*(5), 1000–1016, doi: 10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2. *Cited on pages 18, 19, 27, and 118.*

Thompson, D. W. J., S. Solomon, P. J. Kushner, M. H. England, K. M. Grise, and D. J. Karoly (2011), Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change, *Nature Geosci*, *4*(11), 741–749. *Cited on pages 21 and 35.*

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. *Cited on pages 15 and 63.*

Toth, Z., and E. Kalnay (1997), Ensemble forecasting at NCEP and the breeding method, *Monthly Weather Review*, *125*(12), 3297–3319. *Cited on page 2.*

Troccoli, A., and T. N. Palmer (2007), Ensemble decadal predictions from analysed initial conditions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *365*(1857), 2179–2191. *Cited on pages 38, 55, and 128.*

Turner, J., and J. Overland (2009), Contrasting climate change in the two polar regions, *Polar Research*, *28*(2), 146–164, doi: 10.1111/j.1751-8369.2009.00128.x. *Cited on pages 1, 18, and 20.*

Turner, J., J. C. Comiso, G. J. Marshall, T. A. Lachlan-Cope, T. Bracegirdle, T. Maksym, M. P. Meredith, Z. Wang, and A. Orr (2009), Non-annular atmospheric circulation change induced by stratospheric ozone depletion and its role in the recent increase of Antarctic sea ice extent, *Geophys. Res. Lett.*, *36*(8). *Cited on pages 21 and 35.*

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. *Cited on pages 23, 62, 65, 72, 95, 124, and 127.*

van Leeuwen, P. J. (2009), Particle filtering in geophysical systems, *Monthly Weather Review*, *137*(12), 4089–4114, doi: 10.1175/2009MWR2835.1. *Cited on pages 9, 10, 11, 69, 70, and 99.*

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. *Cited on page 70.*

Vannitsem, S., and C. Nicolis (2008), Dynamical properties of model output statistics forecasts, *Monthly Weather Review*, *136*(2), 405–419. *Cited on page 55.*

Vaughan, D., G. Marshall, W. Connolley, C. Parkinson, R. Mulvaney, D. Hodgson, J. King, C. Pudsey, and J. Turner (2003), Recent rapid regional climate warming on the Antarctic peninsula, *Climatic Change*, *60*(3), 243–274, doi: 10.1023/A:1026021217991. *Cited on page 20.*

Velicogna, I. (2009), Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE, *Geophysical Research Letters*, *36*(19), L19,503, doi: 10.1029/2009GL040222. *Cited on pages 22, 95, and 101.*

Vera, C., M. Barange, O. P. Dube, L. Goddard, D. Griggs, N. Kobysheva, E. Odada, S. Parey, J. Polovina, G. Poveda, B. Seguin, and K. Trenberth (2010), Needs assessment for climate information on decadal timescales and longer, in *World Climate Conference - 3*, vol. 1, pp. 275–286. *Cited on page 2.*

Voldoire, A., E. Sanchez-Gomez, D. Salas y Mélia, B. Decharme, C. Cassou, S. Sénési, S. Valcke, I. Beau, A. Alias, M. Chevallier, M. Déqué, J. Deshayes, H. Douville, E. Fernandez, G. Madec, E. Maisonnave, M. P. Moine, S. Planton, D. Saint-Martin, S. Szopa, S. Tyteca, R. Alkama, S. Belamari, A. Braun, L. Coquart, and F. Chauvin (2012), The CNRM-CM5.1 global climate model: description and basic evaluation, *Climate Dynamics*. *Cited on pages 42 and 134.*

Volodin, E., N. Dianskii, and A. Gusev (2010), Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, *Izvestiya Atmospheric and Oceanic Physics*, *46*(4), 414–431. *Cited on page 134.*

Watanabe, M., T. Suzuki, R. O'ishi, Y. Komuro, S. Watanabe, S. Emori, T. Takemura, M. Chikira, T. Ogura, M. Sekiguchi, K. Takata, D. Yamazaki, T. Yokohata, T. Nozawa, H. Hasumi, H. Tatebe, and M. Kimoto (2010), Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity, *Journal of Climate*, *23*(23), 6312–6335. *Cited on page 135.*

Watanabe, S., T. Hajima, K. Sudo, T. Nagashima, T. Takemura, H. Okajima, T. Nozawa, H. Kawase, M. Abe, T. Yokohata, T. Ise, H. Sato, E. Kato, K. Takata, S. Emori, and M. Kawamiya (2011), MIROC-ESM: model description and basic results of CMIP5-20c3m experiments, *Geosci. Model Dev. Discuss.*, *4*(2), 1063–1128. *Cited on pages 43 and 135.*

Wilks, D. S. (2006), *Statistical Methods in the Atmospheric Sciences*, *International Geophysics Series*, vol. 91, 2 ed., Academic Press. *Cited on pages 13 and 14.*

Worby, A. P., C. A. Geiger, M. J. Paget, M. L. Van Woert, S. F. Ackley, and T. L. DeLiberty (2008), Thickness distribution of Antarctic sea ice, *Journal of Geophysical Research: Oceans*, *113*(C5), doi: 10.1029/2007JC004254. *Cited on page 17.*

Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng (2012), A decadal prediction case study: Late twentieth-century North Atlantic ocean heat content, *Journal of Climate*, *25*(15), 5173–5189. *Cited on page 2.*

Yuan, X. (2004), ENSO-related impacts on Antarctic sea ice: a synthesis of phenomenon and mechanisms, *Antarctic Science*, *16*(04), 415–425. *Cited on page 19.*

Yuan, X., and D. G. Martinson (2001), The Antarctic dipole and its predictability, *Geophys. Res. Lett.*, *28*(18), 3609–3612. *Cited on page 19.*

Yukimoto, S., H. Yoshimura, M. Hosaka, T. Sakami, H. Tsujino, M. Hirabara, T. Y. Tanaka, M. Deushi, A. Obata, H. Nakano, Y. Adachi, E. Shindo, S. Yabu, T. Ose, and A. Kitoh (2011), Meteorological Research Institute-Earth System Model Version 1 (MRI-ESM1) - Model Description, *Tech. Rep. 64*, Meteorological Research Institute, Japan. *Cited on pages 43 and 135.*

Zanna, L., and E. Tziperman (2008), Optimal surface excitation of the thermohaline circulation, *Journal of Physical Oceanography*, *38*, 1820–1830. *Cited on page 2.*

Zhang, J. (2007), Increasing Antarctic sea ice under warming atmospheric and oceanic conditions, *Journal of Climate*, *20*(11), 2515–2529. *Cited on pages 18, 22, 35, 62, and 95.*

Zhang, J. (2014), Modeling the impact of wind intensification on antarctic sea ice volume, *Journal of Climate*, *27*(1), 202–214. *Cited on page 25.*

Zhang, S., M. J. Harrison, A. Rosati, and A. Wittenberg (2007), System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies, *Monthly Weather Review*, *135*(10), 3541–3564. *Cited on page 8.*

Zhang, Y. L., and Y.-Q. Yu (2011), Analysis of decadal climate variability in the tropical Pacific by coupled GCM, *Atmospheric and Oceanic Science Letters*, *4*(4), 204–208. *Cited on page 134.*

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. *Cited on pages 23, 25, 62, 63, 65, 79, 94, 95, 96, and 104.*

Zunz, V., H. Goosse, and S. Dubinkina (2014), Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multidecadal timescales, *Climate Dynamics (submitted in revised version). Cited on pages 10, 13, 15, 95, 96, 115, and 121.*

Zunz, V., and H. Goosse (2014), Influence of meltwater input on the skill of decadal forecast of sea ice in the Southern Ocean, *The Cryosphere Discussions*, *8*(4), 3563–3602, doi: 10.5194/tcd-8-3563-2014. *Cited on pages 10 and 15.*