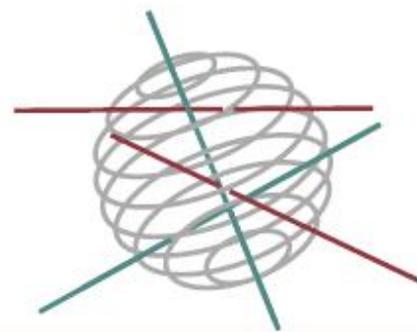


# SSD

SCIENCE FOR A SUSTAINABLE DEVELOPMENT



## Understanding and Predicting Antarctic sea ice variability at the decadal timescale

### “PREDANTAR”

H. GOOSSE, S. CLOSE, S. DUBINKINA, F. MASSONNET, V. ZUNZ,  
S. VANNITSEM, B. VAN SCHAEYBROECK, A. BARTH, M. CANTER



ENERGY



TRANSPORT AND MOBILITY



AGRO-FOOD



HEALTH AND ENVIRONMENT



CLIMATE



BIODIVERSITY



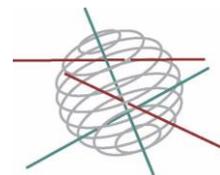
ATMOSPHERE AND TERRESTRIAL AND MARINE ECOSYSTEMS



TRANSVERSAL ACTIONS







***Climate, Antarctica***

FINAL REPORT

**UNDERSTANDING AND PREDICTING ANTARCTIC SEA ICE  
VARIABILITY AT THE DECADAL TIMESCALE**

**“PREDANTAR”**

**SD/CA/04A**

**PROMOTORS**

**Prof. Hugues Goosse**

Centre de recherches sur Ta terre et le  
climat Georges Lemaître  
Earth and Life Institute  
Université catholique de Louvain (UCL)  
Place Pasteur, 3  
1348 Louvain-la-Neuve

**Dr. Stéphane Vannitsem**

Institut Royal Météorologique de  
Belgique (IRMB)  
Avenue Circulaire, 3  
1180 Brussels

**Dr. Alexander Barth**

Université de Liège (ULg)  
GHER/AGO/MARE Institut de Physique  
B5a Allée du 6 Août 17  
4000 Liège

**AUTHORS**

Hugues Goosse, Sally Close, Svetlana Dubinkina, François  
Massonnet and Violette Zunz (UCL)  
Stéphane Vannitsem and Bert Van Schaeybroeck (IRMB)  
Alexander Barth and Martin Canter (ULg)



Published in 2015 by the Belgian Science Policy  
Avenue Louise 231  
Louizalaan 231  
B-1050 Brussels  
Belgium  
Tel: +32 (0)2 238 34 11 – Fax: +32 (0)2 230 59 12  
<http://www.belspo.be>

Contact person: Maaïke Vancauwenberghe  
+32 (0)2 238 36 78

Neither the Belgian Science Policy nor any person acting on behalf of the Belgian Science Policy is responsible for the use which might be made of the following information. The authors are responsible for the content.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without indicating the reference :

H. Goosse, S. Close, S. Dubinkina, F. Massonnet, V. Zunz, S. Vannitsem, B. Van Schaeybroeck, A. Barth and M. Canter. ***Understanding and Predicting Antarctic sea ice variability at the decadal timescale "PREDANTAR"***. Final Report. Brussels: Belgian Science Policy 2015 – 75 p. (Research Programme Science for a Sustainable Development)

**Table of content**

Acronyms.....	4
Summary .....	5
1. Introduction.....	9
2. Methodology and results .....	13
2.1. Adaptation of post-processing tools .....	13
2.1.1. Post-processing as a diagnostic tool.....	13
2.1.2. Post-processing of climate data in a static environment .....	15
2.1.3. Post-processing in a changing environment .....	17
2.1.3.1. Studies with simple models .....	17
2.1.3.2. Studies with a climate model.....	18
2.1.4. Post-processing of ensemble predictions.....	20
2.1.4.1. Ensemble post-processing: spread-conserving calibration.....	20
2.1.4.2. Ensemble post-processing: spread-adjusting calibration.....	20
2.1.4.3. Calibration of ensemble predictions in a hydrological context .....	21
2.1.5. Reliability verification of ensemble predictions based on spread.....	21
2.2. Application of advanced data assimilation methods to study Southern Ocean sea ice cover ...	22
2.2.1. Southern Ocean reanalysis using assimilation of sea surface temperature, sea ice concentration and drift .....	22
2.2.1.1. Model and observations .....	22
2.2.1.2. Correcting wind field using sea ice drift.....	23
2.2.1.3. Reanalysis using data assimilation .....	27
2.2.2. Reconstruction of sea ice thickness and volume using data assimilation.....	34
2.2.3. Bias correction and parameter estimation using data assimilation .....	35
2.3. Understanding the changes in the sea ice cover over the last decades.....	40
2.3.1. Comparison of reconstructions and simulations of past changes .....	40
2.3.2. Analysis of the mechanisms ruling decadal variability .....	43
2.3.2.1. Antarctic sea ice variability and mixed layer properties in CMIP5 models.....	43
2.3.2.2. Decadal trends in Antarctic sea ice extent and ice-ocean interactions.....	45
2.4. Projections and predictions .....	47
2.4.1. Analysis of the changes for the end of 21 <sup>st</sup> century .....	47
2.4.2. Improvement of the decadal forecasts .....	48
2.4.2.1. Antarctic sea ice predictability in an idealized framework .....	49
2.4.2.2. Antarctic sea ice predictability in a realistic framework .....	52
3. Policy support .....	57
4. Dissemination and valorization.....	59
5. Publications .....	61
6. Acknowledgements .....	63
7. References .....	65
8. Annexes.....	75

## Acronyms

<b>BC</b>	Bias Correction
<b>CF</b>	Change Factor
<b>CGCM</b>	Coupled General Circulation Models
<b>CLIO</b>	Coupled Large-scale Ice Ocean
<b>CMIP5</b>	Coupled Model Intercomparison Project phase 5
<b>CRPSS</b>	Continuous Rank Probability Skill Score
<b>ECMWF</b>	European Centre for Medium-range Weather Forecasts
<b>EPS</b>	Ensemble Prediction System
<b>EVMOS</b>	Error in Variable Model Output Statistics
<b>EVP</b>	Elastic-Viscous-Plastic rheology
<b>GCM</b>	General Circulation Models
<b>IRMB</b>	Institut Royal Météorologique de Belgique
<b>KS</b>	Kolmogorov-Sinai
<b>LIM</b>	Louvain-la-Neuve sea Ice Model
<b>MLD</b>	Mixed Layer Depth
<b>MOS</b>	Model Output Statistics
<b>MSE</b>	Mean Squared Error
<b>NCEP</b>	National Centers for Environmental Prediction
<b>NEMO</b>	Nucleus for European Modelling of the Ocean
<b>NGR</b>	Non-homogeneous Gaussian Regression
<b>NPPF</b>	Nudging Proposal Particle Filter
<b>NWP</b>	Numerical Weather Prediction
<b>OSI SAF</b>	Ocean and Sea Ice Satellite Application Facility
<b>OSTIA</b>	Operational Sea Surface Temperature and Sea Ice Analysis
<b>OU</b>	Ornstein-Uhlenbeck process
<b>PDF</b>	Probability Density Function
<b>PF</b>	Particle Filter
<b>RCP</b>	Representative Concentration Pathway
<b>RMI</b>	Royal Meteorological Institute
<b>RMS</b>	Root Mean Square
<b>SIC</b>	Sea Ice Concentration
<b>SIE</b>	Sea Ice Extent
<b>SIR</b>	Sequential Importance Resampling
<b>SST</b>	Sea Surface Temperature
<b>UCL</b>	Université Catholique de Louvain
<b>ULg</b>	Université de Liège
<b>VP</b>	Viscous-Plastic rheology

## Summary

The last decades are characterized by contrasting behaviors of the sea ice in the two polar regions of the Earth (Turner and Overland, 2009). While the sea ice in the Arctic has been largely shrinking, the extent of sea ice surrounding Antarctica displays an increase estimated to be between 0.13 and 0.2 million km<sup>2</sup> between November 1978 and December 2012 (Vaughan et al., 2013). The recent study of Eisenman et al. (2014) suggests that the magnitude of this positive trend may have been overestimated due to a change in the algorithm used to process the satellite data. However, even the lowest estimate of the trend in Antarctic sea ice extent over the last decades indicates a slight increase that is rather puzzling in a global warming context.

The evolution of the Antarctic sea ice is driven by the combination of different mechanisms involving both external forcing and internal variability of the system. For instance, the stratospheric ozone depletion (Solomon, 1999) has been pointed out as a potential cause of the increase in sea ice extent. Nevertheless, this hypothesis is not compatible with several recent model analyses (e.g. Bitz and Polvani, 2012; Sigmond and Fyfe, 2010, 2013; Smith et al., 2012) and the response of Antarctic sea ice to ozone depletion may involve complex mechanisms that require further investigations (Ferreira et al., 2015).

The observed evolution of Antarctic sea ice cover may be associated to changes in the atmospheric circulation or in the ocean stratification (e.g., Bitz et al., 2006; Goosse and Zunz, 2014; Goosse et al., 2009; Holland and Kwok, 2012; Kirkman and Bitz, 2010; Landrum et al., 2012; de Lavergne et al., 2014; Lefebvre and Goosse, 2008; Stammerjohn et al., 2008; Zhang, 2007). For instance, the melting of the Antarctic ice shelves (e.g., Pritchard et al., 2012; Rignot et al., 2008; Shepherd et al., 2012; Velicogna, 2009) and the increase in precipitation at high southern latitudes resulting from the enhanced hydrological cycle (e.g., Liu and Curry, 2010) under global warming conditions may lead to a freshening of the surface of the Southern Ocean. This freshening induces a stronger vertical stratification of the ocean that in turn reduces the exchange of heat between the relatively warm intermediate layer and the colder upper layer of the ocean. This reduction of the vertical heat flux in the ocean favors the formation of sea ice at the surface and can thus account for the observed expansion of Antarctic sea ice.

Those changes in oceanic and atmospheric circulation and in Antarctic sea ice cover have been tentatively attributed (at least partly) to the multi-decadal, internally driven, variability of the system (e.g., Gagné et al., 2015; Mahlstein et al., 2013; Polvani and Smith, 2013; Swart and Fyfe, 2012; Zunz et al., 2013). In particular, the observed increase in sea ice extent since 1979 may have been preceded by a large decrease in ice extent during the 1960's (e.g., Cavalieri et al., 2003; Cotté and Guinet, 2007; Curran et al., 2003; Gagné et al., 2015; de la Mare, 1997, 2009). This hypothesis is supported by the few observational data that are available prior to 1979. Nevertheless, the time period for which reliable observations of the Antarctic sea ice are available is too short to properly investigate the internally driven change in sea ice cover. In this context, the results of climate model simulations constitute a complete set of data that can compensate for the lack of observations. Unfortunately, climate models often display large biases in the Southern Ocean for both the mean state and the variability of the system (Arzel et al., 2006; Bracegirdle et al., 2008; Mahlstein et al., 2013; Zunz et al., 2013).

The research activities undertaken within the framework of the PREDANTAR project first aimed at making the most of imperfect models and incomplete observations in order to improve our understanding of the complex mechanisms that rule the changes in Antarctic sea ice and to perform better predictions. Coupled climate models are here used to identify the processes implied in the observed changes in the Southern Ocean. Through the present project, post-processing tools providing an assessment of model errors and corrections were

developed and applied for both simple models and more complex ones. Additionally, data assimilation techniques were implemented in order to obtain optimal reconstructions of the Antarctic sea ice cover. Those reconstructions constitute valuable estimates of the changes in the state of the ice cover in the Southern Ocean over the last 30 years that compensate for the lack of observations over that period. Based on this improved understanding of the drivers of the changes in Antarctic ice cover, a methodology has been developed in order to improve the predictions and the projections of the changes in that region. In particular, the impact of the initialization of prediction simulations with a state obtained thanks to different data assimilation techniques was assessed showing how a reasonable skill in predicting decadal trend in sea ice extent may be achieved.

Calibration or post-processing aims at improving the predictions once they have been issued by the model. The purpose is to correct either the impact of model errors on the predictions and/or the evaluation of the uncertainty associated with these forecasts. These improvements are strongly related with the concepts of resolution and reliability, for which the predictions must be as close as possible to the truth (high resolution) and at the same time, their uncertainty estimation must be a good measure of the error (high reliability). In the context of the project, calibration techniques have been developed for both deterministic and ensemble predictions and have been proven successful provided the correction is performed for skillful forecasts and for short times. Moreover, the linear post-processing approach enables us to disentangle the importance of model errors and their potential origin in short term predictions of the NEMO-LIM coupled ice-ocean model used to build a Southern Ocean re-analyses.

These techniques have also been used in the perspective of the correction of long-term predictions under static or transient external climate forcing in both reduced-order idealized climate models and in an intermediate complexity climate model (LOVECLIM). As it turns out, transient dynamics of the external forcing are affecting considerably the possibility to use post-processing (bias correction or more sophisticated approaches) under strong climate changes, and care should be taken in using the approach that should be evaluated carefully on a case-by-case basis. Furthermore, a simple bias correction has been found to provide the dominant correction for model error for long-term predictions (annual, inter-annual and decadal time scales), while more sophisticated local and non-local techniques do only provide marginal corrections.

In order to contribute to the understanding of the variability in the Southern Ocean, a reanalysis assimilating sea surface temperature, sea ice concentration and sea ice drift has been realised. As the sea ice drift is strongly related to the winds, a specific procedure for the ice drift has been adopted. The correlation between the 3-day mean surface wind field and the ice-drift is strong and this relationship was used to adjust the wind field using pseudo-wind field observations based on sea ice drift data. The wind field corrections have been independently validated to show the efficiency of the approach.

Based on this adjusted wind, a reanalysis using the coupled ice-ocean model NEMO-LIM2 for the period 1985 to 2006 using 50 ensemble members has been performed. The reanalysis was validated using the World Ocean Database. As the focus of the reanalysis is the Southern Ocean, the impact of the assimilation on the ACC (Antarctic Circumpolar Current) was also assessed by comparing the mean sea surface height of the model to the mean dynamic topography derived from various observations. The assimilation improved the position and strength of the ACC.

For such low resolution models a large part of the error is due to a bias. Exploratory approaches have been implemented in order to reduce the model bias by parameter estimation. The general idea to address this problem is to add a stochastic forcing (at first constant in time) to the dynamical equations and to estimate this forcing using data

assimilation. Simple tests with the Lorenz 96 model have confirmed the validity of this approach. Results of this approach with the NEMO-LIM2 model are also encouraging and show a possible way to reduce the bias of low resolution climate models.

The analyses of the model results provided through the 5<sup>th</sup> Coupled Model Intercomparison Project (CMIP5) have highlighted systematic biases in the mean state and in the internal variability of the simulated Antarctic sea ice cover. Nevertheless, a positive trend in ice extent over the last three decades, although being rare among the CMIP5 historical simulations, is compatible with the internal variability simulated by the CMIP5 models. In those models, the heat supplied by both the ocean below and the atmosphere above the sea ice clearly impacts the sea ice cover. The relative contribution of those two mechanisms in determining the sea ice conditions is strongly model-dependent. Additional investigations based on the results of a model of intermediate complexity have allowed identifying a process related to ice-ocean interactions that potentially accounts for many characteristics of the recent observed changes in Antarctic ice cover. This mechanism consists of a stabilization of the water column due to the changes in the seasonal cycle of ice formation.

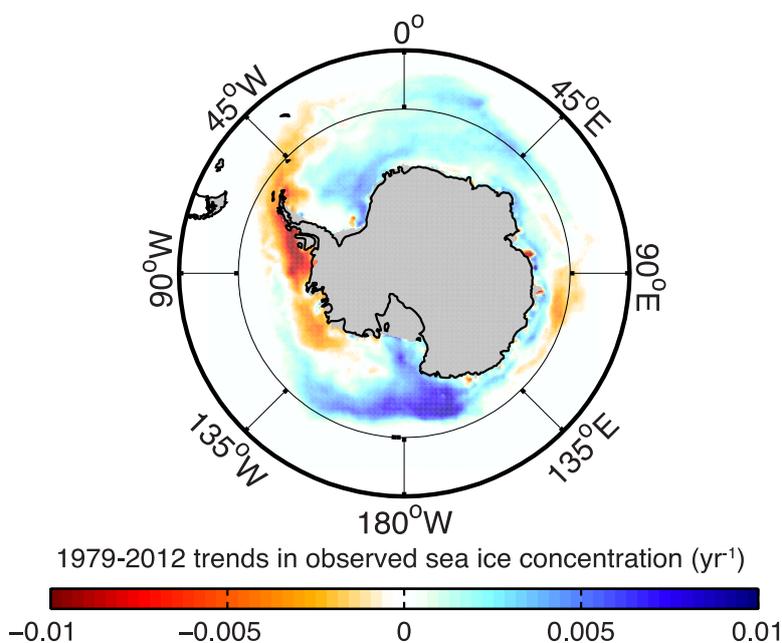
The predictive skill of an Earth-system model of intermediate complexity for the Antarctic ice cover was first assessed in idealized conditions that allow getting rid of the model biases. In this idealized study, nearly no predictability was found for the Antarctic sea ice at interannual timescales, likely because of unpredictable atmospheric processes that dominate the signal at those timescales. Besides, relatively high predictability was highlighted for the trend in sea ice extent at multi-decadal timescales. An adequate initialization of the ocean underlying the ice, achieved thanks to a data assimilation procedure, has been identified as a key element to ensure satisfying predictions of the trends in ice extent. In realistic conditions, our results indicate that the initialization of the system through data assimilation can also improve the simulated trends in ice extent and concentration over the period 1980–2009.

The Antarctic sea ice, although being a relatively small and thin ice blanket over the Southern Ocean, strongly impacts the Antarctic ecosystem and the evolution of the climate at both local and global scale. In particular, the Antarctic sea ice influences the oceanic heat exchanges, the carbon uptake and the sea level rise through interactions with the Antarctic ice sheet. Understanding its behavior and predicting its evolution thus constitute an issue of concern in a sustainable development context. The causes of the recent increase in Antarctic ice extent are still not firmly identified at this stage. Nevertheless, the knowledge related to the mechanisms that rule the Antarctic sea ice variability has been clearly improved thanks to the work carried out in the framework of the PREDANTAR project. This project also allowed testing different techniques, based on post-processing tools and on data assimilation procedure, aimed at improving the reconstructions and the predictions of the Antarctic ice cover. This will strongly contribute to improving future predictions and projections not only for the Southern Ocean but also at global scale. Furthermore, although this work was essentially focused on the sea ice in the Southern Ocean, the post-processing and data assimilation techniques implemented within the framework of PREDANTAR can be used to improve the predictability at decadal timescales in other regions and for other climate variables.



## 1. Introduction

Over the period for which we have consistent satellite records, i.e., since November 1978, the sea ice extent (defined as the surface with the limit of the ice edge, taken as the zone where sea ice concentration reaches 15%) and the sea ice area (defined as the ocean surface covered by sea ice) in the Southern Ocean have slightly increased (Comiso and Nishio, 2008; Parkinson and Cavalieri, 2012). Between November 1978 and December 2012, the increase in ice extent is estimated to be between 0.13 and 0.2 million km<sup>2</sup> (Vaughan et al., 2013). The recent work of Eisenman et al. (2014) suggests that the trend in ice extent may be in reality slightly smaller than the value reported in Vaughan et al. (2013), due to a change in the algorithm used to process the satellite records. Nevertheless, even a slight expansion of the Antarctic sea ice is in clear contrast with the behaviour of its Arctic counterpart (e.g., Turner and Overland, 2009) where the ice extent has decreased over the period 1979–2012 at a rate between 0.45 and 0.51 million km<sup>2</sup> per decade (Vaughan et al., 2013). The overall increase in Antarctic ice extent is the result of an increase in ice concentration in most parts of the Southern Ocean, particularly strong in the Ross Sea sector, and a decrease in some other sectors, especially in the Bellingshausen and Amundsen Seas (Figure 1).



**Figure 1** 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).

The processes responsible for the positive trend in Antarctic sea ice extent observed over the last three decades are still not well known. Several hypotheses have been proposed to explain this positive trend, partly attributing it to changes in the atmospheric circulation or in the oceanic stratification that would impact the sea ice transport and the heat exchanges between the atmosphere, the ocean and the sea ice (Bitz et al., 2006; Goosse and Zunz, 2014; Goosse et al., 2009; Holland and Kwok, 2012; Kirkman and Bitz, 2010; Landrum et al., 2012; de Lavergne et al., 2014; Lefebvre and Goosse, 2008a; Stammerjohn et al., 2008; Zhang, 2007). However, no clear conclusion has been obtained yet regarding the relative importance of various mechanisms.

Another key issue is to determine to what extent the observed changes in the Antarctic sea ice are driven by anthropogenic forcing or are part of the natural multidecadal variability of

the system. In response to the increase in greenhouse gas concentrations in the atmosphere and to the decrease in stratospheric ozone concentration, general circulation models (GCMs) simulate a quite robust increase in the Southern Annular Mode (SAM) index and modifications of the surface pressure in the Bellingshausen-Amundsen sector (e.g., Arblaster and Meehl, 2006; Turner et al., 2009). Their response to the greenhouse gas forcing also implies an increase in precipitation at high latitudes that influences stratification there. In the simulations provided by coupled GCMs, the sea ice response to the increase in greenhouse gas forcing is mainly characterized by a reduction of the extent over the last three decades (e.g., Arzel et al., 2006; Mahlstein et al., 2013; Parkinson et al., 2006; Turner et al., 2013; Zunz et al., 2013). However, the observed positive trend in ice extent is compatible with the natural variability simulated by those GCMs (e.g., Mahlstein et al., 2013; Polvani and Smith, 2013; Zunz et al., 2013).

Models that participated to the CMIP3 (3rd Coupled Model Intercomparison Project) had large biases in the Southern Ocean both for the mean state and the variability of the system (e.g., Arzel et al., 2006; Bracegirdle et al., 2008). Those biases remain present in more recent versions of the coupled climate models involved in the fifth phase of the intercomparison project, CMIP5 (e.g., Turner et al., 2013; Zunz et al., 2013). Those biases raise the question whether the models could reasonably be used to estimate the future changes in the Antarctic sea ice cover, at decadal and longer timescales. Improving the model physics is the most obvious way to reduce those biases. However, this is a long term goal. A complementary way is to try to reduce the error in predictions using available model results and observations. For instance, few attempts have been made to combine the projections provided by different models in order to reduce the uncertainties on future changes, giving to each model projection a weight related to the ability of the model to simulate the present state of the system (e.g., Bracegirdle et al., 2008).

This problem is classical as any model-based forecast of environmental processes progressively degrades due to the presence of both model uncertainties and initial condition errors. This has been studied in detail in meteorology where, for instance, the combined effect of these errors has been recently investigated (Nicolis et al., 2009), with emphasis on the typical time scales at which the mean square error reaches its minimum and on the time at which the contributions of both errors are of similar magnitudes. Both errors are also affecting climate forecasts, at seasonal, interannual and decadal time scales (Meehl et al., 2009). A solution proposed in Numerical Weather Predictions (NWP) to partly correct this decrease of skill, is to use post-processors based on (linear or nonlinear) statistical methods (see e.g., Wilks, 2006). These are usually referred to as Model Output Statistics (MOS) techniques. One of the most popular approaches in the context of weather forecasts consists in building a linear regression between a set of predictors provided by the NWP model and an observable at a certain lead time  $t$ , and to use this statistical relation to perform corrected forecasts of this observable at the same lead time (e.g., Taylor and Leslie, 2005). In this framework, several new approaches have been proposed in recent years for both single and ensemble forecasts (e.g., Vannitsem and Hagedorn, 2011; Vannitsem, 2009). However, such methods have not yet been applied to the climate predictions and more specifically for the Southern Ocean sea ice.

MOS techniques are efficient tools to correct model prediction of future climate. For model simulations of the past climate, observations can be directly used to constrain the model using data assimilation. Ensemble assimilation schemes such as the Ensemble Kalman Filter (EnKF, Evensen, 2003) and the singular evolutive interpolated Kalman (SEIK) filter (Pham, 2001) have been shown to be successful in the context of a coupled ocean-sea ice model (Lisæter et al., 2003, 2007; Rollenhagen et al., 2009). In particular, this approach can contribute to (1) the identification of error sources and (2) the derivation of an improved estimate, not only of the model state but also of the atmospheric forcing fields and uncertain model parameters. A better estimate of the state of the system is essential to analyze the

mechanisms, while optimizing uncertain model parameters will have also the potential to improve model predictions for which there are obviously no observations to constrain the model results.

As briefly presented above, our understanding of the complex mechanisms that rule the changes in Antarctic sea ice is still very fragmentary. Unfortunately, the small amount of observations of some variables like sea ice thickness and the large biases of general circulation models in the Southern Ocean reduce our ability to investigate the dynamics of the system and to propose reliable predictions and projections of the future changes in Antarctic sea ice. In this framework, we will use the term projections for estimates of the changes of the state of the system at the horizon of the 21<sup>st</sup> century while the word prediction will be applied for shorter term forecasts, at decadal scale. For projections, the uncertainties in the model formulation and model parameters as well as in the scenario of future concentration of greenhouse gases, aerosols, various pollutants in the atmosphere, land use, etc, play a large role in explaining the uncertainty range of the estimates. Besides, for prediction, the choice of the scenario plays a smaller role while internal variability (and thus the initial conditions of the forecast) could be a large source of uncertainty (e.g., Hawkins and Sutton, 2011).

The goal of this project is firstly to improve our understanding of the mechanisms responsible for the recent changes in the Antarctic sea ice cover. Secondly, based on this improved understanding, a methodology designed to improve the predictions and the projections of the sea ice changes in the Southern Hemisphere is developed. Both predictions for the next decades and the projections for the end of the 21<sup>st</sup> century are investigated.

These goals are achieved by

- i. Implementing and adapting Model Output Statistics techniques, providing an assessment of model errors in simulated variables, with a specific focus on Antarctic sea ice characteristics (Section 2.1);
- ii. Providing better estimates of the changes in the state of Southern Ocean sea ice cover over the last 30 years thanks to state-of-the-art techniques of data assimilation which are dynamically consistent and enable to obtain additional constraints on the atmospheric forcing of the sea ice cover and on the value of certain poorly known parameters (Section 2.2);
- iii. Analyzing observations, existing simulations with various types of models and new sensitivity studies (Section 2.3);
- iv. Analyzing the projected changes in the Antarctic ice cover over the 21<sup>st</sup> century and assessing the impact of the initialization method on the predictive skill of a climate model for the Antarctic sea ice at decadal timescales (Section 2.4).



## **2. Methodology and results**

### **2.1. Adaptation of post-processing tools**

Within the context of the PREDANTAR project a key challenge was to gain insight into the prediction improvement induced by calibration. Although nowadays climate-model calibration is a standard practice, much insight is still lacking, mostly due to insufficient observations that could be used to assess different calibration approaches for climate purposes. Therefore the issue is approached from a different perspective by use of twin experiments whereby both model predictions and observations are taken from climate models. Moreover a climate model of intermediate complexity, LOVECLIM (Goosse et al., 2010), is used that allows long climate integrations. Such approach allows one to draw firm statistical conclusions. An attempt to calibrate sea-ice model predictions using LOVECLIM based on observations of the satellite era was unsuccessful due to limited predictability.

However, as described in Section 2.1.1, calibration techniques were successfully applied as a diagnostic tool to find model errors using the re-analysis produced within this project. In Section 2.1.2 the limits of post-processing are explored in the absence of climate change, i.e. using constant greenhouse gas forcing. Thereby the emphasis is on development of spatial calibration methods and their skill. In section 2.1.3, on the other hand, a strong change of greenhouse gas concentrations is assumed and assessed whether the assumptions underlying common calibration techniques remain valid under such circumstances.

Calibration or post-processing aims to improve the predictions after they are produced by the model. Therefore it requires a definition of what is a good forecast. In this context, the concepts of resolution and reliability plays a prominent role. This means that the predictions must be as close as possible to the truth (high resolution) but, at the same time, the uncertainty estimation must be a good measure of the error (high reliability). Uncertainty estimation is of prominent importance both for numerical weather prediction and for climatological purposes. For instance initiatives such as the CMIP5 project are aimed to provide uncertainty estimates of the climate change by building a large multi-model ensemble. Different concepts of reliability exist for numerical weather predictions and climatology. In Sections 2.1.4.1 and 2.1.4.2 new calibration methods are presented that combine the different concepts of reliability. In Section 2.1.5 we also propose new verification methods for reliability in case uncertainty estimations are based on a single measure. The same reliability constraints played also an important role in Section 2.1.2 where new calibration methods for spatio-temporal fields are introduced.

#### **2.1.1. Post-processing as a diagnostic tool**

A technique is proposed that is able to identify model errors during the assimilation cycle as it is applied on the reanalysis for sea ice as proposed within the PREDANTAR project and discussed below in Section 2.2.1.3. This work is reported in Barth et al. (2015) that is currently under revision at Ocean Dynamics.

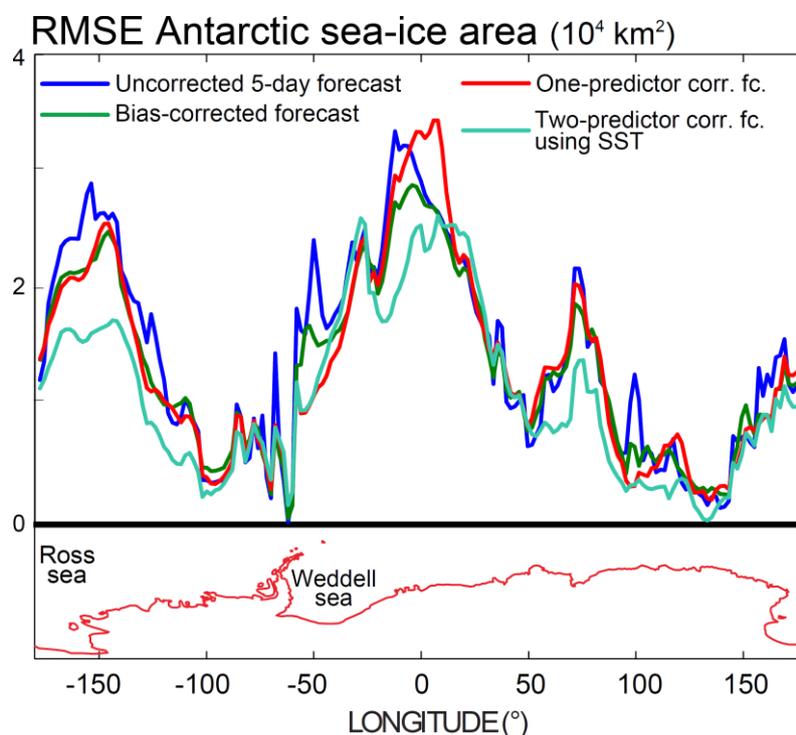
It is known that model errors can be traced by considering how data assimilation pulls the analysis away from the background towards the observations (Klinker and Sardeshmukh, 1992; Rodwell and Palmer, 2007). The problem is approached here from a different perspective.

Following discussions between the different PREDANTAR partners, the idea came up to combine post-processing and data assimilation and to test it on the sea-ice reanalysis produced within the project. Although the main purpose of post-processing or calibration is to

improve the quality of model data, we try to go beyond that and use it as a diagnostic tool for pin-pointing model errors. In Vannitsem and Nicolis (2008) it is argued that Model Output Statistics (MOS) techniques can correct forecasts at small lead times subject to systematic model errors, as opposed to random initial conditions. MOS applies regression between observations (of sea-ice extent) and different model predictors. As a first predictor, the forecasted sea-ice extent is always used. The key point, however, is the consideration of an additional predictor which is strongly correlated to the model error present. In this work we thus sought for additional predictors that strongly correct the forecast. The finding of a good additional predictor may then lead to an increased understanding of the source of error.

The data set considered consists of the ensemble-mean of the NEMO-LIM2 reanalysis (Barth et al., 2015) and the observations are taken from the OSTIA data set (Operational Sea Surface Temperature and Sea Ice Analysis, Roberts-Jones et al., 2012). In the present investigation, the error of the 5-day forecast for the total Antarctic sea ice area is found to be strongly reduced using the SST forecast as a second predictor (see **Figure 2**). Therefore SST is a predictor that is strongly affected by the modelling error. This finding constitutes a first step to the identification of the underlying modeling scheme at the origin of the model error affecting the forecast. Discussions between the project teams led to hypothesis that our findings might have to do with the parameterization of eddy-induced mixing at sub-grid scales, which, in NEMO LIM is done using the Gent-McWilliams scheme (Gent and McWilliams, 1990). This, however, needs further investigation.

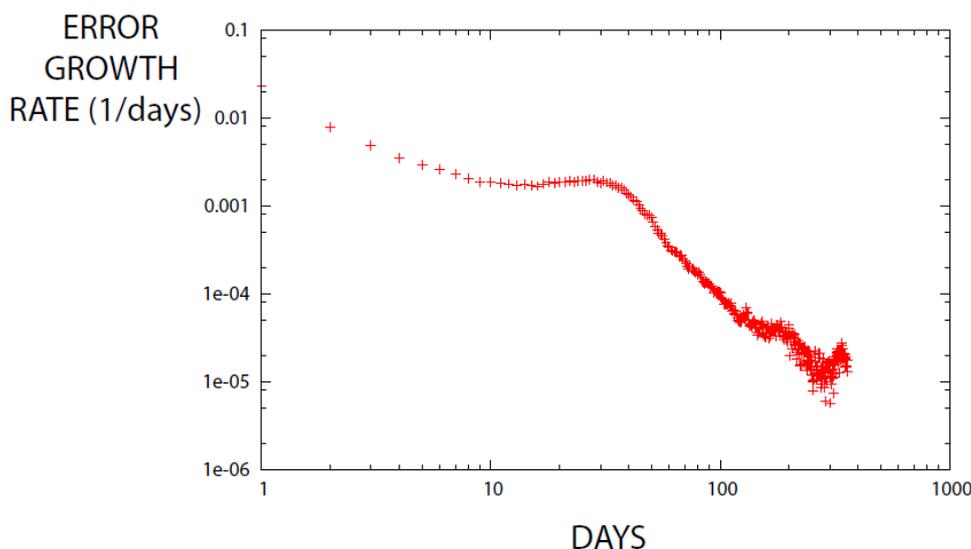
In order to disentangle model error and initial-condition error contributions, our technique was applied both at finite lead time (five days) and at lead time zero (analysis). No significant reduction of error was retrieved for the analysis, indicating that there is no obvious systematic error affecting the sea ice analysis.



**Figure 2** Root Mean Squared Error (RMSE) of the Antarctic sea ice area as a function of longitude for the uncorrected 5-day forecast and different forecasts corrected with post-processing techniques. The longitudinal spacing is  $2^\circ$  and the sea ice area at a certain longitude is the total sea ice area in a range of  $2^\circ$  East from that longitude (all south of  $50^\circ\text{S}$ ).

### 2.1.2. Post-processing of climate data in a static environment

Also, prior to investigating calibration, a limited study was performed of the predictability features of LOVECLIM, more specifically, on the error evolution of the temperature in the ocean as a consequence of small random perturbations of the atmospheric potential-vorticity field. The error growth rate against time is shown in **Figure 3**. It is observed that after the first few days of strong error growth a “linear regime” sets in featuring a constant error growth rate. This period ends after around thirty days, which coincides with the timescale of error saturation in the atmosphere. For longer lead times, nonlinear effects become important and the error growth rate smoothly vanishes.



**Figure 3** The error growth rate of the temperature in the ocean as a function of time using LOVECLIM as a consequence of an initial condition error in the atmosphere.

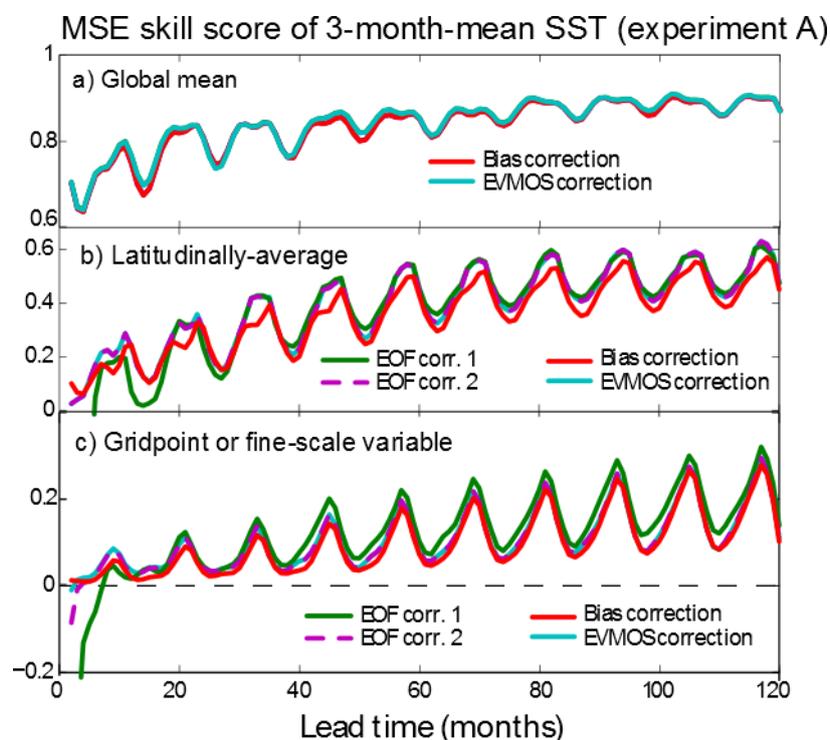
It is a common practice to correct the bias of climate runs as a consequence of errors which are systematic in nature. Three types of errors can be identified: initial condition errors, forcing errors and model errors. Intense ongoing research focuses on improving initialization and forcings that most prominently deteriorate the forecasts at seasonal and decadal timescales, respectively (Meehl et al., 2014). Model errors, however, may be important at all timescales. It is known that in meteorology model errors can be corrected using post-processing (Vannitsem, 2009b) but it remains unclear to what extent these tools can be extended for climate variables.

Based on twin experiments we investigate the following questions: Can we improve upon the bias correction? Are spatially-averaged variables better corrected? Are monthly-averaged variables better corrected than yearly-averaged variables? Is it better to perform calibration based on local data than based on non-local data? What is correction dependence on lead time?

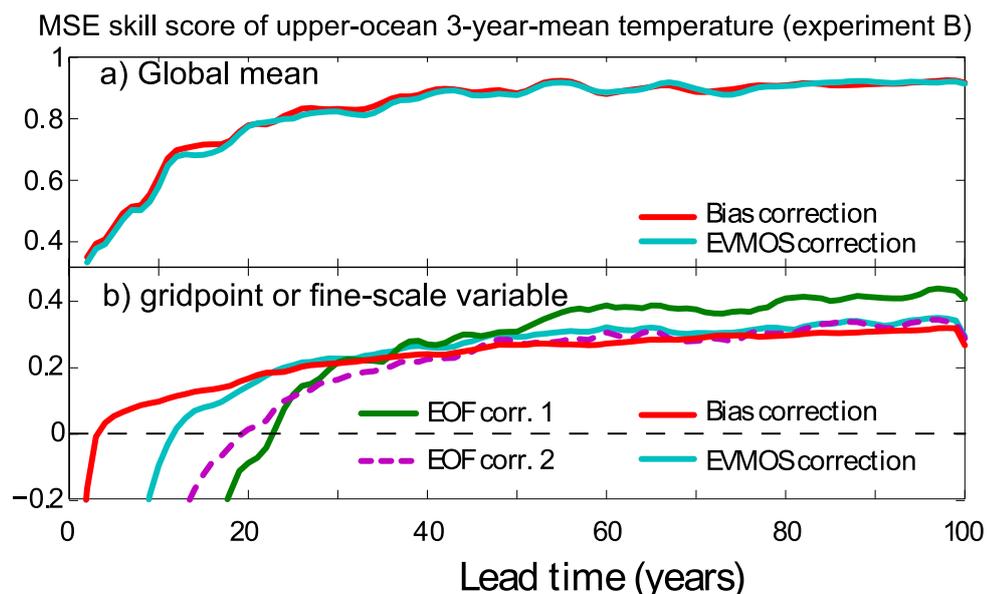
In order to answer these questions a twin experiments is performed using LOVECLIM. Forecast and reference have identical initial conditions but slightly different model parameters. In the two proposed experiments, only deterministic experiments are considered: in experiment A, one thousand predictions with ten years of lead time are conducted using a model error in the radiation scheme. In experiment B, with one hundred forecasts of one hundred years of lead time there is a model error on the drag parameter that couples ocean and atmosphere.

The proposed calibration methods can be divided into *local* correction schemes and *non-local* schemes. The former methods include bias correction and EVMOS and are local in the sense that they act on single quantities separately, without taking into account correlations with other. Two non-local methods are proposed using specific projections of the empirical orthogonal functions (EOF) and can be considered variants of the methods presented in Di Giuseppe et al. (2013). All methods correct both climate mean and variance and were applied on different averages of sea surface temperature (SST). The corrected variables are global means, latitudinal and longitudinal averages, averages over different sectors and finally, the fine-scale variables (grid-point values). Moreover for the purpose of estimating the calibration parameters, different of grid-point subsets are used. More specifically, the regression parameters are assumed identical for all spatial points, for all points of same latitude or longitude, for all points of same sector, and for the same location only.

Verification was performed using cross-validation and analyzing the error growth and its correction on seasonal up to decadal scale using the mean squared error skill score shown in **Figure 4** and **Figure 5**. As expected the skill gain is largest for globally-averaged quantities and smallest for the grid-point variables. However, improvements upon the bias correction are small but most prominent for the fine-scale variables. More specifically for the grid-point variables the first EOF correction method is the best only after a certain period (15 months in **Figure 4c** and 35 years in **Figure 5b**). The second EOF method gives equivalent results to EVMOS for seasonally-averaged SST. Note also that bias corrections are largest in the tropics while additional variance corrections are most pronounced at the poles.



**Figure 4** Mean squared error skill score against lead time for different spatial averages of temperature and for different correction methods. Positive values indicate improvements with respect to the uncorrected predictions.



**Figure 5** Same as **Figure 4**.

From this study it is concluded that improvements upon the bias correction are small overall but largest for local quantities. Skill gain is larger for global quantities and smaller for local quantities and corrections for monthly-averaged and yearly-averaged variables behave qualitatively similar. Unfortunately calibration based on non-local information are only better when taking correlations into account (EOF corrections) and only at larger lead times. Lastly the lead-time dependence of corrections is weak and mostly due to transient climate response.

### 2.1.3. Post-processing in a changing environment

Large discrepancies are found between model climatologies of General Circulation Models for the 20th century and the observed one. Therefore, upon consideration of future climate projections it is common to apply a bias adjustment, or, in other words, to consider only anomalies with respect to the different climatological averages (Meehl et al., 2014). Although useful, one must be aware of the underlying assumptions of bias correction. One such assumption is that the bias is stationary in time and is tested here in Section 2.1.3.1 using low-order models and in Section 2.1.3.2 using LOVECLIM.

#### 2.1.3.1. Studies with simple models

Seasonal and climate forecasts are affected by errors originating from both the initial conditions and the model uncertainty, in particular associated with the coupling between the different climate components. We have first investigated the dynamics of both errors in the context of a low-order moist climate model, and identify the different error-growth regimes depending on the respective amplitudes of both errors. We also relate this dynamics to the Lyapunov instabilities of the system. The coupled atmosphere-ocean (slab) model used has been developed by Lorenz (1984). It is a low-order model containing only a few key processes essential in the dynamics of the climate of a global atmosphere. It is a moist general circulation model including total water as a prognostic variable. The surface is an ocean which exchanges water and heat through evaporation and precipitation. The circulation is driven by solar heating and the thermodynamics of water is included. The model is reduced to 27 variables. An analysis of the post-processing technique has then been

performed when the system is experiencing errors in the coupling parameter between the ocean and the atmosphere.

The statistical and dynamical properties of bias correction and linear post-processing were investigated when the system under interest is affected by model errors and is experiencing parameter modifications, mimicking the potential impact of climate change. The analysis has been first performed for simple typical scalar systems, an Ornstein-Uhlenbeck process (OU) and a limit point bifurcation. It reveals system's specific (linear or non-linear) dependences of biases and post-processing corrections as a function of parameter modifications. The Lorenz '84 model has then be investigated, a low-order model of moist general circulation, incorporating several processes of high relevance in the climate dynamics (radiative effects, cloud feedbacks...), but still sufficiently simple to allow for an extensive exploration of its dynamics.

The analysis of the post-processing technique performed using this model reveals that post-processing can correct forecasts but up to about 10 days. This result suggests that indeed the post-processing can correct the error in this kind of system but it is limited in time for this kind of model error sources. The success is probably limited due to the relatively short time scales that are effectively involved in this system.

Our analysis show that the bias or post-processing corrections display complicate variations when the system experiences temperature climate changes up to a few degrees. This precludes a straightforward application of these corrections from one system's state to another (as usually adopted for climate projections), and increases further the uncertainty in evaluating the amplitudes of climate changes. These results are reported in Vannitsem (2011).

#### *2.1.3.2. Studies with a climate model*

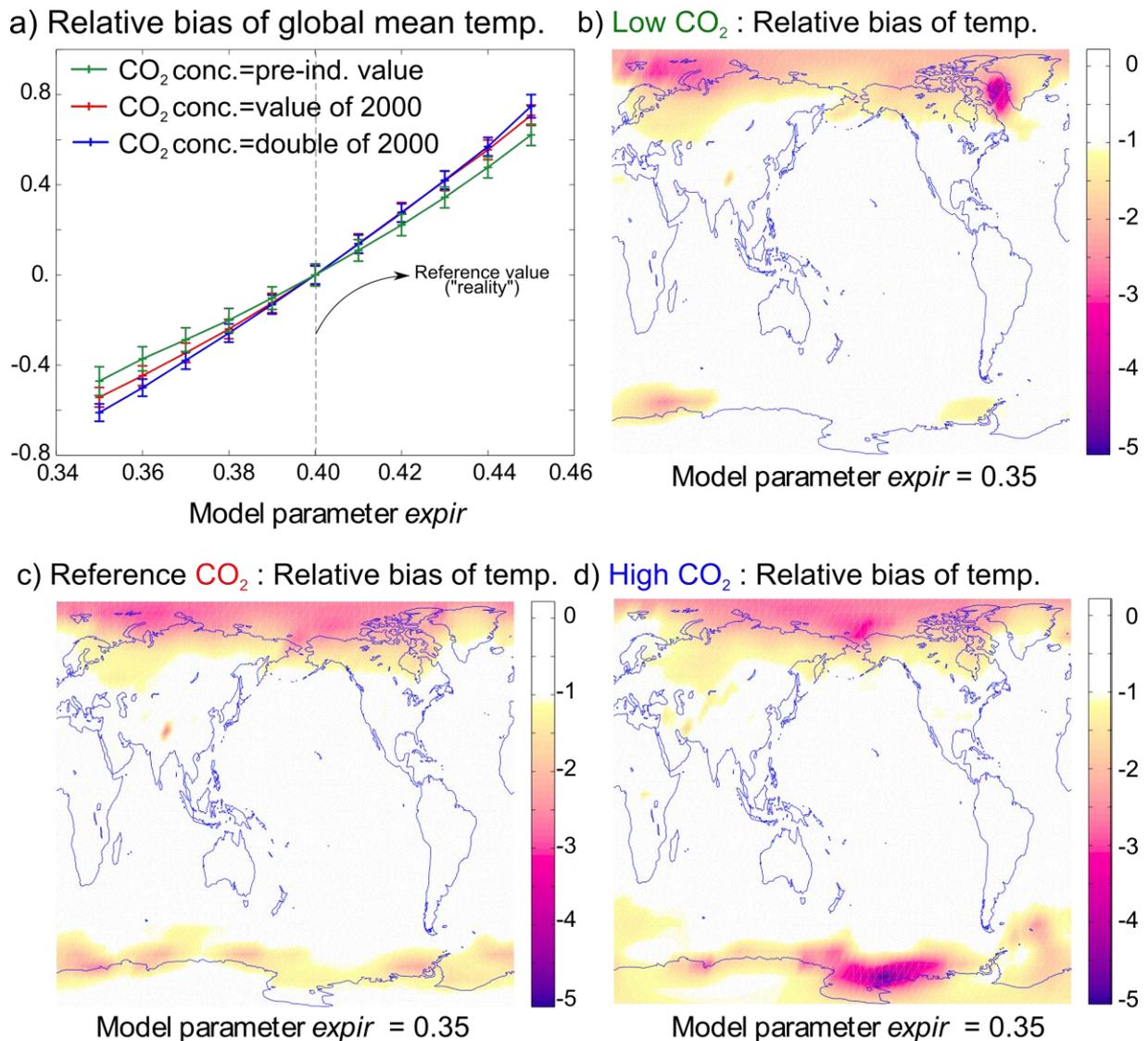
To what extent are the conclusions drawn for low-order models (Section 2.1.3.1) valid for higher-order models? More specifically: Are biases sensitive to climate changes induced by changes of greenhouse-gas concentrations? Are these changes monotonous with the concentration changes? Are bias patterns spatially similar between different climate regimes? Can *transient* climate changes be corrected? Details of this study can be found in Van Schaeybroeck and Vannitsem (2015a).

A twin experiment was set up using the model LOVECLIM whereby the "reference" climate is a model run with a fixed parameter (expir, present in the radiative scheme) while this parameter is varied for the "model" climate. For each parameter value, a 8000-years long model climatology is determined with three different climate forcings, that is, different CO<sub>2</sub> concentrations. For the pre-industrial, the reference and future climate we use respectively a CO<sub>2</sub> concentration of 276 ppm, 370 ppm and 740 ppm.

The calibration techniques used are the simplest available: bias correction (BC) and change factor (CF), which both adjust the climatological mean and have a common underlying hypothesis. More specifically both assume that the bias is independent of climate, i.e. constant as a function of forcing CO<sub>2</sub> concentration.

It is found that this assumption is reasonably well satisfied for global-mean surface temperature and precipitation (see Figure 6a). However, for temperature, spatial bias patterns differ strongly (Figure 6b-d). For almost all other variables, for instance 10m U-V wind and sea surface salinity, the hypothesis underlying BC and CF is strongly violated. Also, bias changes are non-monotonous as a function of the forcing parameter and the model error. This is consistent with the conclusions drawn from the analysis with low-order models.

So far our conclusions were based on the equilibrium climate sensitivity in the sense that the "climate" is taken from equilibrated model runs with constant forcing. Therefore the influence of the transient climate response is neglected. In order to correct the bias under transient climate change, the bias evolution must be predicted. An approach based on linear response theory is presented. Thereby, based on the CO<sub>2</sub> concentration and the climate from one time period (say past 30 years), the climate response to another (future) CO<sub>2</sub> scenario can be reconstructed. However, our results indicate that the assumptions underlying the linear-response theory must be checked case. For instance at intermediate times the theory seems to work reasonably well for global-mean precipitation but does not work for temperature.



**Figure 6** Panel a: the relative bias of global-mean surface temperature as a function of the model error for the three different CO<sub>2</sub> forcings. Here "relative" means it is divided by the climate change signal. The bars show the standard deviations of the annual-mean climatological distributions. The changing slope of the three climates can be considered monotonous as a function of the CO<sub>2</sub> concentration. Although global biases are relatively similar among the three climates, panels b-d show that their spatial patterns differ strongly.

## 2.1.4. Post-processing of ensemble predictions

### 2.1.4.1. Ensemble post-processing: spread-conserving calibration

Since long forecast calibration is a common practice for both numerical weather predictions (NWP) and climate predictions. The simplest calibration method for NWP is called Model Output Statistics (MOS) and is based on statistical features of a prior comparison between model output and observation (Glahn and Lowry, 1972). The most widespread approach LMOS applies ordinary least-squares regression to relate observations and model predictions. It corrects the mean but it tends to degrade the variability. Such degradation happens especially at long lead times when the observation and the prediction become largely uncorrelated. This feature precludes the application of LMOS for ensemble forecast but also for climate projections.

In order to overcome this problem mostly "statistical" calibration approaches were proposed that involve drastic simplifications concerning the statistical properties of ensembles. A strong disadvantage of such "statistical" methods is the requirement of extra processing for reestablishing realistic (spatio-temporal) correlation structures. Within the spirit of the older LMOS methodology and therefore without requirements of such extra procedures, Vannitsem (2009) proposed a new method called EVMOS that can be applied for both deterministic and ensemble forecasts. This method was tested on the ensemble forecasts of ECMWF (Vannitsem and Hagedorn, 2011b). Along the same lines of Van Schaeybroeck and Vannitsem (2011) introduced various post-processing techniques that use linear regression and assessed their quality based on different criteria. More specifically a EVMOS-corrected forecast satisfies the following constraints: the new forecast is unbiased and its variance agrees with the one of the observations. This last criterion is also called *climatological reliability* or marginal calibration.

Note that post-processing software was developed and distributed among the project partners. Also, based on the ECMWF verification the operational implementation of the spread calibration of the ECMWF forecast at RMI, including the use of multiple predictors, has been approved and will soon be provided as a product to the forecasters (Van Schaeybroeck and Vannitsem, 2012). Moreover the techniques were used in two master theses at UCL.

### 2.1.4.2. Ensemble post-processing: spread-adjusting calibration

The different criteria for post-processing and different concepts of reliability led to an approach that is able to impose any (reliability) constraint desired for the calibrated prediction. This allowed introducing a constraint that affects the ensemble-spread reliability, or, in other words, the reliability of the predicted uncertainty estimation. A reliable ensemble forecast is characterized by the fact that the observation may be considered as a member of the ensemble forecast and hence has the same statistical properties including mean and variance. Therefore a general framework is devised by which, based on the different constraints imposed, an optimal correction is performed (Van Schaeybroeck and Vannitsem, 2015b).

Using simple models that exhibit chaotic behavior (the Kuramoto-Sivashinsky equation and the spatially-extended Lorenz '96 system) and using the ECMWF ensemble forecasts for Belgium, these methods were tested and the best among them were selected (Van Schaeybroeck and Vannitsem, 2013; Van Schaeybroeck et al., 2015). Like the EVMOS technique, all proposed methods are so-called *member-by-member* calibration methods, as opposed to the statistical calibration methods, in the sense that the different ensemble members are corrected separately. Therefore if one requires the combined calibration of

forecasts at different stations or lead times, the spatial or temporal correlation structure of the individual members is preserved. Hence no extra treatment for reestablishing realistic (spatio-temporal) correlation structures, rank structures and outliers is required, as opposed to forecasts corrected with "statistical" methods (Schefzik et al., 2013).

Apart from these benefits, in terms of skill, our methods are at the same level as the state-of-the-art statistical methods (Gneiting et al., 2005) and therefore much better than the traditional ensemble-mean correction schemes. Our technique has additional benefits: it avoids the under-correction of ensembles with small spreads, our approach is computationally cheap, it has the correct variance and no underlying Gaussian assumptions such that ensemble skewness and kurtosis are well preserved. Therefore, for practical purposes, member-by-member methods are preferable over statistical ensemble methods.

#### *2.1.4.3. Calibration of ensemble predictions in a hydrological context*

An analysis of the correction of precipitation forecasts has been performed. Extended logistic regression is used to calibrate areal precipitation forecasts over two small catchments in Belgium computed with the European Centre for Medium-range Weather Forecasts (ECMWF) Ensemble Prediction System (EPS) between 2006 and 2010. The parameters of the post-processing are estimated from the hindcast database, characterized by a much lower number of members (5) than the EPS (51). The parameters have therefore to be corrected for predictor uncertainties. They have been fitted on the 51-member EPS ensembles, on 5-member sub-ensembles drawn from the same EPS, and on the 5-member hindcasts. For small ensembles, a simple "regression calibration" method by which the uncertain predictors are corrected has been applied. The different parameter sets have been compared, and the corresponding extended logistic regressions have been applied to the 51-member EPS. The forecast probabilities have then been validated using rain gauge data and compared with the raw EPS. In addition, the calibrated distributions are also used to modify the ensembles of precipitation traces. The post-processing with the extended logistic regression is shown to improve the Continuous Ranked Probability Skill Score relative to the raw ensemble, and the regression calibration to remove a large portion of the bias in parameter estimation with small ensembles. With a training phase limited to a 5-week moving window, the benefit lasts for the first two forecast days in winter and the first five or six days in summer. In general, substantial improvements of the Mean Error and of the Continuous Ranked Probability Score have been put in evidence. This work is reported in Roulin and Vannitsem (2012, 2015).

#### **2.1.5. Reliability verification of ensemble predictions based on spread**

Ensemble forecasts are often reduced to two forecast identities: the "forecast" itself based on ensemble mean, and the uncertainty forecast based on ensemble spread. All other information is then left unused. It is known that ensemble forecasts improve upon deterministic forecast since the ensemble mean provides reduced forecast error as compared to deterministic forecast. However, the usefulness of spread as a measure of the forecast uncertainty is less clear as even the assessment of such usefulness, often called spread-skill verification, is itself still unsettled and forms the subject of intense ongoing research (Christensen et al., 2014; Gritti and Mass, 2007; Hopson, 2014).

Therefore the following issues were addressed: Given only the ensemble spread, what is the best forecast for the uncertainty? Is spread sufficient as the sole measure of uncertainty or is there generally more information in the ensemble? How can one verify if spread is a good predictor of uncertainty? A manuscript (Van Schaeybroeck and Vannitsem, 2014) that is currently under review at *Monthly Weather Review*, describes the details of this work.

A theoretical framework is outlined to forecast the uncertainty or the error of the forecast, given the ensemble spread  $S$  only. Such prediction can be either deterministic or probabilistic. A deterministic prediction implies that the spread is the point-wise forecast for the error while a probabilistic uncertainty prediction is a distribution characterized by one variable (i.e. the spread) only. Three models are introduced for probabilistic uncertainty forecasts that are based on spread only.

Given a series of forecasts for the error and their associated observed errors, verification of such series may then be done depending on the nature of the forecast of the error. The calculation of correlations (Grimit and Mass, 2007; Hopson, 2014) and Mean Squared Errors are common for deterministic forecasts but do not assess all forecast aspects. For probabilistic forecasts the use of standard verification measures like CRPS or Brier score are proposed. Their associated *skill* scores can then be used to compare the spread-based uncertainty forecasts with the full-ensemble forecasts and thus to quantify the presence of information loss when spread is used as the sole uncertainty measure. The application on ECMWF EPS, however, indicates very limited information loss and spread models even improve the uncertainty forecast for small ensemble size.

It is common to perform spread-skill verification by showing ensemble spread against errors both using correlations or linear fitting. It is argued that such analysis is statistically unjustified since it violates the underlying statistical assumptions. However, after a logarithmic transformation of spread and error one can test if spread is good predictor for uncertainty. Application on the EPS of ECMWF shows near-perfect upper-air forecast reliability using geopotential height for lead times between two and seven days.

## **2.2. Application of advanced data assimilation methods to study Southern Ocean sea ice cover**

### **2.2.1. Southern Ocean reanalysis using assimilation of sea surface temperature, sea ice concentration and drift**

Current ocean models have relatively large errors and biases in the Southern Ocean. The aim of this study is to provide a reanalysis from 1985 to 2006 assimilating sea surface temperature, sea ice concentration and sea ice drift. In the following it is also shown how surface winds in the Southern Ocean can be improved using sea ice drift estimated from infra-red radiometers. Such satellite observations are available since the late seventies and have the potential to improve the wind forcing before more direct measurements of winds over the ocean are available using scatterometry in the late nineties. The model results are compared to the assimilated data and to independent measurements (the World Ocean Database 2009 and the mean dynamic topography based on observations). The analysis procedure used to interpolate the observation on the ORCA2 model grid has been published in *Geoscientific Model Development* (Barth et al, 2014) and the reanalysis is currently under revision at *Ocean Modelling* (Barth et al, 2015).

#### *2.2.1.1. Model and observations*

The primitive-equations model used in this study is NEMO (Nucleus for European Modelling of the Ocean, Madec (2008)), coupled to the LIM2 (Louvain-la-Neuve Sea Ice Model) sea ice model (Fichefet and Morales Maqueda, 1997; Timmermann et al., 2005; Bouillon et al., 2009). The global ORCA2 implementation is used, which is based on an orthogonal grid with a horizontal resolution of the order of  $2^\circ$  and 31 z-levels (Massonnet et al., 2013). The model

is forced using air temperature and wind from the NCEP/NCAR reanalysis (Kalnay et al., 1996). Relative humidity, cloud cover and precipitation are based on a monthly climatological mean. The sea surface salinity is relaxed towards climatology with a fresh water flux of  $-27.7$  mm/day times the salinity difference in psu.

Global foundation sea surface temperature from OSTIA (Operational Sea Surface Temperature and Sea Ice Analysis, Roberts-Jones et al., 2012) at an original resolution of  $0.05^\circ$  was reduced to a resolution of  $2^\circ$  by averaging all temperature values within a  $2^\circ$  by  $2^\circ$  grid cell.

Global sea ice fraction from the EUMETSAT Ocean and Sea Ice Satellite application Facility (OSI-SAF Roberts-Jones et al., 2012) was also reduced to a resolution of  $2^\circ$  and assimilated with an error standard deviation of 0.1. The OSTIA sea surface temperature and the OSISAF sea ice fraction are distributed by MyOcean.

Daily sea ice drift from NSIDC (National Snow and Ice Data Center) is also assimilated in the ocean model. The sea ice drift is based on data from the Advanced Very High Resolution Radiometer (AVHRR), Scanning Multichannel Microwave Radiometer (SMMR), Special Sensor Microwave/Imager (SSM/I), and International Arctic Buoy Programme (IABP) data (Fowler, 2003). As the focus of this study is the Antarctic Ocean, only data from the southern hemisphere is used. The error standard deviation for the assimilation is assumed to be 0.1 m/s. The value of this parameter was determined by a series preliminary experiments to find the right balance between correcting as much as possible the sea ice drift without degrading unobserved variables.

### *2.2.1.2. Correcting wind field using sea ice drift*

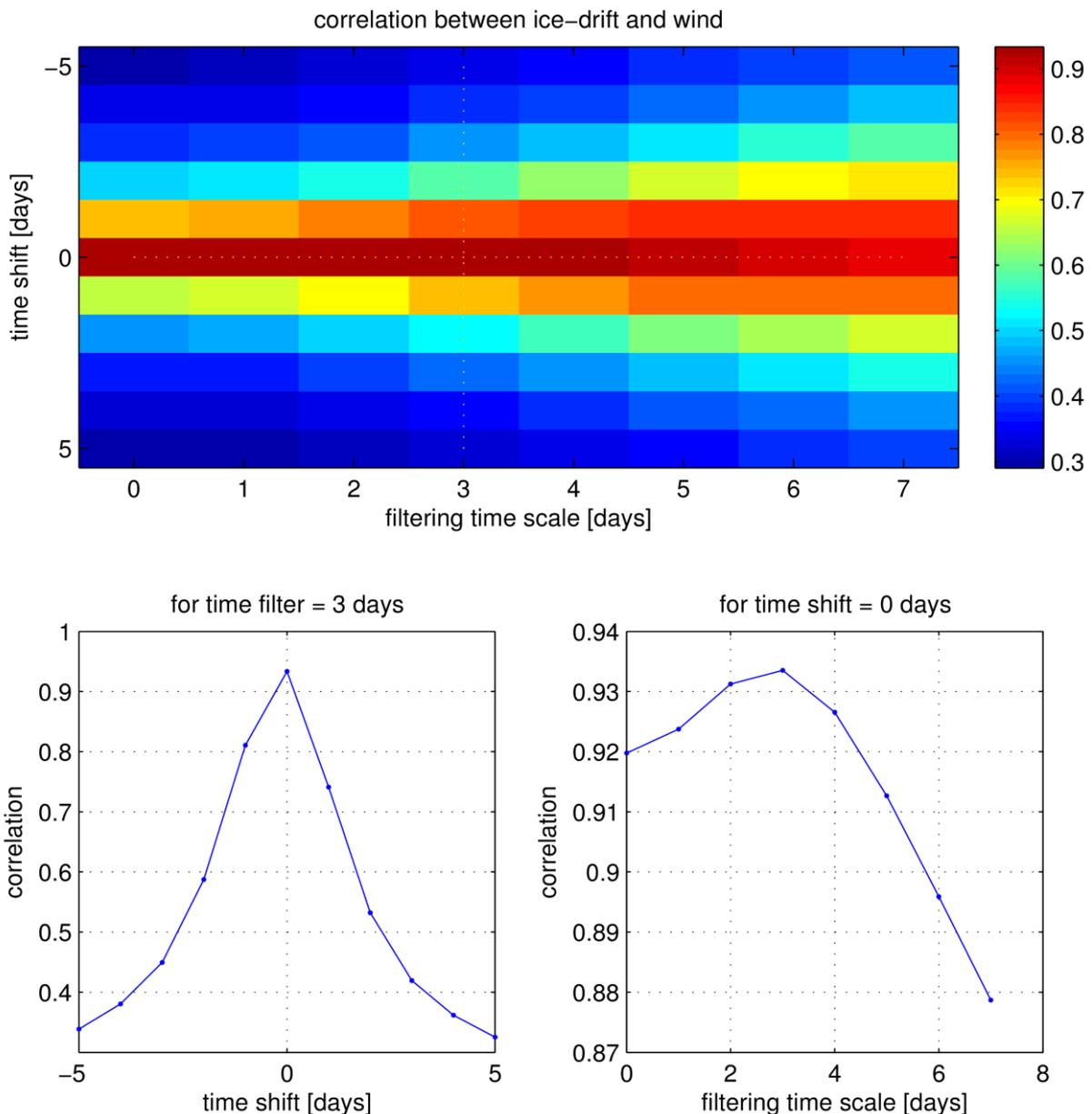
Errors in the sea ice drift can be attributed either to errors in the winds fields, errors in the ocean currents or errors in the response of sea ice to external stresses. As winds and currents have two very different time scales, a two-step approach was adopted. First, the wind fields are adjusted using sea ice drift as described in this section. In a second step the sea ice drift is assimilated into the ocean-sea ice model in order to reduce errors due to the ocean currents.

#### *Relation between sea ice drift and wind*

The model sea ice drift is strongly related to the used wind forcing. To quantify the relationship between sea ice drift and wind fields, the complex correlation coefficient between the daily NEMO-LIM sea ice drift and daily NCEP winds has been computed by introducing the following complex currents (zonal component is the real part and the meridional component is the imaginary part).

In order to maximize the correlation, we correlate the sea ice drift with different transformations of the wind field. More specifically, we use different combinations of lags and filters in time of the wind field. We use a time filter because we anticipate the sea ice drift to have a certain inertia and thus a memory of previous winds. The time lag and the temporal scale of the filter will be determined later. The time filter is implemented using an iterative diffusion scheme using a forward Euler step and a 2nd-order center diffusion operator. The complex correlation coefficient between sea ice drift and filtered and shifted wind fields has been computed. The absolute value of the complex correlation coefficient is maximized by changing the time-lag and time-filter. The complex regression coefficient derives an empirical relationship between the sea ice drift and the wind field. This relationship will be used later for wind field adjustment. The complex correlation and regression coefficients are used instead of the (real) correlation/ regression coefficient derived on the zonal and meridional

component individually because the complex coefficients can represent a rotation by a constant angle between the two vectors (as a result for the Coriolis force) and is thus commonly used to analyze horizontal velocities.



**Figure 7** Magnitude of correlation coefficient (for the year 2000) as function of time lag and filtering time scale (panel a). Panel b show the magnitude of correlation coefficient as a function of the time lag for a filtering time scale sets to 3 days (vertical dotted line in panel a) and panel c show represents the magnitude of correlation coefficient as a function of filtering time scale for a the time lag set to 0 days (horizontal dotted line in panel a).

The correlation analysis showed a strong correlation with magnitude of 0.9363 and a phase of  $-19.52^\circ$  between sea ice drift and 3-day average wind fields (panel (a) of

**Figure 7**). This phase (which is also the phase of the complex regression coefficient) represents the angle between the sea ice drift vector and the wind vector. The maximum value was obtained with no time lag. These results did not confirm the initial expectation of a time lag between wind and sea ice drift as one could assume that the wind (the cause) precedes sea ice drift (the effect). The maximum of the correlation as a function of the time

lag is very well defined while the correlation as a function of the filtering time scale is a bit flatter (panel (b) and (c) of

**Figure 7**). This strong relationship has been used to correct the surface winds. The general approach is to use the regression coefficient to transform the observed sea ice drift as pseudo wind observations and to attempt to improve zonal and meridional wind field components. In particular the following procedure has been adopted to compute the adjusted wind field:

- the first guess wind field is the NCEP reanalysis
- the model is run with this wind field (here for the year 2000)
- the sea ice drift error is calculated by comparing model with observed sea ice drift
- the sea ice drift error is transformed to "wind increment" using the regression coefficient
- "wind increment" is analyzed with the tool `divand` (detailed in the next section) on the ORCA grid and the first guess (the NCEP reanalysis) is added

While other calibration experiment are carried out for the year 1985, the wind field adjustments are first tuned for the year 2000 due to the availability of the Cross-Calibrated Multi-Platform (CCMP) Ocean Surface Wind Vector Analyses (Atlas et al., 2011) which will be used to independently validate the results.

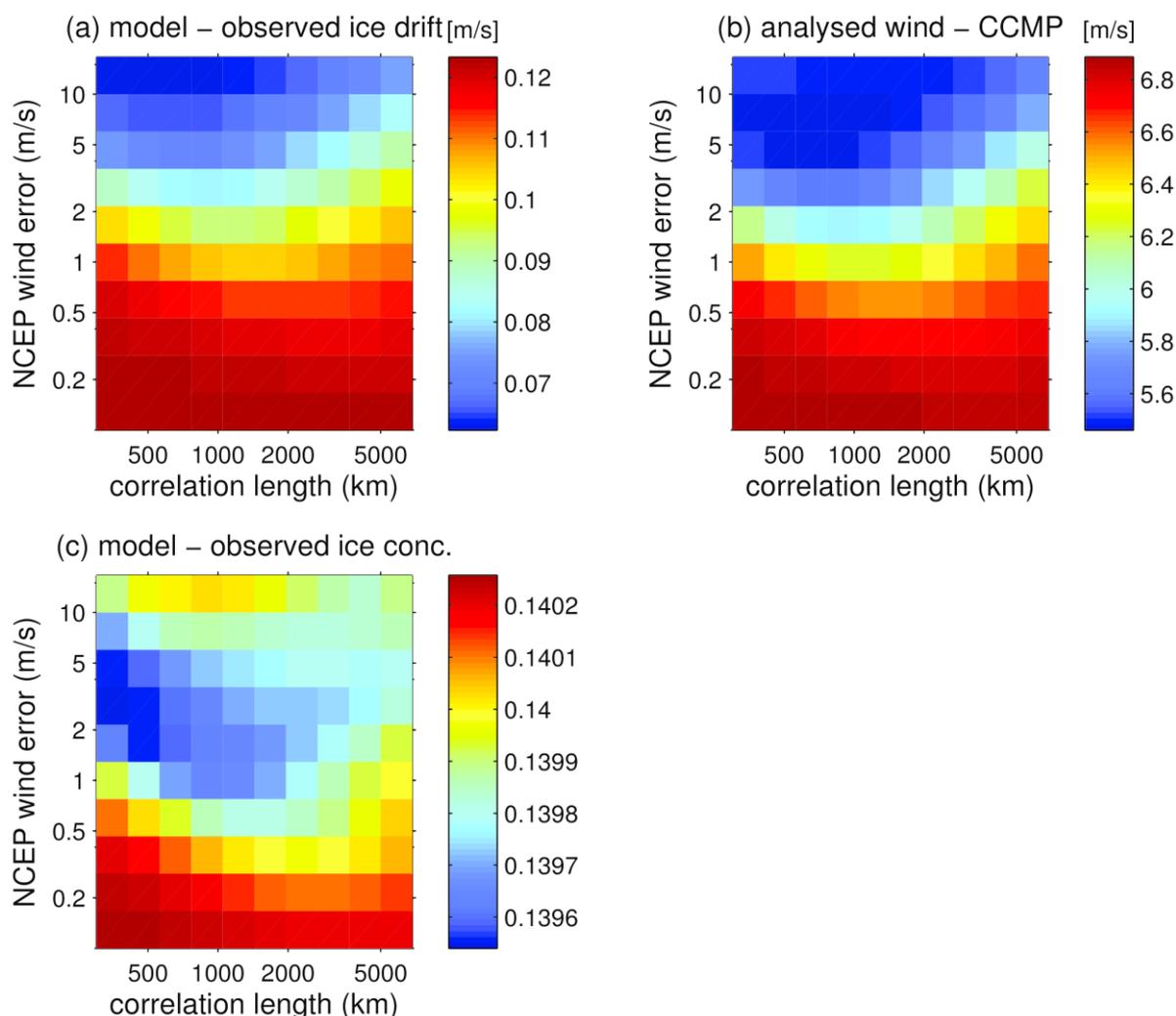
#### *Spatial analysis with `divand`*

The sea ice drift provides only information about the wind field where the model has ice. However, if the sea ice drift indicates that the model should be corrected at a particular location, one can expect that it should also be corrected in a similar way at neighboring grid cells (even if they are not covered by ice). The tool `divand` (Barth et al., 2014) (Data Interpolating Variational Analysis in n-dimensions) is used to spatially interpolate the "wind increment" derived from the sea ice drift on the full ORCA2 grid. An estimation of the background NCEP error standard deviation is necessary to define the data weight. In the present case, the pseudo observation (the wind information derived from the sea ice drift observation) error standard deviation is derived using the error standard deviation of the sea ice drift. The remaining parameter of the analysis, namely the correlation length scale and background NCEP error standard deviation, will be determined in the following. The correlation length controls the distance over which the information from the pseudo-wind observation is extrapolated spatially and the NCEP error standard deviation determines how close the analyzed field has to come to these pseudo-observations.

#### *Calibration*

The correlation length is varied from 300 km to 5000 km and the standard deviation error is varied from 0.1 m/s to 10 m/s. For each parameter 10 values are tested. These values are uniformly distributed in logarithmic space. For each of these parameters, the `divand` analysis is performed and the NEMO-LIM2 model is run simulating the year 2000 with the adjusted wind fields. Figure 8 shows the RMS error between the model sea ice drift obtained using the adjusted wind and the observed sea ice drift. This is not an independent validation since the observed sea ice drift is used to adjust the wind fields. This comparison is rather a confirmation that the adjustment works as expected. The RMS error between the model sea ice drift derived from the original (i.e. non-adjusted) NCEP forcing and ice drift observations is 0.1235 m/s. The ice drift RMS error based on modified wind fields is indeed reduced thanks to the adjustment and, as expected, the lowest RMS error is obtained when using a large value of the standard deviation of the NCEP wind error. In this case, the adjusted wind will thus be forced to come closer to the pseudo-wind observations (based on sea ice drift).

As an independent comparison the CCMP Ocean Surface Wind Vector Analyses (Atlas et al., 2011) based on ERA-40 and observations such as QuickScat for the year 2000 south of 60°S is used. Sea ice drift is not used in the CCMP product. For every tested correlation length scale and NCEP error standard deviation, the RMS error between the adjusted wind and the CCMP wind field is computed (Figure 7, panel b). This comparison shows that the wind field is indeed improved using the sea ice drift measurements. As before, the general tendency is that the RMS difference between the adjusted winds and CCMP winds decreases as the error standard deviation increases. This comparison shows that the optimal range of the correlation length scale is between 500 km and 1500 km.



**Figure 8** RMS difference of the model and observed sea ice drift (panel a), the analyzed winds and CCMP winds (panel b) and model and observed sea ice concentration (panel c) for different values of the correlation length and the NCEP wind error.

The model sea ice concentration obtained by the adjusted wind is also compared to the OSTIA/OSI-SAF observations (Figure 9, panel c). In general the sea ice concentration varies only weakly by changing the parameter of the analysis. This suggests that only a small part of the RMS error in sea ice concentration can be attributed to the wind forcing and that sea ice concentration is mostly driven by thermodynamic forcings. Contrary to the previous comparison the error slightly increases for large values of the NCEP error standard deviation. Overall a large value of the background error improves sea ice drift and reduces the RMS error in comparison with CCMP winds but it degrades the sea ice concentration (if error standard deviation is larger than 5 m/s). Correlation lengths between 500 km and 1000 km

give acceptable results. By combining the results from the different comparisons, the wind fields have been adjusted using a correlation length of 700 km and a background error standard deviation of 2 m/s.

### *2.2.1.3. Reanalysis using data assimilation*

The implemented data assimilation scheme is the Ensemble Transform Kalman Filter (Bishop et al., 2001). In ensemble-based assimilation schemes, the error statistics of the model state vector is estimated by perturbing uncertain aspects of the model. In the present configuration we perturb surface winds (10 m) and surface air temperature (2 m). The adjusted wind from the previous section are used. Atmospheric parameters coming from climatology are not perturbed. The data assimilation scheme employs an ensemble with 50 members. Observations are assimilated every 5 days which is a compromise between available computer resources and maximizing the usage of the observations.

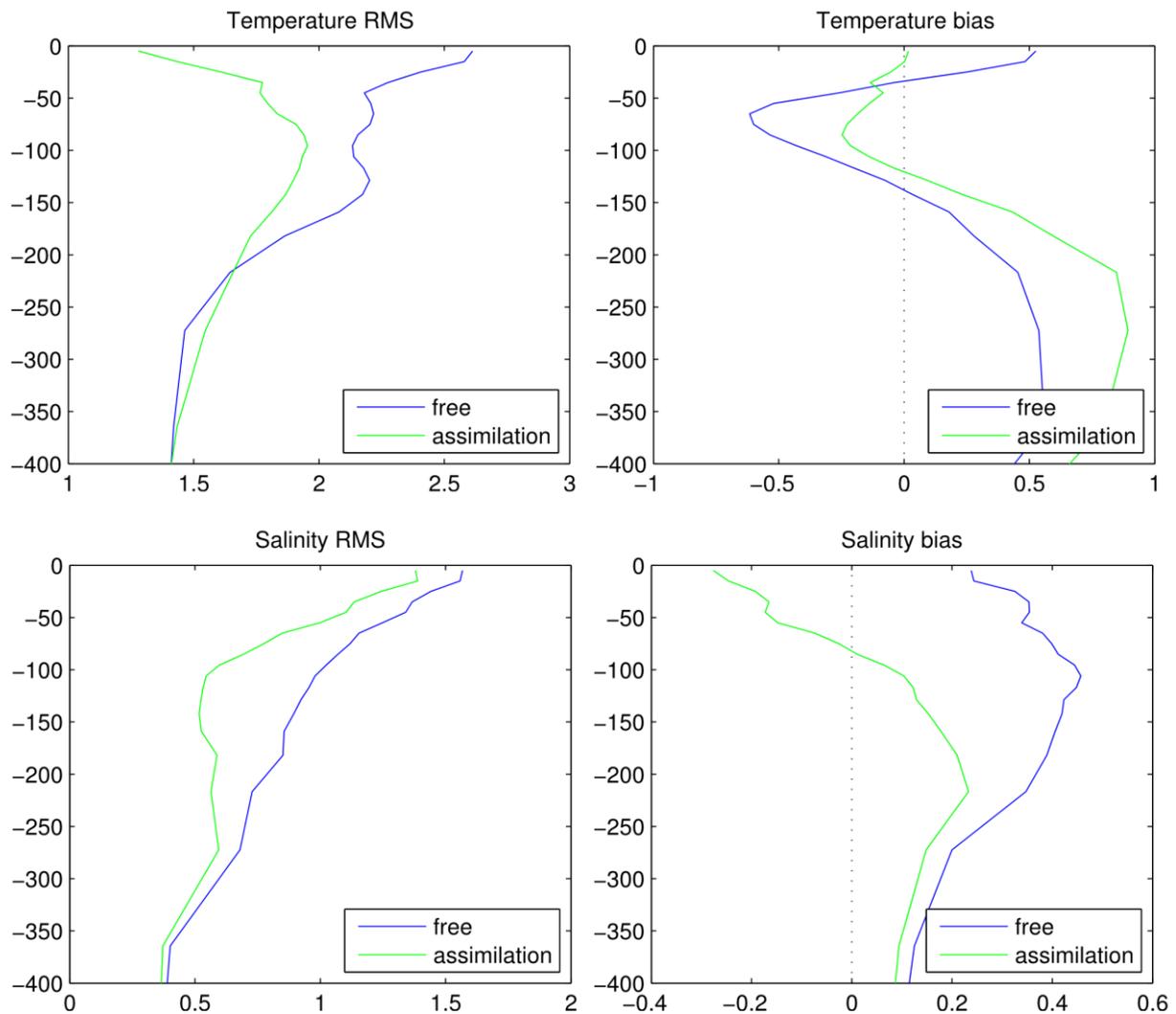
The perturbation scheme is based on a Fourier decomposition (Barth et al., 2011). Only Fourier modes with a time period between 20 and 70 days are used for the perturbations in order to exclude the seasonal variations (which have a large variance and whose amplitude is not representative for its expected error) and short-scale variations which are not the primary focus of this study. Since the perturbation scheme is multivariate, the same range of time scales is used to perturb wind and air temperature.

While the calibration of the assimilation setup was performed on a single year (2000), this section presents the model simulations with data assimilation from 1<sup>st</sup> January 1985 to 31<sup>st</sup> December 2006. The time period was determined to ensure the availability of all used data sets.

### *Validation of the reanalyzed temperature and salinity fields*

All observations of ocean temperature and salinity from the World Ocean Database from the period 1st January 1985 to 31st December 2006 have been collected. The free-running model has the largest temperature error near the surface where the model has the most variability (**Figure 9**). As the model assimilates sea surface temperature, the largest impact of the assimilation is indeed at the surface where the RMS error and bias (which is partly included in the RMS error) are strongly reduced. The RMS error is improved by the assimilation over 200 m depth and the bias over 120 m. Below those depths there is a slight degradation of the temperature which is essentially a systematic error in form of a bias. One possible way forward for improvement of the assimilation scheme could be to include a temperature relaxation toward a climatology to control such error. As the ensemble is generated by perturbing the atmospheric fields, the resulting vertical correlation scale between the surface and the subsurface level is about 100 meters (as calculated by computing the standard deviation averaged over time and horizontal space of the analysis increment). As the error increase at depth is not introduced by the analysis step, it must be introduced by the model reaction to an analyzed initial condition. In fact, it is well known that sequential analysis can produce shocks after restarting the model from an analysis (e.g. Yan et al., 2014).

The model does not assimilate salinity and therefore changes in salinity are only due to the covariance between the observed variables and salinity, and also due to the model adjustment after the analysis. The validation reveals that the assimilation reduces the salinity RMS error and bias everywhere with a diminishing impact at depth. Contrary to the temperature validation, no degradation at depth was observed.



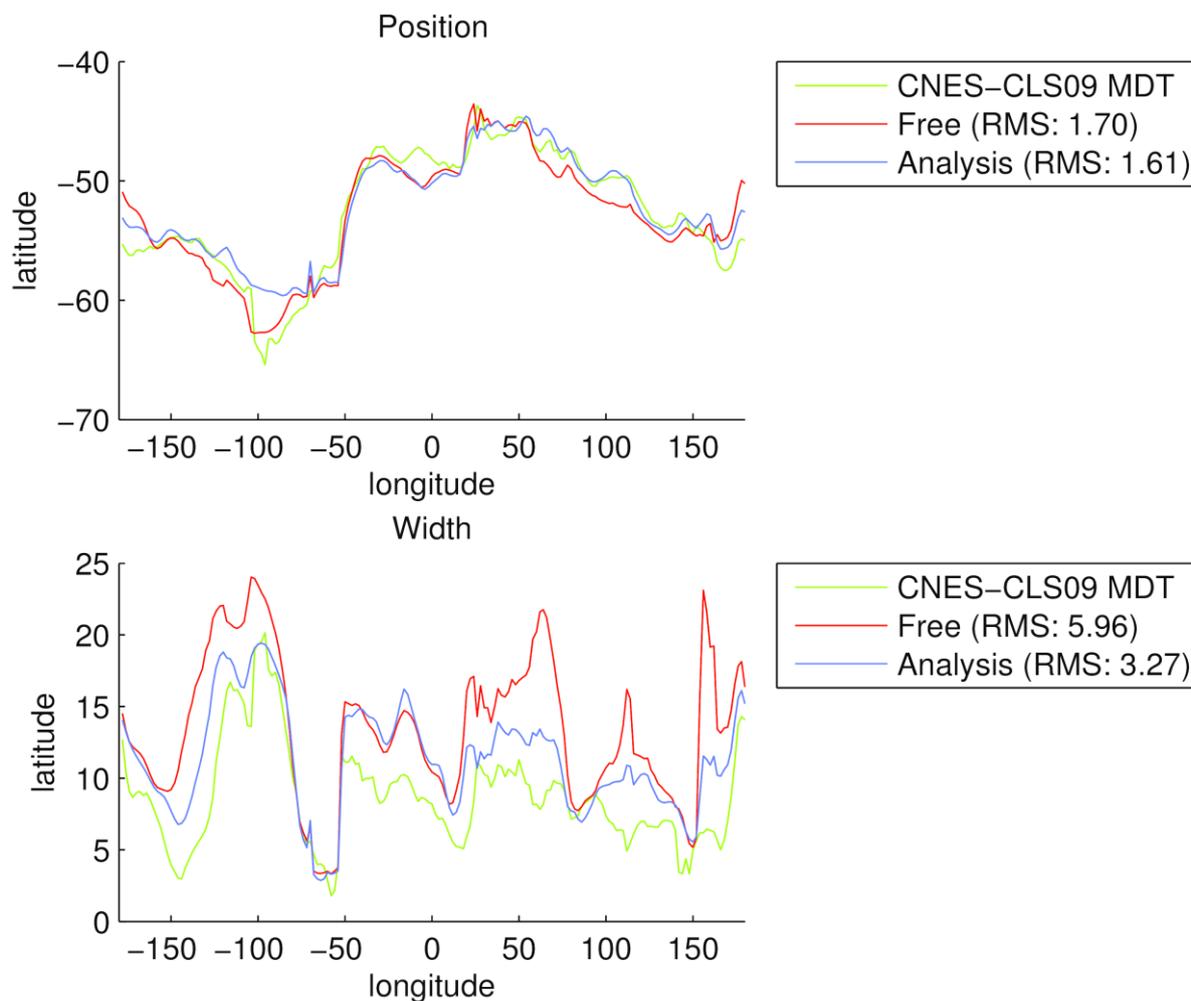
**Figure 9** Validation with World Ocean Database using all observation from 1985-2006. The x-axis is temperature (upper row) or salinity (lower row) and the y-axis is depth.

#### *Validation of reanalyzed mean sea surface height*

The mean model sea surface height was also compared with CNES-CLS09 MDT version 1.1 (Rio et al., 2011). The model sea surface height is related to the mean currents by the geostrophic relationship. The CNES-CLS09 MDT is essentially based on in situ dynamic heights, drifting buoy velocities and the geoid model computed from GRACE (Gravity Recovery and Climate Experiment) data. It is thus an independent data set. As the focus of this study is the southern polar region, the comparison is limited to the area south of 40°S. A constant over this domain has been subtracted to remove any offset which is not dynamically significant. The (centered) RMS error between the free running model and the MDT over this area is 0.218 m which is reduced to 0.165 m between analysis and MDT. The RMS of the 5-day forecast based on the analysis is essentially the same with 0.166 m. Overall the mean SSH gradient is more realistic in the analysis compared to the free model run leading to a more realistic representation of the Antarctic Circumpolar Current. The structure of the gradient is also more realistic in the model run with assimilation, especially in the Amundsen Sea and Ross sea sector.

Assuming a jet with a Gaussian velocity profile, one can determine the characteristics of the polar front by fitting the error function on the mean sea surface height  $h$  (Gille, 1994). This

allows us to identify the position and the width of the front. This fit has been performed on the ORCA2 model grid (also for the CNES-CLS09 MDT) and repeated for all longitudes of the model grid. The model run without assimilation reproduces relatively well the position of the front (**Figure 10**). The RMS error of the position (averaged over all longitudes) is  $1.70^\circ$ . While the assimilation can locally degrade the position of the front, it reduces on average the RMS error to  $1.61^\circ$ . The overall structure of the width of the front agrees with the width determined from the CNES-CLS09 MDT. However, the width in the free model run is overestimated, indicating that the model is too smooth and the ACC (Antarctic Circumpolar Current) is too diffuse. While the width of the front in the analysis is still too large, the assimilation improves its representation and the RMS error is reduced from  $5.96^\circ$  to  $3.27^\circ$ .



**Figure 10** Position and width of the mean SSH front.

#### *Validation of reanalyzed sea ice coverage*

We compare the sea ice coverage data from the free and analyzed NEMO-LIM2 runs with observations from the Operational SST and Sea Ice Analysis (OSTIA) system. We will also compare the NEMO-LIM2 results with data from two other models: the Centro Euro-Mediterraneo sui Cambiamenti Climatici - Climate Model without resolved stratosphere (CMCC-CM), and the same model with a resolved stratosphere (CMCC-CMS). Those models have been chosen as they both use the same ORCA2 grid as NEMO-LIM2.

All sea ice coverage data available from the models are interpolated from their original ORCA2 grid on the grid of OSTIA observations. Using the same procedure, they are then all interpolated on the grid from OSTIA observations. This grid is constantly spaced with a  $2^\circ$

resolution, giving a global coverage of 180 by 90 cells. The data sets cover a period of 21 years, from January 1985 up to December 2005. Data from CMCC-CM(S) models are already monthly averages. Consequently, we took the monthly average for OSTIA observations and data from the NEMO-LIM2 free and analysed runs. We also decided to only consider the southern hemisphere for all the following comparisons since we are only interested in the sea ice coverage in the Southern Hemisphere.

### *Seasonal Cycle*

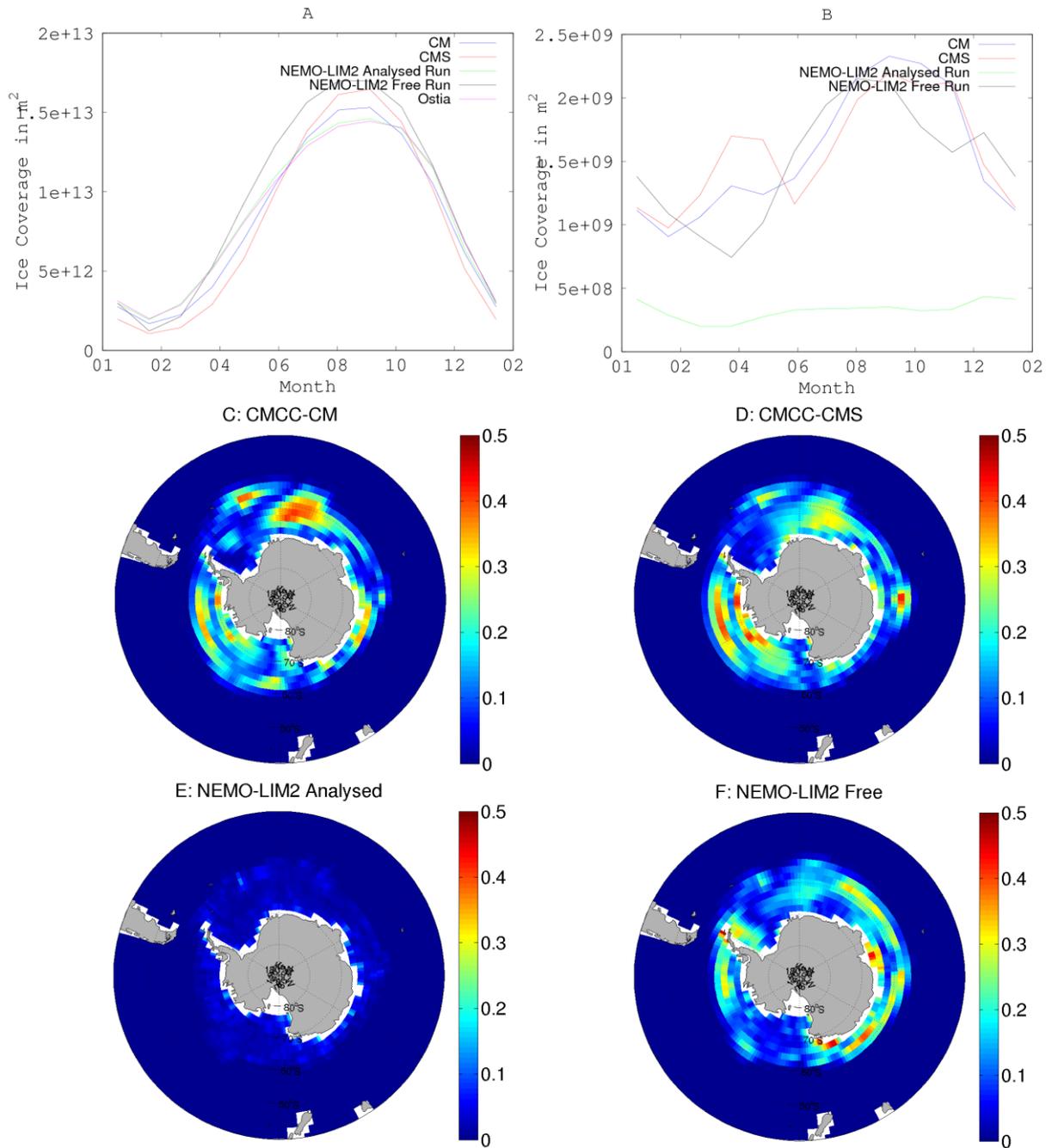
First, we will be looking at the seasonal cycle of the models (Figure 11a). To obtain this figure, we first calculated the monthly sea ice area:

$$SIA_{m,n} = \sum_x \sum_y A_{x,y} SIC_{m,n,x,y}$$

where the indices  $m,n$  refer to months and years respectively, and  $x,y$  to the spatial dimension.  $A$  is the area of the grid cell and  $SIC$  the sea ice concentration. We then averaged it for each month over the considered period, where  $N$  refers to the number of years in the 1985-2005 period:

$$Cycle_m = \frac{1}{N} \sum_n SIA_{m,n}$$

We can clearly see on Figure 11a that all models are globally able to reproduce the mean seasonal cycle of the sea ice area (SIA) over the Southern Ocean. All models underestimate the SIA during the summer period (December-March). Both the CMCC-CM(S) and the NEMO-LIM2 free run clearly tend to overestimate the sea ice area during the winter (July - September). The free run also overestimates the SIA during the autumn, starting from April, and performs better than the CMCC-CM(S) models during the winter. We can also note that the CMCC-CMS systematically performs worse than the CMCC-CM model. Finally, we find out that the assimilated NEMO-LIM2 run sticks to the OSTIA observations, as expected. Interestingly, it slightly underestimates the SIA throughout the whole year. This might be due to the fact that because of the assimilation, the data is smoothed through the whole domain. This tends to slightly reduce the SIC, thus SIA, of the NEMO-LIM2 Assimilated run.



**Figure 11** Each image is for period 1985–2005. (a) Mean monthly seasonal cycle. (b) Mean monthly RMSE. (c) Mean RMSE of CMCC-CM. (d) Mean RMSE of CMCC-CMS. (e) Mean RMSE of NEMO-LIM2 analyzed run. (f) Mean RMSE of NEMO-LIM2 free run.

We should also look at the Root Mean Square Error of the different models with the OSTIA observations, averaged over the 1985-2005 period (Figure 11b). First, we calculated the monthly sea ice area, averaged over the 1985-2005 period:

$$SIA_{m,x,y} = A_{x,y} \frac{1}{N} \sum_n SIC_{m,n,x,y}$$

We then calculated the RMSE with the OSTIA observation over the domain:

$$SIA_{rmse,m} = \sqrt{\sum_x \sum_y (SIA_{m,x,y} - SIA_{obs,m,x,y})^2}$$

Figure 11b represents the mean monthly RMSE of the models compared to the OSTIA observations. As expected, the RMSE of the NEMO-LIM2 analysed run is much lower than the other models, since it assimilates the data from which the RMSE is calculated. However, the NEMO-LIM2 free run performs overall similarly to the CMCC-CM(S) models. We can note that for all the models, the RMSE is at its lowest during the summer months, and at its highest during the winter. The biggest difference between the NEMO-LIM2 free run and the CMCC-CM(S) models is the period from February to May, where the former has a decreasing RMSE, whereas the later ones have an inscreasing RMSE. We note in particular the huge increase in March and April for the CMCC-CMS.

From Figure 11a, we could have thought that the CMCC-CM(S) models would at least perform better during the summer, since they better reproduce the total SIA. However, this is not the case, and the NEMO-LIM2 free run has a RMSE similar to the CMCC-CM(S) models throughout the whole year. This difference could come from the fact that the NEMO-LIM2 free run, though it does not reproduce the correct total SIA, is able to place the ice at better locations than the CMCC-CM(S) models, thus producing a smaller RMSE with OSTIA observations.

This hypothesis is confirmed when looking at the mean spatial RMSE of sea ice concentration (SIC) of the models with the OSTIA observations. It is obtained by calculating the RMSE of the SIC of the models with OSTIA observations, but not averaging over the domain. First, we take the monthly mean state by averaging over the whole considered period:

$$SIC_{m,x,y} = \frac{1}{N} \sum_n SIC_{m,n,x,y}$$

Then, we compute the mean RMSE with OSTIA observations by averaging over a year:

$$SIC_{rmse,x,y} = \sqrt{\frac{1}{12} \sum_m (SIC_{m,x,y} - SIC_{obs,m,x,y})^2}$$

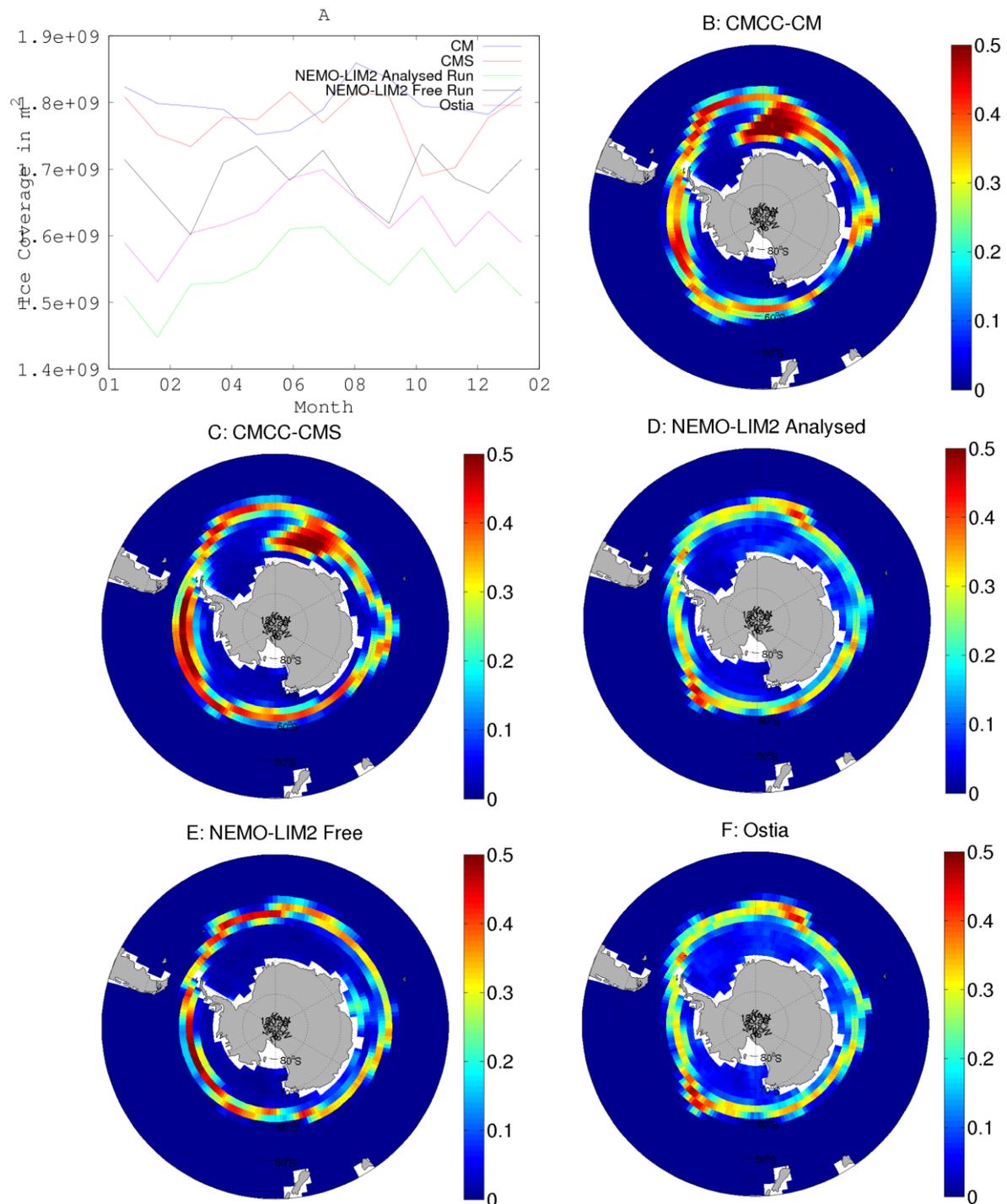
Both CMCC-CM(S) models produce more localised, but larger errors in the sea ice area. Those errors are the strongest around the Lazarev and Riiser-Larsen seas for CMCC-CM (Figure 11c), and in the Amundsen sea for CMCC-CMS (Figure 11d). As expected, the NEMO-LIM2 assimilated run (Figure 11e) performs very well, and has a nearly uniform RMSE over the whole sea ice domain. Finally, the NEMO-LIM2 Free run (Figure 11f) seems to perform rather well, with errors mainly located in the Somov and D'Urville seas, and along the coast of Graham land.

### *Internal Variability*

We will now look at the respective internal variability of all the models. First, we calculated the mean RMS of the model compared to one particular reference year. We then did the same while considering all the years as reference year, and averaged the final result to get a monthly mean internal variability of the model:

$$SIA_{rms,m} = \sqrt{\frac{1}{N} \sum_{n'} \left[ \frac{1}{N} \sum_n \sum_x \sum_y A_{x,y} (SIC_{m,n,x,y} - SIC_{ref,m,n',x,y})^2 \right]}$$

We note from Figure 12a that all the models have the same order of magnitude for their respective internal variability. The variability of CMMC-CM(S) models is at its most 20% higher than that of the OSTIA observations. Interestingly, the NEMO-LIM2 assimilated run has a lower internal variability than the observations, but copies the exact same shape of the observations.



**Figure 12** Each image is for the period 1985–2005. (a) Mean Monthly Internal Variability. (b-f) Spatial internal variability for September of CMCC-CM (b), CMCC-CMS (c), NEMO-LIM2 Free Run (d), NEMO-LIM2 Assimilated Run (e), OSTIA Observations (f).

We can also look at the spatial internal variability of one month in particular. We choose September, when the sea ice area is at its highest. This will enable us to locate the regions where the systematic bias appears in the internal variability. We do so as the previous equation, only not with the spatial sum:

$$SIA_{rms,m,x,y} = \sqrt{\frac{1}{N} \sum_{n'} \left[ \frac{1}{N} \sum_n (SIC_{m,n,x,y} - SIC_{ref,m,n',x,y})^2 \right]}$$

For September, we then obtain the Figure 12b to f.

We note that, for the CMCC-CM(S) models, the area where the internal variability is the highest tend to correspond with the area where the mean RMSE with Ostia observations were the largest (Figure 12b,c). This is especially true for the Lazarev and Riiser-Larsen seas. The NEMO-LIM2 Free and assimilated run seem to much better reproduce the internal variability of the Ostia observations.

As expected, the NEMO-LIM2 analysed run reproduces exactly the behaviour of the OSTIA observation, since it assimilated those data. When comparing the performance of the NEMO-LIM2 free run and the CMCC-CM(S) models, we conclude that though the former one has a worse total sea ice area estimation, it has a better localisation for the ice. Its mean spatial RMSE is thus lower than the one of the CMCC-CM(S) models. All models do reproduce the internal variability of the observations quite correctly.

### 2.2.2. Reconstruction of sea ice thickness and volume using data assimilation

The spatio-temporal structure of the variability of the ice cover in the Southern Ocean reflects changes in ice concentration, drift as well as ice thickness. While sea ice concentration and drift observations can be easily derived from satellite measurements, monitoring the ice thickness is challenging due to the difficulty of satellite signal to penetrate through the sea ice. In situ observations suffer from large spatio-temporal gaps due to the remote location of polar regions. As a consequence, the lack of sea ice thickness observations does not allow providing a long-term view on global and regional changes in Antarctic ice thickness. In this context, a reconstruction of the Antarctic sea ice thickness based on data assimilation of sea ice concentration is particularly valuable. The results discussed in detail in Massonnet et al., (2013) are summarized below and constitute one of the first attempt to reconstruct the Antarctic sea ice thickness and volume using data assimilation.

As in the previous section, the global ocean-sea ice model NEMO-LIM2 forced by atmospheric reanalyses is used. An Ensemble Kalman Filter (EnKF) scheme is implemented in this model in order to assimilate observations of sea ice concentration from the OSI SAF (Eastwood et al., 2011) dataset. The data assimilation procedure consists of two steps and it is described in detail in Mathiot et al. (2012). First, at the forecast step, an ensemble of 25 members is propagated forward in time with the model until observations are available. Second, at the analysis step, model error statistics are estimated from this 25-member sample. Together with observational errors statistics, they are used to weight the model and observational estimates of the system state. These weighted estimates are then combined as to provide a physically consistent analysis with minimum error.

Two model simulations were started in January 1960. In the run ASSIM, the data assimilation procedure is applied, while it is turned off in the run FREE. In a first step, the skill of ASSIM and FREE is assessed using (Cavalieri et al., 1996) sea ice concentration dataset. The sea ice concentration and extent conditions in the different sectors of the Southern Ocean, are better represented in ASSIM than in FREE both at seasonal and interannual timescales (Table 1).

**Table 1** A. 1980–2008 trends of monthly anomalies of sea ice extent in different sectors of the Southern Ocean. For observations (Cavalieri et al., 1996), the  $\pm 2\sigma$  estimate on the trend is also provided. The modeled trends are in bold when they lie in the confidence interval of the observations. B. Correlations between observed and simulated 1980–2008 monthly anomalies of sea ice extent in the corresponding sectors. C. Mean absolute difference of sea ice thickness with respect to the ASPeCt (Worby et al., 2008) dataset between 1980 and 2008. In the table, FREE (ASSIM) refers to the run without (with) assimilation of ice concentration. Table from Massonnet et al. (2013).

Sector	A. Trends sea ice extent %/decade			B. Correlation of monthly anomalies		Mean thickness bias (cm)	
	OBS	FREE	ASSIM	FREE	ASSIM	FREE	ASSIM
Weddell Sea	1.60 ± 2.69	-2.18	<b>2.15</b>	0.57	0.96	29	23
Ind. Ocean	1.93 ± 2.59	4.84	<b>3.22</b>	0.57	0.87	21	17
Pacific Ocean	1.02 ± 3.96	5.71	<b>3.84</b>	0.62	0.88	38	30
Ross Sea	4.46 ± 3.10	<b>7.01</b>	<b>5.66</b>	0.75	0.96	35	31
Amund.-Bel Seas	-5.41 ± 4.35	1.09	<b>-3.01</b>	0.67	0.93	26	17
Southern Ocean	1.44 ± 0.82	2.41	2.66	0.53	0.81	30	25

Given that the data assimilation technique applied here is multivariate, any variable of the state vector is updated as long as it covaries with the assimilated variable(s) (here, the sea ice concentration). In particular, the assimilation of ice concentration data also has positive impacts on the Southern Ocean sea ice thickness in ASSIM (Table 1C). The numbers are obtained by averaging, over all months between 1980 and 2008 and all grid cells falling in the corresponding sectors, the absolute difference (“bias”) between (1) the available ASPeCt (Worby et al., 2008) sea ice thickness estimates interpolated on the model grid and (2) the model sea ice thickness at the corresponding locations. In ASSIM, the biases on sea ice thickness are reduced by ~20% on average compared to FREE. The sea ice thickness provided by the ASSIM run thus forms a fair basis to propose a reconstruction of the Antarctic sea ice volume and thickness. Additionally, this reconstruction of ice volume indicates that sea ice volume and thickness covary well with extent and concentration at the multi-decadal timescale in the Southern Ocean. The results also confirm that the Antarctic sea ice volume displays marked fluctuations up to the decadal time scale.

### 2.2.3. Bias correction and parameter estimation using data assimilation

Data assimilation is a widely recognized tool for state estimation. By augmentation of the state vector, it can also be used to estimate other characteristics, e.g. biases in the model or in the boundary conditions, or even model parameters.

Bias correction techniques (Dee and Da Silva, 1998) have been proposed to estimate the model bias via data assimilation. However, these approaches do not correct the bias in the model. They estimate the bias as a separate field which is subtracted from the model results during assimilation and added back to restart the model. While this task was not initially foreseen in the proposal, we will attempt to treat the bias as an additive parameter (or spatial field) in the model equations and try to estimate this bias with data assimilation parameter estimation methods.

The first step of this work consists in developing and testing the feasibility of this technique on a Lorenz-96 system (Lorenz, 1996). This technique is then applied for calibration of sea ice dynamic parameters in the ocean-sea ice model NEMO-LIM3. The results were first focused on the Arctic, because of the limited spatial availability of the data that we used. This latter work has been published in the *Journal of Geophysical Research: Oceans* (Massonnet et al., 2014).

This part aims at developing a new method of bias correction using data assimilation by trying to estimate the origin of the bias instead of the bias of the model results. First, an estimate of the model’s bias needs to be provided, in particular its possible sources. This

estimation will then provide a basis to create an ensemble of model runs with perturbations around the source of the bias. We use those perturbations as a control variable for data assimilation. With this ensemble, we assimilate observations either from the original run, or from an external source. This allows us to build a stochastic forcing which is directly injected into the model's modified equations. The model is then rerun with the bias correction, and we compare the new ensemble's behavior, to see if we are able to effectively correct the model's bias. The correction of the bias will thus be continuous during the forecast, providing an updated and more reliable analysis. First developed with a twin experiment on a Lorenz '96 model (Lorenz, 1996), this new method is currently being applied and tested on the sea-ice ocean NEMO-LIM model.

### *Lorenz '96 Model*

We first test our approach on a fully controlled mathematical model. In 1963, Edward Lorenz developed a simplified mathematical model aimed at reproducing atmospheric convection. It is notable for having chaotic solutions for certain parameter values and initial conditions (Lorenz, 1963). Originally, it consists of a system of three differential equations. In 1996, it was updated in its 40-variable form, known as the Lorenz '96 model (Lorenz and Emanuel, 1998; Lorenz, 1996). It consists of a circular closed boundaries system with advection and diffusion properties. The system is described by the following equation:

$$\frac{dX_k}{dt} = -X_{k-2}X_{k-1} + X_{k-1}X_{k+1} - X_k + F_k$$

that we slightly modified by taking a spatially changing forcing parameter  $F_k$  instead of a constant one for all the variables.

This model has been widely used to test and improve data assimilation methods, ensemble filters or parameter estimation (Anderson, 2009; van Leeuwen, 2010; Li et al., 2009). Indeed, developing new methodologies relies on multiple specific procedures that need to be tested. This preparation work is better done beforehand on a very small model, which, even if it does not stand comparison with the complexity of realistic models, still enables us to correct the multiple issues we will be facing later on. Even if the Lorenz '96 model is not particularly complex, it still shows similar behavior with the ocean, in particular the chaotic behavior that makes forecasting a real issue.

However, we will use the model in a different way than originally intended. Indeed, many of the previous works based on this model concentrate on the value of each variable during the model run. Since our aim is not to correct the specific value of the variables, but rather correct the bias that affects those variables, we will thus look at the mean value of those variables over a period of time.

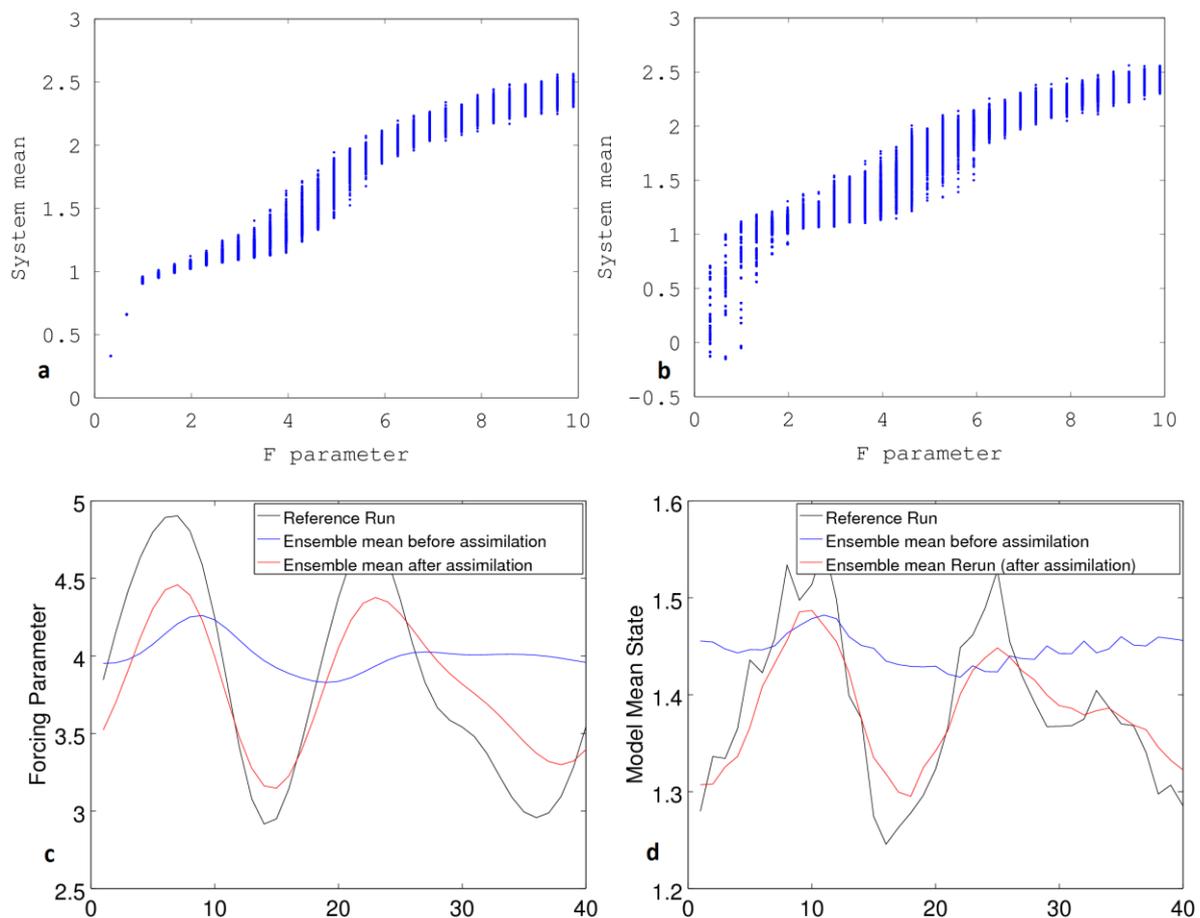
Therefore, we have first looked at the general behavior of the model when launched with a set of different initial conditions, and different  $F_k$  values. We have noted that, even though the model does show a chaotic behavior that highly depends on the initial conditions, the model's mean tends to stabilize itself after a certain amount of time. Lorenz and Emanuel (1998) already noted that if  $F_k < 4$ , the waves can extract energy fast enough to offset the effect of the external forcing. When  $F_k > 4$ , the model becomes completely chaotic over time and shows spatially irregular patterns. Even more, when  $F_k > 15$ , the model becomes totally unstable and collapses.

We have looked at the mean value of the model's variables over a certain period of time. We have noted that there is a significant relationship between the variables mean over time and the forcing parameter  $F_k$ . We choose  $k=1, \dots, 40$ , and a time step of 0.05, which corresponds to about 6 hours in the atmosphere (Lorenz and Emanuel, 1998). We choose 30 evenly

distributed values for  $0 < F_k < 10$ . The model is then run with 15 different initial conditions for each  $F_k$ . We consider the 200 first time steps sufficient for the model to stabilize and dampen the initial conditions. We then take the mean of the model's variables for the last 800 time steps, and average it over the 40 variables to obtain the model's mean state.

Two cases are studied: in the first, the  $F_k$  are constant for all the variables (Figure 13a). In the second, we add a random spatially correlated noise, with standard deviation equals 1, on the forcing parameter in order to obtain a different  $F_k$  for each  $X_k$  (Figure 13b). On both figures,  $F_k = 4$ .

We can clearly see from Figure 13a,b that there is a monotonic relationship between the system's mean and the forcing parameter, whether the later one is, or is not, constant. This supports the working hypothesis that even a fully non-linear system on a local scale can be expected to show a global simple behavior. This also confirms that even though the model's state at a specific point in time depends on the initial conditions, the time average of the model, when ignoring the first 200 time steps, does not depend on the initial conditions anymore. This is important, since our aim is not to predict the exact value of the system at a given point in time. We only aim at correcting the model's forcing parameter and the bias it causes on the model's mean state.



**Figure 13** (a) Lorenz '96 model mean state compared to a constant forcing parameter. (b) Lorenz '96 model mean state compared to a spatially variable forcing parameter. (c) Lorenz '96 model forcing parameter of the observation, ensemble mean and assimilated ensemble mean runs. (d) Lorenz '96 model state temporal mean of the observation, ensemble mean and assimilated ensemble runs.

### Lorenz '96 Twin Experiment

We have tested our method with a Lorenz '96 model twin experiment. As shown before, we can consider the forcing parameter  $F_k$  to be directly linked with the model's mean over a period of time. First, we create a random, but spatially correlated  $F_{k,obs}$  parameter with a mean  $F_{k,obs}' = 4$ , and with a standard deviation of 1. The model is then run once over 1000 time steps, with 15 initial conditions. It is then averaged over time while ignoring the first 200 time steps to avoid the initial conditions to strongly influence the model's mean. We follow the exact same procedure to generate an ensemble of 100 different  $F_{k,ens}$ . Each one of them is run over 1000 time steps, with 15 initial conditions, and averaged without the first 200 time steps. We know that there is an effective bias present on the  $F_{k,obs}$ , and we intend to find it, in order to correct the model's mean. We consider the reference run's mean  $X_{k,obs}$  as the observations. We extend our state vector which consists of the ensemble's model's mean  $X_{k,ens}$  with the ensemble's  $F_{k,ens}$ . Using an ETKF analysis scheme, we obtain a new and updated  $X_{k,ens}$ . We then rerun the model, and expect the ensemble model's mean reruns to improve and come closer to the reference run. The results of this procedure are shown on Figure 13c,d.

The assimilation of observations on the model's mean allowed the correction of the bias on  $F_k$  (Figure 13c). The root mean square error on  $F_{k,ens}$  before assimilation was 0.65371. After the assimilation, it has been reduced to 0.32254 for  $F_{k,ana}$ , and it is already able to reproduce the global shape of the reference run. We need also to look at the model's mean (Figure 13d). The RMS on the ensemble mean  $X_{k,ens}$  is 0.098660. However, we can clearly see that the model's rerun with the assimilated  $F_{k,ana}$  gives much better results. The RMS on  $X_{k,Run2}$  is only 0.037679, and reproduces much better the shape of the observations.

We aimed at correcting the bias on a specific parameter from the Lorenz '96 model. Testing this method with a twin experiment has proven to be successful. Not only did we estimate and improved the ensemble's mean  $F_{k,ens}$ , but we were also able to rerun the model and obtain a much better model's mean state with the  $X_{k,Run2}$ . This means that we can interpret the physical behaviour of this model, and the correction on the bias that we were able to obtain, as a significant improvement on the model's climatology and global behavior, which we did not have before.

The correction of bias with data assimilation parameter estimation having been successfully tested on the Lorenz-96 model, it has then been applied to calibrate sea ice dynamic parameters in the NEMO-LIM3 model (Massonnet et al., 2014). Even though this study is focused on the Arctic, it constitutes insightful results in the framework of PREDANTAR and the main outcomes are thus presented below.

First, biases and their origins need to be investigated. NEMO-LIM is a low resolution model and this resolution can hardly be improved due to the requirement to perform simulations over several decades. This causes a large bias in the area of strong currents that needs to be corrected. Since the currents are strongly related to the heat transportation, they greatly impact the sea surface temperature, and thus the ice concentration.

Unlike for the tests on the Lorenz-96 model, physical constraints are to be applied on the correcting term in order for the model to be realistic. Indeed, spurious gravitational waves or any kind of physically non-existent processes should be avoided. Therefore, the stream function is chosen as a starting point, since it is by construction non divergent. Using the first simulations of NEMO, the yearly-mean turbocline is estimated, which allows to force only the currents at the surface, above the turbocline. Taking the derivatives from the stream function provides us U and V velocity fields, which are directly added to the momentum equation in NEMO.

Preliminary tests show that the model remains stable with forcing terms that are of the same magnitude as the acceleration in a non-forced run. Several different values of correlation length in the forcing term have been tested. There is a spatially coherent covariance between the forcing term and the sea surface temperature, showing that the bias on the current has a direct impact on the sea surface temperature bias.

For the assimilation of parameters, three steps were achieved (Massonnet et al., 2013):

1) *Adaptation of the EnKF code to include a global parameter analysis*

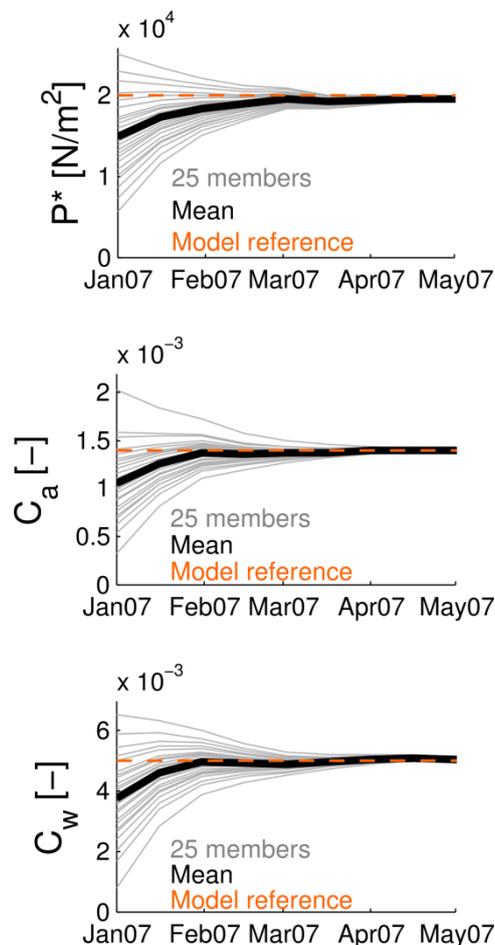
We completed the traditional EnKF state estimation with an analysis of parameters. Because parameters are not dynamically evolving, this latter step can be conducted sequentially, either before or after state estimation.

2) *Calibration of parameters in twin experiments*

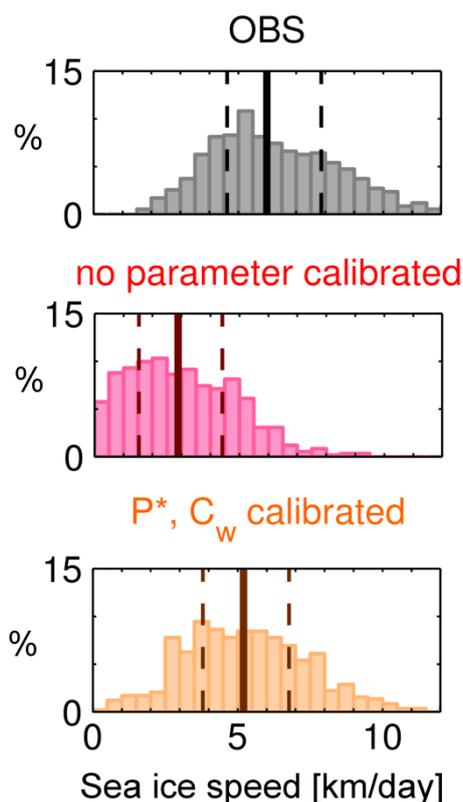
We first assimilated *model* sea ice drift, but starting with an incorrect value of parameters,, in order to test the successful convergence of the ensemble of parameters. **Figure 14** reports the time evolution of the ensemble (here, 25 members) of parameters. The original set of parameters is retrieved within the first year of simulation.

3) *Calibration of parameters with real data*

We then assimilated real data of sea ice drift. We found that the joint estimation of the two parameters  $C_w$  and  $P^*$  yielded the most significant improvements. Simulations with the calibrated values show a clear shift in the statistical distribution of simulated sea ice speeds, towards higher values (Figure 15) and thus in better agreement with observational data.



**Figure 14** Convergence of parameters in the case of "twin experiments"



**Figure 15** Frequency distribution of the observed (top) and modeled (middle and bottom) Arctic sea ice speeds between 2007 and 2012.

## 2.3. Understanding the changes in the sea ice cover over the last decades

### 2.3.1. Comparison of reconstructions and simulations of past changes

Disentangling the contribution of the external forcing and the internal variability in the positive trend in ice extent observed during the last three decades is an important issue. Given that the available observations of the sea ice cover are too sparse in space and time, the climate models constitute adequate tools to gain insight about the evolution of sea ice at multi-decadal timescales. In that context, a pragmatic approach lead us to focus on the question whether the positive trend in Antarctic sea ice extent observed over the last three decades is compatible with a combination of the forced response and the internal variability of the climate system as simulated by current general circulation models. This issue had been raised in several studies (e.g., Arzel et al., 2006; Lefebvre and Goosse, 2008b) but the systematic analysis of the internal variability of the Antarctic sea ice simulated by current general circulation models proposed in the PREDANTAR project is a clear addition to previous knowledge.

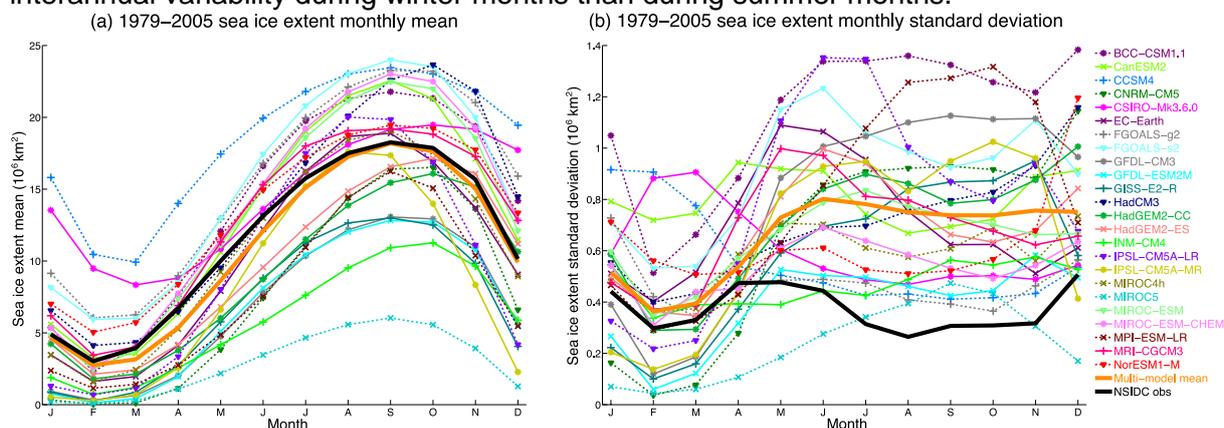
In the framework of PREDANTAR, the results of simulations performed with general circulation models involved in the 5th Coupled Model Intercomparison Project (CMIP5) have been systematically analyzed and compared to the available observations. Those analyses have been discussed in detail in a peer-reviewed article published in the journal *The Cryosphere* (Zunz et al., 2013). The main outcomes are summarized below.

The results of historical simulations performed with 24 general circulation models have been analyzed. The CMIP5 historical simulations generally span the period 1850-2005. Those simulations do not take into account any observations of the atmosphere, sea-ice or ocean state but they are driven by natural and anthropogenic variations in the external forcing. Each

historical simulation is an ensemble simulation whose size varies from one model to the other. The total number of historical simulations analyzed here equals 85.

The analyses were performed over the years 1979-2005, i.e. the time period for which reliable observations of the Antarctic sea ice cover are available. For each model, the seasonal cycle of the sea ice extent has been computed and the interannual variations in ice extent has been estimated on the basis of the standard deviation of each month of the year (Figure 16). On average over all the models, the seasonal cycle of the ice extent is in good agreement with the observations. Nevertheless, the multi-model average hides the wide range of behaviors simulated by each individual model. Indeed, the amplitude of the seasonal cycle as well as the timing of the minimum and maximum of sea ice extent strongly differ from one model to the other and barely fits the observations (Figure 16a).

The monthly standard deviation of the sea ice extent also displays a large discrepancy among the 24 models (Figure 16b). On average over all the models, the monthly standard deviation is higher than the one of the observations, especially during winter months. In February, 15 models display a standard deviation higher than the observed one. In September, all of the 24 models overestimate the standard deviation. In some models, the standard deviation displays a seasonal cycle that results from a significantly higher interannual variability during winter months than during summer months.

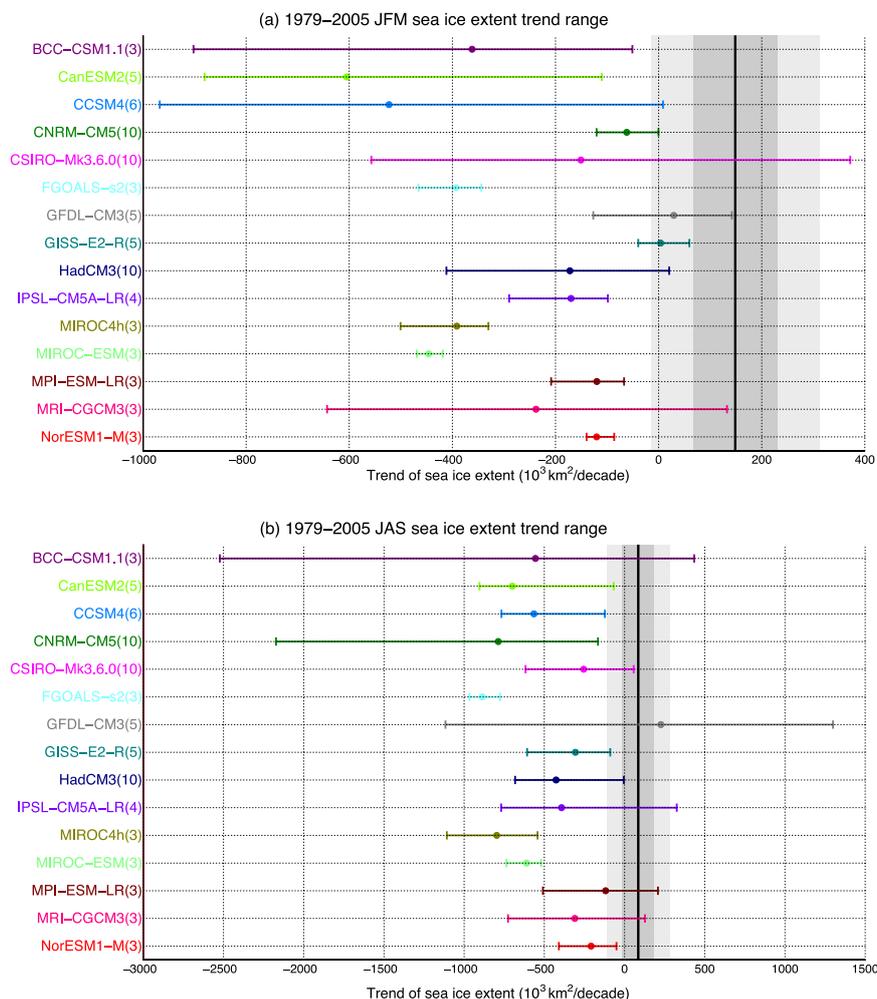


**Figure 16** (a) Monthly mean of Southern Ocean sea ice extent, computed over the period 1979–2005. (b) Standard deviation of detrended Southern Hemisphere sea ice extent, computed over the period 1979–2005 for each month of the year. Colours correspond to the ensemble mean of historical simulations from 24 different models. Orange bold line is the multi-model mean. Black bold line refers to observations (Cavalieri and Parkinson, 2008). Figure from Zunz et al. (2013).

The trends in Antarctic sea ice extent over the period 1979-2005 as simulated by the models have also been analyzed and compared to the observations. For each model, the trends were computed for each member of the ensemble simulation. Given that the interannual variability strongly varies from one season to the other in some models, the analyses of the trends were performed for summer (JFM) and winter (JAS) mean rather than for annual mean.

Observations show that both summer and winter sea ice extent expanded between 1979 and 2005, at a rate of  $149\,000 \text{ km}^2$  and  $86\,000 \text{ km}^2$  per decade respectively. For both seasons, most of the analyzed simulations display decreasing trends in ice extent. Nevertheless, the trends provided by the different members belonging to one model simulation can reach a wide range of values, including positive ones relatively close to the observations (Figure 17). Among all the simulations analyzed, 12 members over 85 have a positive trend between 1979 and 2005 in summer and 10 over 85 have a positive trend in winter. Therefore, positive values such as the observed one appear as relatively rare events, but the observed values are within the range of the internal variability simulated by the models. Nevertheless, it has to be kept in mind that our analyses have also highlighted biases in the interannual variability

simulated by the models. Because of those biases, the link between the internal variability in the Southern Ocean and the observed increase in sea ice extent cannot be confidently established.



**Figure 17** Ensemble mean, minimum and maximum value of the sea ice extent trend for the period 1979–2005 over the whole Southern Ocean for summer (a) and winter (b). The different 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 the maximum value 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). Figure from Zunz et al. (2013).

We went a step further in the investigation of the potential link between the internal variability and the observed positive trend in sea ice extent through the analyses of hindcast simulations performed in the framework of CMIP5. Unlike a historical simulation, a hindcast simulation is initialized through data assimilation of observations, i.e. its initial state is synchronized with the observed state. If the link between the internal variability of the climate system and the observed positive trend in ice extent exist and if the internal variability is predictable, an adequate initialization of the simulation should improve the simulated evolution of the Antarctic sea ice. This aspect has been investigated in the results of 30-year hindcast simulations from 10 models, initialized in 1980. Our analyses did not point out any clear improvement in the simulated ice extent in the hindcast simulations compared to the historical simulations. This lack of improvement could be due to some misrepresentations of

processes in the models but also to the initialization procedure that generally relies on simple methods in the CMIP5 hindcast simulations.

The mechanisms that rule the decadal variability of the Antarctic ice cover have been investigated in the framework of the PREDANTAR project and the results are summarized in Section 2.3.2. Further work has also been devoted to the assessment of the impact of the initialization through more sophisticated data assimilation methods in an idealized and in a realistic framework (Section 2.4.2).

### 2.3.2. Analysis of the mechanisms ruling decadal variability

The Antarctic sea ice is characterized by a large internal variability (e.g., Mahlstein et al., 2013; Polvani and Smith, 2013; Swart and Fyfe, 2013; Zunz et al., 2013). Numerous studies have investigated the response of the sea ice to the dominant modes of atmospheric variability in the Southern Ocean (e.g., Holland and Raphael, 2006; Lefebvre et al., 2004; Liu et al., 2004; Simpkins et al., 2012; Stammerjohn et al., 2008; Zhang, 2007). Besides, the mechanisms responsible for the large Antarctic sea ice variability at multi-decadal timescales have not been firmly identified yet. A part of the research undertaken in the framework of the PREDANTAR project was intended to address this important issue.

#### 2.3.2.1. Antarctic sea ice variability and mixed layer properties in CMIP5 models

The drivers of the Southern Ocean mixed layer variability and their link with the sea ice cover were investigated in simulations performed in the framework of CMIP5. This research work has been presented in detail in a peer-reviewed article published in the *Journal of Geophysical Research: Oceans* (Close and Goosse, 2013).

Simulations performed with 7 CMIP5 models under the representative concentration pathway (RCP) 4.5 scenario were analyzed. This forcing scenario represents a medium mitigation situation (Taylor et al., 2011). In addition, the chosen simulations provide relatively long time series (2006-2100) that are valuable to study processes of multi-decadal lengths. Here, only the first member of each model ensemble simulation was taken into account in the analyses.

To understand the processes that govern ocean-sea ice variability, it is necessary to determine which mechanisms exert the primary controls on the variability of the upper ocean (in particular, on salinity, which dominates the equation of state in the Southern Ocean region). The contributions to the salinity budget from the evaporation-precipitation balance, horizontal (Ekman and residual) advective fluxes, diffusive processes, vertical and lateral entrainment and brine/meltwater input from the sea ice formation cycle are thus estimated using:

$$\frac{\partial S_m}{\partial t} = \frac{(E - P)S_m}{h_m} - u_e \cdot \nabla S_m - u \cdot \nabla S_m + \kappa \nabla^2 S_m - \frac{w_e \Delta S_m}{h_m} - \frac{(u \cdot \nabla h_m) \Delta S_m}{h_m} - \frac{(\rho_i S_i - \rho_0 S_m) F_i}{\rho_0 h_m}$$

where  $S_m$  represents the salinity of the mixed layer,  $E$  denotes evaporation,  $P$  is precipitation,  $u_e$  and  $u$  represent Ekman velocity and residual (non-Ekman) horizontal component of the velocity respectively,  $w_e$  denotes the entrainment velocity,  $\Delta S_m$  is the salinity difference across the base of the mixed layer,  $h_m$  signifies the mixed layer depth,  $\rho_0$  and  $\rho_i$  denote mixed layer density and sea ice density respectively.

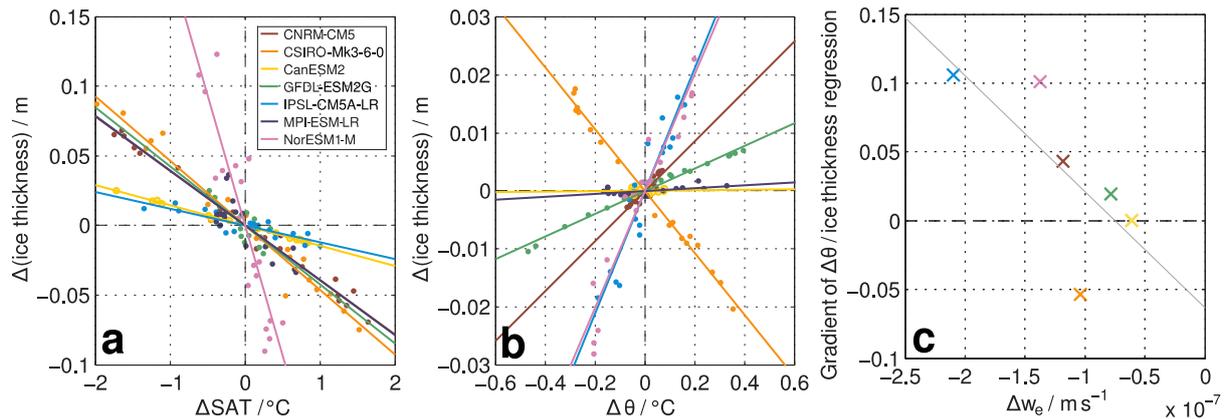
The response of the hydrographic properties of the upper Southern Ocean under the RCP4.5 forcing scenario exhibits a number of robust features. The mixed layer is characterized by a strong freshening tendency, accompanied by a corresponding increase in buoyancy and a weak temperature response in the ice-covered region. At lower latitudes, strong near-surface warming is evident. Whilst the models do not agree well on the magnitude of the changes,

there is good agreement on the sign of both the salinity and density tendencies, with the mixed layer tending to shoal in all of the models. Sea ice loss over the 100 year run is evident in all models except one (GFDL-ESM2G). The upper ocean buoyancy loss occurs at a faster rate than that of the underlying ocean, thus leading to a decoupling of the mixed layer, as noted in the previous CMIP3-based analyses of Capotondi et al. (2012). This decoupling suggests an important role of surface processes in driving the observed freshening tendency, and further credence is added to this notion by the evaluation of the individual terms of the salinity budget.

The magnitudes of the various terms of the salinity budget vary strongly amongst models, and, due both to the small sample size and presence of outliers, the median and interquartile range was used in our analysis in order to reduce bias in the results. The analysis of the individual terms indicates that component of salinity change driven by the evaporation-precipitation balance is very consistently represented amongst the models. When considering the model-median of the absolute values of the various terms, the dominant contributions to the salinity budget south of 60°S are supplied by entrainment and the meltwater input/brine rejection associated with the formation cycle of sea ice. This hypothesis is supported by the vertical distribution of the salinity change since the freshening signal is the strongest adjacent to the continent and weakens as the anomaly is advected northward and subducted into the interior ocean.

The dominant role of entrainment in effecting mixed layer salinity change that is surmised above is consistent with the changes observed in the ocean interior, where increased heat storage at depth is evident in all models. The approximately density-compensated nature of this change in temperature indicates that it is not surface forced, but rather related to interior dynamics. This entrainment-modulated restriction of the supply of heat to the surface is consistent with the previous work of Bitz et al. (2006) and Zhang (2007), with increased (reduced) supply of upwelled oceanic heat limiting (promoting) sea ice growth at the surface.

To permit assessment of the relative dependence of sea ice formation on the heat supplied from the ocean against that of the heat supplied by the atmosphere, time series of the oceanic heat content of the 500-1000m layer have been calculated. Considering the median field spatial distribution, sea ice thickness is significantly correlated with surface air temperature at the 95% level ( $r=-0.87$ ) and heat change at depth at the 90% level ( $r=0.74$ ). A multiple linear regression was performed to determine the dependence of sea ice thickness on the two forcings. The relative influence of the two mechanisms is strongly model-dependent. Whilst all models demonstrate a decrease in sea ice in response to increased surface air temperature (Figure 18), the decrease in sea ice thickness in response to 1°C of warming at 65°S varies from ~18cm (NorESM1-M) to ~1cm (IPSL-CM5A-LR and CanESM2). The response to changes in oceanic heat content is perhaps even more disparate, with one model (CSIRO- MK3.6.0) demonstrating a decrease in sea ice in response to increased heat storage at depth (corresponding to a uniform warming of the upper kilometer of the water column), whilst the remaining six models demonstrate an increasing tendency. The extent of dependency of sea ice thickness on mid-depth/deep ocean heat storage is linked to the change in the rate of entrainment exhibited by the models, with models that exhibit the greatest change in entrainment velocity also showing the strongest relationship between heat storage at depth and ice thickness (Figure 18c).



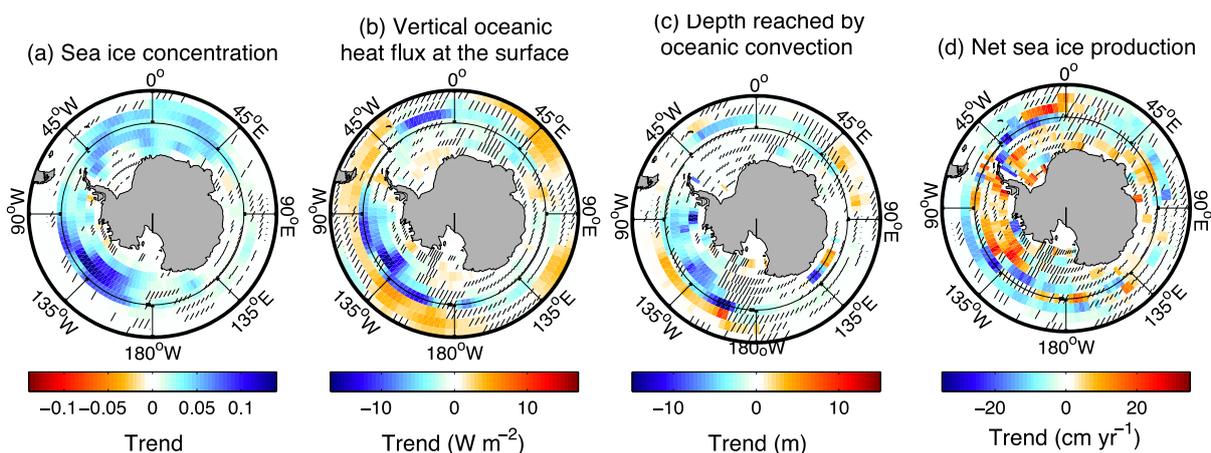
**Figure 18** Change in reconstructed zonal-mean sea ice thickness at 65°S derived from multiple-linear regression between sea ice thickness and (a) surface air temperature and (b) mean potential temperature over the depth range 500–1000 m. Colored dots show pentannual averages (see text). (c) Dependence of the strength of the relationship between potential temperature at depth and ice thickness on entrainment. The gray line shows the linear regression of change in entrainment rate on the gradient of the ice thickness / potential temperature regression. Figure from (Close and Gosse, 2013).

### 2.3.2.2. Decadal trends in Antarctic sea ice extent and ice-ocean interactions

The analyses of the CMIP5 simulations under the RCP4.5 forcing scenario have highlighted the key role played by both the heat supplied by the atmosphere and the entrainment-modulated supply of heat by the ocean below the mixed layer in determining the sea ice conditions at the surface. In the framework of PREDANTAR, we went a step further and we investigated ice-ocean interactions that could lead to an increase in Antarctic sea ice extent similar to the one observed over the last three decades. This has been achieved thanks to the analysis of simulations performed with the climate model LOVECLIM (Gosse et al., 2010), an Earth-system model of intermediate complexity. This study has been published as a peer-reviewed article in the journal *The Cryosphere* (Gosse and Zunz, 2014).

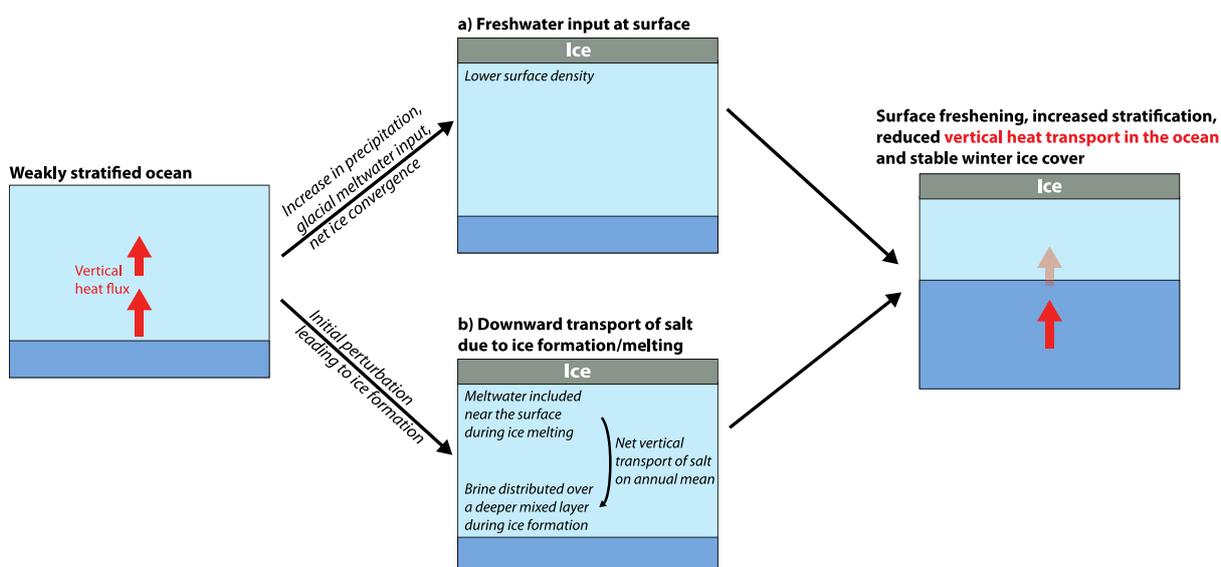
A 5000-year control simulation, using a constant forcing corresponding to pre-industrial conditions, has been performed with LOVECLIM. The last 1000 years of this experiment display a stable climate and were analyzed in order to identify the dominant processes controlling the decadal trends in sea ice extent. To that purpose, 11 periods characterized by an increase in ice extent larger than  $10^5 \text{ km}^2$  per decade for each calendar month, during a minimum of 30 years, were identified. This value was selected as it is close to the increase observed over the period 1979-2010 (Parkinson and Cavalieri, 2012).

The trend in 12 different ocean variables, on average over the 11 selected periods, have been investigated and discussed in detail in Gosse and Zunz (2014). Only 4 of those variables are presented here for brevity (Figure 19). As expected, the periods of large positive trends in ice extent display an increase in ice concentration, uniformly distributed around the continent (Figure 19a). Those periods are also characterized by a decrease of the vertical oceanic fluxes in ice-covered area (Figure 19b). This reduction in the upward heat flux is associated with a decrease in the depth reached by convection (Figure 19c). During the periods of increase in ice extent, the net sea ice production increases inside the pack (Figure 19d) because of the colder air and the lower oceanic flux. This increase in ice production results in a net northward export of ice that melts once it reaches lower southern latitudes, resulting in a net sea ice production decreases close to the ice edge.



**Figure 19** Trends in annual means averaged over the 11 periods showing a large increase in Antarctic sea ice extent scaled to represent 30 year changes of (a) ice concentration, (b) vertical oceanic heat flux at the ocean surface (positive upward,  $Wm^{-2}$ ), (c) depth reached by oceanic convection (m) and (d) net sea ice production (production minus melting) ( $cm\ year^{-1}$  water equivalent). The hatched areas represent the regions for which the average trend over the 11 periods is not significantly different at the 95% level from the mean trend in periods of identical length but not showing a large increase in ice extent. Figure from Goosse and Zunz (2014).

The sea ice transport of ice thus plays an obvious role in the expansion of the pack and the stabilization of the water column close to the ice edge (Figure 20a). Besides, the stabilization of the water column inside the pack was attributed to a positive feedback associated with the seasonal cycle of ice formation (Figure 20b). In winter, brine is released through the formation process of ice and it is mixed over a deep layer. On the contrary, during summer, the melting of ice releases freshwater that is included in a shallow layer. On average over a year, the cycle of ice formation and melting thus results in a net downward transport of salt and the ocean gets more stratified. Because of this stronger stratification due to the presence of sea ice, more heat is stored at depth in the ocean and the vertical oceanic heat flux is reduced, which contributes to maintain a higher ice extent. This mechanism can be illustrated by a simple 3-level model (see Goosse and Zunz, 2014).



**Figure 20** Schematic representation of the stabilization of the Southern Ocean by sea ice processes. Figure from Goosse and Zunz (2014).

The relationship between the increase in ice extent and the atmospheric circulation has also been investigated through the analysis of the trend in geopotential height at 800 hPa in the individual periods characterized by a large increase in ice extent (see Goosse and Zunz,

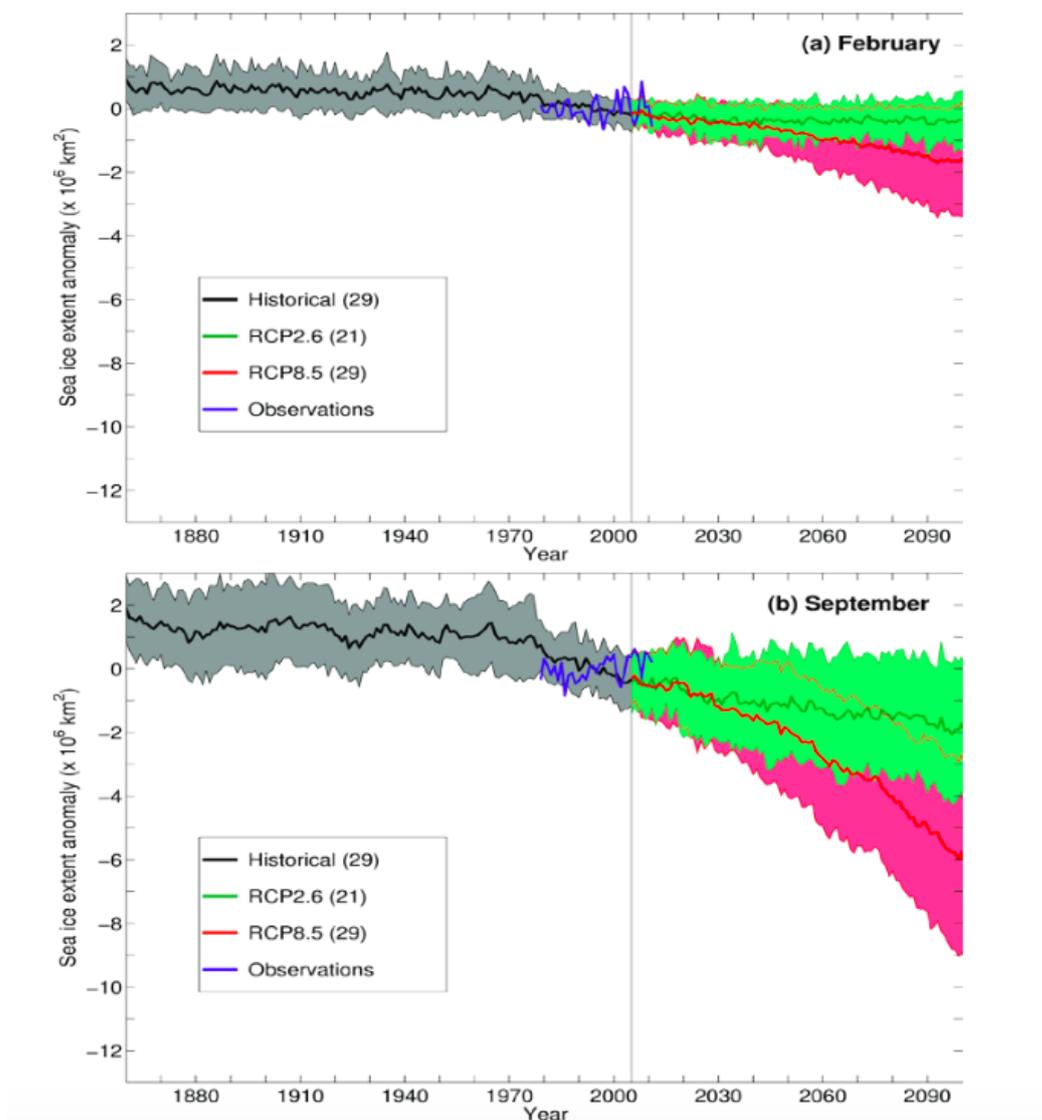
2014). Nevertheless, as the spatial patterns of the trend in geopotential height are very different in the 11 periods, the increase in ice extent could not be associated to one specific atmospheric circulation pattern. Besides, many regional changes in ice concentration can be attributed to the atmospheric circulation, as pointed out in numerous studies (e.g., Goosse et al., 2009; Holland and Kwok, 2012; Lefebvre and Goosse, 2008a; Liu et al., 2004; Simpkins et al., 2012).

In order to assess the compatibility between the ice-ocean positive feedback identified in the long control simulation and the behaviour of the Antarctic ice cover over the last 30 years, a simulation performed with LOVECLIM and constrained by surface air temperature observations through a data assimilation procedure were analyzed. This latter analysis indicate that, at least in some sectors the simulated results over the last 30 years are compatible with a shallower mixed layer and an increase in heat content below the mixed layer (see Goosse and Zunz, 2014). The mechanism proposed to explain the natural variability in LOVECLIM is thus consistent with our simulation with data assimilation. Nevertheless, this encouraging result does not constitute a definitive proof that the stabilization of the water column associated with ice is the dominant mechanism that control the recent changes as other processes may also contribute to positive ice-ocean feedback.

## **2.4. Projections and predictions**

### **2.4.1. Analysis of the changes for the end of 21<sup>st</sup> century**

The results provided by 29 general circulation models involved in CMIP5 were investigated. Changes in sea ice extent over the period 1850-2005 in historical simulation and over the period 2006-2100 in projection simulations under the forcing scenarios RCP2.6 and RCP8.5 were analyzed. Antarctic sea ice extent simulated by the 29 CGCMs analyzed shrink in response to the 21<sup>st</sup> century greenhouse gas loading (Figure 21). The decrease of the ensemble mean ranges from 15 % in scenario RCP2.6 to 52% in scenario RCP8.5 in summer and between 9% and 25% in September for those two scenarios, but high uncertainties remain. These uncertainties reflect both the strong decadal variations in single model realizations and a large inter-model scatter. Additionally, the confidence in these long-term projections is to be questioned as the same models poorly simulate the recent trends of sea ice extent in Antarctica



**Figure 21** Anomalies in Southern Hemisphere sea ice extent in (a) February and (b) September. The thick blue lines are the observed anomalies. The thick black lines are the anomalies of the multi-model mean of 29 CMIP5 models over the historical period, up to 2005. The red (resp. green) thick lines are the anomalies of the multi-model mean of the CMIP5 models under the high-emission RCP8.5 (resp. low-emission RCP2.6) scenarios. The number of models used for averaging is indicated in parentheses. Shading denotes  $\pm 1$  standard deviation obtained from the model distribution around the multi-model mean.

Because of the very limited skill of the current systems for decadal predictions in the Southern Ocean, no attempt has been performed to use them without adequate corrections as tools to estimate future changes.

#### 2.4.2. Improvement of the decadal forecasts

The analyses of the CMIP5 historical simulations presented in Section 2.3.1 indicated that the observed increase in Antarctic sea ice extent is compatible with the internal variability simulated by the models. Besides, systematic biases in the internal variability represented by those models were pointed out. Nevertheless, if the link between positive observed trend in ice extent and the internal variability of the system is confirmed and if this variability is predictable, an adequate initialization of the simulation should improve the modeled trend in ice extent.

Recent idealized studies have pointed out potential predictability in the Southern Ocean (Pohlmann et al., 2009), in particular for the position of the ice edge (Holland et al., 2013). In a realistic setup, our analyses of the hindcast simulations performed with the models involved in CMIP5 indicated that the initialization methods implemented in those models do not systematically improve the simulated evolution of the Antarctic ice extent. As those initialization procedures generally rest on simple data assimilation methods, the improvement that could be achieved thanks to more sophisticated data assimilation had to be examined.

In the framework of PREDANTAR, research has been undertaken in order to systematically assess how the predictability of Antarctic sea ice depends on the data assimilation method that is used to initialize the model simulation. The analyses were based on the results provided by the model of intermediate complexity LOVECLIM. This research work was presented in detail in 3 peer-reviewed articles (Dubinkina and Goosse, 2013; Zunz and Goosse, 2015; Zunz et al., 2015) and is summarized below.

#### 2.4.2.1. Antarctic sea ice predictability in an idealized framework

The analyses were first performed in an idealized framework. This approach consists of using pseudo-observations instead of actual observations for both the initialisation and the verification of the hindcasts. The pseudo-observations are obtained from a reference simulation performed with the same model as the one used to generate the hindcasts.

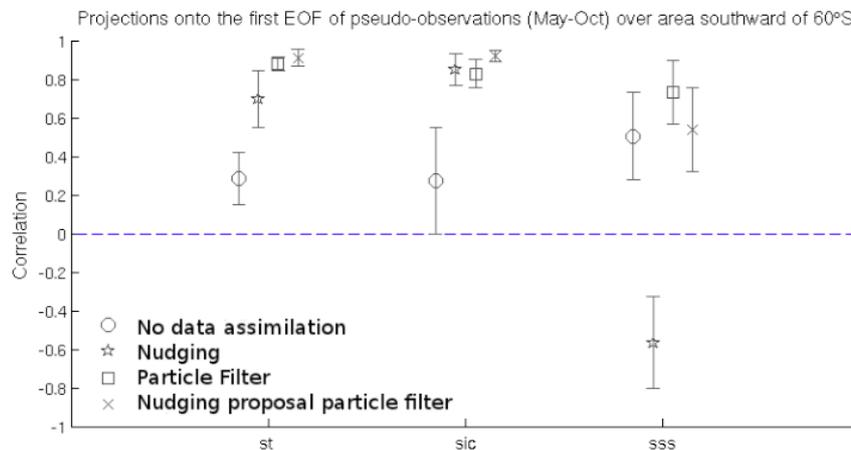
Three data assimilation methods were implemented in the model LOVECLIM (Dubinkina and Goosse, 2013): a simple one, based on a nudging procedure, and two more sophisticated ones, based on a particle filter (PF or SIR for Sequential Importance Resampling particle filter) and on a combination of a particle filter and a nudging (the nudging proposal particle filter, NPPF). Here, those three methods assimilate anomalies of surface air temperature. We assess these methods for reproducing the climate of the high latitudes of the Southern Hemisphere during the past 150 years. All the simulations consist of 96-member ensemble.

The nudging (e.g., Kalnay, 2007) is a data assimilation technique commonly used in decadal climate prediction studies. It consists of adding to the model equations a term that pulls the solution toward the (pseudo-) observations. In LOVECLIM, the nudging term corresponds to an additional heat flux between the atmosphere and the ocean  $Q = \gamma (T_{mod} - T_{obs})$ .  $T_{mod}$  and  $T_{obs}$  are the monthly mean surface air temperature simulated by the model and from the (pseudo-) observations respectively.  $\gamma$  determines the relaxation time and equals  $120 \text{ W m}^{-2} \text{ K}^{-1}$ .

The particle filter is an ensemble data assimilation technique that consists of the following steps (Dubinkina et al., 2011; van Leeuwen, 2009). Starting from a set of different initial conditions, an ensemble of simulations is propagated forward in time with the model for a period of prescribed duration, here 3 months. A member of the ensemble (called particle) differs from another only due to slightly different initial conditions. After the propagation step, a weight is attributed 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 3 months using the model, and the whole procedure is repeated until the end of the period of interest.

All three data-assimilation methods provide with good estimations of surface air temperature and of sea ice concentration, with the nudging proposal particle filter obtaining the highest

correlations with the pseudo-observations. When reconstructing variables that are not directly linked to the pseudo-observations such as atmospheric circulation and sea surface salinity, the particle filters have equivalent performance and their correlations are smaller than for surface air temperature reconstructions but still satisfactory for many applications. The nudging, on the contrary, obtains sea surface salinity patterns that are opposite to the pseudo-observations, which is due to a spurious impact of the nudging on vertical exchanges in the ocean, see Figure 22 and Dubinkina and Goosse (2013).



**Figure 22** Correlations between first principal components of the pseudo-observations and projections of the model simulations onto the corresponding first EOFs of the pseudo-observations for different variables: st is for surface temperature, sic is for sea ice concentration, sss is for sea surface salinity. EOFs are computed for May–October of twenty-four 21-yr periods over the area southward of 60°S. The circle is the mean correlation for simulations without data assimilation; the star is the mean correlation for the model simulations using the nudging; the square is the mean correlation for the model simulations using the particle filter; the cross is the mean correlation for the model simulations using the nudging proposal particle filter. Error bars correspond to one standard deviation.

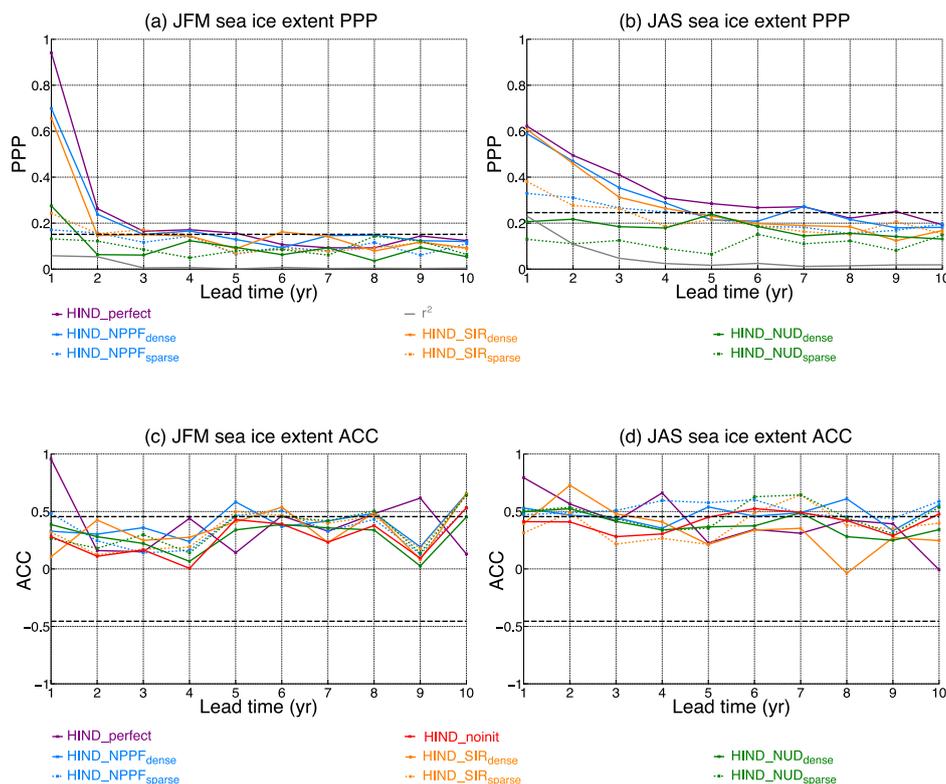
The impact of the three data assimilation methods on the quality of decadal predictions in an idealized framework was then assessed. Every 5 years between 1900 and 1990, the values of the state variables of the model are extracted from the data assimilation simulations and are used to initialise 30-year long hindcast (i.e. a forecast over a past period) simulations. Each hindcast consists of a 96-member ensemble simulation. Two datasets were used as pseudo-observations: one in which data are available everywhere (dense pseudo-observations) and the other one for which some data have been voluntarily removed (sparse-pseudo-observations), in order to mimic the lack of observations encountered in real conditions.

The quality of the prediction was assessed through two indices. (1) The prognostic potential predictability (PPP) tells us how the members belonging to one ensemble are spread, in comparison to the climatological variance. A PPP close to 1 indicates a small spread of the ensemble and is interpreted as a low uncertainty on the ensemble mean. On the contrary, a PPP close to or smaller than 0 corresponds to a large uncertainty. (2) The correlation between the hindcasts ensemble means and the corresponding pseudo-observations was used to assess the accuracy of the prediction. The predictive skill provided by different initialisation methods was assessed for the sea ice extent at interannual timescale (from one month to 10 years ahead) and at multi-decadal timescale (from 10 to 30 years ahead). Similar diagnostics were also applied to the ice edge location (see Zunz et al., 2015).

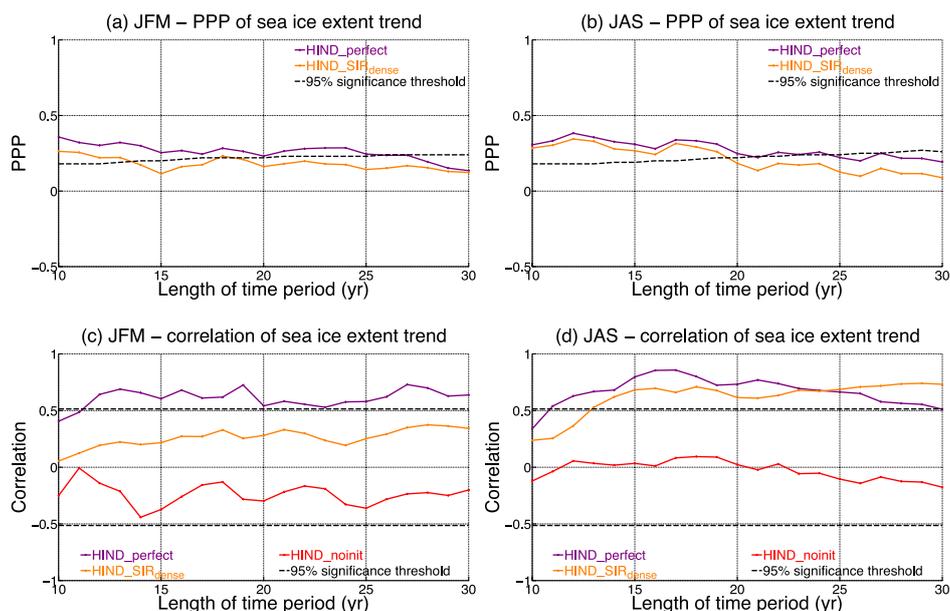
For any data assimilation methods used here to initialise the hindcasts, the use of dense pseudo-observations systematically leads to a higher PPP than if sparse data are assimilated, at both interannual (Figure 23a,b) and multi-decadal (Figure 24a,b) timescales. This is due to the stronger constraint applied on the ensemble through the initialisation with

dense data, implying a slower scatter of the members during the integration. At interannual timescale, the PPP of summer sea ice extent decreases sharply during the first 2 years of integration and falls below the 95% level of significance around the 3rd year, for any initialisation method (**Figure 23a**). The PPP of winter sea ice extent remains statistically significant during more than 3 years for the hindcasts initialized with dense pseudo-observations through the particle filters (**Figure 23b**). The larger PPP of the sea ice extent at interannual timescales in winter is provided by the ocean. Thanks to its inertia, it can store the information provided at the initialisation and brings them back to the surface during winter, when the vertical mixing is stronger due to the surface cooling and the brine rejection triggered by the sea ice formation.

At interannual timescales, a slight gain of correlation of the sea ice extent between the hindcasts and the pseudo-observations is provided by the initialisation with pseudo-observations during the first 4 years of integration (**Figure 23c,d**). In addition, the correlation is slightly larger in winter than in summer. Nevertheless, the correlation barely outstrips 0.5 and most of it is ensured by the external forcing. At multi-decadal timescales, the correlation of the trend in sea ice extent between the hindcasts and the pseudo-observations is clearly much higher in the hindcasts initialised with pseudo-observations (**Figure 24c,d**). The larger correlation found in initialised hindcasts at multi-decadal timescales is due to a better initialisation of the ocean below the sea ice when pseudo-observations are taken into account (not shown).



**Figure 23** 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. Figure from Zunz et al. (2015).



**Figure 24** 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. This autocorrelation depends on the length of the time period used to compute the trends.

#### 2.4.2.2. Antarctic sea ice predictability in a realistic framework

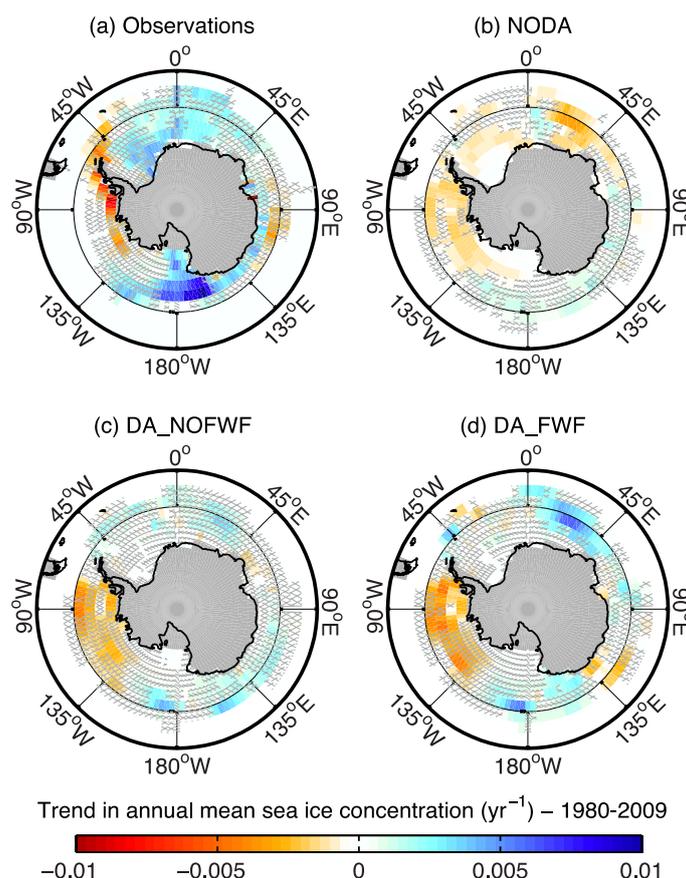
Having assessed the predictive skill that can be achieved in an idealized framework thanks to different initialization procedures, we logically undertook some work to check whether our results remain valid in realistic conditions (Zunz and Goosse, 2015). Given that our idealized study demonstrated that the predictive skill for the sea ice extent is higher at multi-decadal timescale than at interannual timescale, we focused on the predictability of 30-years trends in ice extent and concentration in a realistic framework. Reliable satellite observations of the Antarctic ice cover being available from 1979 onwards only, our diagnostics were limited to that time period.

In addition to the impact of the initialization method, the role played by a freshwater input in the observed trend in ice extent and concentration was also investigated. Indeed, several recent studies have proposed a link between the melting of the Antarctic ice sheet and the observed expansion of ice cover in the Southern Ocean (e.g., Bintanja et al., 2013; Hellmer, 2004; Swingedouw et al., 2008). Nevertheless, deriving the magnitude of the freshwater input derived from the Antarctic ice sheet mass imbalance requires data that are not available for the whole period spanned by our simulations. Furthermore, the estimate of the changes in mass imbalance for future projections would require a comprehensive representation of the polar ice sheets in climate models. The present study is based on results provided by the model LOVECLIM in a configuration that does not include an ice sheet component. Nevertheless, an estimate of the magnitude of the freshwater input was obtained through the data assimilation procedure.

Two simulations with data assimilation based on a nudging proposal particle filter (DA\_NOFWF and DA\_FWF) were first carried out with LOVECLIM over the period 1850-

2009. The assimilated data are surface air temperature anomalies from the HadCRUT3 dataset (Brohan et al., 2006). The data assimilation procedure is applied every three months. Both DA\_NOFWF and DA\_FWF consist of an ensemble of 96 simulations. In DA\_FWF, an additional random freshwater flux following an autoregressive process was applied on each member of the ensemble. The additional freshwater flux is computed every three months, i.e. at the same frequency as the data assimilation (for details, see Zunz and Goosse, 2015). This additional freshwater flux increases the range of the solutions reached by the ensemble and can randomly pull some of the members towards the observation, leading to an improvement of the efficiency of the particle filter. A weighted (i.e., taking into account the weight of each particle) average of the additional freshwater flux received by each particle of the ensemble provides an a posteriori estimate of the freshwater that is more likely to provide a state compatible with the observations. The results of the two simulations with data assimilation DA\_NOFWF and DA\_FWF were compared to the results of a simulation that does not include any information from the observations (NODA) and to the available observations.

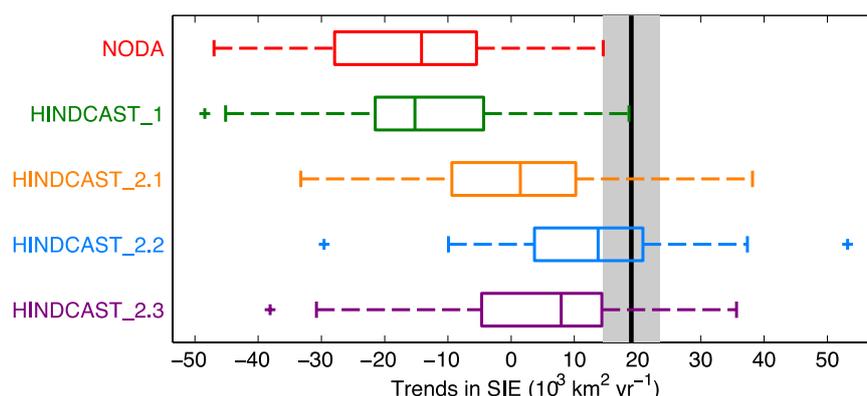
The spatial distribution of the observed trend in ice concentration is characterized by a decrease in ice extent in the Bellingshausen and Amundsen seas and an increase elsewhere, particularly strong in the Ross Sea (Figure 25a). This spatial structure is not reproduced by the simulation NODA (Figure 25b). Beside, in the two simulations with data assimilation, a spatial pattern of the trends in ice concentration similar to the observed one emerges (Figure 25c,d).



**Figure 25** Trend in yearly mean sea ice concentration between 1980 and 2009, shown for (a) the observations (Comiso, 1999), (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. Figure from (Zunz and Goosse, 2015).

The observed trend in ice extent over the period 1980-2009 derived from the version 2 of the Bootstrap data (Comiso, 1999) equals  $19.0 \times 10^3 \text{ km}^2 \text{ yr}^{-1}$ . The simulation NODA provides a trend in ice extent over that period equal to  $-15.5 \times 10^3 \text{ km}^2 \text{ yr}^{-1}$ , a negative value that is mostly due to the simulated response to the external forcing. Besides, the two simulations with data assimilation provide trends in ice extent that are much closer to the observed one,  $-3.0 \times 10^3 \text{ km}^2 \text{ yr}^{-1}$  in DA\_NOFWF and  $-2.8 \times 10^3 \text{ km}^2 \text{ yr}^{-1}$  in DA\_FWF, though those simulated trends are still slightly negative. In addition, the distribution of the trends in ice concentration is also in better agreement with the observations in the two simulations with data assimilation compared to the simulation without data assimilation. In the simulation DA\_FWF, the analyses of the ocean heat and salt contents in the top 100 m of the ocean and in the layer between  $-100$  and  $-500$  m indicates that the ocean in the 1970's is characterized by a warm and salty surface layer, a cold intermediate layer and strong vertical mixing. From the early 1980's onward, this state evolves towards a stabilization of the ocean column involving a fresher and cooler ocean surface that favors the formation of sea ice.

Hindcast simulations were then performed with LOVECLIM over the period 1980-2009. One hindcast was initialized from a state extracted from the simulation DA\_NOFWF (HINDCAST\_1) and three hindcasts from the simulation DA\_FWF. Those latter three hindcasts differ from each other in the additional freshwater flux they received: no additional freshwater flux (HINDCAST\_2.1), time evolving (HINDCAST\_2.2) and constant freshwater flux (HINDCAST\_2.3). The hindcast initialized from DA\_NOFWF fails in simulating a trend in ice extent close to the observations. Besides, the three hindcasts initialized from DA\_FWF provide trend in ice extent and concentration that roughly fit the observations (Figure 26).



**Figure 26** Boxplots showing the trends in annual mean sea ice extent in ensemble simulations. The edges of the boxes are the 1<sup>st</sup> and 3<sup>rd</sup> quartile while the central mark in the box is the median. The beginning (end) of the whisker corresponds to the 1<sup>st</sup> (3<sup>rd</sup>) quartile minus (plus) 1.5 times the interquartile range. The crosses denote the outliers. The black vertical line corresponds to the observed value of the trend (Comiso, 1999), surrounded by one standard deviation shown as the grey shaded rectangle.

In DA\_FWF, the state of the ocean in the late 1970's is characterized by a relatively warm and salty surface layer and a cold intermediate layer (for details, see Zunz and Goosse, 2015). The surface layer then freshens and cools down during the following years, leading to an increase in ice formation at the surface. This specific state in the late 1970's was likely achieved thanks to the combination of the data assimilation with the additional random freshwater flux. Besides, the additional freshwater flux does not seem essential during the following years since a hindcast simulation initialized with an adequate state can reproduce a positive trend in ice extent without any additional freshwater flux (HINDCAST\_2.1 in Figure 26). Overall, our results confirm the conclusion of recent studies about the important role played by the ocean in controlling the state of the sea ice at the surface (e.g., Bintanja et al., 2013; Hellmer, 2004; de Lavergne et al., 2014; Swingedouw et al., 2008). Here, we went a step further and pointed out that an adequate initialization of the ocean could improve the simulated trend in ice extent and concentration over the past three decades. Those results

are rather encouraging and open perspectives to predict the Antarctic sea ice over the next decades. Nevertheless, it has to be kept in mind that our tests in realistic conditions were carried out over the only 30-year period for which reliable observations of the sea ice cover are available. Furthermore, some of the key ice-ocean processes may be missing or misrepresented in climate models used to perform decadal predictions and this issue should be addressed in future work.



### 3. Policy support

Hugues Goosse is contributing author of the IPCC AR5, in charge in particular of the future changes in the sea ice cover and in the Southern Ocean.

Hugues Goosse is co-author of the update of the ACCE report (Antarctic Climate Change and the Environment ) performed in December 2012 : Turner J., N. Barrand, T. Bracegirdle, P. Convey, D. Hodgson, M. Jarvis, A. Jenkins, G. Marshall, H. Roscoe, J. Shanklin, J. French, H. Goosse, M. Guglielmin, J. Gutt, S. Jacobs, C. Kennicutt, V. Masson-Delmotte, P. Mayewski, F, Navarro, S. Robinson, T. Scambos, M. Sparrow, K. Speer, C. Summerhayes, D. Thompson, A. Klepikov, 2014. Antarctic Climate Change and the Environment – An Update. Polar Record 50, (3) 237-259 DOI: 10.1017/S0032247413000296.

Those reports will be important elements for the scientific basis of any decision making related to future climate changes.

Hugues Goosse is member of the 'OCNexus Statement drafting group experts' that has been asked by the European Marine Board to prepare a document describing the critical role of ocean science in responding to climate change. This document will be presented at the European Parliament on 21 October 2015 in the context of the upcoming COP21 United Nation Climate Conference that will be held in Paris in December 2015.



#### **4. Dissemination and valorization**

In addition to the point raised in the section 3 'policy support', the results obtained specifically within the framework of PREDANTAR have been reported in detail in more than 20 scientific articles published in peer-reviewed journal (see Section 5). The results achieved during this project were also regularly presented in more than 50 oral and poster contributions in international conferences. Information related to PREDANTAR is archived in the website of the project ([www.climate.be/PREDANTAR](http://www.climate.be/PREDANTAR)). A workshop presenting the main results of the project and discussing the perspectives, with 20 participants from Belgium, France, United Kingdom and Norway was held May 21-22 2015 in Brussels. The project website also proposes an outreach section where the mechanisms related to the interactions between the sea ice and the ocean are explained in a movie showing simple experiments and a report describing the main conclusions of the project in a way adapted to a large audience. This report is presented as an annex to this document.



## 5. Publications

### 2011

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

Vannitsem, S.: Bias correction and post-processing under climate change, *Nonlinear Processes in Geophysics*, 18, 911–924, [doi:10.5194/npg-18-911-2011](https://doi.org/10.5194/npg-18-911-2011), 2011.

Vannitsem, S. and Hagedorn, R.: Ensemble forecast postprocessing over Belgium: Comparison of deterministic-like and ensemble regression methods, *Meteorol. Appl.*, 18, 1, 94–104, [doi:10.1002/met.217](https://doi.org/10.1002/met.217), 2011.

Van Schaeybroeck, B. and Vannitsem, S. Post-processing through linear regression, *Nonlin. Processes Geophys.*, 18, 147-160 [doi:10.5194/npg-18-147-2011](https://doi.org/10.5194/npg-18-147-2011), 2011.

### 2012

Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F., and Vancoppenolle, M.: Better constraints on the sea-ice state using global sea-ice data assimilation, *Geoscientific Model Development*, 5, 1501–1515, [doi:10.5194/gmd-5-1501-2012](https://doi.org/10.5194/gmd-5-1501-2012), 2012.

Roulin, E. and Vannitsem, S.: Postprocessing of Ensemble Precipitation Predictions with Extended Logistic Regression Based on Hindcasts, *Monthly Weather Review*, 140, 874–888, [doi: 10.1175/MWR-D-11-00062.1](https://doi.org/10.1175/MWR-D-11-00062.1), 2012.

Van Schaeybroeck, B. and Vannitsem, S. Toward post-processing of ensemble forecasts based on hindcasts. *Publications scientifiques et techniques Royal Meteorological Institute Belgium*, 61, 2012.

### 2013

Close, S. E. and Goosse, H.: Entrainment-driven modulation of Southern Ocean mixed layer properties and sea ice variability in CMIP5 models, *Journal of Geophysical Research: Oceans*, [doi:10.1002/jgrc.20226](https://doi.org/10.1002/jgrc.20226), 2013.

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

Goosse, H., Roche, D. M., Mairesse, A., and Berger, M.: Modelling past sea ice changes, *Quaternary Science Reviews*, 79, 191–206, [doi: 10.1016/j.quascirev.2013.03.011](https://doi.org/10.1016/j.quascirev.2013.03.011), 2013.

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

Van Schaeybroeck, B. and Vannitsem, S.: Reliable Probabilities Through Statistical Post-processing of Ensemble Forecasts, in: *Proceedings of the European Conference on*

Complex Systems 2012, edited by Gilbert, T., Kirkilionis, M., Nicolis, G., and Vannitsem, S., vol. III, pp. 347–352, [doi:10.1007/978-3-319-00395-5\\_45](https://doi.org/10.1007/978-3-319-00395-5_45), 2013.

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

## 2014

Barth, A., Beckers, J.- M., Troupin, C., Alvera-Azcárate, A., and Vandenbulcke, L.: divand-1.0: n-dimensional variational data analysis for ocean observations, *Geoscientific Model Development*, 7, 225–241, [doi:10.5194/gmd-7-225-2014](https://doi.org/10.5194/gmd-7-225-2014), 2014.

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

Massonnet, F., Goosse, H., Fichet, T. and Counillon, F.: Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter, *J. Geophys. Res. Ocean.*, 119(7), 4168–4184, 2014.

Vannitsem, S.: Dynamics and predictability of a low-order wind-driven ocean– atmosphere coupled model, *Climate Dynamics*, 42, 1981–1998, [doi:10.1007/s00382-013-1815-8](https://doi.org/10.1007/s00382-013-1815-8), 2014.

Van Schaeybroeck, B. and Vannitsem, S., 2014. A probabilistic approach to forecast the uncertainty with ensemble spread, *Monthly Weather Review* (in revision), 2014.

## 2015

Barth, A., Canter, M., Van Schaeybroeck, B., Vannitsem, S., Massonnet, F., Zunz, V., Mathiot, P., Alvera-Azcárate, A., and Beckers, J.-M.: Assimilation of sea surface temperature, sea ice concentration and sea ice drift in a model of the Southern Ocean, *Ocean Modelling* (in revision), 2015.

Roulin E., and Vannitsem S., Post-processing of medium-range probabilistic hydrological forecasting: impact of forcing, initial conditions and model errors, *Hydrol. Process.*, 29, 1434–1449, [doi:10.1002/hyp.10259](https://doi.org/10.1002/hyp.10259), 2015.

Van Schaeybroeck, B. and Vannitsem, S. Ensemble post-processing using member-by-member approaches: theoretical aspects, *Quarterly Journal of the Royal Meteorological Society*, 141: 807–818, [doi:10.1002/qj.2397](https://doi.org/10.1002/qj.2397), 2015a.

Van Schaeybroeck, B. and Vannitsem, S. Assessment of calibration assumptions under strong climate changes, *Geophysical Research Letters* (in revision), 2015b.

Van Schaeybroeck, B., Mailier, P. and Vannitsem, S. Ensemble post-processing in Belgium based on reforecasts, *Weather and Forecasting* (in preparation), 2015.

Zunz, V., Goosse, H., and Dubinkina, S.: Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multi-decadal timescales, *Climate Dynamics*, 44, 2267–2286, [doi:10.1007/s00382-014-2344-9](https://doi.org/10.1007/s00382-014-2344-9), 2015.

Zunz, V. and Goosse, H.: Influence of freshwater input on the skill of decadal forecast of sea ice in the Southern Ocean, *The Cryosphere*, 9, 541–556, [doi:10.5194/tc-9-541-2015](https://doi.org/10.5194/tc-9-541-2015), 2015.

## 6. Acknowledgements

The authors thank the Belgian Federal Science Policy (Research Program on Science for a Sustainable Development). The authors acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modelling groups for producing and making available their model output. For CMIP the US Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals. H. Goosse is Senior Research Associate with the Fonds National de la Recherche Scientifique (F.R.S.–FNRS-Belgium), Alexander Barthis a F.R.S.–FNRS Research Associate and Violette Zunz is a Research Fellow with the Fonds pour la formation à la Recherche dans l'Industrie et dans l'Agronomie (FRIA-Belgium). Computational resources have been provided by the supercomputing facilities of the Université catholique de Louvain (CISM/UCL) and the Consortium des Equipements de Calcul Intensif en Fédération Wallonie Bruxelles (CECI) funded by the Fond de la Recherche Scientifique de Belgique (F.R.S.–FNRS-Belgium).



## 7. References

- Anderson, J. L.: Spatially and temporally varying adaptive covariance inflation for ensemble filters, *Tellus A*, 61(1), 72–83, 2009.
- Arblaster, J. M. and Meehl, G. A.: Contributions of External Forcings to Southern Annular Mode Trends, *J. Clim.*, 19(12), 2896–2905, 2006.
- Arzel, O., Fichet, T. and Goosse, H.: Sea ice evolution over the 20th and 21st centuries as simulated by current AOGCMs, *Ocean Model.*, 12(3--4), 401–415, 2006.
- Bintanja, R., van Oldenborgh, G. J., Drijfhout, S. S., Wouters, B. and Katsman, C. A.: Important role for ocean warming and increased ice-shelf melt in Antarctic sea-ice expansion, *Nat. Geosci*, 6(5), 376–379, 2013.
- Bitz, C. M. and Polvani, L. M.: Antarctic climate response to stratospheric ozone depletion in a fine resolution ocean climate model, *Geophys. Res. Lett.*, 39(20), doi:10.1029/2012GL053393, 2012.
- Bitz, C. M., Gent, P. R., Woodgate, R. A., Holland, M. M. and Lindsay, R.: The Influence of Sea Ice on Ocean Heat Uptake in Response to Increasing {CO<sub>2</sub>}, *J. Clim.*, 19(11), 2437–2450, 2006.
- Bracegirdle, T. J., Connolley, W. M. and Turner, J.: Antarctic climate change over the twenty first century, *J. Geophys. Res.*, 113(D3), D03103, doi:10.1029/2007JD008933, 2008.
- Brohan, P., Kennedy, J. J., Harris, I., Tett, S. F. B. and Jones, P. D.: Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850, *J. Geophys. Res.*, 111(D12), doi:10.1029/2005JD006548, 2006.
- Capotondi, A., Alexander, M. A., Bond, N. A., Curchitser, E. N. and Scott, J. D.: Enhanced upper ocean stratification with climate change in the CMIP3 models, *J. Geophys. Res. Ocean.*, 117(C4), n/a–n/a, 2012.
- Cavalieri, D. J. and Parkinson, C. L.: Antarctic sea ice variability and trends, 1979--2006, *J. Geophys. Res.*, 113(C7), doi:10.1029/2007JC004564, 2008.
- Cavalieri, D. J., Parkinson, C. L., Gloersen, P. and Zwally, H. J.: Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS Passive Microwave Data., 1996.
- Cavalieri, D. J., Parkinson, C. L. and Vinnikov, K. Y.: 30-Year satellite record reveals contrasting Arctic and Antarctic decadal sea ice variability, *Geophys. Res. Lett.*, 30(18), doi:10.1029/2003GL018031, 2003.
- Christensen, H. M., Moroz, I. M. and Palmer, T. N.: Evaluation of ensemble forecast uncertainty using a new proper score: application to medium-range and seasonal forecasts, *Q. J. R. Meteorol. Soc.*, (January), n/a–n/a, doi:10.1002/qj.2375, 2014.
- Close, S. E. and Goosse, H.: Entrainment-driven modulation of Southern Ocean mixed layer properties and sea ice variability in CMIP5 models, *J. Geophys. Res. Ocean.*, doi:10.1002/jgrc.20226, 2013.

Comiso, J.: Bootstrap Sea Ice Concentrations from Nimbus-7 SMMR and DMSP SSM/I-SSMIS. Version 2, January 1980 to December 2009, 1999.

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

Cotté, C. and Guinet, C.: Historical whaling records reveal major regional retreat of Antarctic sea ice, *Deep Sea Res. Part I Oceanogr. Res. Pap.*, 54(2), 243–252, doi:10.1016/j.dsr.2006.11.001, 2007.

Curran, M. A. J., van Ommen, T. D., Morgan, V. I., Phillips, K. L. and Palmer, A. S.: Ice Core Evidence for {Antarctic} Sea Ice Decline Since the 1950s, *Science (80-. )*, 302(5648), 1203–1206, doi:10.1126/science.1087888, 2003.

Dee, D. P. and Da Silva, A. M.: Data assimilation in the presence of forecast bias, *Q. J. R. Meteorol. Soc.*, 124(545), 269–295, 1998.

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

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

Eastwood, S., Larsen, K. R., Lavergne, T., Nielsen, E. and Tonboer, R.: Global sea ice concentration reprocessing - product user manual., 2011.

Eisenman, I., Meier, W. N. and Norris, J. R.: A spurious jump in the satellite record: has Antarctic sea ice expansion been overestimated?, *Cryosph.*, 8(4), 1289–1296, doi:10.5194/tc-8-1289-2014, 2014.

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

Ferreira, D., Marshall, J., Bitz, C. M., Solomon, S. and Plumb, A.: Antarctic Ocean and Sea Ice Response to Ozone Depletion: A Two-Time-Scale Problem, *J. Clim.*, 28(3), 1206–1226, 2015.

Gagné, M.-È., Gillett, N. P. and Fyfe, J. C.: Observed and simulated changes in Antarctic sea ice extent over the past 50 years, *Geophys. Res. Lett.*, 42(1), 2014GL062231, 2015.

Gent, P. R. and McWilliams, J. C.: Isopycnal Mixing in Ocean Circulation Models, *J. Phys. Oceanogr.*, 20(1), 150–155, 1990.

Di Giuseppe, F., Molteni, F. and Tompkins, A. M.: A rainfall calibration methodology for impacts modelling based on spatial mapping, *Q. J. R. Meteorol. Soc.*, 139(674), 1389–1401, doi:10.1002/qj.2019, 2013.

Glahn, H. R. and Lowry, D. A.: The Use of Model Output Statistics (MOS) in Objective Weather Forecasting, *J. Appl. Meteorol.*, 11(8), 1203–1211, 1972.

Gneiting, T., Raftery, A. E., Westveld, A. H. and Goldman, T.: Calibrated Probabilistic Forecasting Using Ensemble Model Output Statistics and Minimum CRPS Estimation, *Mon. Weather Rev.*, 133(5), 1098–1118, doi:10.1175/MWR2904.1, 2005.

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

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

Goosse, H., Brovkin, V., Fichefet, T., Haarsma, R., Huybrechts, P., Jongma, J., Mouchet, a., Selten, F., Barriat, P.-Y., Campin, J.-M., Deleersnijder, E., Driesschaert, E., Goelzer, H., Janssens, I., Loutre, M.-F., Morales Maqueda, M. a., Opsteegh, T., Mathieu, P.-P., Munhoven, G., Pettersson, E. J., Renssen, H., Roche, D. M., Schaeffer, M., Tartinville, B., Timmermann, a. and Weber, S. L.: Description of the Earth system model of intermediate complexity LOVECLIM version 1.2, *Geosci. Model Dev.*, 3(2), 603–633, doi:10.5194/gmd-3-603-2010, 2010.

Grimit, E. P. and Mass, C. F.: Measuring the Ensemble Spread–Error Relationship with a Probabilistic Approach: Stochastic Ensemble Results, *Mon. Weather Rev.*, 135(1), 203–221, doi:10.1175/MWR3262.1, 2007.

Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in projections of regional precipitation change, *Clim. Dyn.*, 37(1-2), 407–418, doi:10.1007/s00382-010-0810-6, 2011.

Hellmer, H. H.: Impact of Antarctic ice shelf basal melting on sea ice and deep ocean properties, *Geophys. Res. Lett.*, 31(10), doi:10.1029/2004GL019506, 2004.

Holland, M. and Raphael, M.: Twentieth century simulation of the southern hemisphere climate in coupled models. Part II: sea ice conditions and variability, *Clim. Dyn.*, 26(2), 229–245, 2006.

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

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

Hopson, T. M.: Assessing the Ensemble Spread–Error Relationship, *Mon. Weather Rev.*, 142(3), 1125–1142, doi:10.1175/MWR-D-12-00111.1, 2014.

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

Kirkman, C. H. and Bitz, C. M.: The Effect of the Sea Ice Freshwater Flux on Southern Ocean Temperatures in CCSM3: Deep-Ocean Warming and Delayed Surface Warming, *J. Clim.*, 24(9), 2224–2237, 2010.

Klinker, E. and Sardeshmukh, P. D.: The Diagnosis of Mechanical Dissipation in the Atmosphere from Large-Scale Balance Requirements, *J. Atmos. Sci.*, 49(7), 608–627, doi:10.1175/1520-0469(1992)049<0608:TDOMDI>2.0.CO;2, 1992.

De la Mare, W. K.: Abrupt mid-twentieth-century decline in Antarctic sea-ice extent from whaling records, *Nature*, 389(6646), 57–60, 1997.

De la Mare, W. K.: Changes in Antarctic sea-ice extent from direct historical observations and whaling records, *Clim. Change*, 92(3), 461–493, 2009.

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

De Lavergne, C., Palter, J. B., Galbraith, E. D., Bernardello, R. and Marinov, I.: Cessation of deep convection in the open Southern Ocean under anthropogenic climate change, *Nat. Clim. Chang.*, 4(4), 278–282, 2014.

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

Van Leeuwen, P. J.: Nonlinear data assimilation in geosciences: an extremely efficient particle filter, *Q. J. R. Meteorol. Soc.*, 136(653), 1991–1999, 2010.

Lefebvre, W. and Goosse, H.: An analysis of the atmospheric processes driving the large-scale winter sea ice variability in the Southern Ocean, *J. Geophys. Res.*, 113(C2), C02004, doi:10.1029/2006JC004032, 2008a.

Lefebvre, W. and Goosse, H.: Analysis of the projected regional sea-ice changes in the Southern Ocean during the twenty-first century, *Clim. Dyn.*, 30(1), 59–76, 2008b.

Lefebvre, W., Goosse, H., Timmermann, R. and Fichefet, T.: Influence of the Southern Annular Mode on the sea ice-ocean system, *J. Geophys. Res. Ocean.*, 109(C9), C09005, 2004.

Li, H., Kalnay, E. and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, *Q. J. R. Meteorol. Soc.*, 135(639), 523–533, 2009.

Lisæter, K. A., Rosanova, J. and Evensen, G.: Assimilation of ice concentration in a coupled ice-ocean model, using the Ensemble Kalman filter, *Ocean Dyn.*, 53(4), 368–388, doi:10.1007/s10236-003-0049-4, 2003.

Lisæter, K. A., Evensen, G. and Laxon, S.: Assimilating synthetic CryoSat sea ice thickness in a coupled ice-ocean model, *J. Geophys. Res. Ocean.*, 112(C7), n/a–n/a, doi:10.1029/2006JC003786, 2007.

Liu, J. and Curry, J. A.: Accelerated warming of the Southern Ocean and its impacts on the hydrological cycle and sea ice, *Proc. Natl. Acad. Sci.*, 107(34), 14987–14992, 2010.

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

Lorenz, E. N.: Deterministic Nonperiodic Flow, *J. Atmos. Sci.*, 20(2), 130–141, 1963.

Lorenz, E. N.: Formulation of a Low-Order Model of a Moist General Circulation, *J. Atmos. Sci.*, 41(12), 1933–1945, doi:10.1175/1520-0469(1984)041<1933:FOALOM>2.0.CO;2, 1984.

Lorenz, E. N.: Predictability: A problem partly solved, in Proc. Seminar on predictability, vol. 1., 1996.

Lorenz, E. N. and Emanuel, K. A.: Optimal sites for supplementary weather observations: Simulation with a small model, *J. Atmos. Sci.*, 55(3), 399–414, 1998.

Mahlstein, I., Gent, P. R. and Solomon, S.: Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations, *J. Geophys. Res. Atmos.*, 118, 1–6, doi:10.1002/jgrd.50443, 2013.

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

Massonnet, F., Goosse, H., Fichefet, T. and Counillon, F.: Calibration of sea ice dynamic parameters in an ocean-sea ice model using an ensemble Kalman filter, *J. Geophys. Res. Ocean.*, 119(7), 4168–4184, 2014.

Mathiot, P., König Beatty, C., Fichefet, T., Goosse, H., Massonnet, F. and Vancoppenolle, M.: Better constraints on the sea-ice state using global sea-ice data assimilation, *Geosci. Model Dev.*, 5(6), 1501–1515, doi:10.5194/gmd-5-1501-2012, 2012.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D. and Stockdale, T.: Decadal Prediction: {Can it be skillful?}, *Bull. Am. Meteorol. Soc.*, 90(10), 1467–1485, 2009.

Meehl, G. a., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M., Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vecchi, G. and Yeager, S.: Decadal Climate Prediction: An Update from the Trenches, *Bull. Am. Meteorol. Soc.*, 95(2), 243–267, doi:10.1175/BAMS-D-12-00241.1, 2014.

Nicolis, C., Perdigo, R. a. P. and Vannitsem, S.: Dynamics of Prediction Errors under the Combined Effect of Initial Condition and Model Errors, *J. Atmos. Sci.*, 66(3), 766–778, doi:10.1175/2008JAS2781.1, 2009.

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

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

Pham, D. T.: Stochastic Methods for Sequential Data Assimilation in Strongly Nonlinear Systems, *Mon. Weather Rev.*, 129(5), 1194–1207, 2001.

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

Polvani, L. M. and Smith, K. L.: Can natural variability explain observed Antarctic sea ice trends? New modeling evidence from CMIP5, *Geophys. Res. Lett.*, 40(12), 3195–3199, doi:10.1002/grl.50578, 2013.

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

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

Roberts-Jones, J., Fiedler, E. K. and Martin, M. J.: Daily, global, high-resolution SST and sea ice reanalysis for 1985-2007 using the OSTIA system, *J. Clim.*, 25(18), 6215–6232, doi:10.1175/JCLI-D-11-00648.1, 2012.

Rodwell, M. J. and Palmer, T. N.: Using numerical weather prediction to assess climate models, *Q. J. R. Meteorol. Soc.*, 133(622), 129–146, doi:10.1002/qj.23, 2007.

Rollenhagen, K., Timmermann, R., Janjić, T., Schröter, J. and Danilov, S.: Assimilation of sea ice motion in a finite-element sea ice model, *J. Geophys. Res. Ocean.*, 114(C5), n/a–n/a, 2009.

Roulin, E. and Vannitsem, S.: Postprocessing of Ensemble Precipitation Predictions with Extended Logistic Regression Based on Hindcasts, *Mon. Weather Rev.*, 140(3), 874–888, doi:10.1175/MWR-D-11-00062.1, 2012.

Roulin, E. and Vannitsem, S.: Post-processing of medium-range probabilistic hydrological forecasting : impact of forcing , initial conditions and model errors, *Hydrol. Process.*, 29(July 2014), 1434–1449, doi:10.1002/hyp.10259, 2015.

Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E. and Taylor, K. E.: 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, 2000.

Van Schaeybroeck, B. and Vannitsem, S.: Post-processing through linear regression, *Nonlinear Process. Geophys.*, 18(2), 147–160, doi:10.5194/npg-18-147-2011, 2011.

Van Schaeybroeck, B. and Vannitsem, S.: Toward post-processing ensemble forecasts based on hindcasts \*, *Publ. Sci. Tech. R. Meteorol. Inst. Belgium*, 61, 2012.

Van Schaeybroeck, B. and Vannitsem, S.: Reliable probabilities through statistical post-processing of ensemble predictions, *Proc. Eur. Conf. Complex Syst. 2012*, Ed. by Gilbert, Kirkilionis, Nicolis, Vannitsem, 347–352, doi:10.1007/978-3-319-00395-5\_45, 2013.

Van Schaeybroeck, B. and Vannitsem, S.: A probabilistic approach to forecast the skill with ensemble spread, *Mon. Weather Rev.*, submitted, 2014.

Van Schaeybroeck, B. and Vannitsem, S.: Assessment of calibration assumptions under strong climate changes, *Geophys. Res. Lett.*, submitted, 1–24, 2015a.

Van Schaeybroeck, B. and Vannitsem, S.: Ensemble post-processing using member-by-member approaches: theoretical aspects, *Q. J. R. Meteorol. Soc.*, 141(April), 807–818, doi:10.1002/qj.2397, 2015b.

Van Schaeybroeck, B., Mailier, P. and Vannitsem, S.: Ensemble post-processing in Belgium based on reforecasts, *Weather Forecast.*, (in preparation), 2015.

Schefzik, R., Thorarinsdottir, T. L. and Gneiting, T.: Uncertainty Quantification in Complex Simulation Models Using Ensemble Copula Coupling, *Stat. Sci.*, 28(4), 616–640, doi:10.1214/13-STS443, 2013.

Shepherd, A., Ivins, E. R., A, G., Barletta, V. R., Bentley, M. J., Bettadpur, S., Briggs, K. H., Bromwich, D. H., Forsberg, R., Galin, N., Horwath, M., Jacobs, S., Joughin, I., King, M. A., Lenaerts, J. T. M., Li, J., Ligtenberg, S. R. M., Luckman, A., Luthcke, S. B., McMillan, M., Meister, R., Milne, G., Mougintot, J., Muir, A., Nicolas, J. P., Paden, J., Payne, A. J., Pritchard, H., Rignot, E., Rott, H., Sørensen, L. S., Scambos, T. A., Scheuchl, B., Schrama, E. J. O., Smith, B., Sundal, A. V., van Angelen, J. H., van de Berg, W. J., van den Broeke, M. R., Vaughan, D. G., Velicogna, I., Wahr, J., Whitehouse, P. L., Wingham, D. J., Yi, D., Young, D. and Zwally, H. J.: A Reconciled Estimate of Ice-Sheet Mass Balance, *Science* (80-), 338(6111), 1183–1189, doi:10.1126/science.1228102, 2012.

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

Sigmond, M. and Fyfe, J. C.: The {Antarctic} Sea Ice Response to the Ozone Hole in Climate Models, *J. Clim.*, 27(3), 1336–1342, doi:10.1175/JCLI-D-13-00590.1, 2013.

Simpkins, G. R., Ciasto, L. M., Thompson, D. W. J. and England, M. H.: Seasonal Relationships between Large-Scale Climate Variability and {Antarctic} Sea Ice Concentration, *J. Clim.*, 25(16), 5451–5469, 2012.

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

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

Stammerjohn, S. E., Martinson, D. G., Smith, R. C., Yuan, X. and Rind, D.: Trends in Antarctic annual sea ice retreat and advance and their relation to El Niño Southern Oscillation and Southern Annular Mode variability, *J. Geophys. Res.*, 113(C3), doi:10.1029/2007JC004269, 2008.

Stroeve, J. C., Kattsov, V., Barrett, A., Serreze, M., Pavlova, T., Holland, M. and Meier, W. N.: Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations, *Geophys. Res. Lett.*, 39(16), doi:10.1029/2012GL052676, 2012.

Swart, N. C. and Fyfe, J. C.: Observed and simulated changes in the Southern Hemisphere surface westerly wind-stress, *Geophys. Res. Lett.*, 39(16), L16711, 2012.

Swart, N. C. and Fyfe, J. C.: The influence of recent Antarctic ice sheet retreat on simulated sea ice area trends, *Geophys. Res. Lett.*, 40(16), 4328–4332, 2013.

Swingedouw, D., Fichefet, T., Huybrechts, P., Goosse, H., Driesschaert, E. and Loutre, M.-F.: Antarctic ice-sheet melting provides negative feedbacks on future climate warming, *Geophys. Res. Lett.*, 35(17), doi:10.1029/2008GL034410, 2008.

Taylor, A. a. and Leslie, L. M.: A Single-Station Approach to Model Output Statistics Temperature Forecast Error Assessment, *Weather Forecast.*, 20(6), 1006–1020, doi:10.1175/WAF893.1, 2005.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A.: An Overview of {CMIP5} and the Experiment Design, *Bull. Am. Meteorol. Soc.*, 93(4), 485–498, 2011.

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

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

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

Vannitsem, S.: A unified linear Model Output Statistics scheme for both deterministic and ensemble forecasts, *Q. J. R. Meteorol. Soc.*, 135(644), 1801–1815, doi:10.1002/qj.491, 2009a.

Vannitsem, S.: A unified linear Model Output Statistics scheme for both deterministic and ensemble forecasts, *Q. J. R. Meteorol. Soc.*, 135(644), 1801–1815, doi:10.1002/qj.491, 2009b.

Vannitsem, S.: Bias correction and post-processing under climate change, *Nonlinear Process. Geophys.*, 18(6), 911–924, doi:10.5194/npg-18-911-2011, 2011.

Vannitsem, S. and Hagedorn, R.: Ensemble forecast post-processing over Belgium: comparison of deterministic-like and ensemble regression methods, *Meteorol. Appl.*, 18(1), 94–104, doi:10.1002/met.217, 2011a.

Vannitsem, S. and Hagedorn, R.: Ensemble forecast post-processing over Belgium: comparison of deterministic-like and ensemble regression methods, *Meteorol. Appl.*, 18(1), 94–104, doi:10.1002/met.217, 2011b.

Vannitsem, S. and Nicolis, C.: Dynamical Properties of Model Output Statistics Forecasts, *Mon. Weather Rev.*, 136(2), 405–419, doi:10.1175/2007MWR2104.1, 2008.

Vaughan, D. G., Comiso, J. C., Allison, I., Carrasco, J., Kwok, R., Mote, P., Murray, T., Paul, F., Ren, J., Rignot, E., Solomina, O., Steffen, K. and Zhang, T.: Observations: Cryosphere, 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., 2013.

Velicogna, I.: Increasing rates of ice mass loss from the Greenland and Antarctic ice sheets revealed by GRACE, *Geophys. Res. Lett.*, 36(19), L19503, doi:10.1029/2009GL040222, 2009.

Wilks, D. S.: Comparison of ensemble-MOS methods in the Lorenz '96 setting, *Meteorol. Appl.*, 13(03), 243, doi:10.1017/S1350482706002192, 2006.

Worby, A. P., Geiger, C. A., Paget, M. J., Van Woert, M. L., Ackley, S. F. and DeLiberty, T. L.: Thickness distribution of Antarctic sea ice, *J. Geophys. Res. Ocean.*, 113(C5), doi:10.1029/2007JC004254, 2008.

Zhang, J.: Increasing {Antarctic} Sea Ice under Warming Atmospheric and Oceanic Conditions, *J. Clim.*, 20(11), 2515–2529, 2007.

Zunz, V. and Gousse, H.: Influence of freshwater input on the skill of decadal forecast of sea ice in the Southern Ocean, *Cryosph.*, 9(2), 541–556, doi:10.5194/tc-9-541-2015, 2015.

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

Zunz, V., Gousse, H. and Dubinkina, S.: Impact of the initialisation on the predictability of the Southern Ocean sea ice at interannual to multi-decadal timescales, *Clim. Dyn.*, 44(7-8), 2267–2286, 2015.



## 8. Annexes

Annexe 1: Minutes

Annexe 2: Publications

Annexe 3: Understanding and predicting Antarctic sea ice variability at the decadal timescale

The annexes are available on the website

[http://www.belspo.be/belspo/SSD/science/pr\\_climate\\_en.stm](http://www.belspo.be/belspo/SSD/science/pr_climate_en.stm)