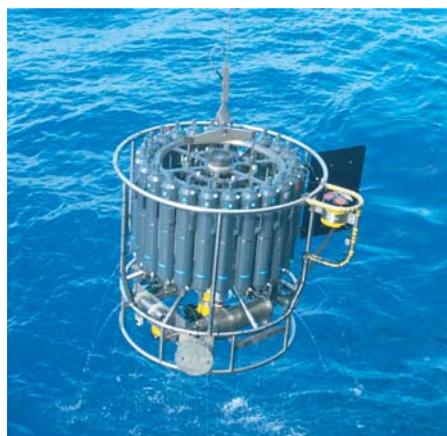




Impact of initialization procedures on the predictive skill
of a coupled ocean-atmosphere model and
related mechanisms for predictability

Iuliia Polkova



Hinweis

Die Berichte zur Erdsystemforschung werden vom Max-Planck-Institut für Meteorologie in Hamburg in unregelmäßiger Abfolge herausgegeben.

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

Die "Berichte zur Erdsystemforschung" führen die vorherigen Reihen "Reports" und "Examensarbeiten" weiter.

Notice

The Reports on Earth System Science are published by the Max Planck Institute for Meteorology in Hamburg. They appear in irregular intervals.

They contain scientific and technical contributions, including Ph. D. theses.

The Reports do not necessarily reflect the opinion of the Institute.

The "Reports on Earth System Science" continue the former "Reports" and "Examensarbeiten" of the Max Planck Institute.



Anschrift / Address

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg
Deutschland

Tel.: +49-(0)40-4 11 73-0
Fax: +49-(0)40-4 11 73-298
Web: www.mpimet.mpg.de

Layout:

Bettina Diallo, PR & Grafik

Titelfotos:

vorne:

Christian Klepp - Jochem Marotzke - Christian Klepp

hinten:

Clotilde Dubois - Christian Klepp - Katsumasa Tanaka

Impact of initialization procedures on the predictive skill
of a coupled ocean-atmosphere model and
related mechanisms for predictability

Iuliia Polkova

Hamburg 2014

Iuliia Polkova
Max-Planck-Institut für Meteorologie
und
Institut für Meereskunde
Universität Hamburg
Bundesstrasse 53
20146 Hamburg

- Korrigierte Fassung -

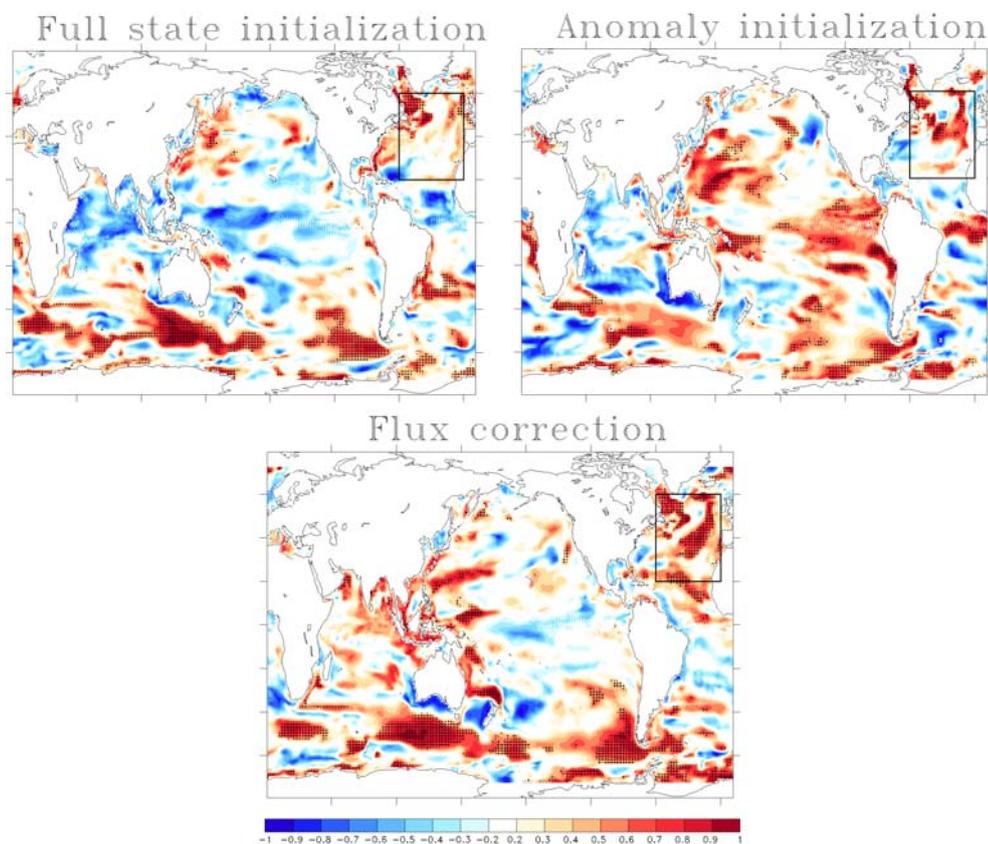
Als Dissertation angenommen vom Fachbereich Geowissenschaften
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
der Universität Hamburg

Gutachter:
Prof. Dr. Detlef Stammer
und
Dr. Armin Köhl

Tag der Disputation:
31. Januar 2014

Prof. Dr. Christian Betzler
Leiter Fachbereich Geowissenschaften

Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model and related mechanisms for predictability



Iuliia Polkova

Hamburg 2014

Cover Figure: Sensitivity of the predictive skill to three initialization procedures
Correlation for detrended SST hindcasts at lead time yr2-5

Abstract

The sensitivity of the predictive skill of a decadal climate prediction system is investigated with respect to details of the initialization procedure. For this purpose, the coupled ocean-atmosphere UCLA/MITgcm climate model is initialized using the following three different initialization approaches: full state initialization (FSI), anomaly initialization (AI) and full state initialization employing heat flux and freshwater flux corrections (FC). The ocean initial conditions are provided by the GECCO state estimate (the German contribution to Estimating the Circulation and Climate of the Ocean project), from which ensembles of decadal hindcasts are started every 5 years from 1961 to 2001. To evaluate the performance of the initialized hindcasts, they are compared with the persistence forecasts and an ensemble of twentieth-century simulations (un-initialized hindcasts).

The study is divided in two main parts. In the first part the primary focus is to estimate the predictive skill for differently initialized hindcasts of sea surface temperature (SST), sea surface height (SSH) and the Atlantic meridional overturning circulation (AMOC). The predictive skill for SST, SSH and AMOC is assessed against the GECCO synthesis using anomaly correlation coefficient and root-mean-squared-error skill score. In regions with a deep mixed layer the predictive skill for SST anomalies remains significant for up to a decade in the FC experiment. By contrast, FSI shows less persistent skill in the North Atlantic and AI does not show high skill in the extratropical Southern Hemisphere, but appears to be more skillful in the tropics. In the extratropics, the improved skill is related to the ability of the FC initialization method to better represent the mixed layer depth, and the highest skill occurs during wintertime. The correlation skill for the spatially averaged North Atlantic SSH hindcasts remains significant up to a decade only for FC. The North Atlantic MOC initialized hindcasts show high correlation values in the first pentad while correlation remains significant in the following pentad too for FSI and FC. Overall, for the current setup, the FC approach appears to lead to the best results, followed by the FSI and AI procedures.

An extended analysis of predictive skill for SSH hindcasts from different initialization experiments is performed in the second part. The analysis employs a method that allows to distinguish different contributions to steric SSH changes, namely those related to density changes imposed by temperature or salinity anomalies beneath the mixed layer (thermosteric and halosteric heave terms) and those related to processes of density compensated temperature–salinity changes (spice term). The spice and heave contributions in the mixed layer are not separated (mixed layer term). The patterns of the predictive skill suggest significant improvement of initialization which is related to the thermosteric heave term (in the subtropical Pacific, western North Atlantic), thermosteric mixed layer term (in the subtropical Atlantic) and spice term (in the eastern and subpolar North Atlantic and Southern Ocean). These contributions imply useful predictive skill as they occur in

the regions of large interannual variability and show improvement beyond the persistence skill. In the North Atlantic, all the initialized hindcasts seem to correctly represent the thermosteric SSH changes in the mixed layer, while for the quantities below the mixed layer, the AI hindcasts appear to have no predictive skill, in contrast to FSI and FC.

Zusammenfassung

Die allgemeine Zielsetzung dieser Doktorarbeit besteht darin, die Empfindlichkeit der dekadischen Vorhersagen des UCLA/MITgcm Modells in Abhängigkeit verschiedener Initialisierungsstrategien, insbesondere bei vollständiger Initialisierung (FSI), bei Anomalie-Initialisierung (AI) und bei der Flux-Korrektur (FC), zu untersuchen. Die Anfangsbedingungen sind der GECCO-Ozeansynthese (the German contribution to Estimating the Circulation and Climate of the Ocean project) entnommen. Die initialisierten Hindcasts (retrospektive Experimente) für den Zeitraum von 1961 bis 2001 werden in 5-Jahres-Schritten gestartet. Um die Qualität der initialisierten Hindcasts zu bewerten, wurden sie mit Persistenzprognosen und einem Ensemble von nichtinitialisierte Klimasimulationen verglichen.

Die Studie gliedert sich in zwei Hauptteile. Im ersten Teil wird die Vorhersagegüte der verschiedenen initialisierten dekadischen Vorhersagen der Meeresoberflächentemperatur (SST), des Meeresspiegels (SSH) und der atlantischen meridionalen Umwälzzirkulation (AMOC) geschätzt. Die Vorhersagegüte von SST, SSH und AMOC wird mit den GECCO-Daten evaluiert. Der Vergleich erfolgt anhand des Korrelationskoeffizienten und des mittleren quadratischen Fehlers. In Regionen mit einer tiefreichenden Mischungsschicht bleibt die Vorhersagegüte für SST-Anomalien bis zu einem Jahrzehnt signifikant. Im Gegenteil dazu zeigt sich bei der FSI-Methode für den Nordatlantik eine wenig stabile Vorhersagegüte und bei der AI für die außertropische südliche Hemisphäre keine signifikante Vorhersagegüte. In den Außertropen beruht die bessere Vorhersagegüte auf der besseren Darstellung der Mischungsschichttiefen durch die FC-Methode, und die beste Vorhersagegüte wird im Winter erreicht. Nur die FC-Methode führt zu guten Vorhersagen für Nordatlantik SSH für einen Zeitraum von bis zu zehn Jahren. Die für den nördlichen AMOC berechneten Hindcasts zeigen eine gute Korrelation während den ersten fünf Jahren. Für FSI und FC bleibt die Korrelation auch für die folgenden fünf Jahre signifikant. Für den gewählten Experimentrahmen zeigt die FC-Methode die besten Resultate, gefolgt von FSI und AI.

Eine erweiterte Analyse der Vorhersagegüte für SSH Hindcasts für verschiedene Initialisierungs-Strategien ist im zweiten Teil erfolgt. Die Analyse basiert auf einer Methode, die es erlaubt, die zugrundeliegenden Prozesse von sterischer SSH-Änderungen zu unterscheiden. Sterische SSH-Änderungen sind in die folgenden Terme aufgeteilt: Dichteveränderungen aufgrund von Temperatur- oder Salzgehaltsänderungen unterhalb der Mischungsschichttiefe (Heave Term) und dichtekompensierte Temperatur- und Salzgehaltsänderungen (Spice Term). Die Heave und Spice Terme in der Mischungsschichttiefe wurden nicht aufgeteilt (Mixed-layer Term). Die Vorhersage des Heave Terms (wegen Temperaturänderungen im subtropischen Pazifik und westlichen Nordatlantik), des Mix-layer Terms (wegen Temperaturänderungen im subtropischen Atlantik) und des Spice Terms (im östlichen und subpolaren Nordatlantik und Südliche Ozean) zeigt eine Verbesserung

durch die Initialisierung. Diese Komponenten zeigen nutzbare Vorhersagegüten aufgrund guter Korrelation in Regionen mit großer zwischenjährlicher Variabilität und eine bessere Vorhersagegüte als die Persistenzprognosen.

Acknowledgement

I am especially thankful to my scientific adviser Detlef Stammer for his support during my PhD study, for motivating with the research and encouraging me to publish. I also wish to thank my co-adviser Armin Köhl for plenty of useful discussions and putting up with endless comments and questions on experimental design and analysis methods. And one more time thanks to both for many enlightening conversations and ideas. I would also like to thank Johanna Baehr for chairing my advisory panel, useful discussions and many helpful advices.

I am grateful to the staff of the Remote Sensing and Assimilation Group from the Institute of Oceanography and the Max Planck Institute for Meteorology for always having the doors open and for their willingness to help and having a most enjoyable PhD time. I am especially thankful to Neeraj Agarwal for his help to get started with the coupled model and to set up the model runs, and for discussion of the results. I would also like to acknowledge providing climatology of 300-year spin-up run for constructing initial conditions for the anomaly initialized experiments. I wish to thank Nuno Serra and Nikolay Koldunov for valuable help with computing environment, for their advices and for always having the right words to cheer me up. Very special thanks to Camille Marini for her help with Matlab scripts, inspiration to finish my PhD study and extremely positive thinking. I would also like to thank Mark Carson and Benjamin Runkle for proofreading of the third Chapter of this thesis. I thank also Holger Pohlmann, Daniela Matei, Johanna Baehr and Jochem Marotzke and two anonymous reviewers who helped to improve Chapter 3 significantly during the review process for the Climate Dynamics.

I would like to thank Max Plank Institute for Meteorology for financial support and for giving me the opportunity to expand my knowledge about the Earth's climate system. All model simulations were made in the Deutsches Klimarechenzentrum (DKRZ). Thanks IT support group of CIS. Special thanks also go to the Office of the graduate school SICSS for many useful courses, meetings and retreats.

Many, many thanks to my husband Norman just for everything: his spirit, all the support words for any situation, taking care of me, reading and commenting on the first drafts of the thesis and paper. I wish to thank my mum and grandma, my family in Ukraine and here in Hamburg for their love and support of my decisions. A special thank to my former adviser Anatolij Lytovchenko and to Victor Dolin for the inspiration to become a scientist and motivation to start this PhD study. Many thanks to my friends and office mates that I have shared my time with: Reema, Xueyuan, Marlene, Natalia, Dima, Vika, Yulia and many others.

Contents

Abstract	i
Zusammenfassung	iii
Acknowledgement	v
1 Introduction	1
1.1 Motivation	1
1.2 Achievements in the field of decadal predictions	1
1.3 Goals of the thesis	6
1.4 Structure of the thesis	7
2 Methodology for Exploring Decadal Predictability	9
2.1 Model and climate parameters	10
2.1.1 Description of a coupled model	10
2.1.2 Parameters to study decadal predictability	11
2.2 Experiments	15
2.2.1 Spin-up run	15
2.2.2 Twentieth-Century Simulations	16
2.2.3 Assimilation runs and initialized hindcasts	20
2.3 Initialization of decadal climate predictions	20
2.3.1 Initial conditions (GECCO)	21
2.3.2 Full state initialization	22
2.3.3 Anomaly initialization	23
2.3.4 Full state initialization employing flux correction	23
2.4 Predictive Skill Estimation	26
2.4.1 Trend removal	26
2.4.2 Bias correction	28
2.4.3 Verification Metrics	29
2.4.4 Persistence	30
2.4.5 Assessment of Significance Level	31

3	Impact of Initialization Procedures on Predictive Skill of a CGCM	33
3.1	Introduction	33
3.2	Methodology	35
3.3	Initialized Decadal SST Hindcasts	40
3.3.1	Global distribution of SST predictive skill	40
3.3.2	The North Atlantic SST predictability	42
3.3.3	Possible mechanism responsible for the NA SST predictability	45
3.4	Initialized Decadal SSH Hindcasts	46
3.5	Initialized Decadal AMOC Hindcasts	49
3.6	Summary and Discussion	51
4	Mechanisms Associated with Decadal Predictability	57
4.1	Introduction	57
4.1.1	Motivation	57
4.1.2	Observed SSH changes and its contributions	58
4.1.3	Regional SSH changes and SSH forcing factors	60
4.1.4	Some candidate mechanisms for decadal variability	61
4.2	Steric SSH decomposition analysis	63
4.3	Variability of steric SSH changes and its contributions from GECCO	65
4.4	Predictability of thermosteric and halosteric SSH changes	68
4.5	Predictability of SSH heave, spice and mixed-layer terms	72
4.5.1	Predictive skill in different initialization experiments	72
4.5.2	Indication of mechanisms for predictability	76
4.6	Concluding remarks	79
5	Conclusions and outlook	83
5.1	Thesis overview and main conclusions	83
5.1.1	Improving decadal predictions through different initialization procedures	83
5.1.2	Sensitivity of the results	84
5.1.3	Processes giving rise to predictive skill	85
5.2	Recommendations and future work	87
	Acronyms	91
	Bibliography	93
	Declaration	107

Chapter 1

Introduction

1.1 Motivation

Climate change projections indicate that global average temperature is expected to continue increasing, so do ocean heat content and sea level (IPCC IPCC, 2013). Whereas historical temperature records suggest that our planet is warmer than it was in 1970s, observation records over the last decade have detected the slowdown of warming (Knight et al., 2009; Met Office Hadley Centre, 2013). One of the hypotheses is that natural internal variability is responsible for the pause. However, longer periods of observation would be needed in order to distinguish the natural climate variability from the longer-term climate change. On the other hand, the state-of-the-art climate models represent powerful tools to predict and analyze such climate phenomena. Moreover, modeling studies by Knight et al. (2009), Meehl et al. (2011) and Guemas et al. (2013) have shown that the models do contain necessary mechanisms of climate variability and are able to capture the slowdown in global temperature rise over the last decade. Though climate predictions still contain uncertainties related to internal variability, model errors and external forcing, Hawkins and Sutton (2009) speculate that proper initialization of climate predictions with the observations of current climate state has the potential to narrow the uncertainty from internal variability (Fig. 1.1) and by enabling comparison of climate predictions with observations, it allows to identify model errors and consequently reduce model-related uncertainty.

1.2 Achievements in the field of decadal predictions

In the late 1990s climate research community introduces a novelty – decadal climate predictions which mean to predict climate fluctuations for the nearest future (up to 10 years) starting from the observed state of the climate system. This new field intends to bridge a gap between seasonal forecasts and long-term climate change projections and is expected to find implications for climate change adaptation strategies, agricultural infrastructure management, investments planning, etc. The decadal predictability research began with

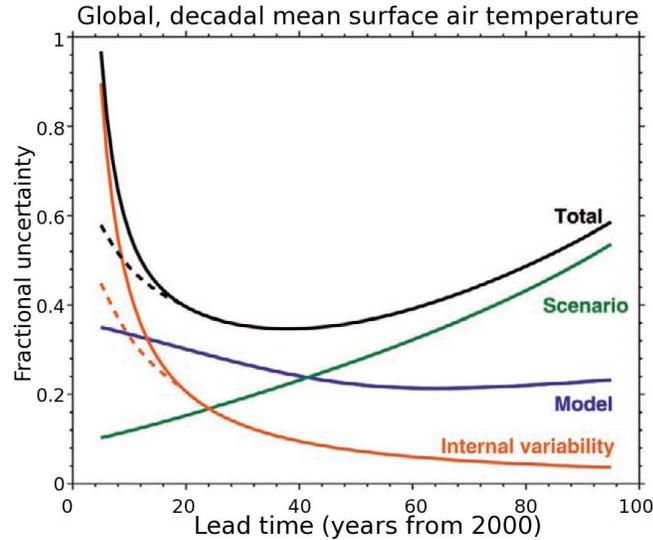


Figure 1.1— The relative importance of each source of uncertainty in decadal mean surface air temperature predictions is shown by the fractional uncertainty (the 90% confidence level divided by the mean prediction), for the global mean, relative to the warming since the year 2000 (i.e., a lead of zero years). The dashed lines indicate reductions in internal variability, and hence total uncertainty, that may be possible through proper initialization of the predictions through assimilation of ocean observations (Smith et al., 2007). Adapted from Hawkins and Sutton (2009).

“perfect model experiments” (Griffies and Bryan, 1997; Boer, 2000; Collins, 2002; Collins et al., 2006). The lack of available historical observation records for assessment of predictability stipulated usage of numerical model simulations (Griffies and Bryan, 1997). To quantify predictability with a “perfect model” approach, a set of integrations with a coupled model is usually carried out, starting from atmospheric conditions chosen randomly from the model’s climatology, but with identical oceanic initial conditions. Thereafter, predictability of the system is assessed as a “rate of separation” of model solutions to give an indication of the error growth rate (Boer, 2000). This approach does not compare the model simulations to observations, and is used to assess the ability of the model to reproduce itself given an amount of uncertainty in initial conditions. Whereas continuously new observed climate data become available, this method is being widely applied. For instance, recent studies by Branstator and Teng (2010) and Teng et al. (2011) use “perfect model” approach to analyze initial-value predictability versus predictability of the forced response in the Atlantic basin.

The “perfect model” studies have shown that the Atlantic meridional overturning circulation (AMOC) could be predicted for about a decade (Collins et al., 2006; Teng et al., 2011). Whereas global mean surface temperature is predictable only one year ahead, on a regional scale a wide range of potential predictability is found, for instance, the North Atlantic yields predictability for about five years (Griffies and Bryan, 1997; Collins, 2002; Collins et al., 2006).

Along with “perfect model” experiments, another potential predictability approach has been studied by Boer (2000, 2004), Pohlmann et al. (2004) and Boer and Lambert (2008) and others, and termed as “diagnostic” potential predictability approach. These studies are based on variance analysis that identifies regions where long timescale variance is a useful fraction of the total variance and mean to indicate that prediction is potentially possible. In this respect, potential predictability is found at mid to high latitudes, in particular over the North Atlantic and the Southern Ocean (Boer, 2004). In contrast to the approach discussed previously, “diagnostic” potential predictability analysis can be applied to both real and modeled systems. Boer (2011) states that potential predictability provides information about the upper limit of predictability.

Above-mentioned approaches utilize the terms (potential) predictability and (potential) predictive skill. Predictability, as defined in Glossary of Meteorology, is “the extent to which future states of the climate system may be predicted based on knowledge of current and past states of the system” (Glickman, 2000). Boer (2000) explains the term potential predictive skill as a “statistical and indirect concept where the argument is that the observed or modeled variability of some mean quantity is greater than can be accounted for by sampling error given the noise in the system”. In general, potential predictive skill can be assessed for both observation dataset and decadal climate prediction system. For the latter, potential predictive skill can be assessed in terms of error growth, correlation (Collins, 2002), EOFs (Griffies and Bryan, 1997), relative entropy (Branstator and Teng, 2010), etc. Moreover, as shown by Hawkins et al. (2011) predictive skill can also be assessed for statistical methods that are used for analyzing the quality of decadal predictions. Eventually, as more observational records became available, research groups began to assess predictive skill of a climate system from a set of predictions for the past climate (hindcasts) and their comparison with the observational counterpart. When predictions are verified against real observations, verification metrics signify actual predictive skill of the system (Smith et al., 2007). However, it is still difficult to estimate statistically reliable predictive skill for many climate variables, simply because of short observational records’ time (e.g., for sea surface height and AMOC). Tebaldi and Knutti (2007) state that models can also be evaluated using reanalysis data. The study by Boer et al. (2013) compared actual predictive skill (also called forecast skill) and potential predictive skill for decadal predictions and concluded that the latter one does not imply actual skill, meaning that if models have deficiencies in simulating processes in some regions, this will result in low forecast skill where significant potential predictive skill might still exist.

The studies by Smith et al. (2007), Keenlyside et al. (2008) and Pohlmann et al. (2009) started a new era of decadal prediction experiments, namely predictions initialized with the observed state of the climate system. Meehl et al. (2009, accepted, 2013) drew attention of the climate modeling community to the issues in focus regarding decadal climate predictions, and two of the main questions that are still not entirely resolved

- how to initialize decadal predictions and
- how to evaluate decadal prediction skill.

Attempting to answer these questions, the decadal climate research community adopted much of the experience from the seasonal-to-interannual forecast practice (Goddard et al., 2012). Thus, the initialization strategy that has now been widely used for decadal predictions is well known from seasonal-to-interannual forecasting, namely full state initialization (FSI). Because of model's drift, FSI has a significant drawback for predictive skill in seasonal-to-interannual and, subsequently, in decadal prediction systems. In seasonal forecasting the common practice is to remove the drift by means of a posteriori bias-correction (Stockdale, 1997), while Meehl et al. (2009) indicated that this might also remove some part of useful signal that is masked as bias. Alternative ways of initialization of decadal climate predictions that account for drift are anomaly initialization method (AI) and flux correction (FC). Though, it is not yet clear which of the initialization techniques is the most appropriate for decadal climate prediction because all three initialization techniques may have some technical problems. For instance, at the beginning (on short lead times) initialized predictions may face a problem of initialization shock due to an imbalance between the initial conditions and model dynamics, this may cause abrupt changes or non-linear interaction between model drift and the evolution of model state. Other issues related to different initialization strategies for decadal predictions were discussed earlier by Pierce et al. (2004), Goddard et al. (2013), Magnusson et al. (2012a,b) and Smith et al. (2013).

Because initialization of decadal prediction embrace not only the method of initialization but also the initial conditions, it is important to consider the quality of initial conditions and their possible impact on forecast quality (Doblas-Reyes et al., 2011a). As was mentioned before initialization shock is a problem of initialization methods but it is expected to be reduced if the initial conditions are derived in the coupled data assimilation approach which aims to incorporate observational information from the ocean, atmosphere, ice and land surface into a coupled model (Sugiura et al., 2008). So far initial conditions were mostly derived from atmospheric or oceanic general circulation models. In principle all components of the climate system should be initialized (in studies by Smith et al. (2007), Troccoli and Palmer (2007), Doblas-Reyes et al. (2011a) and Magnusson et al. (2012a) both the ocean and atmosphere were initialized). But since most of the long-term memory resides in the ocean, it is not uncommon to initialize only ocean component of the coupled model (Keenlyside et al., 2008; Pohlmann et al., 2009; Matei et al., 2012). The studies by Carton and Santorelli (2008), Stammer et al. (2009), Lee et al. (2010) and Munoz et al. (2011) give overviews of the performance of available ocean state estimates. Evaluating different ocean synthesis products on potential applicability to provide initial conditions for decadal forecasting, Kröger et al. (2012) came to a conclusion that the choice of the product as the source of initial conditions should be based on an

agreement between climate variations obtained after assimilating the ocean state estimate into the model and variations of the original state estimate.

The quality of initialized decadal predictions is often addressed by using deterministic metrics (Smith et al., 2007, 2013; Pohlmann et al., 2009; Matei et al., 2012). A shortcoming of these methods is that information about prediction uncertainty is not available. Tebaldi and Knutti (2007) suggest that the best verification approach should include multiple diagnostics of performance that identify uncertainty, possible correlations, and limitations in the model's performance. In this respect, in the verification framework for decadal climate predictions Goddard et al. (2013) proposed to additionally use probabilistic metrics, like continuous ranked probability skill score, which measures improvement of probabilistic forecast relative to a reference forecast. Recent studies by Corti et al. (2012) and Hazeleger et al. (2013a) investigated predictive skill of decadal predictions in terms of the Brier skill score, which enable to estimate level of closeness of forecast probabilities to the observed frequencies, as well as level of correspondence of forecast probabilities to climatological probability of the event.

A large datasets of results for near-term climate predictions is now available from the WCRP Coupled Model Intercomparison Project - Phase 5 (CMIP5; IPCC IPCC, 2013), which also addresses the decadal timescales. CMIP5 initialized experiments show that the initialization of decadal prediction systems with the recent state of the climate system provides regional improvement of decadal predictions for temperature and, additionally to the large ocean regions where the externally forced hindcasts (un-initialized) have high predictive skill, the initialized hindcasts show skill over the North Atlantic, Indian Ocean and the southeast Pacific. Branstator and Teng (2010) and Boer et al. (2013) found out that information from initial conditions seems to be more essential in the first years to a decade and becomes undetectable after about a decade with the external forcing taking over subsequently. Though the skill might get invalidated in the case of an explosive volcanic eruption (Doblas-Reyes et al., 2011b), which is currently not possible to predict, but large volcanic eruption can effect oceanic temperature on decadal timescale (Delworth et al., 2005). The predictive skill for CMIP5 initialized predictions of precipitation seems to be very poor (Goddard et al., 2013). Whereas AMOC initialized hindcasts appear to be skillful on decadal timescales (Pohlmann et al., 2012). Numerous studies on predictive skill of initialized hindcasts have demonstrated that state-of-the-art-models are able to capture processes which provide skill on decadal timescales. Observed and modeled climate modes of decadal variability like AMO, NAO, PDO and ENSO (e.g., Kerr, 2000; Knight et al., 2005; Hurrell and Van Loon, 1997; Mantua and Hare, 2002; Power et al., 2006, and others) are widely considered as potential sources of decadal climate predictability in the initialized predictions (Matei et al., 2012; Mochizuki et al., 2010; Doblas-Reyes et al., 2013). Also the AMOC variability shows high predictive skill and is believed to have impact on the sea surface temperature (SST) variability in the North Atlantic. Matei et al. (2012) suggest

that predictive skill for SST and OHC fluctuations at the beginning of the forecasts is due to persistence, while at longer lead times the skill might be a consequence of initializing the AMOC variability. The results from the study by Pohlmann et al. (2009) showed that the AMOC leads the North Atlantic SST variability by about 5 years. However, the study by Hazeleger et al. (2013b), based on the multi-model approach, suggest that coupled climate models appear to have different relationships between the AMOC and temperature (and salinity) in the subpolar North Atlantic. Therefore, the precise physical mechanisms in high predictability regions are yet to be completely understood.

To summarize, significant contributions have been made to estimate potential and actual predictive skill of decadal climate prediction systems. It has been shown that initialization accounts for internal natural variability and provides predictive skill for important climate variables at decadal timescales. So far, primarily the ocean component of the coupled models was initialized, though initialization of other climate system components such as sea ice and land surface may also contribute to predictive skill on decadal timescales. In prediction studies, anomaly initialization or/and full state initialization were mostly applied, but it is still unclear which of the initialization methods is the most appropriate for decadal predictions. In order to skillfully predict climate fluctuations on decadal timescales, further research is required to study the benefits from different initialization techniques. For this, the understanding of underlying physical processes responsible for high predictive skill of coupled climate models is very important.

1.3 Goals of the thesis

This thesis aims to investigate the sensitivity of retrospective decadal predictions performed with the UCLA/MITgcm to three ocean initialization approaches: full state initialization, anomaly initialization and flux correction. Firstly, we are going to evaluate the regions and duration of high predictive skill for differently initialized hindcasts of sea surface temperature (SST), Atlantic meridional overturning circulation (AMOC) and sea surface height (SSH). In order to find out whether initialization provides any useful information, predictive skill for the initialized hindcasts is compared with the skill for twentieth-century simulations (un-initialized hindcasts) and low-skill persistence forecasts. The performance of different initialization strategies is compared to answer the question whether any of them show an advantage over the other strategies and if so, then why? Secondly, we are interested to understand possible underlying mechanisms giving rise to high predictive skill. For this, the decomposition analysis, which allows to distinguish different contributions to SSH changes associated with density anomalies, is used.

1.4 Structure of the thesis

The remaining thesis consists of four Chapters:

Chapter 2 describes the methodology for studying decadal climate predictability. This includes the description of the UCLA/MITgcm coupled model and initial conditions which are sampled from the GECCO ocean state estimate, the experimental set-up, details of implementation of three initialization approaches, procedures of bias correction, de-trending, predictive skill estimation and significance level assessment.

Chapter 3 provides an evaluation of the predictive skill for the initialized SST, SSH and AMOC hindcasts in terms of correlation coefficients and root mean square error skill score. The evaluation of the skill is mainly performed with respect to the GECCO ocean state estimate. The predictive skill for the North Atlantic SST and SSH with possible mechanisms for the North Atlantic SST predictability are also discussed in this Chapter.

Chapter 4 explores processes which contribute to predictive skill for SSH in different initialization experiments. In this Chapter we analyze steric contributions to SSH changes imposed by temperature (thermosteric SSH changes) and salinity (halosteric SSH changes) anomalies, and estimate predictive skill for these terms. Thereafter, the analysis of predictive skill for three contributions derived from both thermosteric and halosteric SSH changes, namely those imposed by temperature or salinity anomalies within and beneath the mixed layer, and those related to processes of density compensated temperature-salinity anomalies (advection on isopycnals) is carried out.

Finally, a conclusion and outlook are given in **Chapter 5**.

Chapter 3 of this thesis has been published in Climate Dynamics: Polkova, I., Köhl, A., Stammer, D. (2013), Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model, *Clim. Dyn.* DOI: 10.1007/s00382-013-1969-4.

Chapter 2

Methodology for Exploring Decadal Predictability

This chapter outlines the details of the research methods used to address the key questions posed in Chapter 1.1. The methodology to explore decadal climate variability and predictability incorporates a range of tools:

- the UCLA/MITgcm coupled ocean-atmosphere model

All the experiments in this study were performed in the coupled climate model framework. The MITgcm ocean model was used for both GECCO ocean state estimate (the source of initial conditions in this study) and as a component of the coupled model for the initialized hindcasts. Model, data and description of tested climate parameters such as sea surface temperature (SST), sea surface height (SSH) and Atlantic meridional overturning circulation (AMOC) are discussed in Section 2.1.

- Experiments and initialization approaches

The experimental setup is described in Section 2.2. Section 2.3 explains an implementation of different initialization strategies for decadal predictions: full state initialization, anomaly initialization and flux correction.

- Methods for diagnostics and verification of decadal predictions

The procedures of bias correction of the initialized hindcasts, de-trending, persistence forecast, assessment of predictive skill through deterministic metrics such as correlation coefficients and root mean squared error skill score including significance test are given in Section 2.4.

- Method for identification of underlying mechanisms

A method that provides some insight into underlying mechanisms giving rise to high predictive skill will be discussed latter in Section 4.2.

2.1 Model and climate parameters

2.1.1 Description of a coupled model

The climate model used in this study is the University of California, Los Angeles coupled general circulation model (UCLA CGCM), which consists of the UCLA atmospheric general circulation model (UCLA AGCM; see www.atmos.ucla.edu/~mechoso/esm/agcm.html) coupled with the Massachusetts Institute of Technology ocean general circulation model (MIT OGCM; see mitgcm.org). The performance of UCLA CGCM is described for ENSO forecasts (Cazes-Boezio et al., 2008). The model was also used recently by Stammer et al. (2011) to simulate the global climate response to Greenland ice sheet melting.

UCLA AGCM

The UCLA AGCM is a state-of-the-art model with advanced parameterization of the major physical processes in the atmosphere. It is a finite difference model that integrates the primitive equations of the atmosphere. The model's horizontal and vertical discretization are arranged as "C" and Lorenz grids, respectively (The model divides the atmosphere into a rectangular grid of layered vertical columns). The parameterization of cumulus convection, including its interaction with the planetary boundary layer (PBL) uses prognostic version by Arakawa and Schubert (1974) and described by Pan and Randall (1998). Surface heat fluxes are calculated following the bulk formula proposed by Deardorff (1972) and modified by Suarez et al. (1983). The model includes parameterization of prognostic cloud liquid water and ice (Köhler, 1999). The parameterization of PBL processes are based on the mixed-layer approach of Suarez et al. (1983), as revised by Li et al. (2002), and upgraded to multi-layer by Konor et al. (2009). The parameterization of radiation processes is performed according to Harshvardhan et al. (1987, 1989), and includes the effects of cumulus, ice and PBL clouds.

For this thesis the UCLA AGCM version 7.1 with a horizontal resolution of 2.5° longitude and 2° latitude, and 29 vertical layers was used. The AGCM is coupled to the first-generation Simplified Simple Biosphere model (SSiB; Xue et al., 1991) that has three soil layers and one vegetation layer. Ma et al. (2010) discussed the upgraded land surface processes that yield a significant improvement of UCLA AGCM performance. The model uses data for vegetation type from Dorman and Sellers (1989); Xue et al. (1996) to specify monthly climatological land surface properties. The ozone mixing ratios are prescribed as two-dimensional (latitude and height) according to monthly climatology from Li and Shine (1995). This version of the model does not include aerosol effects and changes in solar activity. The distributions of greenhouse gases, sea ice, and ocean surface albedo are prescribed as monthly mean observed climatology. A more detailed AGCM description is given at <http://www.atmos.ucla.edu/~mechoso/esm/agcm.html>.

MIT OGCM

Here we use the same version of the MIT OGCM that was employed in the study by Stammer et al. (2011). The model solves the primitive equations (Adcroft et al., 2008) such that, at any time t , the prognostic variables for the flow are the horizontal velocity (u, v), potential temperature (θ) and salinity (S), and the sea surface height (η), which is diagnosed using a linearized implicit scheme (Dukowicz and Smith, 1994).

The MIT OGCM has a quasi-global domain 80°S to 80°N with all lateral boundary conditions closed. The resolution of the model is $1^\circ \times 1^\circ$ with the refinement of the meridional resolution equatorward of 30°S and 30°N , such that at the equator it equals $1/3^\circ$. The meridional and vertical resolution in the tropics provides realistic representation of tropical currents and related thermal structure (Lee et al., 2002). The OGCM uses 46 levels with thickness ranging from 10 m in the top 150 m, and gradually increasing to 400 m near the maximum model depth of 5815 m. The total number of grid points is 360 by 224 by 46, zonally, meridionally, and vertically, respectively.

The model employs the K-Profile Parameterization (KPP) vertical mixing scheme of Large et al. (1994) and the isopycnal mixing schemes of Redi (1982) and of Gent and McWilliams (1990) with surface tapering (Large et al., 1997). Surface freshwater fluxes are applied as real freshwater fluxes. The fields between the AGCM and OGCM are exchanged between each other once in a day. No-slip bottom, free-slip lateral, and free surface boundary conditions are employed. Laplacian diffusion and friction are used, with the exception that horizontal friction is biharmonic. Isopycnal diffusivity and isopycnal thickness diffusivity is $100 \text{ m}^2\text{s}^{-1}$. Vertical diffusivity is $10^{-5} \text{ m}^2\text{s}^{-1}$. Horizontal eddy viscosity coefficient for mixing of momentum laterally is $10^{-4} \text{ m}^2\text{s}^{-1}$, biharmonic viscosity coefficient for mixing of momentum laterally is $10^{13} \text{ m}^4\text{s}^{-1}$.

The sea-ice distribution is prescribed according to an observed monthly climatology in the atmospheric model, in the ocean at the northern and southern boundaries the sea ice mask is used to reset the temperature to the value -1.9°C .

The model's bathymetry is based on ETOPO5 (Data Announcement 88MGG-02, Digital relief of the Surface of the Earth, NOAA, National Geophysical Data Center, Boulder, Colorado, 1988).

2.1.2 Parameters to study decadal predictability

In this Subsection, the choice of the climate variables used to study decadal predictability, and their prognostic or diagnostic evaluation in the coupled model is explained.

Sea surface temperature

Being the "easiest" ocean variable to observe, SST has the longest observation records among the other ocean quantities. SST is crucial for studying the ocean as well as the atmosphere since it controls the exchanges of heat flux between these two components.

We analyze initialized SST hindcasts for comparison reasons, because many decadal predictability studies primarily estimate the skill for either SST or surface air temperature (Pohlmann et al., 2009; Mochizuki et al., 2010; Matei et al., 2012; Magnusson et al., 2012a,b). These studies suggest improved SST skill due to initialization, e.g. in the North Atlantic, North Pacific and Indian Ocean. Another reason for the analysis of SST in this study is the use of heat flux correction, which aims to improve model SST that is, through physical connections, expected to lead to better representation of other variables.

The time evolution of temperature is evaluated prognostically in the MIT OGCM (Adcroft et al., 2008). We then analyze monthly and annual mean SST fields from the model.

Atlantic meridional overturning circulation

The Atlantic meridional overturning circulation (AMOC) variability is considered to contribute to climate variations in the Atlantic Ocean and over the northwest Europe on decadal timescales (Hurrell et al., 2006). We also analyze the AMOC for this thesis because to date very little is known how different initialization approaches impact AMOC evolution. Some modeling studies discovered weakening of AMOC in response to increased external forcing (Gregory et al., 2005), however observation records are too short to verify this. By contrast, Menary et al. (2013) pointed out a diversity in models' responses to historical forcing and that more realistic representation of the anthropogenic aerosol forcing (both direct and indirect effects) might result in a weaker downward trend of AMOC or even a strengthening of AMOC. A recent study by Pohlmann et al. (2012) showed that different ocean syntheses exhibit similar behavior of AMOC at 45°N: an upward trend from 1960s to 1990s, followed by the downward trend. The AMOC fluctuations were highly correlated with observed variations of North Atlantic Oscillation, the strength of the North Atlantic subpolar gyre, Atlantic SST dipole and Labrador Sea convection, and provided predictability up to 6 years due to initialization. Whereas the twentieth-century integrations (un-initialized experiments) had a tendency to show downward trend all over the analyzed period.

The AMOC property is computed as the mean meridional mass transport streamfunction of latitude and depth integrated zonally in the Atlantic basin:

$$\Psi(y, z) = \int_{-H}^z \int_{east}^{west} v(x, y, z) dx dz \quad (2.1)$$

where v is the monthly mean meridional velocity field, $v = v(x, y, z, t)$, H denotes a depth profile along zonal section.

Sea surface height

The SSH is a useful property to analyze because it can act as a proxy for other climate properties, like ocean temperature, OHC and ocean circulation. Leuliette and Wahr (1999)

reported high correlation between SST and SSH records due to heat storage near the surface of the ocean. SSH changes are directly related to the temperature changes in a volume of seawater, (when warm/cold water expands/contracts, the sea level rises/falls), and the averaged temperature changes are in turn proportional to OHC. The ocean stores large amounts of heat that is distributed around the world by ocean currents, and the non-uniform redistribution of heat (and salt) horizontally and vertically through air-sea exchange and the ocean circulation affects the spatial variability of SSH (Bindoff et al., 2007).

The state-of-the-art models contain mechanisms for simulating internal climate variability, which additionally to forced response to variations in the Earth’s energy budget affect SSH changes (Delworth and Knutson, 2000; Sturges and Hong, 2001; Zhang and Church, 2012). Using ”perfect model approach”, Griffies and Bryan (1997) showed that potential predictive skill for SSH might be higher than that for SST, in particular, when the multi-decadal variations of the thermohaline circulation were active, the first and second EOF patterns for the North Atlantic SSH showed potential predictability time scales on the order of 10–20 years, whereas the first EOF of SST had potential predictability time scales of 5–7 years. Motivated by these results, we analyze the initialized SSH hindcasts, in particular to investigate whether they capture the processes and simulate correctly variability associated with SSH predictability on decadal timescales.

Ocean general circulation models usually simulate dynamic topography which is the SSH taken with respect to a motionless state and is therefore an expression of the ocean’s large-scale circulation. When referring to the dynamic topography in this thesis, we will use the term SSH. For comparison of model SSH with observations, one should keep in mind the other contributions to SSH changes (Fig. 2.1), so that the verification dataset reflects the dynamic topography only.

The SSH fields from the MIT OGCM are diagnosed using an implicit free-surface scheme, which is described by Dukowicz and Smith (1994). The model also uses real freshwater flux formulation and the Boussinesq approximation and does not include sea ice model.

The evolution equation of sea level in non-Boussinesq models is obtained by integration of the mass continuity equation over the depth of the ocean:

$$\frac{\partial \eta}{\partial t} = (P - E + R) - \nabla \cdot \mathbf{U} - \int_{-H}^{\eta} \frac{1}{\rho} \frac{d\rho}{dt} dz, \quad (2.2)$$

where η is the SSH, $P - E + R$ is the freshwater flux (precipitation - evaporation + river runoff), $\nabla \cdot \mathbf{U}$ denotes the convergence of vertically integrated horizontal currents and ρ is density of seawater. Thus, the first term in the equation on the right hand side denotes boundary mass flux across the ocean surface and affects SSH through changes in boundary forcing, the second and third terms are the sea level tendency and steric effect. Therefore, Eq. 2.2 suggests that SSH changes result from changes in dynamic

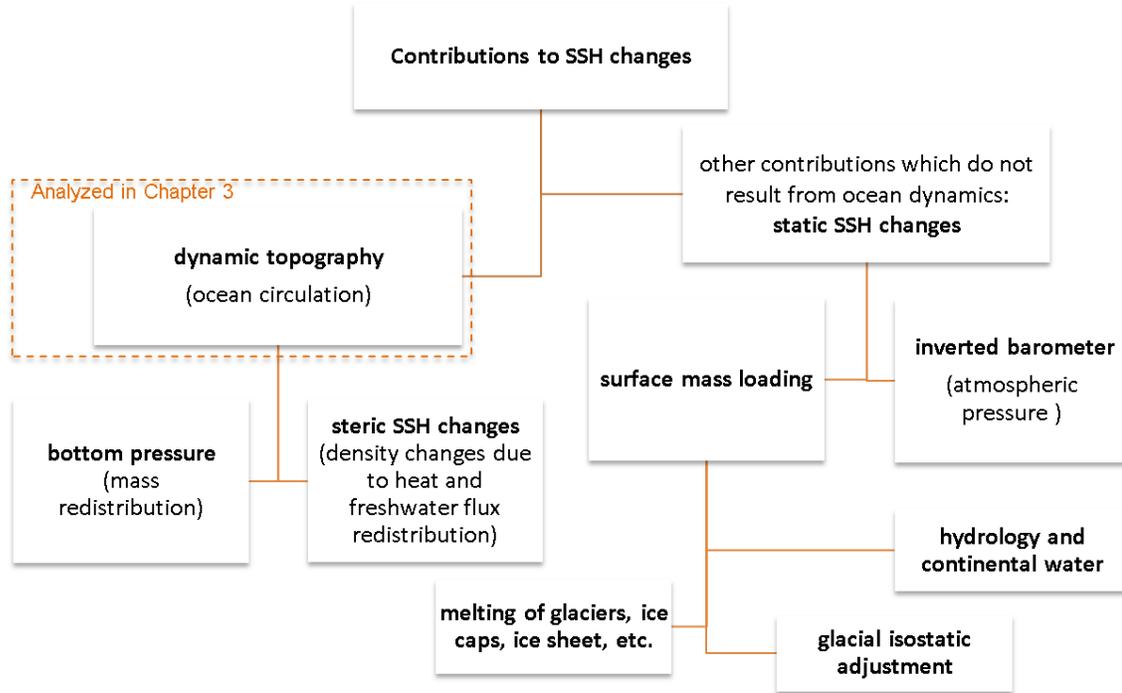


Figure 2.1— Possible contributions (reviewed by Church et al., 2010) to SSH changes and contributions (dynamic topography) simulated by the coupled model.

fields within the ocean by horizontal motion and upward expansion/contraction of water column in response to local density changes.

In the Boussinesq fluid, the mass conservation ($dM = \rho dV = \text{const}$) is replaced by volume conservation ($dV = dM/\rho = \text{const}$) which implies that the model does not allow total volume changes and consequently global mean SSH changes. An example with a local heat flux helps to understand this difference and how it affects regional SSH changes. If the ocean is heated at the surface, in non-Boussinesq model (and in reality) SSH changes result from the density changes (third term of the Eq. 2.2); the sea level will rise as a result of an increased ocean volume (and decreased density). In the Boussinesq model, in order to maintain constant volume, the density and mass must decrease. Loss of mass implies negative bottom pressure anomaly that leads to a large-scale adjustment (geostrophic adjustment). As a result, the volume will be transported to the regions where mass changes occur. In this case, this is the second term in Eq. 2.2 which pushes the sea level upward. In order to account for the non-Boussinesq effect in global mean sea level change, η can be corrected to allow ocean mass to be conserved (for more details see Greatbatch, 1994; Ponte, 1999).

2.2 Experiments

This study includes the following experiments (see also Tab. 2.1):

- 300-year spin-up run (climatology from the last 100 years of spin-up run is used in the anomaly initialized experiments),
- an ensemble of twentieth-century integrations (20C) used as a reference run for comparison with the initialized hindcasts,
- 3 assimilation runs (one for each initialization strategy: full state initialization (FSI-ASSIM), anomaly initialization (AI-ASSIM) and flux correction (FC-ASSIM)) used as a source of atmospheric initial condition for the initialized hindcasts and
- 3 sets of initialized hindcasts (one for each initialization strategy: FSI-HIND, AI-HIND and FC-HIND).

Table 2.1— Summary of experiments.

Experiments	Initialization and forecast period	Initial conditions	Forcing*	Realizations
20C run	1900-2011, 1910-2011, 1920-2011	From spin-up run	GHG	3
FSI-ASSIM	monthly, 1952-2001	GECCO full state	GHG	1
FSI-HIND	at the end of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000	GECCO full state	GHG	3
AI-ASSIM	monthly, 1952-2001	GECCO anomalies +	GHG	1
AI-HIND	at the end of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000	+ model climatology	GHG	3
FC-ASSIM	monthly, 1952-2001	GECCO full state +	GHG	1
FC-HIND	at the end of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000	+ flux correction	GHG	3

*As greenhouse gas (GHG) forcing, only CO₂ concentrations are applied to all the model runs.

All coupled model simulations were carried out on a high performance computing systems Tornado and Blizzard operated by the German Climate Computing Centre (Deutsches Klimarechenzentrum GmbH, DKRZ).

2.2.1 Spin-up run

We used a climatology of 300-year spin-up run with constant greenhouse gas (GHG) concentration from the year 1980 for constructing initial conditions in the anomaly initialized experiments (see AI-ASSIM and AI-HIND in Subsec. 2.3.3). In the current experiment,

the coupled model was initialized from a state of rest with Levitus temperature and salinity climatology (Levitus and Boyer, 1994; Levitus et al., 1994). All three assimilation runs were started subsequently with the atmospheric state from the end of the spin-up run.

2.2.2 Twentieth-Century Simulations

An ensemble integration of the 20C simulations was performed starting from the same initial condition in 1900, 1910, 1920 until 2011. The model was forced by annually and globally averaged concentrations of CO₂ (Fig. 2.2; Robertson et al. (2001); Tans and Keeling (NOAA/ESRL, Scripps Institution of Oceanography, 2011).)

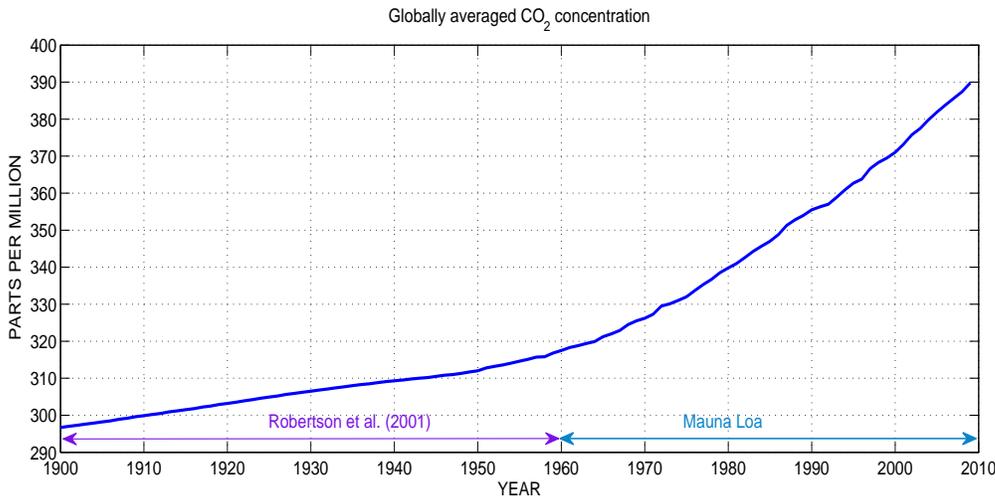


Figure 2.2— Time series of the globally averaged annual mean CO₂ concentration (ppm).

The time series of globally averaged annual mean SST anomaly are shown in Fig. 2.3 a. There are some discrepancies between the behavior of the 20C ensemble mean and the HadISST observations¹. The interdecadal variability and trend of the GECCO SST anomalies are also somewhat different from HadISST until the mid-1970s, with better agreement starting from the mid-1980s. Over the period 1990-2000, the 20C SST variability is more consistent with GECCO and HadISST, though the 20C SSTs are cooler by about 0.2°C. The correlation coefficients between the annual mean SST anomalies from 20C and GECCO (HadISST) are shown in Fig. 2.3b. When pentadal and decadal means are analyzed, correlation between 20C and GECCO (HadISST) goes up, suggesting good agreement in long-term behavior rather than interannual variability.

¹HadISST is the sea surface temperature dataset developed at the Met Office Hadley Centre for Climate Prediction and Research. The data set consists the monthly mean global fields of SST on a 1° latitude-longitude grid from 1871 to 2011 (Rayner et al., 2003).

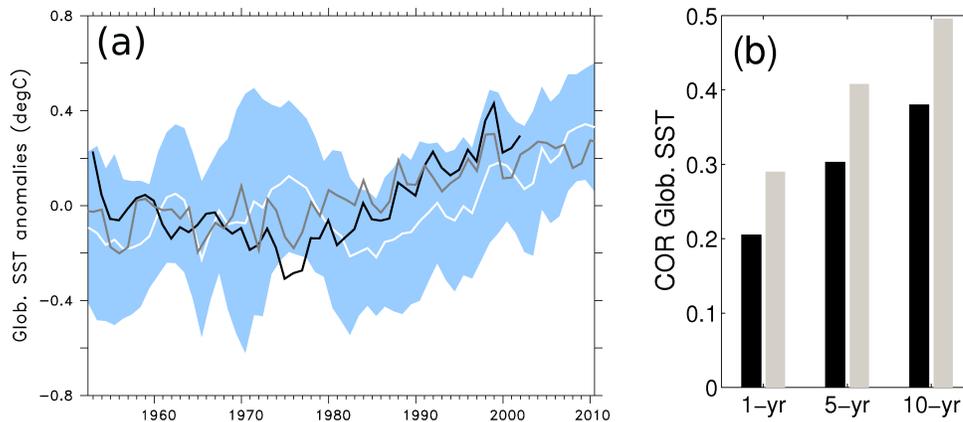


Figure 2.3— (A) Time series of the globally averaged annual mean SST anomaly from 20C run (white), HadISST (gray) and GECCO (black). Ensemble spread in the 20C ensemble is shown in light blue. (B) Correlation coefficients computed for annual, pentadal and decadal mean SST.

Model bias

Model bias is an important issue for climate predictions because an error in the mean state can affect the variability of the coupled model. The typical SST bias of the UCLA CGCM computed with respect to HadISST is shown in Fig. 2.4. Overall the model is warmer than observation, except for the tropical Pacific where it has a prominent cold bias that extends too far to the west. Strong warm biases are observed at the western boundaries of the North Pacific and the North Atlantic with the bias extended to the Labrador Sea, at the western coast of Africa and along the Antarctic Circumpolar current. The SST cold bias in the South Pacific at about 15°S amounts up to 5 °C and the warm biases in some regions in the northern North Atlantic and the northern North Pacific come up to 8°C.

issue for climate predictions because an error in the mean state can affect the variability of the coupled model.

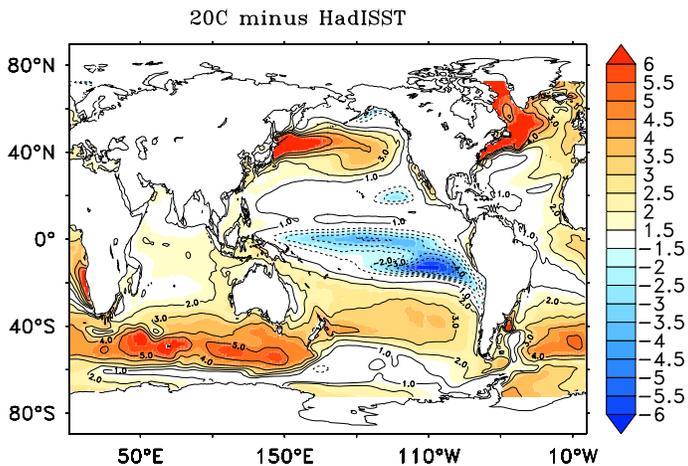


Figure 2.4— Typical SST bias (°C) from the UCLA/MIT CGCM computed as the difference between the 20C run and HadISST over the period 1952-2001.

The SST bias can be caused by the deficiencies to represent physical processes both in the atmospheric and ocean models and can get amplified through coupled feedbacks. Fig. 2.5 shows the freshwater flux, heat flux and wind speed from the atmospheric model

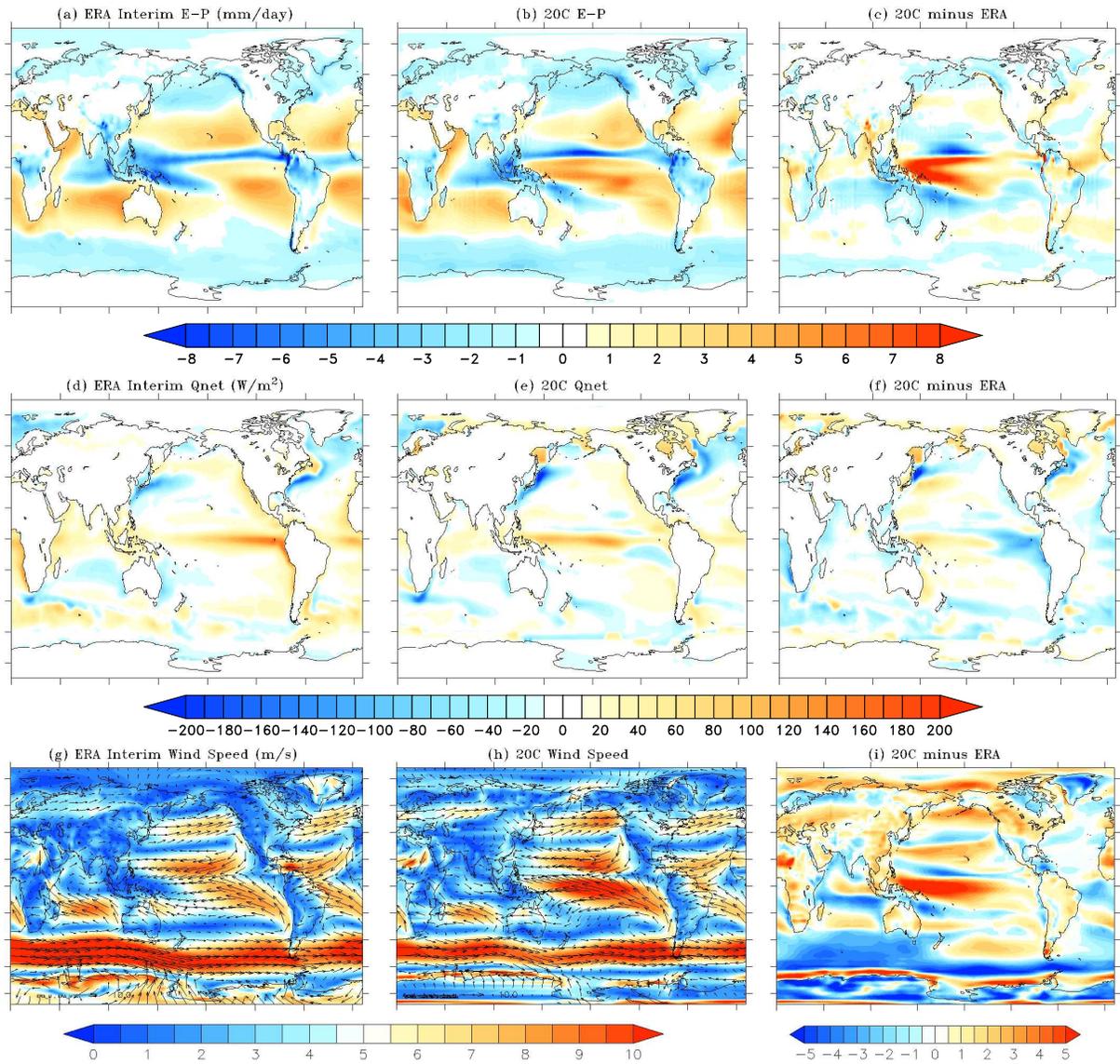


Figure 2.5— Comparison of atmospheric net freshwater flux and heat flux, and annual wind speed patterns from ERA Interim (left panel) and 20C run (middle panel). Right panel shows a difference between 20C and ERA Interim for the time mean over 1989 to 2011.

compared the fluxes from the ERA-Interim reanalysis². There is a strong bias in freshwater flux in the western tropical Pacific (Fig. 2.5 c), which signifies that the UCLA CGCM suffers from a common for current coupled models bias, like double intertropical convergence zone (ITCZ). The ocean model has a problem of wrong position of the west-

²ERA-Interim global atmospheric reanalysis is produced by the European Centre for Medium Range Weather Forecasts (Simmons et al., 2007). The data set has $0.75^\circ \times 0.75^\circ$ spatial resolution from 1989 to 2012.

ern boundary currents, which might cause a dipole heat flux bias along western boundary currents (Fig. 2.5 f). In the Southern Ocean the model simulates weaker, in comparison with ERA-Interim, winds which may result in zonally extended warm SST bias (Fig. 2.5 h). Also the mixing processes in the ocean are rather reduced there (Fig. 2.6) and the heat from the atmosphere might not penetrate deep enough into the ocean leading to the sub-surface temperature increase.

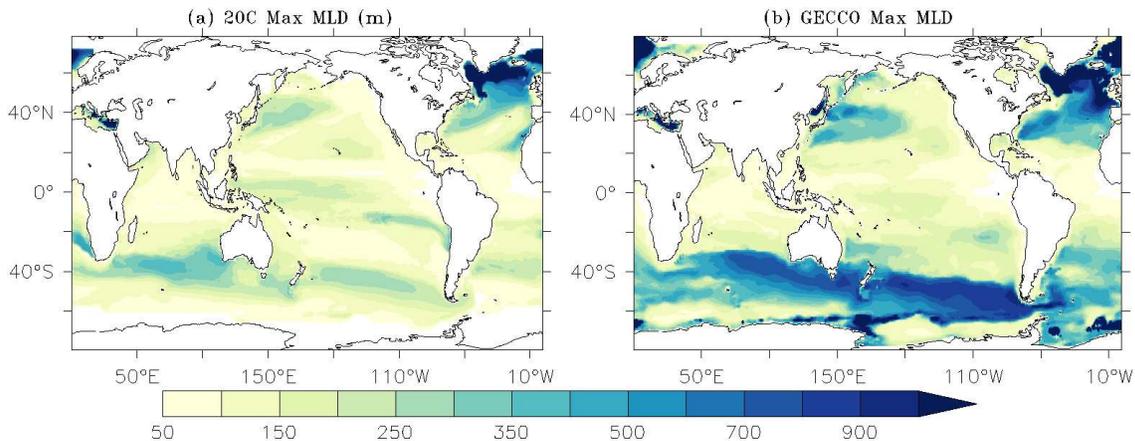


Figure 2.6— Maximum mixed layer depth in meters derived from monthly mean fields over 1952-2001 from (a) the 20C run and (b) the GECCO ocean state estimate (which is discussed in Subsection 2.3.1).

Warm SST bias at the eastern South Atlantic is believed to be related to the lack of low clouds (Barthelet et al., 1998). This type of bias is usually inherent in both eastern South Pacific and South Atlantic oceans. In the current CGCM, this feature is obvious in the eastern South Atlantic, while in the eastern South Pacific the colder underlying SSTs with the minimum around 110°W can be associated with less penetration of shortwave radiation due to increased cloudiness of the spurious southern ITCZ (Mechoso et al., 1995; Yu and Mechoso, 1999), in this region the negative heat flux bias is apparent.

The secondary minimum of the elongated SST bias in the equatorial Pacific is observed at 150°W. Coupled models usually tend to underestimate zonal SST gradient along the equatorial Pacific to the west and overestimate to the east with the upwelling shifted westward from the observed location. Mechoso et al. (1995) attributed cold equatorial SST bias to the deficiencies of the models to simulate winds, mixed layer and upwelling correctly. According to Figs. 2.5 and 2.6, the UCLA CGCM tends to simulate too strong easterlies and has deeper mixed layer in the western tropical Pacific, compared to the mixed layer from the GECCO ocean synthesis. This may suggest that cool SST anomalies in the eastern equatorial Pacific lead to intensification of zonal trade winds to the west and further cooling and westward spreading of SST through increased wind stirring and latent heat loss.

2.2.3 Assimilation runs and initialized hindcasts

We performed three 50-year long assimilation runs, which cover the period 1952-2001 for three assimilation runs employing initialization methods discussed in Subsec. 2.3.2 – 2.3.4 (FSI-ASSIM, AI-ASSIM and FC-ASSIM). The assimilation runs are performed by re-initializing the ocean component of the coupled model at intervals of one month with the GECCO ocean state estimate discussed in Subsec. 2.3.1. Re-initialization of the long-term continuous integration is a simple data assimilation method and implies subdividing long integrations into short-term (in our case monthly) periodic integrations. The method was employed in weather forecasting to eliminate systematic error growth in long integrations, as a downscaling approach in regional climate models (Lo et al., 2008) and also as an adjustment of an initial condition to obtain a simulation in agreement with the remotely-sensed derived observations (Moulin et al., 1998). Re-initialization is to some extent an equivalent of nudging approach, they both aim to keep the model close to the observed climate. The difference is that nudging is based on adding a difference between modeled and observed state to a prognostic equation and applied at each time step of integration, while re-initialization is applied once per re-initialization interval.

The resulting assimilation runs provide atmospheric initial conditions for the following initialized decadal hindcasts. The ensembles of initialized hindcasts are started from the 31st of December 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000 with 3 realizations for each starting date (following the CMIP5 protocol in which a minimum of 3 ensemble members is required for the initialized experiments; Taylor et al., 2012). Hence, we obtained 27 hindcasts over the period 1961-2010 for each initialization scheme. The ensemble members were generated using lag initialization approach by shifting the atmospheric state by two days relative to the assimilation run (these are the atmospheric states from December 27, 29 and 31). The initial conditions for the ensemble members differ only in their atmospheric state, ocean initial conditions are unchanged.

2.3 Initialization of decadal climate predictions

Internal variability is the natural variability of the climate system that occurs in the absence of external forcing and is addressed by initialization procedures. Decadal climate predictions, especially designed to study internal variability, are used to make predictions for the next few years starting from the recent state of the climate system. To date, different approaches exist to initialize decadal predictions, which differ in the way of implementation and dealing with model drift. In the following sections we introduce the initial conditions and three initialization methods used across the study, namely full state initialization, anomaly initialization and full state initialization employing flux correction. First two are frequently applied in the decadal prediction studies, while the benefits of using the latter one for the initialized predictions is less understood.

2.3.1 Initial conditions (GECCO)

In this study the GECCO ocean state estimate was used as the ocean initial conditions (3-D temperature, salinity, zonal and meridional velocities, and SSH fields) for decadal climate predictions. The ocean synthesis fields are available from the German contribution to Estimating the Circulation and Climate of the Ocean project and represent the 50-yr period over 1952-2001 (Köhl and Stammer, 2007, 2008). GECCO has a quasi-global domain, 1° horizontal resolution and 23 vertical layers. This ocean state estimate is a modification of the 11-yr ECCO ocean synthesis (Köhl et al., 2007) and uses the adjoint method to bring the MIT OGCM model in agreement with observations, initial temperature and salinity conditions and time-dependent surface fluxes of momentum, heat and freshwater. GECCO combines most of the World Ocean Circulation Experiment observations available for the synthesis period. Input data are described more detailed in the ECCO Report (Köhl et al., 2006).

It is important for this study to consider the quality of the GECCO synthesis as a source of initial conditions and its possible impact on decadal prediction quality. Köhl and Stammer (2007, 2008) analyzed GECCO with respect to long-time change of the sea level and AMOC. They found a general increase in the AMOC at 25°N over the period 1962-2001. The correlation analysis revealed the four different processes that contribute to the GECCO AMOC variability, two of them do not involve time lags, namely the Ekman drift and coastal upwelling, and the other two are Rossby and Kelvin wave-based processes. In terms of the SSH variability, the results are in a good agreement with the AVISO³ observations. The study lacked salinity data in the Southern Ocean. The "estimation of the global mean freshwater flux from sparse salinity data was not well conditioned" in the data assimilation approach and provides an unrealistically large bias (Köhl and Stammer, 2008). Pohlmann et al. (2009) used GECCO synthesis for initializing decadal hindcasts. They found that GECCO SST is much higher in the Southern Hemisphere in the 1950s and too low (high) in the Pacific in the 1970s (1990s) when comparing with HadISST. Though, GECCO is constructed with the ocean circulation model dynamically consistent with most of the observations, it is still limited by the formulation of the numerical model.

Comparisons of different ocean state estimates including GECCO were performed by Carton and Santorelli (2008), Stammer et al. (2009), Lee et al. (2010) and Munoz et al. (2011). Carton and Santorelli (2008) documented differences between analyzed ocean products and historical observations in terms of upper ocean heat content (OHC) decadal variability. For the global average OHC, GECCO demonstrated cooling until mid-1970s, a multi-decadal period of warming thereafter, with some disagreement to other ocean state estimates in the general pattern of OHC in the 1970s. Stammer et al. (2009) claimed that this ocean state estimate was the only dynamically self-consistent solution (obtained

³AVISO SSH data set is a global gridded product at latitude-longitude resolution 1/3° from 1992 available from the Ssalto/Duacs project and distributed by Archiving, Validation and Interpretation of Satellite Oceanographic Data (AVISO) project, with support from Cnes (<http://www.aviso.oceanobs.com/duacs/>)

using four-dimensional variational data assimilation) among the compared products and identified this period as the one with enhanced observation data errors. Lee et al. (2010) suggested that the coarse resolution (0.5° - 2°) is another important limiting factor of consistency among the analyzed ocean products. Munoz et al. (2011) and Kröger et al. (2012) analyzed the AMOC variability of the ocean products. Munoz et al. (2011) observed large disagreement among actual ocean state estimates with respect to the interannual variability of the Atlantic meridional heat transport around latitudinal bands centered at 30° S and 30° N, and to the long-term mean of the maximum AMOC. Kröger et al. (2012) analyzed three ocean synthesis products as potential initial conditions for decadal predictions. The results demonstrated that GECCO AMOC variability obtained after assimilation of the estimate into the ECHAM5-MPIOM does not resemble variability of the original ocean state estimate. Thus, we expect that using in this study the same ocean model (although, the grid is still different) for initialized experiments as was used for estimating the initial conditions (from GECCO) might resolve this issue.

2.3.2 Full state initialization

In the FSI experiment the MIT OGCM is initialized with the full field monthly mean GECCO state (full state implies both mean seasonal cycle and departures from it):

$$IC_{FSI} = \langle temp(x, y, z), salt(x, y, z), u(x, y, z), v(x, y, z), \eta(x, y) \rangle, \quad (2.3)$$

where IC_{FSI} represents the ocean initial state for temperature, salinity, zonal and meridional velocities, and sea surface height at grid points x, y, z in the zonal, meridional and vertical directions, respectively.

The grid (meridional and vertical) and topography of GECCO and the coupled model differ, therefore linear interpolation was applied to the GECCO fields.

Pros and cons

- + Uses initial conditions that represent the actual observed climate state, so that the initialization approach aims to start the model close to "real-world attractor" (Troccoli and Palmer, 2007; Magnusson et al., 2012a,b; Smith et al., 2013).
- At the beginning of the forecast can have initial shock, which is a rapid adjustment due to imbalance between initial conditions and model dynamics (Balmaseda and Anderson, 2009). This may lead to errors in the forecast, which are greater than the errors at longer lead time.
- Forecasts drift away from their initial state towards the model's climate due to development of systematic errors and loss of predictability (Pierce et al., 2004). Development of biases is attributed to the underrepresentation of some physical processes both in ocean and atmospheric components.

- Relatively expensive, as a posteriori bias correction requires large enough set of hindcasts to better estimate the mean drift.
- Bias correction procedure might be inappropriate if the bias is non-stationary (not the same for hindcasts as well as for the real-time forecasts; Kumar et al., 2012).

2.3.3 Anomaly initialization

In the AI experiment, MIT OGCM is initialized with the GECCO anomalies:

$$IC_{AI} = \langle temp^*(x, y, z), salt^*(x, y, z), u^*(x, y, z), v^*(x, y, z), \eta^*(x, y) \rangle, \quad (2.4)$$

where IC_{AI} represents the ocean initial state and indices * imply that the initial field is obtained by adding monthly GECCO anomalies to model’s climatology which is computed from the last hundred years of a long spin-up run with constant CO₂ concentrations from the year 1980. The GECCO anomalies are calculated with respect to the 50-year mean.

Pros and cons

- + Avoids the model drift (Schneider et al., 1999).
- + By avoiding model drift, the lead-time dependent bias correction is not required, instead a long integration to estimate the model climatology is needed for constructing initial conditions and estimating the bias. This makes the procedure of bias correction less expensive. (Of note, the sampling period used for the observed climatology should be consistent with that used for the model climatology.)
- Does not guarantee reduction of initialization shock because the structure of the observed anomalies may not be consistent with the model mean state (Magnusson et al., 2012a; Goddard et al., 2013).
- The disadvantage of anomaly initialization approach is that large errors are allowed in the climate mean state and this can affect internal variability (Schneider et al., 1999). Moreover, findings by Robson (2010) suggest that errors in the assimilated density anomalies, which arise from nonlinearities in the equation of state, may affect climate in regions sensitive to density changes, like the subpolar gyre in the North Atlantic.
- Potential drifts (non-stationary biases) due to an incorrect model response to anthropogenic or natural external forcing are possible (Goddard et al., 2013).

2.3.4 Full state initialization employing flux correction

Initialization approach based on empirical corrections of heat, freshwater or momentum fluxes is another application to deal with model drift, which represents a workaround strategy until model systematic errors are solved. In the FC experiments here, the coupled

model was initialized from the GECCO fields (Eq. 2.3). Additionally, a monthly climatological correction of the surface heat and freshwater fluxes during model integration was applied. This was performed in three steps:

- Step 1. First, a 10-year run was carried out, in which SST and sea surface salinity (SSS) of the coupled model were relaxed toward the monthly climatology of GECCO with a timescale of 30 days. For this, the surface forcing terms from the potential temperature and salinity equations were modified by adding externally supplied fields to the fluxes that come from the atmospheric model (Adcroft et al., 2008):

$$\begin{aligned} F_{SST} &= -\lambda(SST - SST_{clim}^G) + \frac{1}{C_p \rho_0 \Delta z_s} Q, \\ F_{SSS} &= -\lambda(SSS - SSS_{clim}^G) + \frac{S_0}{\Delta z_s} (E - P - R). \end{aligned} \quad (2.5)$$

where F_{SST} and F_{SSS} are surface forcing fields, SST_{clim}^G , SSS_{clim}^G are GECCO climatological relaxation fields used in conjunction with the nudging coefficient $\lambda = \frac{1}{T}$ that controls the magnitude of the flux correction (the smaller relaxation time T , the stronger correction, here $T = 30$ days), $\Delta z_s = 10$ m is the top ocean layer thickness, $\rho_0 = 1000$ kg/m³ is the reference density, $S_0 = 35$ ppt is the reference salinity, and $C_p = 3994$ J/kg°C is the specific heat capacity. Net surface heat flux Q including shortwave radiation in W/m², and freshwater flux $E - P$ in kg/m²/s are provided by the atmospheric model. R is river runoff.

- Step 2. The flux correction terms were constructed from the climatology of SST and SSS obtained from a 10-year run:

$$\begin{aligned} \Delta SST &= SST_{clim}^{10yr-run} - SST_{clim}^G, \\ \Delta SSS &= SSS_{clim}^{10yr-run} - SSS_{clim}^G. \end{aligned} \quad (2.6)$$

- Step 3. Eventually, the assimilation run and initialized hindcasts (FC-ASSIM and FC-HIND) were carried out employing the corrected terms that were added to the fields of heat flux and freshwater flux as in Eq. 2.5:

$$\begin{aligned} F_{SST} &= -\lambda \Delta SST + \dots \\ F_{SSS} &= -\lambda \Delta SSS + \dots \end{aligned} \quad (2.7)$$

Climatological correction for February and December in terms of heat and freshwater fluxes is shown in Fig. 2.7. In addition to this, annual mean flux correction terms are described in Chapter 3. In contrast to the freshwater flux adjustments, the heat flux terms have vigorous seasonal variability and the corrections are strongest in the summer hemisphere.

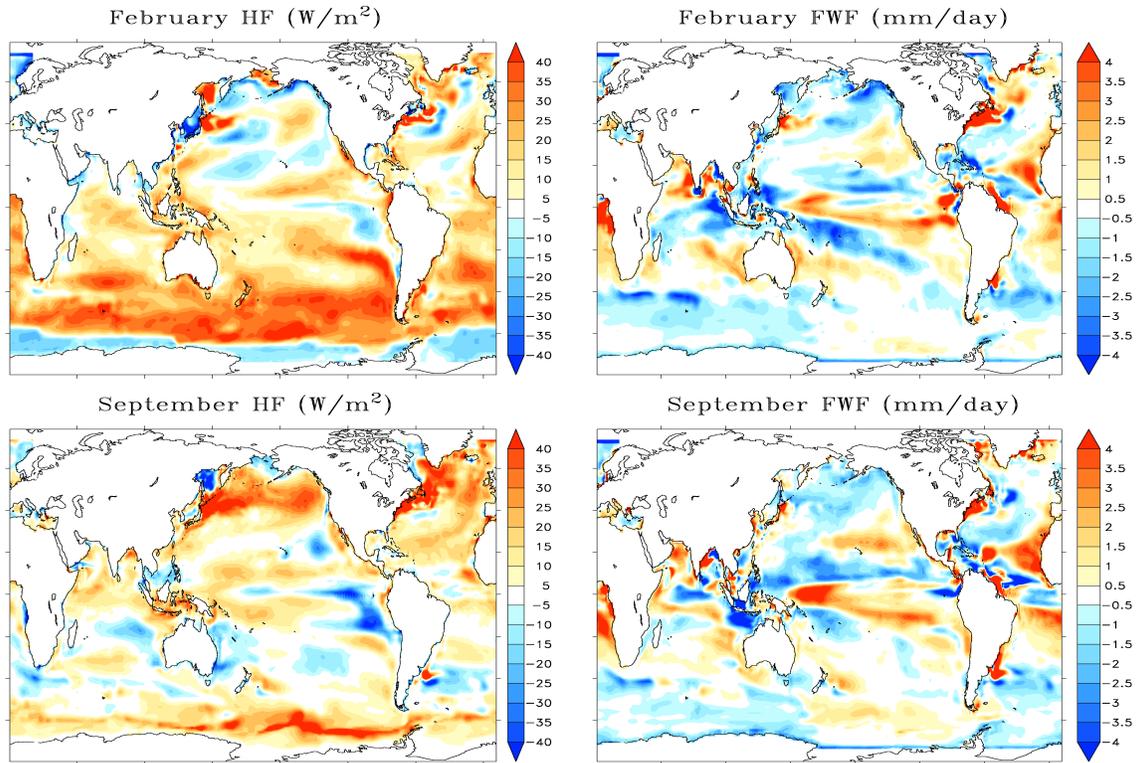


Figure 2.7— AI-ASSIM is by 1.3°C warmer than GECCO. Examples of flux correction terms for February (upper panel) and September (middle panel).

Pros and cons

- + The same as FSI, uses initial conditions which represent the actual observed climate state, so that the initialization approach aims to start and keep the model close to "real-world attractor" (Rosati et al., 1997; Magnusson et al., 2012a,b), therefore reducing the model drift.
- + By improving the mean state and the seasonal cycle of SST, might improve the amplitude of SST interannual variability, and through physical connections it might lead to better representation of other variables, e.g., wind stress, precipitation, etc. (Manganello and Huang, 2009; Magnusson et al., 2012a,b).
- Difficulties to construct correction terms that are relevant for future climate and the future evolution of the atmospheric state might be affected by the artificial flux terms. Some disadvantages for studying coupled climate feedbacks when applying flux correction are discussed by Neelin and Dijkstra (1995). Marotzke and Stone (1995) pointed out that the transient behavior of a coupled model that employs flux correction might be erroneous.

- Flux correction schemes are model-dependent. Spencer et al. (2007) suggest that heat flux corrections can cause unexpected effects on both the mean state and variability of coupled models and could be used with caution for seasonal forecasting. This contradicts the mentioned above advantage of improved mean and variability.
- As FSI and AI, might have potential drifts (non-stationary biases) due to an incorrect model response to anthropogenic or natural external forcing (Magnusson et al., 2012a).

The difference between initialized hindcasts for the globally averaged SST is illustrated in Fig. 2.8. FSI-HIND show strong drift in first years of hindcasts. To remove unwanted drift in the predicted anomalies, we perform a lead-time dependent bias correction, which is explained latter in Subsec. 2.4.2. Unlike in FSI-HIND, the AI-HIND hindcasts do not drift and evolve close to the anomaly initialized assimilation run. An offset towards warm SSTs is discussed in Subsec. 2.2.2. FC-HINDs do not drift as strongly as FSI-HIND, although there is still quite a large offset from the assimilation run, which is not the case in the AI-HIND. The drift in FC-HIND might result from the relatively weak relaxation time. Therefore, FC-HINDs are also subject to applying bias correction.

In terms of assimilation runs, Fig. 2.9 shows the amount of time-mean SST difference between corresponding assimilation run and original GECCO synthesis. The SST bias in FC-ASSIM is significantly improved in many regions except for the western boundary currents and the Southern Ocean, and a cold bias in the equatorial Pacific. The FSI-ASSIM and FC-ASSIM bias for the global SST amounts to 0.52 and 0.38 °C, respectively.

2.4 Predictive Skill Estimation

Preparing the hindcasts

Predicted anomalies are estimated as departures of the initialized hindcasts from the average condition (1952-2001) of the corresponding assimilation run. The same time period was used for computing GECCO and 20C anomalies. For the verification metrics we use the annual mean fields of anomalies.

2.4.1 Trend removal

Decadal climate predictions are initialized with the observed state of the climate system which also includes information about external forcing. Current climate represents internal variability imprinted on top of the long-term global warming trend. Previous studies show that a great fraction of skill for the global SST comes from the external forcing. While for regional climate parameters, like the North Atlantic SST, it is believed that the impact from radiative forcing will be small at least in the first few years of decadal predictions, which are likely to be dominated by internal natural variability of the climate

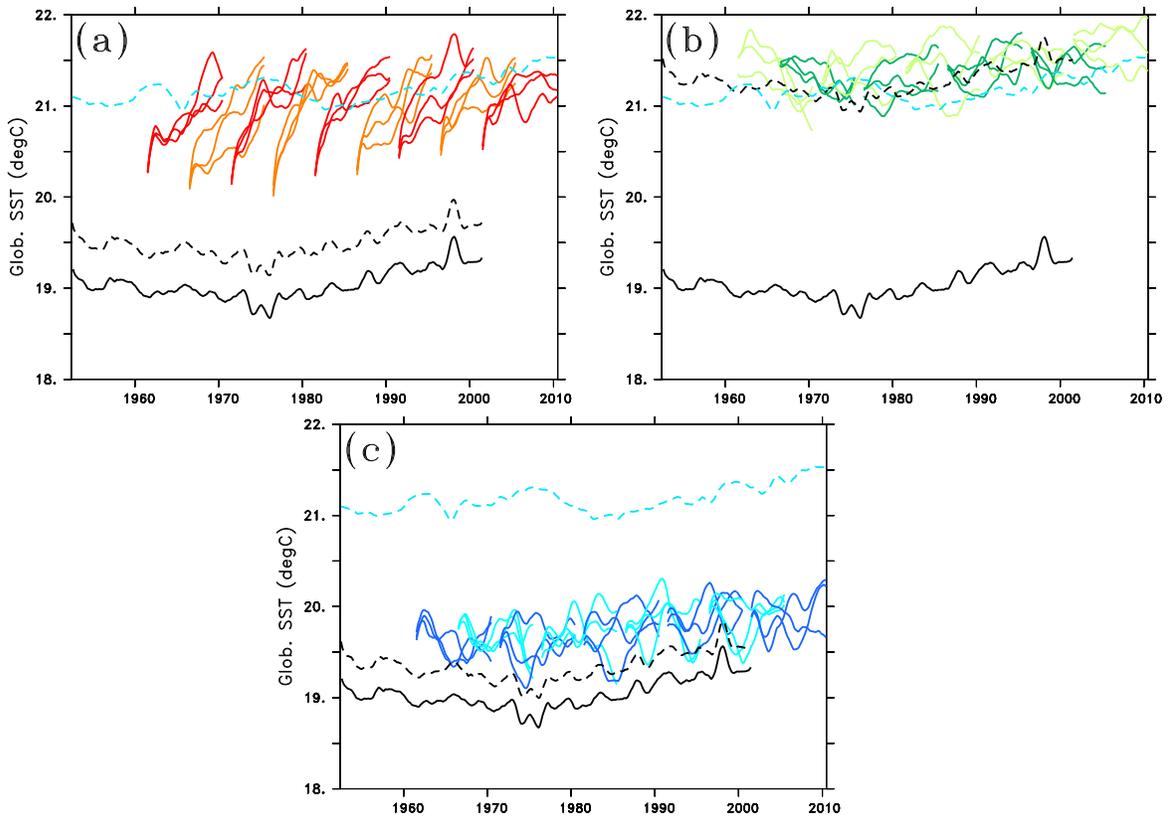


Figure 2.8— Time series of the global mean SST anomalies ($^{\circ}\text{C}$) for GECCO (black solid), assimilation runs (black dashed) and initialized hindcasts: FSI-HIND (red and orange), AI-HIND (green and light green) and FC-HIND (blue and light blue). 12-month running mean is applied. The offset in FSI-HIND and FC-HIND from GECCO shows an amount of drift in the first year.

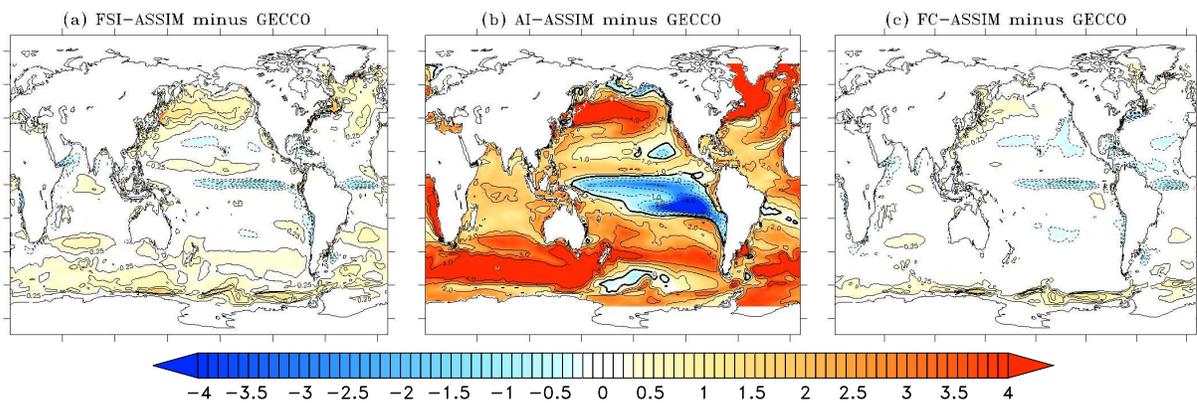


Figure 2.9— (Upper panel) Time mean (1952-2001) SST bias ($^{\circ}\text{C}$) computed with respect to GECCO from (a) FSI-ASSIM, (b) AI-ASSIM and (c) FC-ASSIM.

system (Branstator and Teng, 2010). Thus, in order to investigate how much improvement three different initialization procedures can provide, predictive skill of the model will be analyzed with respect to the detrended fields. We do not remove the hindcasts trend from 10-year simulations because this period might be too short to separate trends from decadal variability. Therefore, the long term trend of GECCO is subtracted from the predicted SST, SSH and AMOC fields.

Because trend in AMOC hindcasts appears to be opposite to that due to external forcing, we assume the former one might be related to the long-term variability. Thus we also analyze the non-detrended AMOC hindcasts. The linear trends for SST, SSH and AMOC are calculated from GECCO fields over the initialization period 1961-2001, using the least squares approach by fitting linear regression model (von Storch and Navarra, 1999) as follows:

$$\alpha_0 = \frac{1}{N} \sum_{t \in N} o_t - \alpha_1 \left(\frac{1}{N} \sum_{t \in N} t \right) \text{ and } \alpha_1 = \frac{\sum_{t \in N} \left(o_t - \frac{1}{N} \sum_{t \in N} o_t \right) \left(t - \frac{1}{N} \sum_{t \in N} t \right)}{\sum_{t \in N} \left(t - \frac{1}{N} \sum_{t \in N} t \right)^2}$$

and the detrended hindcasts are obtained as

$$o_t^D = o_t - (\alpha_0 + \alpha_1 t). \quad (2.8)$$

where o_t represents the SST (or the SSH) GECCO anomalies field (x, y, t) . Coefficients $\alpha_0 = \alpha_0(x, y)$ and $\alpha_1 = \alpha_1(x, y)$ are computed for each grid point, N is a dimension size. Assuming the variable is normally distributed as $N(\alpha_0 + \alpha_1 t, \sigma^2)$, o_t^D represents detrended data.

The drawback of this method is that if the trends of decadal hindcasts and GECCO are different, then by removing the GECCO trend from hindcasts, the latter might still have some remaining trend. In the following assessment of hindcast's performance this approach may lead to underestimation of the skill (when the hindcasts with remaining trend are compared against the observations with no trend). The alternative approach is to subtract own trends of GECCO and hindcasts over lead years (e.g., detrend time series that contain fields at particular lead year). Similar approach was also suggested by Kharin et al. (2012) as time-dependent trend adjustment and an alternative to bias-correction procedure.

2.4.2 Bias correction

An a posteriori bias correction procedure is commonly applied in seasonal forecasting (Stockdale, 1997) and can be computed from a series of initialized hindcast over all starting dates or in a cross-validated way. Smith et al. (2013) argued that, given a small sample size (as in CMIP5 initialized experiments), the latter approach may lead to underestimation

of predictive skill, while mean bias-correction tends to overestimate the skill. Hence, the authors suggested to use both methods, stating that the "true skill" lies between these two estimates. Both methods imply dependency of bias on lead year but not on the starting date. In this thesis we mostly apply mean bias correction as the sample size is rather small and it is important to use all the starting dates to better estimate the bias.

The mean bias correction is performed as follows. Let $H = h_{ist}, \dots, h_{IST}$ represent predicted anomaly, where i identifies a particular ensemble member in an ensemble of I simulations, $s = 1, \dots, S$ identifies starting date for initialized hindcasts, and t identifies a particular hindcast year (lead year). The initialized ensembles for a particular initialization scheme consists of $S = 9$ starting dates, $I = 3$ simulations per starting date that are $T = 10$ years in length. We correct the predictions for the lead-time dependent bias (b_t) computed with respect to GECCO. The GECCO synthesis goes only until 2001, therefore the number of integrations that participate in bias estimation is $N = 21$ (or $S = 7$ cases):

$$b_t = \sum_{n=1}^N (h_{nt} - o_{nt})/N, \quad (2.9)$$

$$h_{nt}^* = h_{nt} - b_t,$$

where o_{nt} is an observed (in our case GECCO) counterpart of the hindcast h_{nt} , and h_{nt}^* is a bias corrected initialized hindcast. This procedure is applied to non-detrended (AMOC) hindcasts. To account for trend in the SST, SSH and AMOC hindcasts, we compute the bias with respect to detrended GECCO fields:

$$b_t^D = \sum_{n=1}^N (h_{nt} - o_{nt}^D)/N, \quad (2.10)$$

$$h_{nt}^{D*} = h_{nt} - b_t^D.$$

where index (D) implies that computations take into account a linear trend. The bias corrected hindcasts for the North Atlantic SST, AMOC and North Atlantic SSH as well as their mean biases are shown in Fig. 3.5, 3.10, 3.12, respectively.

2.4.3 Verification Metrics

The verification metrics are applied to the ensemble mean prediction averages:

$$h_{st} = \frac{1}{I} \sum_{i=1}^{I=3} h_{ist}. \quad (2.11)$$

where the input for computations are bias corrected hindcasts ($h_{ist} = h_{nt}^{D*}$ for detrended time series, and $h_{ist} = h_{nt}^*$ for non-detrended ones).

The predictive skill can be assessed by comparing the modeled predictions with the ob-

servations or with reanalysis data (Tebaldi and Knutti, 2007; Goddard et al., 2013). While for a single given diagnostic or variable, the model performance is difficult to estimate, previous studies indicate that the differences are more distinct if multiple diagnostics are considered. Hence, the skill of the model is measured using the anomaly correlation as a function of lead time (COR) and root-mean-square-error skill score ($RMSS$). Correlation with GECCO is computed as follows:

$$COR_t = \frac{\sum_s h_{st} o_{st}}{\sqrt{\sum_s h_{st}^2 \sum_s o_{st}^2}}. \quad (2.12)$$

where h_{st} is the ensemble mean of predicted anomalies at lead time t or the subsample of the 20C ensemble mean, o_{st} is observed anomaly. The 20C ensemble mean covers the period 1920-2010; for estimating predictive skill, it is sub-sampled to have the same prediction intervals as the initialized hindcasts. We note that GECCO is used as the verification dataset for SST, SSH and MOC. Additionally, for evaluation of SST skill, we use the UK Met Office Hadley Centre observations dataset (HadISST; Rayner et al. (2003)).

The analysis of $RMSS$ addresses the question of the relative accuracy of initialized hindcasts with respect to un-initialized hindcasts. The skill score is defined as follows:

$$RMSE_{HIND(\text{or } 20C)} = \sqrt{\frac{1}{S} \sum_s (h_{st} - o_{st})^2}, \quad (2.13)$$

$$RMSS = 1 - \frac{RMSE_{HIND}}{RMSE_{20C}},$$

where $RMSE_{HIND}$ and $RMSE_{20C}$ are the root-mean-squared-errors of initialized and un-initialized hindcasts (h_{st}), respectively, computed with respect to GECCO (o_{st}). COR and $RMSS$ are computed either for each grid point or for spatial averages.

2.4.4 Persistence

To find out whether initialized hindcasts show any predictive skill due to initialization, they are compared with the low-skill persistence forecasts. The damped persistence forecasts are calculated from the annual GECCO anomaly fields, starting from the year that precedes the initialization year. For this we use the first-order autoregression or AR(1) model (or damped persistence, von Storch and Navarra, 1999):

$$g_t = \alpha g_{t-1} + N_t, \quad (2.14)$$

where g_t represents the GECCO data over the period 1960-2001, N_t is a white noise process (here $N_t = 0$). The autoregression parameter α is calculated as the lag-1 autocorrelation coefficient from the GECCO time series over the period 1960-2001.

This method is more appropriate for comparing performance of initialized hindcasts and persistence at every lead year. When the skill is computed for the averages over lead years 2 to 5 (yr2-5) and years 6 to 9 (yr6-9), the persistence forecasts are calculated from the averages over 4 years, which precede the initialization date. For instance, for the first start date 1961, the hindcasts at lead time yr2-5 (1962-1965) are compared with the persistence forecast at lag-5 (1957-1960) and the hindcasts at yr6-9 (1966-1969) are compared with the persistence lag-9 (1957-1960).

2.4.5 Assessment of Significance Level

An estimation of the significance level for *COR* and *RMSS* is based on a nearest-neighbor bootstrap method proposed by Goddard et al. (2013). This method is good for a small-sized set of integrations because it does not need assumptions about distributions, and for the integrations which are non-independent in space or time because this dependency is preserved by re-sampling the data in blocks of length based on autocorrelation function.

A verification metric as a function of lead time is computed for ensemble-mean initialized hindcasts h_s , where s identifies a starting date. Hence, the non-parametric bootstrap algorithm is as follows:

Step 1A. First, we carry out a re-sampling with replacement of S starting dates for both hindcasts and their counterpart observation time-series which gives us bootstrapped samples: h_{is}^B and o_s^B . Given the likely time-dependence (e.g. due to anthropogenic forcing) between the starting dates, block moving bootstrap of neighboring 5-year pairs is applied.

For example, for a set of starting dates:

$$s^h = 1^{(1960)} \quad 2^{(1965)} \quad 3^{(1970)} \quad 4^{(1975)} \quad 5^{(1980)} \quad 6^{(1985)} \quad 7^{(1990)} \quad 8^{(1995)} \quad 9^{(2000)}$$

a re-sampled set of starting dates may look like:

$$s^{h^B} = s^{o^B} = \underline{2}^{(1965)} \quad 3^{(1970)} \quad \underline{6}^{(1985)} \quad 7^{(1990)} \quad \underline{3}^{(1970)} \quad 4^{(1975)} \quad \underline{2}^{(1965)} \quad 3^{(1970)} \quad \underline{1}^{(1960)}$$

where underline denotes the dates that are picked randomly and the following dates are assigned in a pair. The sample of dates for observations is the same as for hindcasts.

Step 1B. Thereafter, we resample with replacement the ensemble members, for instance:

$$i^h = \begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \end{array} \quad \text{after re-sampling:} \quad i^{h^B} = \begin{array}{cccccccc} 3 & 2 & 1 & 1 & 1 & 2 & 3 & 2 & 3 \\ 1 & 1 & 2 & 3 & 2 & 1 & 3 & 2 & 1 \\ 1 & 3 & 2 & 2 & 1 & 3 & 2 & 3 & 2 \end{array}$$

and construct an ensemble mean h_s^B of bootstrapped hindcasts.

Step 2. Compute verification statistics (*COR*, *RMSS*) based on bootstrapped samples h_s^B (Step 1B) and o_s^B (Step 1A).

Step 3. Repeat Steps 1 and 2 a large number of times, say 1000, to generate a probability distribution of skill scores:

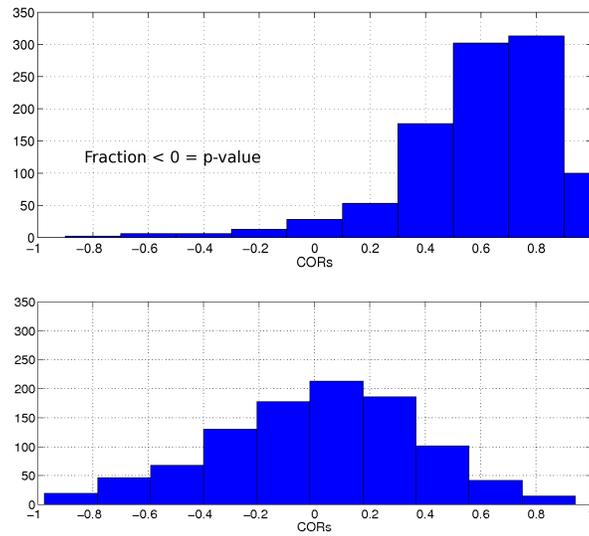


Figure 2.10— Example of the histograms of the re-sampled correlations between the FC-INIT and GECCO for the lead year 2 (upper panel) and lead year 10 (lower panel).

Step 4. A fraction of negative values serves as p value for the test and is compared to the significance level α : if the p value $\leq \alpha$, COR ($RMSS$) is significant for the $(1 - \alpha) \times 100$ confidence level.

Chapter 3

Impact of Initialization Procedures on Predictive Skill of a Coupled Ocean–Atmosphere Model⁴

3.1 Introduction

Insight into the predictability of the climate system at decadal timescales can be gained from initializing coupled climate models (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009; Mochizuki et al., 2010). In addition to the external forcing, initializing coupled climate models accounts for internal variability of the climate system, such as El Niño–Southern Oscillation, fluctuations in the meridional overturning circulation (MOC) and changes in ocean heat content. Branstator and Teng (2010) reported the dominance of the initial conditions for potential predictability of the North Atlantic upper ocean heat content for up to a decade, with subsequent control by the external forcing. The initialization of dynamical climate models with the current state of the climate system is expected to further improve the skill of near-term climate predictions (Climate Model Intercomparison Project Phase 5: Doblas-Reyes et al. (2011b)). However, the full benefits gained from using present day initial conditions for climate predictions are not yet determined, nor do we know the best practices for climate model initialization procedures.

In general terms, the initialization of decadal climate predictions refers to a start-up of a coupled climate model given the recent climate state as initial conditions. In principle, all climate components should be initialized; however, since most of the climate memory resides in the ocean, it is not uncommon for decadal predictions to primarily initialize the ocean model component using ocean information provided by some kind of ocean state estimate. To date, numerous ocean synthesis products exist, which in principle all could be used as initial conditions (see http://icdc.zmaw.de/easy_init_ocean.html for detailed

⁴Polkova I., Köhl A. and Stammer, D. Impact of initialization procedures on the predictive skill of a coupled ocean–atmosphere model. *Climate Dynamics* DOI 10.1007/s00382-013-1969-4

information). Examples in the context of initialized decadal prediction experiments are provided by Smith et al. (2007, 2013), Keenlyside et al. (2008), Pohlmann et al. (2009), Mochizuki et al. (2010), Matei et al. (2012), Magnusson et al. (2012a) and Hazeleger et al. (2013a). Comparisons of different ocean state estimates were performed by Carton and Santorelli (2008), Stammer et al. (2009), Lee et al. (2010) and Munoz et al. (2011).

While differences exist between available ocean syntheses, a guiding principle for their use as initial conditions in a coupled model should be that the initial ocean state is as close as possible to the numerical details of the ocean component of the model, so as to minimize initial drift of the coupled model and thereby improve the forecast skill. However, this choice may not guarantee a low drift because uncoupled and coupled models tend to establish different climate and the ocean synthesis additionally introduces changes to the forcing that reduces the climate bias from which the coupled model can not benefit. Unless initial conditions are estimated with the coupled model itself used for the climate prediction, model drift cannot be avoided altogether. To minimize drift, the right choice of the initialization procedure is essential. Because decadal climate predictions are an initial value problem, some initialization approaches used for seasonal to interannual forecasting are potentially applicable to initialize predictions at decadal timescales as well (Goddard et al., 2012). Among them are full state initialization, anomaly initialization and flux correction.

The strengths and shortcomings of the full state and anomaly initializations were reviewed previously by Meehl et al. (2009, accepted, 2013) and applied by Troccoli and Palmer (2007), Doblas-Reyes et al. (2011a), Smith et al. (2007, 2013), Pohlmann et al. (2009), Mochizuki et al. (2010) and others. An alternative initialization approach that accounts for model drift employs flux correction. The flux correction scheme was introduced by Sausen et al. (1988) for a simple non-linear ocean-atmosphere model. Applications of the flux correction in state-of-the-art coupled climate models have undergone further development and differ significantly from application to application (adjustment of heat, freshwater and momentum fluxes). The study by Shackley et al. (1999) reviews aspects of using flux correction in coupled ocean-atmosphere general circulation models. Magnusson et al. (2012a), focusing mainly on systematic error, recently discussed these three initialization methods in the context of seasonal to decadal forecasts without identifying a clear preference to any of them.

The goal of this work is to investigate the sensitivity of decadal predictive skill of the UCLA/MITgcm coupled climate model to three different ocean initialization approaches, notably the full state initialization (FSI), anomaly initialization (AI) and full state initialization employing flux correction (FC). In contrast to several previous studies, we initialize the ocean component with the ocean synthesis provided by the project “German contribution to Estimating the Circulation and Climate of the Ocean” (GECCO; Köhl and Stammer (2008)), which was produced using the same MITgcm ocean model that is also a part of the coupled system. While still not ideal in the sense that we did not estimate the initial state using the coupled model, we assume that using the same ocean model for esti-

minating the initial conditions as well as for the prediction of the coupled model is already a step in this direction. The present work focuses specifically on assessing which of the three initialization procedures lead to the best predictive skill for the North Atlantic sea surface temperature (NA SST) and the Atlantic meridional overturning circulation (AMOC). We also analyze sea surface height (SSH) hindcasts to learn whether it is possible to improve SSH predictions through initialization procedures.

The remaining paper is organized as follows. In Section 3.2 we provide a description of the coupled ocean-atmosphere general circulation model, initialization procedures, experiments and methods used to estimate the predictive skill of the model. In Section 3.3 we compare the skill of the three initialization approaches for SST with focus on the North Atlantic region. We also consider possible mechanisms responsible for the NA SST predictability. The SSH predictive skill is addressed in Section 3.4 and in Section 3.5 AMOC initialized hindcasts are discussed. Finally Section 3.6 concludes with a summary and discussion.

3.2 Methodology

The Coupled Model

The climate model used in this study is the University of California, Los Angeles coupled general circulation model (UCLA CGCM), which consists of the UCLA atmospheric general circulation model (UCLA AGCM; see www.atmos.ucla.edu) coupled with the Massachusetts Institute of Technology ocean general circulation model (MIT OGCM; see mitgcm.org). The performance of UCLA CGCM is described for ENSO forecasts (Cazes-Boezio et al., 2008). The model was used recently by Stammer et al. (2011) to simulate the global climate response to Greenland ice sheet melting. The UCLA AGCM is a state-of-the-art model with advanced parameterizations of the major physical processes in the atmosphere. In the present study, we use the UCLA AGCM version 7.1 with a horizontal resolution of 2.5° longitude and 2° latitude, and 29 vertical layers. The MIT OGCM has a quasi-global domain 80°S to 80°N and uses 46 levels. The resolution is $1^\circ \times 1^\circ$ with the refinement of the meridional resolution equatorward of 30°S and 30°N , such that at the equator it equals $1/3^\circ$.

We initialize the coupled model with the GECCO fields. The GECCO ocean synthesis product is a modification of the first ECCO ocean synthesis and uses the adjoint to the MIT OGCM model with a quasi-global domain, 1° horizontal resolution and 23 vertical layers (Köhl and Stammer, 2008). Decadal hindcasts were initialized previously with the GECCO synthesis by Pohlmann et al. (2009), Kröger et al. (2012) and Matei et al. (2012).

Initialization Methods

In the FSI and FC experiments, UCLA CGCM is initialized with the 3-D GECCO temperature, salinity, zonal and meridional velocities, and sea surface height (SSH). In addition,

in the FC experiment monthly climatological mean corrections to the surface heat and freshwater fluxes are applied. We calculate these fluxes from the relaxation terms obtained from a 10-year run, in which SST and sea surface salinity (SSS) of the model were relaxed toward the monthly climatology of GECCO, with a timescale of 30 days. The correction terms were constructed from the climatology of the associated heat and freshwater fluxes. Fig. 3.1 displays annual mean of the flux correction terms. The model in general tends to exhibit a warm SST bias and negative SSS bias (not shown). Positive surface heat flux correction reduces SST of the model, and negative surface freshwater flux correction increases SSS. For instance, heat flux correction of about 15 to 40 W/m^2 is employed in the Southern Ocean, and about 25 W/m^2 is added to the Labrador Sea, along the eastern side of Greenland and the Kuroshio Current. The freshwater flux correction values are in general between -1 to 1 mm/day and reach up to 4 mm/day on the eastern side of the subtropical gyres and up to -4 mm/day in the western equatorial Pacific Ocean and Mid-Indian Basin.

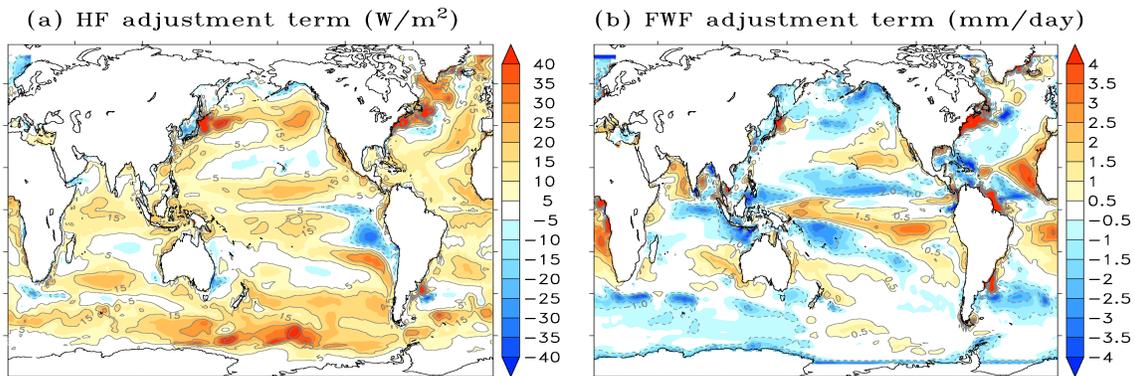


Figure 3.1— Annual mean of heat flux (a) and freshwater flux (b) adjustment terms.

In the AI experiment, the model is initialized from anomalies of the 3-D GECCO temperature, salinity, zonal and meridional velocities, and SSH. The GECCO anomalies with respect to the 50-year mean are added to the coupled model's climatology, which is computed from the last hundred years of a long spin-up run with constant CO_2 concentrations centered in the period that covers initialized experiments (1961-2010). This may introduce a small remaining drift due to differences in the external forcing.

Experiments

To test each of the above initialization schemes, an assimilation run and a series of initialized hindcasts are carried out. Before starting the assimilation runs, the coupled model was brought close to equilibrium during a 300-year spin-up run with constant greenhouse gas (GHG) concentration from the year 1980. In the assimilation experiments, the GECCO ocean state was introduced to the coupled model by a monthly re-initialization procedure,

which is an equivalent of nudging with the relaxation time scale 30 days. By doing this the atmosphere is brought to the state that is in agreement with GECCO. As ocean initial conditions for the assimilation runs and further for initialized hindcasts, we always use the GECCO temperature, salinity, zonal and meridional velocities, and SSH fields. The resulting assimilation runs cover the period 1952-2001 and are used as a source of atmospheric initial conditions for the following initialized decadal hindcasts. Ensembles of initialized hindcasts are started at the end of 1960, 1965, 1970, 1975, 1980, 1985, 1990, 1995 and 2000. Each ensemble contains three realizations that are generated by shifting the atmospheric state by two days relative to the assimilation run.

Hereafter, the assimilation runs using FSI, AI, or FC, are referred to as FSI-ASSIM, AI-ASSIM and FC-ASSIM, respectively. Accordingly, three sets of initialized coupled integrations are named as FSI-HIND, AI-HIND and FC-HIND. All experiments are summarized in Tab. 2.1.

We also performed ensemble integrations of the twentieth-century simulations starting from the same initial condition in 1900, 1910, 1920 till 2011, where the model is forced by annually and globally averaged concentrations of GHG (natural variations of the forcing, namely changes in solar irradiance and volcanic aerosols, were ignored). GHG concentrations are provided by Robertson et al. (2001) and Tans and Keeling (NOAA/ESRL, Scripps Institution of Oceanography, 2011). In the following we will refer to this experiment as 20C. Due to including only CO₂ external forcing, the 20C run obtains considerably lower skill than, for instance, in study by Pohlmann et al. (2009) and the correlation with observations for the globally averaged annual mean SST is only 0.24 in comparison to their value which is slightly larger than 0.6.

Time series of the spatially averaged NA SST and NA SST anomalies for the assimilation runs and the ensemble mean 20C simulation are shown in Fig. 3.2 a, b. The NA SST from the assimilation runs resembles well the interannual variability of GECCO. In addition, FC-ASSIM NA SST is the closest to the GECCO mean state. Anomaly correlation coefficients for NA SST between GECCO and assimilation runs are high, 0.98 for all three experiments. By contrast, the 20C run performs poorly in simulating the NA SST variability, the correlation with GECCO is only 0.21. Figs. 3.2 c, d show that the AMOC at 26.5°N of the three assimilation runs agrees with the GECCO estimate in terms of its interannual variability, although the AI-ASSIM AMOC reveals a strong upward trend. Kröger et al. (2012) also found an amplification of the AMOC trend at 25.6°N in their anomaly assimilation run. The same was true for the AMOC variability, which was highly increased in almost the entire basin and also for the other two ocean syntheses. The authors speculated about the strong effect of the anomaly initialization method on the transient behavior of the AMOC. Thus, depending on where anomalies enter the coupled model, they might interfere with existing variability. We also found that the evolution of the AMOC is influenced by the deficiencies of the initialization methods to correctly insert information associated with the flow through the narrow passages in the Caribbean Basin, whereas at high latitudes the AMOC trend and variability are in good agreement with

GECCO for all the initialization experiments (not shown). The correlation coefficients for the AMOC at 26.5°N are 0.85 for FSI-ASSIM, 0.83 for AI-ASSIM and 0.83 for FC-ASSIM. The variability of the overturning stream function in the 20C ensemble mean does not resemble the variability of GECCO. Due to opposing trends, the 20C AMOC at 26.5°N over the period 1952-2001 is negatively correlated with GECCO by about -0.65. This result is consistent with the finding of Pohlmann et al. (2009). Previous studies also suggest that a common feature of coupled model 20C simulations is a reduction in the AMOC when only the GHG radiative forcing is prescribed (Gregory et al., 2005). However, a recent study by Menary et al. (2013) indicates that the 20C integrations tend to simulate a weaker trend or even a strengthening of the AMOC, when they include a more realistic representation of the anthropogenic forcing taking into account aerosols (with both direct and indirect effects).

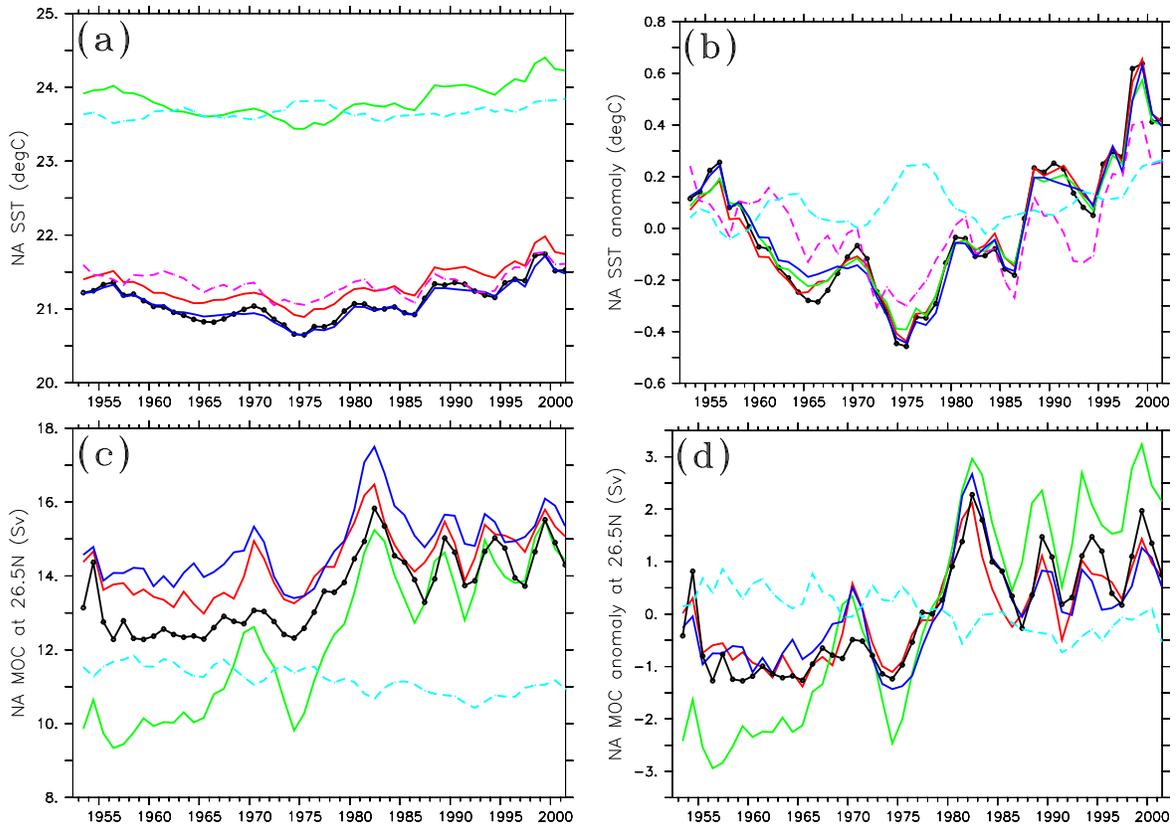


Figure 3.2— Time series of the North Atlantic annual-mean SST (a) and SST anomalies (b) from FSI-ASSIM (red), AI-ASSIM (green), FC-ASSIM (blue), GECCO (black), HadISST (dashed purple) and the 20C ensemble mean (dashed light blue). Time series of the annual-mean Atlantic MOC at 26.5°N (c) and Atlantic MOC anomalies at 26.5°N (d).

Skill Metrics

The predicted anomalies are estimated as departures of the initialized (un-initialized) hindcasts from the time-average condition over 1952-2001 of the corresponding assimilation run (20C run). Long term trends are calculated from the GECCO fields over the initialization period 1961-2001 and removed from the SST and SSH fields. The uninitialized hindcasts were also detrended. This step is done in order to focus on the predictive skill due to initialization rather than the total signal which contains also the external forcing response. Because there appears to be no trend due to external forcing in the AMOC case, we compute the skill for the non-detrended AMOC hindcasts. However, to avoid possible build-in skill, especially in the AI-HIND where the amplification of the AMOC trend is evident, we analyze in the following both the non-detrended and detrended AMOC hindcasts. Hereafter, for each starting date from 1961 to 1991, the lead-time dependent bias is derived with respect to the (detrended) GECCO data, and the overall mean bias is then removed from a set of hindcasts over all starting dates. A posteriori bias correction is based on the common procedure used in seasonal forecasting (Stockdale, 1997), recently discussed in more detail for decadal predictions by Magnusson et al. (2012a) and Smith et al. (2013). Following the recommendation by Goddard et al. (2013), we analyze the ensemble mean prediction averages over three ensemble members for each starting date.

The verification metrics are computed from annual mean fields of the SST, SSH and AMOC anomalies. The predictive skill of the model is measured using the anomaly correlation coefficients as a function of lead time (COR). The lead time here is the length of time between the starting date and the year of initialized hindcasts. The question about the relative accuracy of initialized hindcasts with respect to externally forced hindcasts is addressed in the analysis of the root-mean-square error skill score (RMSS), defined as

$$RMSS = 1 - \frac{RMSE_{INIT}}{RMSE_{20C}}. \quad (3.1)$$

Root-mean-squared errors of the initialized ($RMSE_{INIT}$) and un-initialized ($RMSE_{20C}$) hindcasts are given as the square root of the average squared differences between the predicted anomalies and GECCO anomalies.

The significance level estimation for the verification metrics is based on the nearest-neighbor bootstrap approach, which uses moving-blocks resampling with replacement algorithm applied to initialized/uninitialized simulations and observations in order to estimate the probability of obtaining high skill scores by chance. In the presence of temporal auto-correlation (trends due to external forcing, long-term variability), instead of resampling individual years, the blocks of consecutive years are resampled to preserve dependent structure in the sample. Here, the nearest-neighbor bootstrap considers five-year pairs. We have implemented the algorithm, which was described in details by Goddard et al. (2013) and especially designed for the CMIP5 set-up with every five-year initialization.

Throughout this study, the skill assessment for all quantities is mainly done against the

GECCO state estimate. We compare the skill of the initialized hindcasts with the persistence forecasts, computed at one-year lag for the comparison with initialized hindcasts at the first lead year, and 5- and 9-year lags of 4-year averaged (detrended) GECCO fields for the comparison with initialized hindcasts at lead time yr2-5 and yr6-9, respectively.

Since the evaluation of skill must be performed against observed anomalies and it is not clear how close the GECCO synthesis represents the real ocean state, we use additionally the UK Met Office Hadley Centre SST product (HadISST; Rayner et al. (2003)) and the AVISO SSH maps. However, because of the relatively short observational records for SSH, we use the AVISO SSH fields only for a qualitative comparison without calculating skill scores.

3.3 Initialized Decadal SST Hindcasts

3.3.1 Global distribution of SST predictive skill

Previous studies indicate that information provided by the initial conditions becomes insignificant with time. To judge the predictive skill of the initialized decadal SST hindcasts and define how fast the hindcasts “forget” about their initial state, we compare them first against the GECCO synthesis. Fig. 3.3 shows the regions with high predictive skill in terms of the correlation coefficients (COR) for detrended annual mean SST in the cases of initialized hindcasts: FSI-HIND, AI-HIND and FC-HIND and the two reference hindcasts: 20C run and persistence.

In the initialized experiments, the SST COR is significant in the first lead year over wide areas of the ocean, particularly in the North Atlantic, the central South Pacific, the western North Pacific, the Indian Ocean and the extratropical Southern Hemisphere. The high skill obtained during the first lead year is partly due to the closeness of the hindcasts to the GECCO initial conditions, which is also shown by the high skill of the persistence. For lead time yr2-5 and yr6-9, a significant reduction of the correlation is evident. FSI-HIND and FC-HIND show more pronounced skill in the Southern Ocean at lead time yr2-5 and yr6-9 than other hindcasts. For the North Atlantic, FSI-HIND and AI-HIND show some skill in the subpolar gyre region at lead time yr2-5, while for FC-HIND the region with high skill is even more extended. Persistence forecasts also show skill in these regions for lead time yr2-5. Collins (2002) reported that the North Atlantic and the Southern Ocean regions demonstrated potential predictability for the surface air temperature on decadal time scales. AI-HIND shows skill in the equatorial Atlantic, which persists also at lead time yr6-9. For this period both AI-HIND and FC-HIND appear to be more skillful in the North Atlantic than all other hindcasts. There is some skill in the eastern tropical Pacific at lead time yr2-5 for AI-HIND and yr6-9 for FSI-HIND and in the eastern South Pacific at lead time yr6-9 for FC-HIND. FSI-HIND shows high correlation values in the Kuroshio-Oyashio extension region over lead time yr6-9. Mochizuki et al.

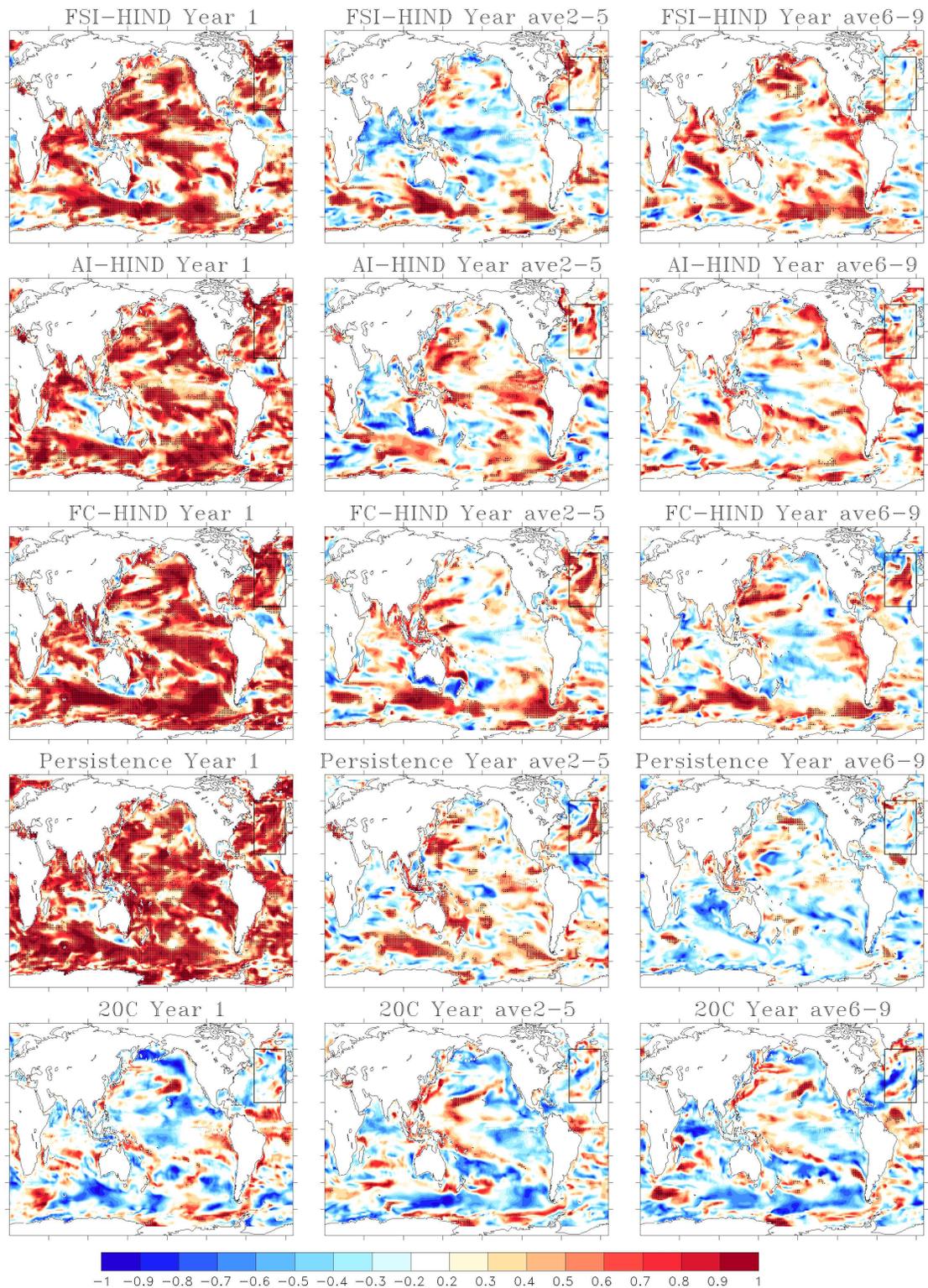


Figure 3.3— SST COR between detrended GECCO and initialized hindcasts for the first lead year (left column), lead time yr2-5 (middle column) and yr6-9 (right column). The hatched regions indicate significant skill at the 90% level according to a bootstrap procedure.

(2010) reported that their initialized Pacific Decadal Oscillation hindcasts are significantly correlated with observations of SST and temperature in the upper 300m over almost a decade in the Kuroshio-Oyashio extension and the subtropical oceanic frontal regions. The authors determined these regions as the centers of action of the Pacific Decadal Oscillation, which might thus be predictable on decadal timescales. The 20C run shows almost no skill for the given hindcast periods.

When comparing the initialized hindcasts with HadISST observations, the areas with high correlation are reduced, particularly in the extratropical Southern Hemisphere (not shown). In the North Atlantic the skill for FC-HIND at lead time yr2-5 is rather extended to the tropics. At lead time yr6-9 FSI-HIND and AI-HIND show some skill over larger areas in the eastern Pacific than FC-HIND does. The relatively low skill of the initialized hindcasts is mainly due to disagreement between GECCO and HadISST.

The distribution of the SST RMSS values is shown in Fig. 3.4. Positive values of the RMSS imply that initialized hindcasts outperform the 20C run, while negative values indicate that the skill of the initialized hindcasts is worse than that of the externally forced hindcasts. In general terms, the patterns are in agreement with the areas of high skill computed in terms of SST COR. For lead time yr6-9, FSI-HIND shows more skill in the Pacific Ocean and Indian Ocean than other initialized hindcasts do.

3.3.2 The North Atlantic SST predictability

Motivated by the results shown in Figs. 3.3 and 3.4, which indicate some skill over the North Atlantic basin, we analyze the time series of the spatially averaged NA SST anomalies. Fig. 3.5 a shows the mean model bias from ensemble-averaging the difference to GECCO. In most cases the tropical band of latitudes extended from the equator till 20°N in the North Atlantic basin did not obtain significant predictive skill, therefore we exclude this region from spatial averaging of the NA SST. At first, we intended to correct the bias only for FSI-HIND. Further analysis of the FC-HIND SST reveals that the model drift is not entirely removed. This is a consequence of the relaxation approach which requires some bias to generate correction fluxes. An incremental update procedure as proposed by Bloom et al. (1996) can resolve this problem but a smaller relaxation time may be a further option to reduce systematic model errors. Moreover, flux correction changes only the buoyancy fluxes, while wind stress forcing may still be biased. Therefore, correction of the residual drift is still required for FC-HIND. For comparison reasons, AI-HIND is also bias-corrected. For FSI-HIND, the systematic error grows over the first five years and saturates at about 2°C. The FC-HIND drift appears to be more moderate and levels out around the value of 0.7°C after the first four years. The AI-HIND has a warm bias of about 3°C for all lead years.

The time series of the spatially averaged (50°W-10°W; 20°N-60°N) bias corrected NA SST anomalies are shown in Fig. 3.5 b, c, d. The trends in FSI-HIND and FC-HIND

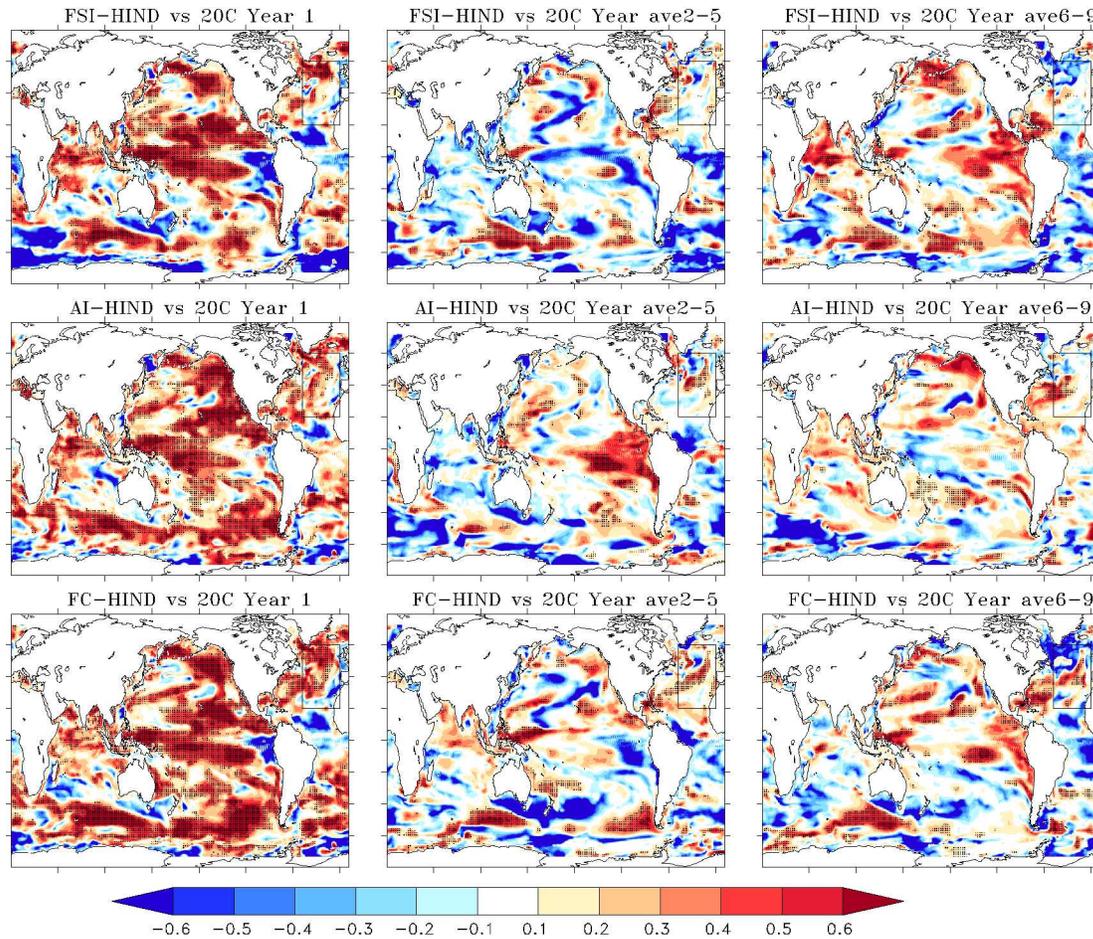


Figure 3.4— SST RMSS for the first lead year (the first column), lead time yr2-5 (the second column) and yr6-9 (the third column) for the detrended data. The hatched regions indicate significant skill at the 90% level according to a bootstrap procedure.

resemble most closely the trend in GECCO. The interannual variability of the NA SST in FC-HIND agrees better with GECCO than those from FSI-HIND and AI-HIND. For the last hindcast interval started in 2001, the AI-HIND NA SST anomalies stay closer to the observed SST anomalies than the other two experiments do. Hindcast quality statistics for the detrended spatially averaged NA SST anomalies are shown in Fig. 3.6. For lead time up to yr2-5, both metrics indicate significant skill only for the FC-HIND experiment. The FC-HIND NA SSTs are significantly correlated with GECCO until lead year six, while the NA SST COR already drops below the significance level after the third year in FSI-HIND and after the second lead year in AI-HIND (not shown). The NA SST COR for the 20C run shows no significant skill for the entire forecast period. The correlation coefficient for the persistence forecast is significant only in the first lead year.

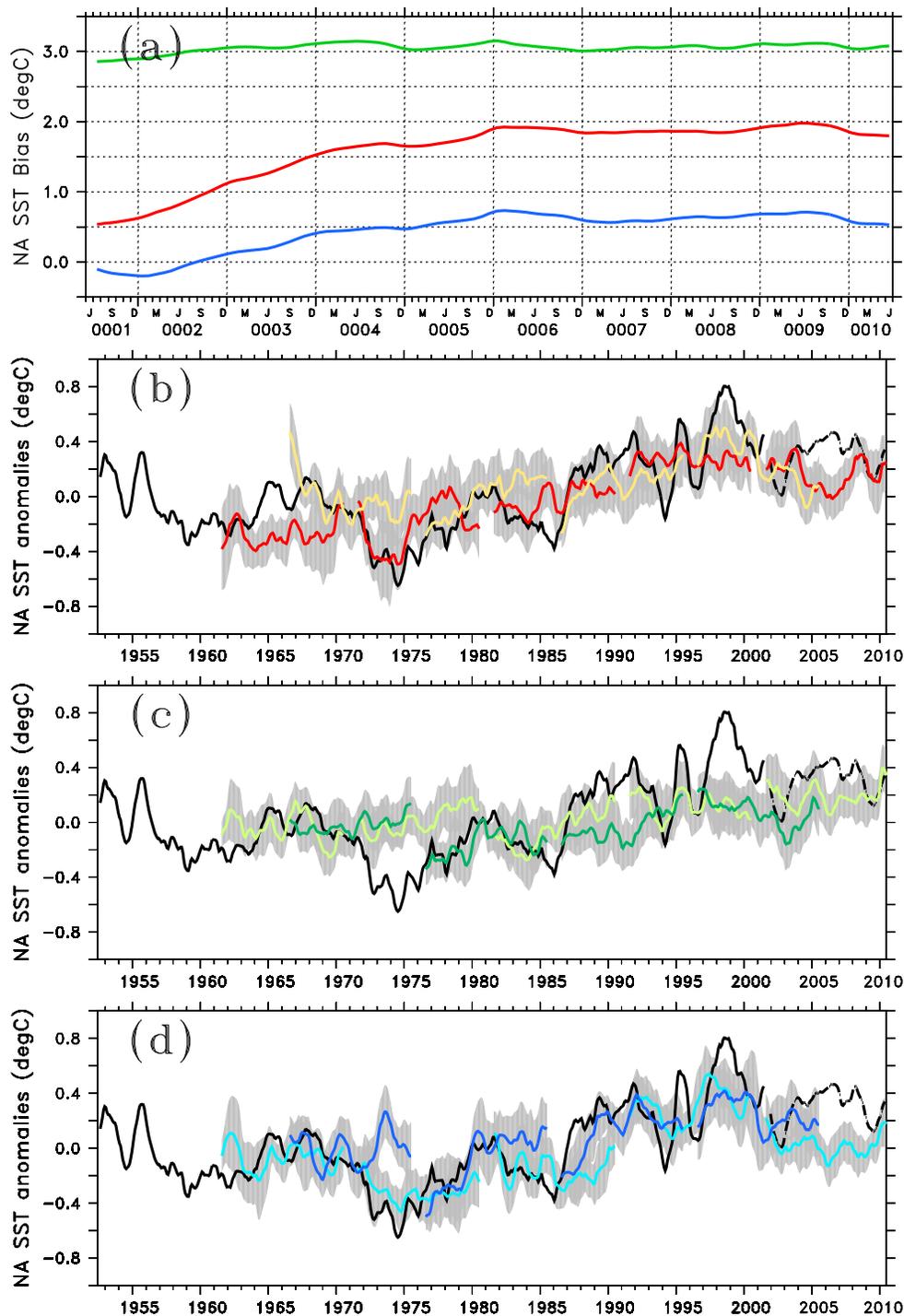


Figure 3.5— (A) Lead-time dependent North Atlantic SST bias averaged over the region 50°W - 10°W , 20°N - 60°N from FSI-HIND (red), AI-HIND (green) and FC-HIND (blue). Time series of the North Atlantic SST anomalies from GECCO (solid black), HadISST (dashed black; 2001–2010) and bias-corrected initialized hindcasts: FSI-HIND (b), AI-HIND (c) and FC-HIND (d). 12-month running mean is applied. Grey shading implies ensemble spread.

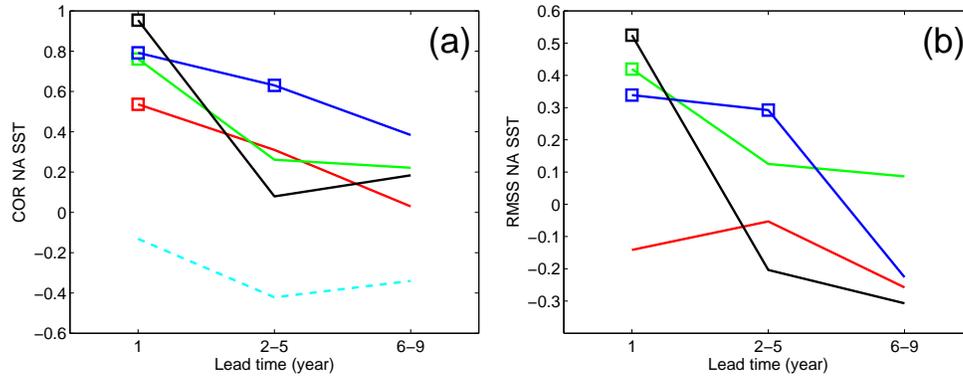


Figure 3.6— Anomaly correlation coefficient (a) and root mean square error skill score (b), both are computed with respect to GECCO, for the North Atlantic SST averaged over the region 50°W - 10°W , 20°N - 60°N from FSI-HIND (red), AI-HIND (green), FC-HIND (blue), the 20C ensemble mean (dashed light blue) and the persistence forecast (black). Squares imply significant correlation at the 90% confidence interval estimated with a bootstrap procedure. Skill scores are computed for the SST time series with the long-term trend taken out.

3.3.3 Possible mechanism responsible for the NA SST predictability

It can be expected that regionally high predictive skill in the North Atlantic and in the extratropical Southern Hemisphere may be linked to the seasonal cycle of the mixed layer depth (Frankignoul and Hasselmann, 1977; Alexander and Deser, 1994; Ciasto and Thompson, 2009). The mixed layer is essentially deeper in winter than in summer. Fig. 3.7 shows the regions of maximum mixed layer depth derived from GECCO and the initialized hindcasts for the first lead year. The winter SST anomalies of the initial conditions in the first lead year reflect the anomalies of the temperature of the deep mixed layer and skill persist throughout wintertime. Thereafter, these anomalies remain beneath the shallow summer mixed layer and are re-entrained into the mixed layer during the following winter. High correlation regions for SST (Fig. 3.3) are clearly related to areas of deep winter mixed layers and the higher skill, for instance, in the extratropical Southern Hemisphere for FSI-HIND and FC-HIND is related to a better agreement of their mixed layer depths with GECCO. The periodic shielding of the deep winter mixed layer by the shallow seasonal mixed layer implies a damped periodic oscillation of skill reflecting a good agreement with GECCO in wintertime. Indeed, the skill of the spatially averaged monthly mean NA SST hindcasts shown in Fig. 3.8, demonstrates high correlation and low RMSE during the winter months January, February and March, while low correlation and high RMSE are seen during the summer months June, July and August. This result is in agreement with previous studies by Marshall et al. (1993) and Qiu and Huang (1995) which imply that the ocean has a selective memory to winter initial conditions.

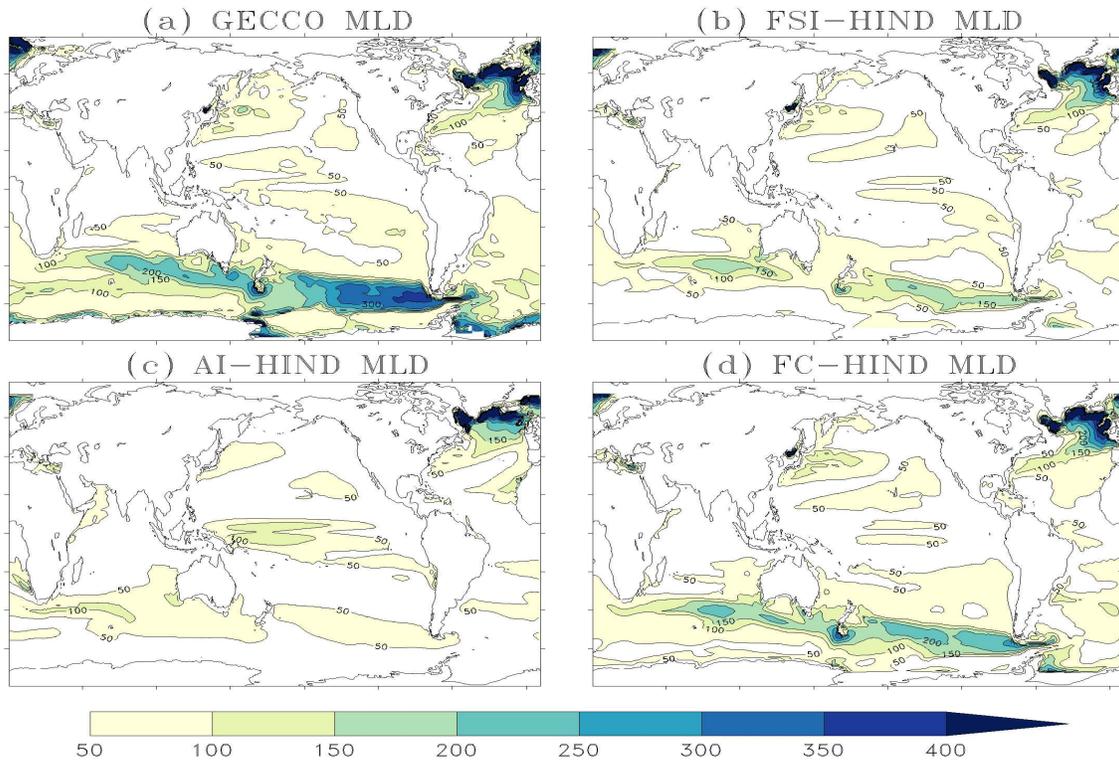


Figure 3.7— Annual climatological mean of the deep mixed layer in the first lead year for (a) GECCO, (b) FSI-HIND, (c) AI-HIND and (d) FC-HIND. The mixed layer depth is computed using a potential density criterion (0.125 in sigma units).

3.4 Initialized Decadal SSH Hindcasts

As indicated above, much of the hindcast skill of SST anomalies is believed to be related to temperature anomalies in the deep winter mixed layer, which therefore also relate to heat content anomalies. Much of the low frequency sea level variability outside of shallow or polar areas is associated with steric sea level changes (Vinogradova et al. (2007)). The GECCO synthesis shows good agreement with the AVISO data spanning the period 1993–2001, and the study by Köhl and Stammer (2007) indicates that during these recent years most of the regional sea level trend is due to the thermosteric sea level, therefore also reflecting heat content changes. The clear link between variations of SST and SSH, also seen by Leeuwenburgh and Stammer (2001), motivates the analysis of predictability for decadal SSH hindcasts.

We base the computation of the skill for initialized SSH hindcasts on GECCO data because routine SSH observations with satellite altimeters became available only in 1993. The distribution of COR as a function of lead time is given in Fig. 3.9. We observe high

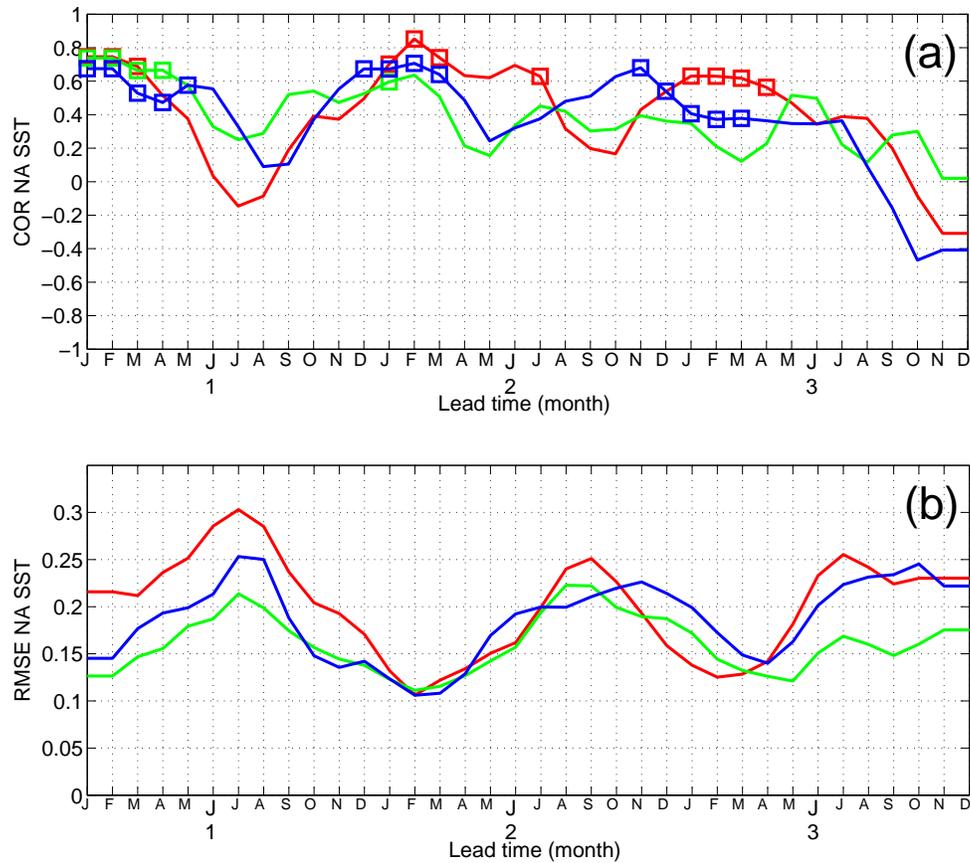


Figure 3.8— Anomaly correlation coefficient (a) and root mean square error (b), both are computed with respect to GECCO, for the North Atlantic monthly-mean SST averaged over the region 50°W - 10°W , 20°N - 60°N from FSI-HIND (red), AI-HIND (green), FC-HIND (blue). Squares imply significant correlation at the 90% confidence interval estimated with a bootstrap procedure. 3-month running mean is applied. Skill scores are computed for the SST time series with the long-term trend taken out

correlation in the first lead year over slightly larger areas of the ocean than for SST COR (Fig. 3.3). In general, the SSH COR patterns are concordant with those of SST COR; additionally, at lead time yr2-5 high correlation values are seen in the Arabian Sea. Areas of significant correlation in the North Atlantic and the Southern Ocean are more extended than for SST COR.

The mean bias and time series of the spatially averaged (50°W - 10°W ; 20°N - 60°N) NA SSH are shown in Fig. 3.10. FSI-HIND and FC-HIND start with a cold bias in the first lead year. Thereafter, the drift shows a different tendency in these experiments. There is no saturation phase such as in the SST bias. FSI-HIND shows a very strong negative bias beyond 5 years. The AI-HIND NA SSH has a positive bias after the first lead year that constantly increases such that the drift amounts to 1 cm per decade. Like for AI-HIND, the FC-HIND bias also shows an upward trend. The NA SSH anomalies after bias correction are rather flat, being in the range ± 1 cm. For Fig. 3.10 we removed the global

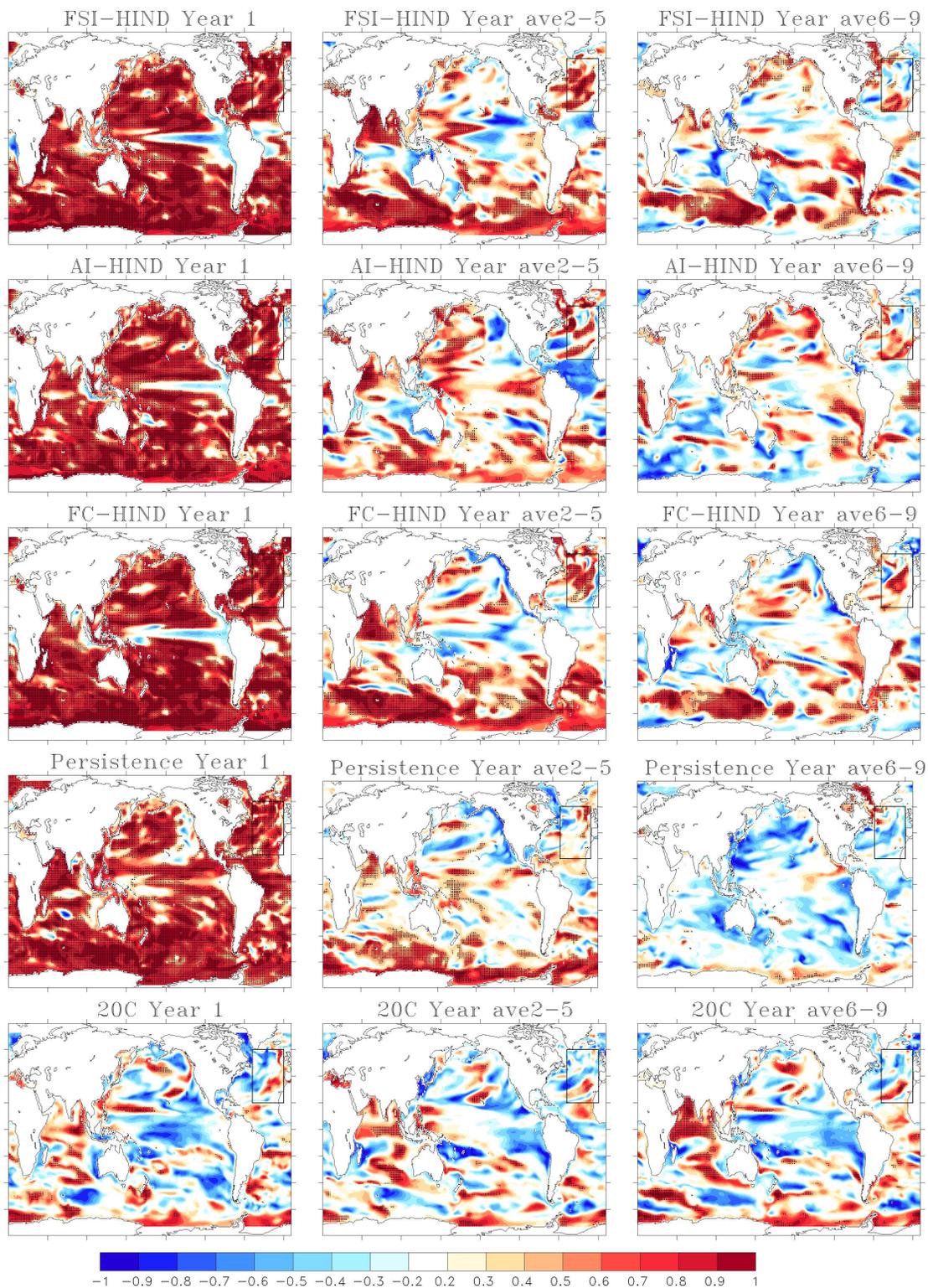


Figure 3.9— SSH COR between the detrended GECCO data and hindcasts for the first lead year (left column), lead time yr2-5 (middle column) and yr6-9 (right column). The hatched regions indicate significant skill at the 90% level according to a bootstrap procedure.

trend from AVISO data, because the Boussinesq approximation used by the model implies constant ocean volume.

Fig. 3.11 shows COR and RMSS for the spatially averaged NA SSH. In terms of correlation, all the initialized hindcasts show high significant skill up to lead time yr2-5. The correlation coefficients between the FC-HIND NA SSH hindcasts and GECCO stay significant also at lead time yr6-9. The correlation coefficients computed for each lead year remain significant up to three years for FSI-HIND, five years for AI-HIND and nine years for FC-HIND (not shown). High RMSS is evident for the persistence forecast, FC-HIND and AI-HIND for the first year, hereafter AI-HIND shows an improvement over the uninitialized hindcasts (although not significant) at lead time yr2-5 and yr6-9, and FC-HIND for yr6-9. One reason of a low RMSS connected with a high COR in FC-HIND at lead time yr2-5 might be a remaining drift that is not corrected properly. COR is not affected by the drift because lead years are correlated, note that AI-HIND has highest RMSS and probably the smallest drift.

3.5 Initialized Decadal AMOC Hindcasts

Variability of the AMOC is widely considered to play an important role in driving NA SST variations (Marshall et al., 2001) and demonstrates potential predictability on interannual to decadal time scales (Collins et al., 2006). The mean bias and the time series of the AMOC anomalies at 26.5°N (at 1000m) are shown in Fig. 3.12. The FSI-HIND and FC-HIND AMOC hindcasts drift towards the model's preferred state at about 10-11 Sv (not shown). This is also the range at which the AMOC at 26.5°N occurs in the 20C ensemble mean. The bias of both FSI-HIND and FC-HIND starts from a negative value of about 1 Sv, which decreases to close to neutral values during years 3 to 4, but increases rapidly thereafter over the rest of the time interval. The AI-HIND AMOC has a negative bias in the range of -1 to -1.5 Sv. After adjusting the bias, the AMOC anomalies of FSI-HIND and FC-HIND are in good agreement with GECCO (Fig. 3.12 b, d). AI-HIND shows some large maxima before 1980 that are not in GECCO AMOC, which mainly stays within the AI-HIND ensemble spread (Fig. 3.12 c).

Predictive skill in terms of correlation coefficients for non-detrended and detrended AMOC is summarized in Figs. 3.13 and 3.14, respectively. South of the equator, high correlation between the hindcasts and GECCO remains significant up to a decade in all the initialized experiments. A reduction of skill in the first lead year is evident in the region between the equator and 20°N in the FSI-HIND (Fig. 3.13) and is more pronounced in all the initialized hindcasts (Fig. 3.14) and also the persistence when detrended data are used. The decrease of the skill is related to deficiencies of the initialization methods to correctly insert information associated with the flow through narrow passages passing through the Caribbean Sea into the Gulf of Mexico and through the Strait of Florida into the North Atlantic. This can already be noticed in the assimilation runs (not shown).

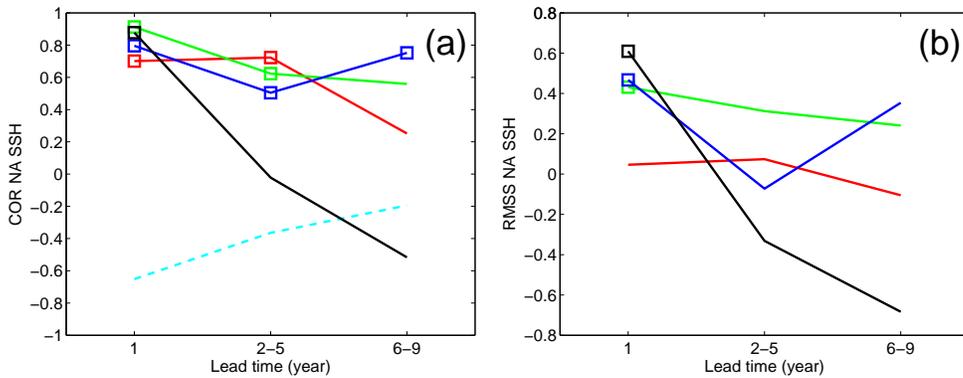


Figure 3.11— Anomaly correlation coefficient (a) and root mean square error skill score (b), both are computed with respect to GECCO, for the for the North Atlantic SSH averaged over the region 50°W - 10°W , 20°N - 60°N from FSI-HIND (red), AI-HIND (green), FC-HIND (blue), the 20C ensemble mean (dashed light blue) and the persistence forecast (black). Squares imply significant correlation at the 90% confidence interval estimated with a bootstrap procedure. Skill scores are computed for the SSH time series with the long-term trend taken out.

According to Fig. 3.13 b, c, all three approaches lead to skillful AMOC predictions during the first pentad over the whole North Atlantic up to 55°N , while only the FSI-HIND and FC-HIND obtain significant skill during the second pentad (with the exception of the 42°N - 48°N latitudinal band). Due to the opposing trend, the 20C Atlantic MOC is negatively correlated with GECCO. For the detrended AMOC hindcasts the skill is somewhat decreased but the overall significantly skillful regions for the initialized hindcasts remain the same. The correlation coefficients for non-detrended AMOC at 26.5°N computed for each lead year remain significant up to five years for FSI-HIND, six years for AI-HIND and nine years for FC-HIND (not shown).

3.6 Summary and Discussion

The goal of this study was to estimate the impact of different initialization strategies on the predictive skill of sea surface temperature (SST), sea surface height (SSH) and the Atlantic meridional overturning circulation (AMOC). The predictive skill was tested against the GECCO synthesis and the HadISST product for SST, while the predictive skill for SSH and the AMOC was assessed exclusively against the GECCO synthesis. All the experiments were performed using the UCLA/MITgcm coupled ocean-atmosphere model.

In the FSI approach, the coupled climate model tends to drift away from the initial conditions as the model state becomes centered on the modeled climatological state. In this case it is necessary to apply a lead-time dependent bias correction (Stockdale, 1997; Doblas-Reyes et al., 2011a; Magnusson et al., 2012a). Early on, Meehl et al. (2009) consider the FSI method as less efficient because predictable signals of small amplitude can be masked by model biases. However, more recently, Smith et al. (2013) show that FSI

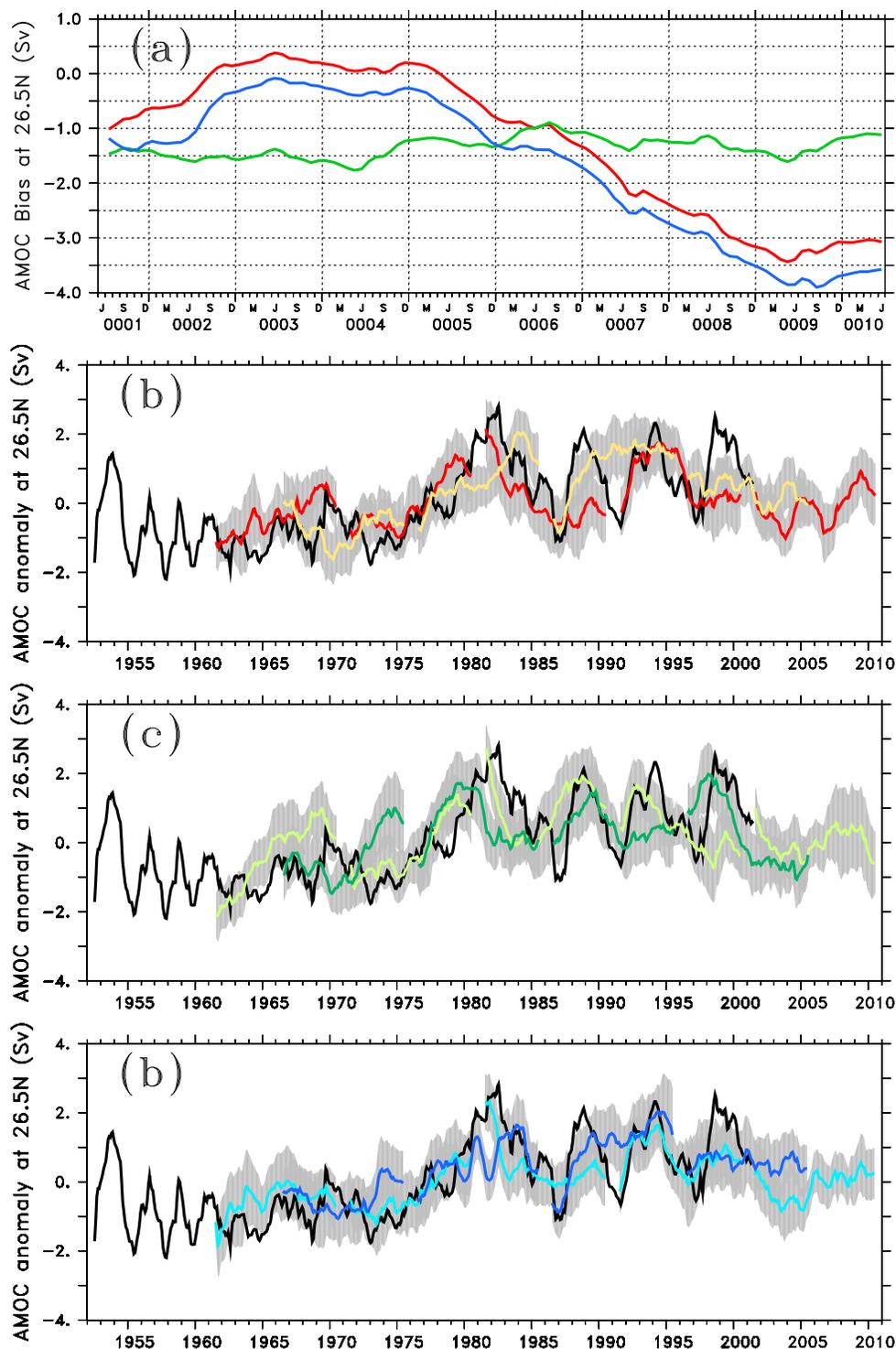


Figure 3.12— (A) Lead-time dependent Atlantic MOC bias at 26.5°N (1000m) from FSI-HIND (red), AI-HIND (green) and FC-HIND (blue), respectively. Time series of the Atlantic MOC anomalies from GECCO (black) and bias-corrected initialized hindcasts: FSI-HIND (b), AI-HIND (c) and FC-HIND (d). 12-month running mean is applied. Grey shading implies ensemble spread.

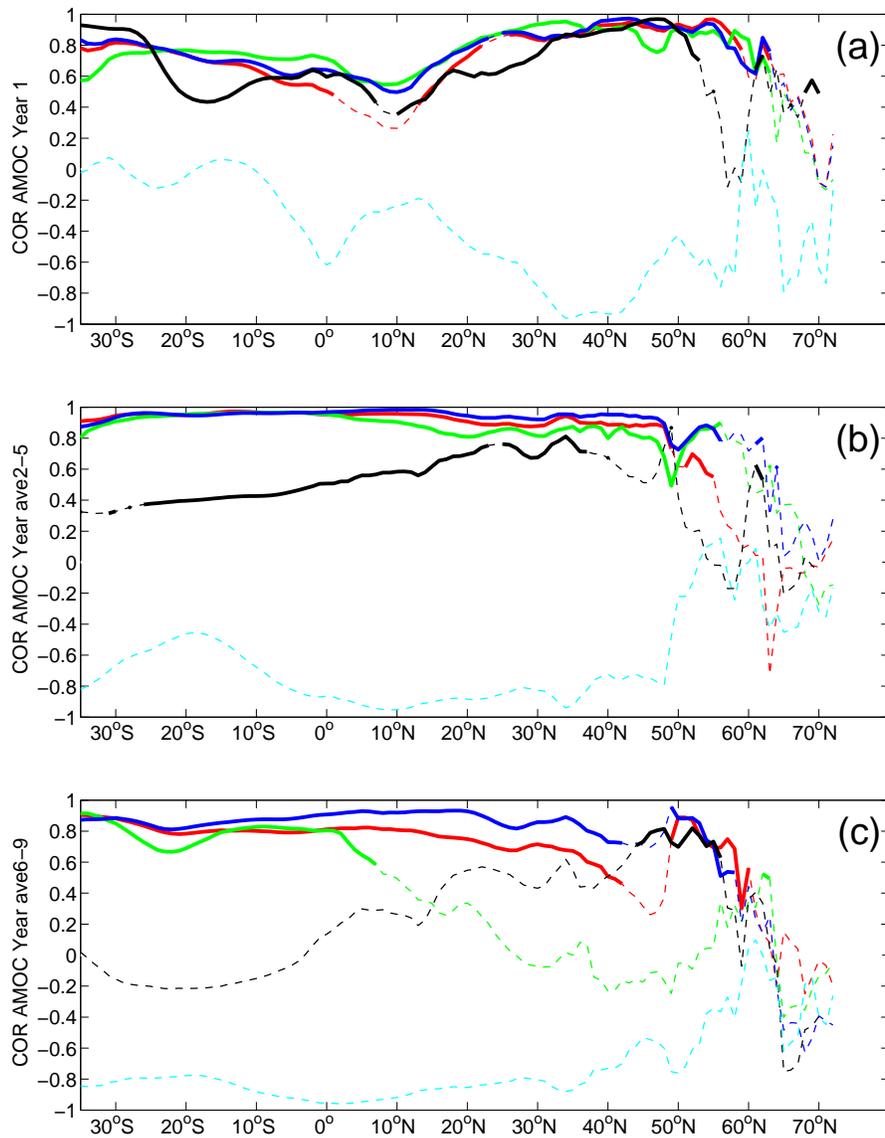


Figure 3.13— Anomaly correlation coefficient for the Atlantic MOC at 1000m between GECCO and initialized hindcasts: FSI-HIND (red), AI-HIND (green), FC-HIND (blue), the 20C ensemble mean (light blue) and the persistence forecast (black). CORs are calculated for the first lead year (a), lead time yr2-5 (b) and yr6-9 (c). The hindcasts that do not obtain significant skill at 90% significance level are shown in dashed curves.

can be advantageous on seasonal timescales, while on multi-year time scales FSI and AI lead to a comparable predictive skill. For the detrended North Atlantic SST and North Atlantic SSH hindcasts, averaged over the region 50°W-10°W, 20°N-60°N, our results show predictive skill up to 3 years. For the non-detrended AMOC the FSI results show high predictive skill between 15°N and 55°N up to lead time yr2-5.

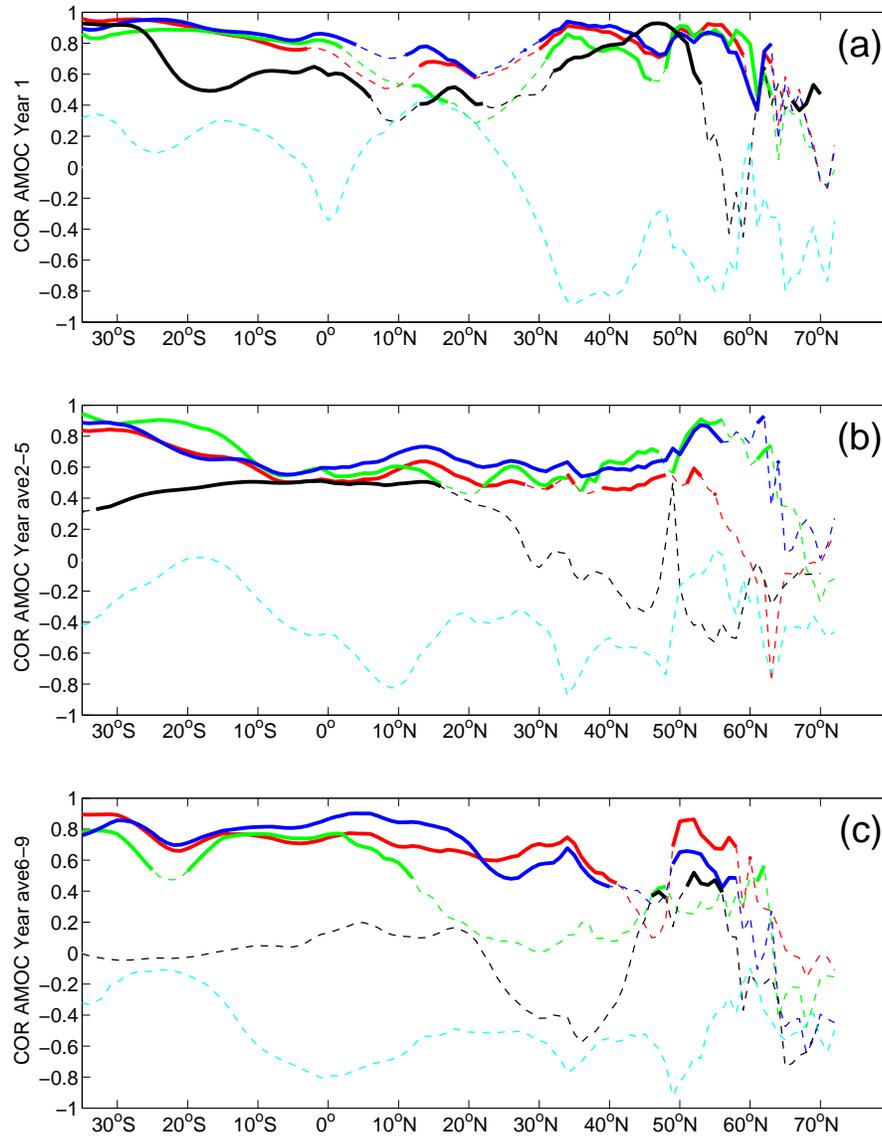


Figure 3.14— Anomaly correlation coefficient for the detrended Atlantic MOC at 1000m between GECCO and initialized hindcasts: FSI-HIND (red), AI-HIND (green), FC-HIND (blue), the 20C ensemble mean (light blue) and the persistence forecast (black). CORs are calculated for the first lead year (a), lead time yr2-5 (b) and yr6-9 (c). The hindcasts that do not obtain significant skill at 90% significance level are shown in dashed curves.

The AI approach prevents the coupled climate model's drift by starting the model close to its mean state. Here, the initial condition is a combination of the GECCO anomaly and the climatological state of the coupled model. An analogous method was implemented previously, e.g. by Smith et al. (2007), Pohlmann et al. (2009), Mochizuki et al. (2010) and Matei et al. (2012), who applied relaxation of the ocean's temperature and salinity to the analyses of the observed anomalies superimposed on the model's climate. For the

detrended spatially averaged North Atlantic SST, our approach shows predictive skill for 2 lead years, and 5 lead years for the North Atlantic SSH. By contrast, Pohlmann et al. (2009) reported predictive skill up to a decade for the eastern North Atlantic SST. However, the authors analyzed the non-detrended SST hindcasts. Therefore, a large fraction of their SST skill is related to the external forcing. One reason why the anomaly initialized SST hindcasts obtain skill lower than previously assessed is that the mean state was constructed from the last hundred years of the long spin-up run with constant CO₂ concentrations centered in the period that covers initialized experiments (1961-2010). A further reason is the low skill of the 20C run in simulating the observed global mean SST evolution during the last 50 years due to underrepresentation of the GHG forcing, and a lack of volcanic forcing or solar variations, which also indicates that the contribution to the skill due to the forced response is lower. In terms of the non-detrended AMOC, our AI results appear to be skillful between 15°N and 55°N up to 5-6 lead years. While Pohlmann et al. (2009) reported about the significant skill only for the higher latitudes, the skill obtained for a larger band of latitudes here might be due to the advantage of using the same ocean model for initialized hindcasts as was used for obtaining the GECCO ocean synthesis.

The FC approach corrects the SST and sea surface salinity fields towards the GECCO climatology. It improves the skill for SST by suppressing climate model drift, although not completely and an a posteriori correction of a drift is still required. For the detrended spatially averaged North Atlantic SST, our results show predictive skill up to 6 years and up to 9 years for the spatially averaged North Atlantic SSH. For the non-detrended AMOC between 15°N to 55°N the skill remains significant at lead time yr2-5 and yr6-9.

In summary, the FC initialized hindcasts yield the best predictive skill for SST and SSH over parts of the North Atlantic and the extratropical Southern Hemisphere. By contrast, FSI shows less persistent skill in the North Atlantic, and AI does not show high skill in the extratropical Southern Hemisphere. In these regions, the predictive skill is related to the ability of the initialized model to correctly represent the seasonal cycle of the mixed layer depth, especially during wintertime, and the FC approach appears to be most successful in this respect. As a consequence, hindcast quality statistics for the space-time averages for the North Atlantic shows that the FC method provides an improvement over AI and FSI, although still requires drift correction for the initialized hindcasts. In contrast to FSI and FC, the AI SST hindcasts appear to be more skillful in the tropics. In study by Magnusson et al. (2012a), for lead time yr2-5, all the initialization approaches showed some skill in the tropical Pacific with the highest skill in the AI experiments. Both detrended and non-detrended North Atlantic MOC initialized hindcasts show high correlation values in the first pentad (with some exceptions around the latitude of the Caribbean) while correlation remains significant in the following pentad for FSI and FC. A remaining drift in the predicted AMOC and SST anomalies in FC-HIND can possibly be further removed by applying an additional momentum flux correction or a 3-D flux correction.

Performing decadal hindcasts on more frequent initialization intervals is necessary for robust verification metrics (van Oldenborgh et al., 2012; Smith et al., 2013). Moreover, the

low number of ensemble members (in our case three, the minimum number of realizations suggested for the CMIP5 experimental setup; Taylor et al. (2012)) might also contribute to some variations of the statistics. The coupled model and the GECCO synthesis do not contain an ice component, which also could contribute to reduced model skill (Pohlmann et al., 2009).

Chapter 4

SSH Predictability and Mechanisms Associated with Decadal Predictability

4.1 Introduction

4.1.1 Motivation

In Sections 3.3 and 3.4 we identified patterns of high predictive skill for sea surface temperature (SST) and sea surface height (SSH). Similarities in patterns of skill for SST and SSH suggest common underlying processes. Griffies and Bryan (1997) speculated that in terms of potential predictability, SSH as a property that represents subsurface features of the ocean can be more skillful than SST, which is directly coupled to atmospheric variability. Our results also demonstrate longer predictability timescales for SSH, in contrast to SST. These findings motivate us to further investigate what are the possible mechanisms for high predictive skill of analyzed climate variables?

As was shown in Section 3.3.3, predictive skill for SST in regions of a deep winter mixed layer is related to re-emergence of SST anomalies from one winter to another. This damped oscillating processes yields predictability until winter SST anomalies are eventually mixed down into the deep ocean or damped back to the atmosphere (e.g., Frankignoul and Hasselmann, 1977; Alexander and Deser, 1994). For differently initialized SST hindcasts this mechanism provides predictive skill up to 2–4 winter seasons in the North Atlantic Ocean, with the flux corrected hindcasts having the longest skill. The timescales for re-emergence mechanism are consistent with findings of Deser et al. (2003).

This mechanism is based on persistence of ocean heat content (OHC) anomalies within mixed layer in winter-time and below mixed layer in highly stratified environment in summer-time. Because OHC changes affect SSH, persistence of OHC anomalies can also contribute to predictability of SSH changes. This Chapter is mainly focused on the analysis of SSH changes from different initialization experiments and tackles the following ques-

tions: what processes contribute to high predictive skill for SSH and why do differently initialized hindcasts lead to different duration of predictive skill for SSH?

4.1.2 Observed SSH changes and its contributions

Global mean sea level is considered as an indicator of climate change and has been seen as a major threat to low-lying coastal areas (e.g., Church et al., 2010).

The estimates of SSH changes are usually based on reconstructions from satellite altimetry (allowing SSH changes to be measured globally in parallel with the in situ measurements of thermosteric SSH since 1992) and tide gauges. Church et al. (2011) newly revised the estimates of different contributions to global mean sea-level rise, according to which the first largest contribution is the ocean thermal expansion (thermosteric SSH) with the value updated from about 23% of the observed global rate (IPCC AR4; Bindoff et al., 2007) to about 40%; the melting of glaciers and ice caps have been recognized as the second largest contribution (Fig. 4.1). The revised assessment includes a new estimate, the rate of sea-level changes due to groundwater depletion (Konikow, 2011). The uncertainties in the estimates of different contributions, which mainly arise from incompleteness of ocean observations and lack of data for the deep-ocean contribution, are still very large, for instance, in study by Church et al. (2011), the uncertainty in the global estimate of sea-level rise is 27%. The estimated halosteric (due to salinity changes) contribution seems to be relatively small, 1% to the globally averaged total SSH changes and 10% to the steric SSH changes (Antonov et al., 2002; Ishii et al., 2006). In addition to forced variations in the Earth's energy budget, SSH changes are

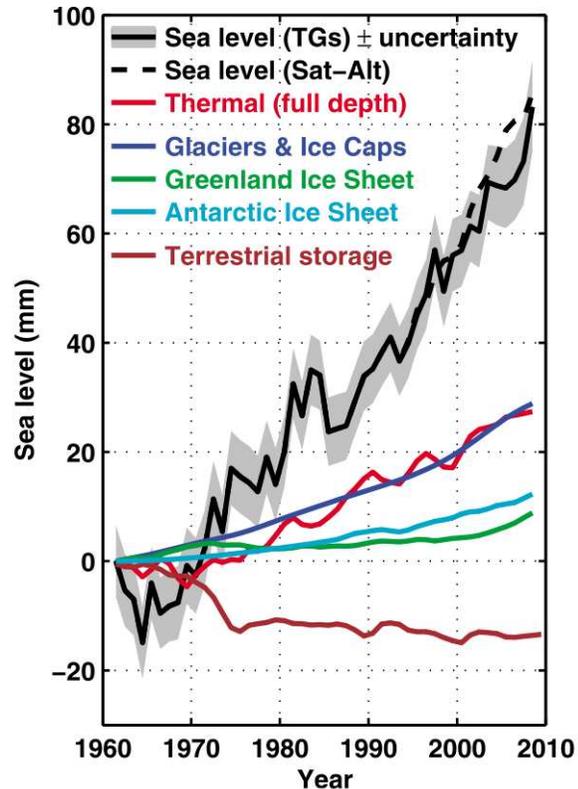


Figure 4.1— The global sea-level budget from 1961 to 2008. The observed sea-level using coastal and island tide gauges (solid black line with grey shading indicating the estimated uncertainty) and using TOPEX/Poseidon/Jason-1&2 satellite altimeter data (dashed black line). The two estimates have been matched at the start of the altimeter record in 1993. Adapted from Church et al. (2011).

also influenced by internal climate variability (Delworth and Knutson, 2000; Zhang and Church, 2012). While it is difficult to distinguish SSH forced response from internal variability, because of relatively sparse observations including short records of altimeter data, some studies indicate that natural fluctuations of global-scale SSH changes represent significantly smaller contribution to global SSH changes than the anthropogenic forcing of the Earth's energy budget (Huber and Knutti, 2011; Santer et al., 2011).

However, estimates of regional SSH changes can be very different from global ones because sea-level variability exhibits strong regional patterns, which are often associated with climate modes (Stammer et al., 2013). As was mentioned in Section 2.1.2, the UCLA OGCM does not allow studying other contributions than those related to the ocean dynamics. To point out the contributions which are specifically analyzed in this Chapter, we expand Fig. 2.1 that shows the major contributions to regional SSH changes (Fig. 4.2).

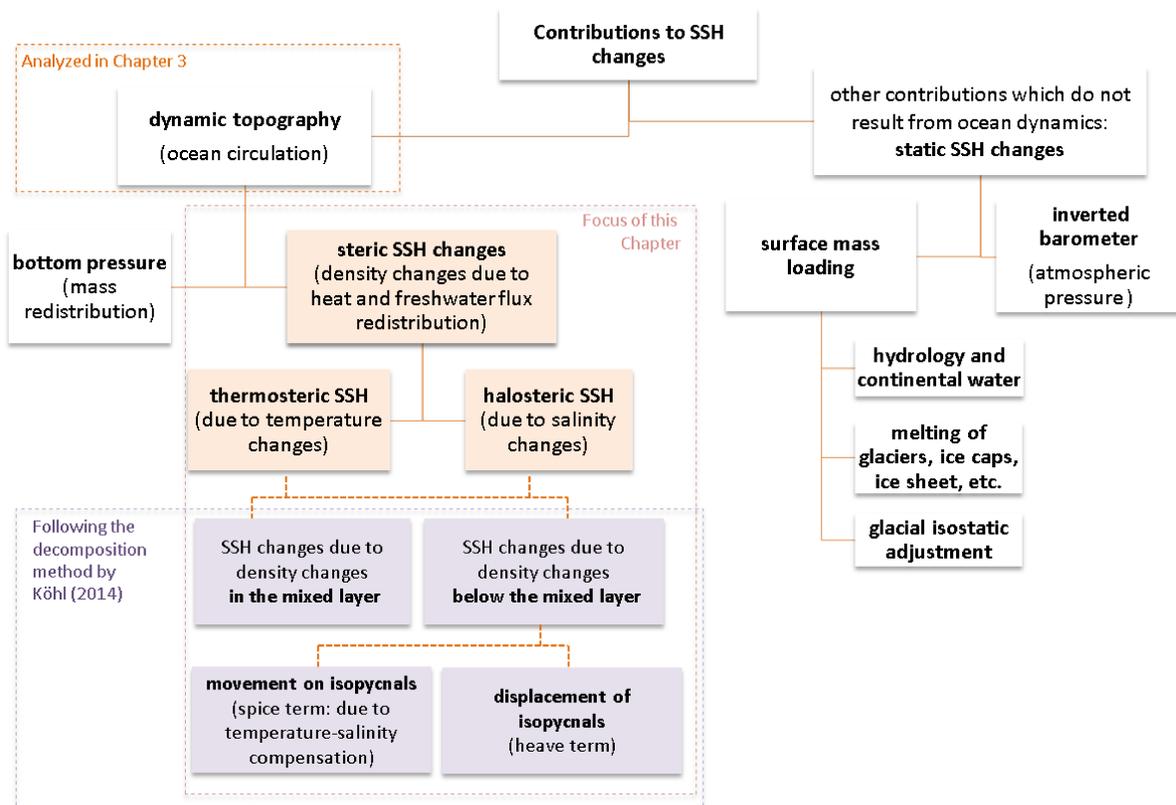


Figure 4.2— Contributions to SSH changes. The colored boxes represent the parameters which are covered in this Chapter. The decomposition of SSH changes into contributions related to density changes is discussed in Subsection 4.2.

Some modeling studies advocate that regional SSH changes are caused by density (baroclinic) changes in the ocean, with small contributions from changes in the barotropic circu-

lation at mid to high latitudes (e.g., Lowe and Gregory, 2006). Unlike in global estimates, halosteric contributions play important role for regional SSH variability. For instance, in the North Atlantic halosteric and thermosteric contributions can be nearly of the same magnitude with opposite sign (Lowe and Gregory, 2006; Landerer et al., 2007). Banks and Gregory (2006) related interior heat redistribution to the advection of temperature anomalies along isopycnals; this is also accompanied by changes in the ocean circulation (convection and high-latitude sinking). Therein, the North Atlantic is characterized by a vigorous overturning circulation accompanied by downward mixing of heat from the surface along isopycnals as well as by diapycnal diffusion, formation of the North Atlantic Deep Water and exchanges of water properties with the rest of the abyssal oceans. The local SSH changes due to mass redistribution occur in high latitudes (Vinogradova et al., 2007) and along the coastal areas as a result of bottom pressure increase across shallow shelf areas, following ocean thermal expansion (Landerer et al., 2007).

4.1.3 Regional SSH changes and SSH forcing factors

In the last decade much progress has been made in understanding the factors controlling regional SSH changes (Church et al., 2010; Stammer et al., 2013). Based on this, the following forcing factors and responses were identified:

- forcing: changes in surface wind stress and buoyancy,
 - mixed layer depth changes
 - * lateral and vertical mixing,
 - * local density changes (as warm/cold and fresh/salty water expands/contracts, SSH rises/falls);
 - advection of heat and salt
 - * Sverdrup response to Ekman pumping (e.g., water is pumped down into the ocean, which leads to a deepening of the thermocline and an increase in local SSH);
 - density changes
 - * wave-dynamics, like first baroclinic Rossby waves (SSH changes along Rossby waves propagation pathways);
- forcing: changes in (real) mass flux through freshwater flux (P–E+R).
 - river runoff,
 - sea ice and ice sheet melting.

In reality the forcing factors may perform in conjunction, as well as there are local and non-local forcing effects and response of the ocean through SSH changes can be a complex combination of different processes.

4.1.4 Some candidate mechanisms for decadal variability

Most of the recent studies on statistical assessment of forecast skill for decadal climate predictions are focused on identifying whether initialized predictions are able to capture observed decadal variability and if so, then in which regions and for how long? In order to gain confidence in predictive skill, the next step is to understand which physical processes give rise to high predictive skill across climate models. Below some candidate mechanisms, which might be relevant to predictive skill at decadal timescale, are reviewed.

In general, internal climate variability is characterized by a variety of timescales. The timescales of underlying processes define the relevance of different mechanisms to regional variability, and hence control the duration of predictive skill. On decadal timescales, climate modes, like North Atlantic Oscillation, Atlantic Multidecadal Oscillation, Pacific Decadal Oscillation, El-Niño Southern Oscillation and others, are widely considered as potential sources of predictability. Thought, there are still many unknowns, like what is triggering these phenomena and how do they interact together? For instance, Parker et al. (2007) stressed on possible combination of different influences that can cause nontrivial regional response patterns with enhanced or masked anthropogenic effect. Moreover, finding plausible physical mechanisms responsible for predictive skill is quite challenging because mechanisms of simulated decadal variability can differ across different climate models (Murphy et al., 2010).

Much effort to identify decadal variability mechanisms has been made in 1990s (for details, see Anderson and Willebrand, 1996), before climate modeling community was able to run large sets of initialized hindcasts and could verify them against available observations. In this respect, Dijkstra and Ghil (2005) review some mechanisms of low-frequency variability related to wind and buoyancy forcing. Thus, according to the candidate mechanism proposed by Weaver and Sarachik (1991) decadal variability is characterized by advection of salinity and temperature anomalies from the region of their origin, between the subtropical and subpolar gyres, to the eastern ocean boundary, where deep water is formed. The time needed to complete the advection path is estimated to be about eight years. The other mechanism introduced by Latif and Barnett (1996) is related to ocean adjustment with some lag to the change in wind stress curl. The ocean adjusts to this atmospheric forcing through changes in the strength of the subtropical gyre and northward heat advection by the western boundary current, this leads to SST changes which in turn can reinforce wind stress changes. Thus, adjustment allows continuous oscillations and might be predictable several years ahead.

A recent study by Weijer et al. (2013), addresses the role of Rossby basin modes in decadal variability of North Pacific Ocean. This involves long-wave basin modes that have a damping timescale longer than the basin-crossing time for the energy carried westward by the long Rossby waves to be returned to the eastern boundary of the North Pacific Ocean. Cessi and Louazel (2001) and Primeau (2002) found out that, for the certain shape of ocean basin, the modes decay rate is related to the partition of energy that can

be either dissipated in the western boundary layer or transmitted back to the eastern boundary through a forced gravity wave response. Though, it might be expected that at decadal timescales there will be no predictability associated with wind forcing, because of white spectrum of wind forcing at low frequencies, Capotondi and Alexander (2001) and Cessi and Louazel (2001) have shown some evidence for decadal variability, stating that when basin modes are coupled to the wind stress, they can become sustained instead of being damped.

In this Chapter 4 we investigate the origin of steric SSH changes (see Fig. 4.2) and their predictability in the context of different initialization strategies. Therefore, the structure of Chapter 4 is organized as follows:

Section 4.2 introduces a decomposition approach for steric SSH changes into the *mixed-layer term* (thermosteric and halosteric SSH changes due to temperature and salinity anomalies in the mixed layer, respectively), *heave term* (thermosteric and halosteric SSH changes due to displacements of isopycnals; from previous studies this term is believed to be governed by wind forcing) and *spice term* (associated with advection of density compensated temperature-salinity anomalies on surfaces of constant density).

In **Section 4.3** we investigate interannual variability of steric SSH changes and its contributions, which are derived from the GECCO verification dataset.

Section 4.4 contains an overview of regions with high predictive skill for thermosteric and halosteric components in different initialization experiments.

The skill for the mixed-layer term, heave term and spice term is discussed in **Section 4.5**. Finally, concluding remarks are presented in **Section 4.6**.

4.2 Steric SSH decomposition analysis

Regional steric SSH changes can be derived diagnostically (Gill and Niller, 1973). Firstly, we construct density from temperature and salinity time-varying fields using the UNESCO international equation of state (IES80) and, secondly, compute steric SSH changes, by integrating density anomalies over a column of sea water:

$$\Delta\eta = -\frac{1}{\rho_0} \int_{-H}^0 \left(\rho(T, S, z) - \overline{\rho(T, S, z)} \right) dz, \quad (4.1)$$

where $\Delta\eta$ is steric SSH changes, ρ is density as a function of potential temperature T , salinity S and depth z , ρ_0 is referenced to 1027 kg/m³, H is the ocean depth. The overbar denotes the time-mean value, density anomalies are computed with respect to the long-term mean over 1952-2001. For the initialized hindcasts the time mean is derived from the corresponding to the initialization technique assimilation run.

To distinguish the contributions of temperature and salinity anomalies to steric SSH changes, the time-varying temperature and salinity, respectively, are replaced by their time mean as follows:

$$\begin{aligned} \Delta\eta_T &= -\frac{1}{\rho_0} \int_{-H}^0 \left(\rho(T, \bar{S}, z) - \overline{\rho(T, \bar{S}, z)} \right) dz, \\ \Delta\eta_S &= -\frac{1}{\rho_0} \int_{-H}^0 \left(\rho(\bar{T}, S, z) - \overline{\rho(\bar{T}, S, z)} \right) dz, \end{aligned} \quad (4.2)$$

where $\Delta\eta_T$ and $\Delta\eta_S$ are the thermosteric and halosteric SSH changes, respectively, \bar{T} and \bar{S} are the time-mean values of the temperature and salinity fields, respectively. For the initialized hindcasts, time mean is calculated from the corresponding to the initialization technique assimilation run (over 1952-2001). Thus, the idea behind the thermosteric and halosteric SSH changes is: when water warms up, the density drops, consequently the volume of the water expands and the sea level goes up; changes in salinity occur as a result of added or removed amount of freshwater flux, thus freshening of the water column will result in density decrease and sea level rise.

Köhl (accepted, 2014) proposed a scheme for diagnostic of thermosteric and halosteric SSH changes, it aims to detect the origin of regional SSH changes. The concept of this approach is as follows: tracer anomalies subduct into the deep ocean along the surfaces of constant densities (isopycnals). Temperature and salinity anomalies that are compensated on a given isopycnal surface are referred to as spice anomalies (e.g., Munk, 1981; Schneider, 2000; Yeager and Large, 2004). These anomalies do not change density and are advected like passive tracers with the mean flow. While subducting, tracer anomalies induce vertical deviations of isopycnals. Vertical and lateral isopycnal displacements are termed as heave anomalies and are induced by wind forcing (Doney et al., 2007). Heave anomalies can also propagate from the place of origin through planetary wave dynamics (Schneider, 2000).

This approach uses an assumption that the influence of mixing processes is small. To satisfy this assumption, the contributions of thermosteric and halosteric SSH changes are computed above the maximum of the mixed layer depth (mixed-layer term); and below the maximum of the mixed layer depth they are decomposed into the SSH changes due to displacement of isopycnals (heave term) and due to movement along isopycnals (spice term).

Therefore, the decomposition is performed as follows: Steric SSH anomalies can be calculated as $\Delta\eta = -\frac{1}{\rho_0} \int_{-H}^0 \Delta\rho dz$. The displacement aligned to the time mean density gradient (heave term) can be calculated by multiplying the density anomalies $\Delta\rho$ by $\frac{\nabla\bar{\rho}}{|\nabla\bar{\rho}|} \frac{\nabla\bar{\rho}}{|\nabla\bar{\rho}|}$. To account separately for temperature and salinity perturbations, re-write the density gradient as $\nabla\bar{\rho} = -\alpha \nabla\bar{T} + \beta \nabla\bar{S}$, where $\alpha = -\partial\rho/\partial T$ and $\beta = \partial\rho/\partial S$ are the thermal expansion (“minus” sign) and haline contraction (“plus” sign) coefficients, respectively. To account for the adiabatic compressibility instead of potential density ρ , use neutral density γ . Then heave SSH contributions can be calculated as:

$$\eta_T^{heave*} = -\frac{1}{\rho_0} \int_{-H}^0 \alpha \frac{\nabla\bar{T} \nabla\bar{\gamma}}{\|\nabla\bar{\gamma}\|^2} \Delta\rho dz,$$

$$\eta_S^{heave*} = -\frac{1}{\rho_0} \int_{-H}^0 \beta \frac{\nabla\bar{S} \nabla\bar{\gamma}}{\|\nabla\bar{\gamma}\|^2} \Delta\rho dz.$$

Taking into account the assumption about small impact of mixing processes, the integral in Eq. (4.3) is split into two parts: below and above the maximum mixed layer depth (MLD). This gives “pure” heave term and the mixed-layer term, respectively.

Therefore, the “pure” heave term is calculated as:

$$\eta_T^{heave} = -\frac{1}{\rho_0} \int_{-H}^{-maxMLD} \alpha \frac{\nabla\bar{T} \nabla\bar{\gamma}}{\|\nabla\bar{\gamma}\|^2} \left(\rho(T, S, z) - \rho(\bar{T}, \bar{S}, z) \right) dz,$$

$$\eta_S^{heave} = -\frac{1}{\rho_0} \int_{-H}^{-maxMLD} \beta \frac{\nabla\bar{S} \nabla\bar{\gamma}}{\|\nabla\bar{\gamma}\|^2} \left(\rho(T, S, z) - \rho(\bar{T}, \bar{S}, z) \right) dz.$$
(4.3)

And the mixed-layer term is calculated as:

$$\eta_T^{MLD} = -\frac{1}{\rho_0} \int_{-maxMLD}^0 \left(\rho(T, S, z) - \rho(\bar{T}, \bar{S}, z) \right) dz,$$

$$\eta_S^{MLD} = -\frac{1}{\rho_0} \int_{-maxMLD}^0 \left(\rho(T, S, z) - \rho(\bar{T}, \bar{S}, z) \right) dz.$$
(4.4)

The spice term is then calculated as:

$$\eta_T^{spice} = -\eta_S^{spice}, \text{ therefore } \eta^{spice} = \eta_T - \eta_T^{heave} - \eta_T^{MLD}, \quad (4.5)$$

where $\nabla\bar{\gamma}$, $\nabla\bar{T}$ and $\nabla\bar{S}$ are gradients of neutral density, temperature and salinity, respectively, with $\nabla = \frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}$. Overbar denotes time-mean fields. η_T^{heave} and η_S^{heave} are the heave terms for thermosteric and halosteric SSH changes, respectively. The maximum MLD field is calculated as the largest monthly mean MLD field over the period 1952-2001. η_T^{MLD} and η_S^{MLD} are the mixed-layer terms for thermosteric and halosteric SSH changes, respectively. η_T^{spice} and η_S^{spice} are the spice terms for thermosteric and halosteric SSH changes, respectively, because they result from density compensating temperature-salinity anomalies, the spice term can be calculated either from thermosteric or halosteric SSH changes.

Computing the density gradients, one should pay attention to the choice of reference pressure (depth). To trace a water parcel from one place to another, the dependence of density on pressure should be removed. Using potential density for calculating density gradients does not account for the adiabatic compressibility or pressure dependency of the haline and thermal expansion coefficients. It is therefore in practice to use isopycnals with a reference pressure near the investigated depth range (Lynn and Reid, 1968) or neutral surfaces (McDougall, 1987). Neutral surfaces are surfaces along which particles can move without changing their potential energy. Small displacements of a water parcel on neutral surfaces at constant potential temperature and salinity result in the same density of water parcel as the surrounding water.

Spice contributions do not have a density signature and will not be seen in the total steric SSH changes, however any temperature or salinity anomalies affect thermosteric and halosteric SSH. Both the thermosteric and halosteric spice terms are always equal but opposite in sign. In the special case, given no movement of isopycnals, the thermosteric and halosteric SSH changes will be purely due to density compensating temperature-salinity anomalies (or spice contribution) and vice versa if water mass characteristics on isopycnals remain unchanged, the steric SSH changes will result only from the heave thermosteric and halosteric components.

4.3 Variability of steric SSH changes and its contributions from GECCO

To get an impression about amplitude and distribution of steric SSH variability and steric SSH contributions, we analyze interannual variability derived from the GECCO verification dataset. Information about patterns of SSH variability is also necessary in order to understand whether the initialized hindcasts yield any predictive skill in high variability regions. For this purpose, Fig. 4.3 demonstrates the GECCO SSH variability derived

over the period 1990-2001 for different contributions to steric SSH changes obtained using Eqs. (4.1) – (4.5). The regions of high steric SSH variability are evident in the tropical Pacific and Mid-Indian Basin, extratropical Southern Hemisphere and subpolar region of the North Atlantic Ocean, and are consistent with those previously reported by Köhl (accepted, 2014) for the GECCO2 product. Also Piecuch and Ponte (2011), analyzing origin of interannual variability for steric SSH changes over the period 1993-2004 in the ECCO ocean state estimate, concluded that high variability was mostly happening in the shallow tropics with an exception in the Southern Ocean and the subpolar North Atlantic. High SSH variability in the tropical Pacific Ocean is related to climate modes, like El-Niño Southern Oscillation in the Pacific Ocean and Indian Ocean, with the contribution from Indian Ocean zonal mode. The tropical oceans are dominated by dynamics of the equatorial current system.

Piecuch and Ponte (2011) suggested that 70% of SSH variability can be explained by advection processes associated with shallow wind-driven processes, and vertical advection in the extra-equatorial region driven by Ekman pumping. And up to 90% of variability can be explained with oceanic transports including advection and diffusive mixing. In addition to these contributions, Köhl (accepted, 2014) investigated the role of temperature-salinity compensation processes for interannual steric SSH variability for the period 1993-2011 and pointed out its importance in the Atlantic and Indian Oceans. Doney et al. (2007) suggest that, for the upper-ocean (0–400m), in the tropics and subtropics temperature variability results from heave term, and salinity variability is dominated by the spice term; for the North Atlantic both temperature and salinity variability is related to the heave term in the subtropics, and spice term in the subpolar region.

In summary, the patterns of GECCO SSH variability are in agreement with previously reported and imply that

- the total steric SSH variability pattern is dominated by thermosteric SSH variability, which has somewhat larger values than that of the total signal. This is related to the role of halosteric SSH changes, which compensate thermosteric SSH changes and thereby reduce the total signal (e.g., Köhl, accepted, 2014);
- the thermosteric SSH variability pattern is dominated by heave component in the tropical Pacific, Indian Ocean and Atlantic subtropical gyres; and the mixed-layer term in the subpolar and subtropical mode water regions (as defined by Talley, 1999), and along the equator in the Pacific Ocean;
- the halosteric SSH variability pattern is dominated to a great extent by the spice component with mixed-layer-related SSH variability in the subpolar region of the North Atlantic and Southern Ocean; the heave contributions are evident in the subpolar North Pacific.

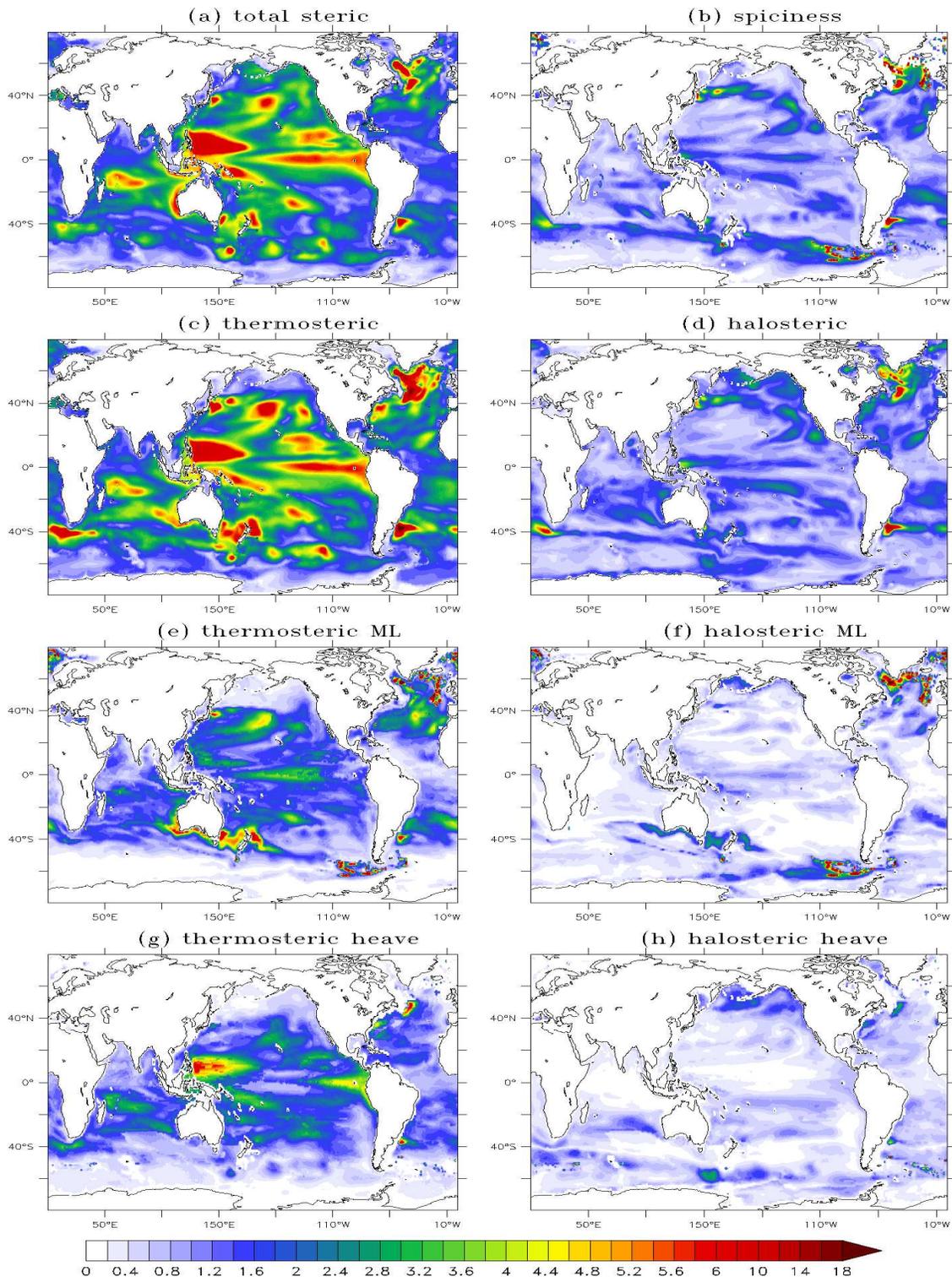


Figure 4.3— Interannual variability (1990-2001) of steric SSH (a, cm) and its components: thermosteric (c) and halosteric (d) SSH changes, both of which are further decomposed into contributions within the mixed layer (e, f) and below the mixed layer into spice (b) and heave terms (g, h). The spice thermosteric and halosteric SSH variability share the same patterns.

4.4 Predictability of thermosteric and halosteric SSH changes

Predictive skill for steric SSH and its thermosteric and halosteric contributions

This section gives an overview of regions with high predictive skill for steric, thermosteric and halosteric SSH changes. The correlation values in the first lead year are high for all initialization approaches due to closeness of hindcasts to GECCO initial state (not shown). The distribution of significant skill for steric SSH terms at lead time yr2-5 is shown in Fig. 4.4. The skill in different colors refers to: significant correlation (Sec. 2.4.5) between hindcasts and GECCO (in orange color) with overlaid significant correlation between hindcasts and GECCO in the regions of substantial interannual variability derived from GECCO (in green color) and significant skill of the persistence forecasts (in red color).

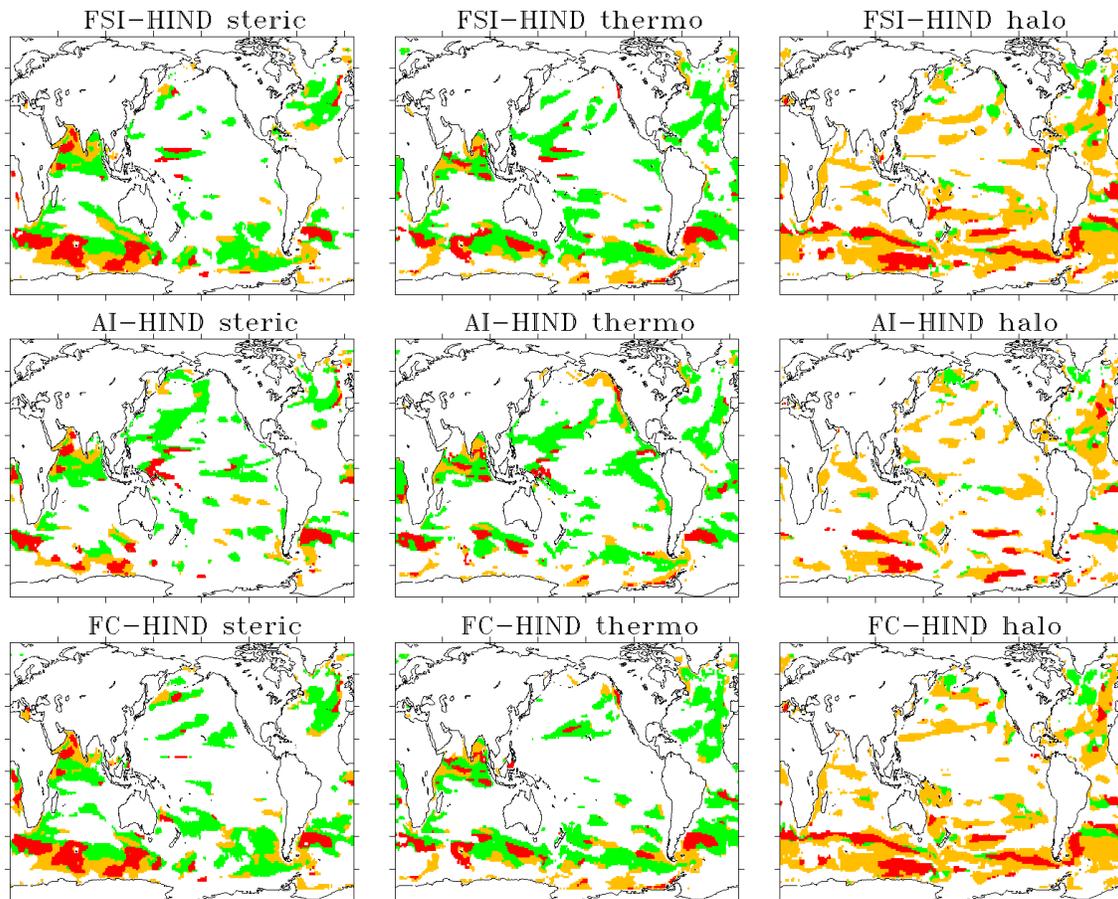


Figure 4.4— Significant COR skill for steric, thermosteric and halosteric SSH changes computed between detrended GECCO and initialized hindcasts: FSI-HIND (left), AI-HIND (middle) and FC-HIND (right) for lead time yr2-5. Orange area stands for significant COR with no interannual variability, green for significant COR over the regions with $\text{STD} > 1.5$ cm, and red for significant COR from persistence.

The patterns of skill for steric SSH changes resemble those for SSH changes (dynamic topography analyzed in Chapter 3, Fig. 3.9) in the Indian Ocean, western tropical Pacific, Southern Ocean and the North Atlantic. FSI-HIND and FC-HIND demonstrate very similar correlation distribution in these regions.

The correspondence in patterns of skill for steric SSH changes and its contributions indicate where predictability of steric SSH changes might come from. Hence, high correlation values in the Indian Ocean (for all the initialized hindcasts) and western tropical Pacific (for FSI-HIND and AI-HIND) to a large extent relate to the thermosteric contribution (green color). Overall, the amplitude of the halosteric SSH signal and its interannual variability (Fig. 4.3) are smaller than those for thermosteric SSH, therefore, the skill for the halosteric term is mostly shown in orange color, except for high latitudes, namely, in the northern North Pacific and subpolar North Atlantic. In the eastern North Atlantic and the extratropical Southern Hemisphere the correlation values from both contributions are significant but mostly the thermosteric term signifies “useful” skill. By “useful” skill we mean significant skill in the regions with high interannual variability (green color).

In the tropical North Atlantic (for all the initialized hindcasts) and in some areas of the eastern North Pacific (thermosteric terms of AI-HIND and FC-HIND), decomposed components show the skill which is not present in the total steric signal. This might be related to compensated temperature-salinity anomalies (spice term) that have no effect on steric SSH.

In terms of the performance of different initialization strategies, in contrast to FSI-HIND and FC-HIND, AI-HIND demonstrates higher correlation values in the western tropical Pacific, which are mostly coming from the thermosteric contribution, and lower skill in the eastern North Atlantic and Southern Ocean. The persistence skill is mostly shown in the Southern Ocean and Indian Ocean (red color).

To sum up, a qualitative comparison of skill patterns for the initialized SSH changes (Fig. 3.9) and its steric contribution suggests that the SSH skill is mainly dominated by the steric SSH changes, which in some regions appear to be predictable on decadal timescale, and patterns of predictive skill have more regions of significant correlation values than the persistence forecasts. The remaining contribution to SSH changes from bottom pressure (mass redistribution) and associated predictive skill was not investigated. Both thermosteric and halosteric contributions demonstrate large-scale patterns of high predictive skill, whereas thermosteric term appears to be skillful in the regions where interannual variability occurs.

Predicting SSH changes knowing its contributions

To get an idea about the amplitude of different SSH contributions, Fig. 4.5 demonstrates an example of the time series for SSH changes (dynamic topography), steric SSH changes and its temperature and salinity contributions averaged over the North Atlantic (NA).

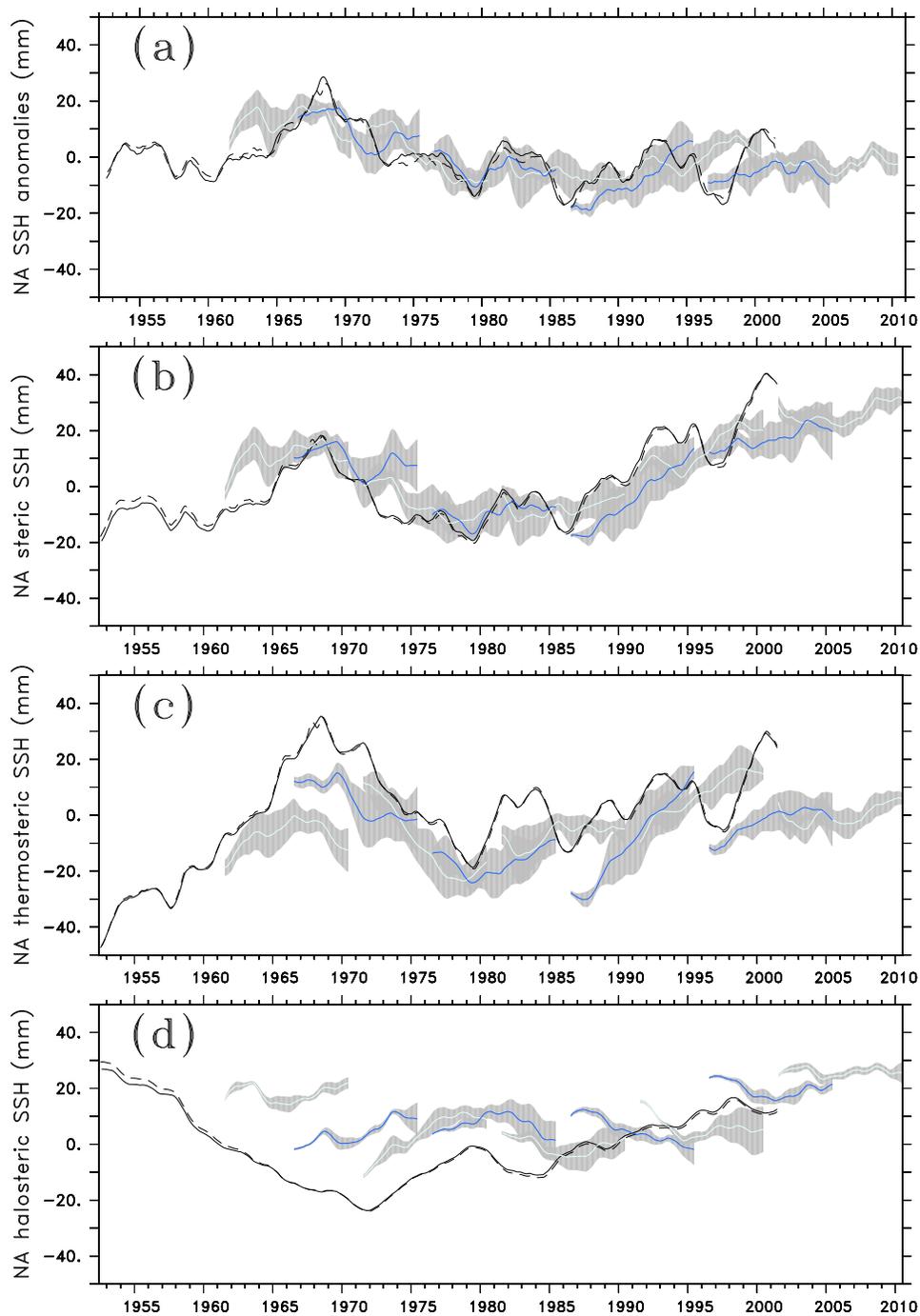


Figure 4.5— Time series of the North Atlantic (50°W-10°W, 20°N-60°N) SSH anomalies (dynamic topography; a), steric (b), thermosteric (c) and halosteric contributions (d) from GECCO (solid black), assimilation run (dashed black) and bias-corrected initialized hindcasts FC-HIND (blue). Grey shading implies ensemble spread.

When comparing interannual variability of GECCO SSH changes with that of steric SSH changes a clear resemblance is evident. The amplitude of the GECCO thermosteric and halosteric terms suggests the equal importance of these two contributions in the North Atlantic Ocean, at the same time, the interannual variability is larger for the thermosteric contribution. These two contributions also oppose each other, with warm/salty and cold/fresh combinations on top of a warming/freshening trend since 1980s. The amplitude of the total steric SSH signal is smaller than for its individual components. The FC-HIND SSH changes and steric SSH changes follow closely their GECCO counterparts and are in good agreement with each other. For the thermosteric and halosteric SSH changes more discrepancies are found between FC-HIND and GECCO, and the halosteric hindcasts tend not to vary much with time.

The correlation coefficients between the initialized steric SSH changes and GECCO SSH changes (dynamic topography); and between the initialized thermosteric SSH changes and GECCO SSH changes (dynamic topography) are shown in Fig. 4.6 a. The initialized steric SSH skillfully predict GECCO SSH changes, up to yr2-5 for FSI-HIND and AI-HIND and up to yr6-9 for FC-HIND. These findings are consistent with the results for SSH skill analyzed in Chapter 3 (Fig. 3.11), indicating that contribution from bottom pressure to SSH changes in this region is negligible. By contrast, Fig. 4.6 b shows that when thermosteric SSH changes are used to predict GECCO SSH changes, the FC-HIND are no longer skillful at lead time yr6-9, also for AI-HIND the skill is reduced at lead time yr2-5.

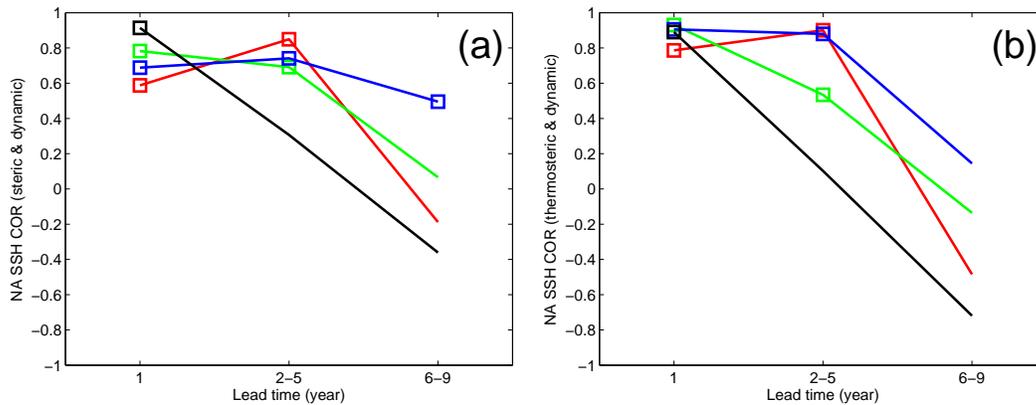


Figure 4.6— Anomaly correlation between (a) steric SSH changes (from the initialized hindcasts) and GECCO SSH changes (dynamic topography), and between (b) thermosteric SSH changes and GECCO SSH changes. The North Atlantic SSH anomalies are averaged over 50°W - 10°W , 20°N - 60°N and derived from FSI-HIND (red), AI-HIND (green) FC-HIND (blue) and the persistence forecast (black). Squares imply significant correlation at the 90% confidence interval estimated with a bootstrap procedure. Skill is computed for the SSH time series with the long-term trend taken out.

The correlation skill for thermosteric SSH changes between the initialized hindcasts and GECCO drops faster than the skill for steric SSH changes (not shown); this is already expected from Fig. 4.5 c, d. These results indicate that the steric contribution appears to

be more predictable than its individual components. The spice term, which results from temperature-salinity compensation processes, does not affect steric SSH changes but is present in the thermosteric and halosteric SSH changes, and might be a reason for the reduced skill.

4.5 Predictability of SSH heave, spice and mixed-layer terms

4.5.1 Predictive skill in different initialization experiments

To track further indications of processes which provide predictability of steric SSH changes, the patterns of correlation skill for the thermosteric and halosteric contributions at lead time yr2-5 and 6-9yr are analyzed. Figs. 4.7 and 4.8 show only significant correlation (Sec. 2.4.5) in orange color with overlaid skill from the persistence forecast in red color. In green color those regions are shown where the model has skill and where the interannual variability exists, therefore implying “useful” skill from the initialization. An interesting result is that the regions which seemed to have predictive skill for thermosteric SSH (Fig. 4.4) show skill due to persistence for single contributions (e.g., Indian Ocean). And regions with “useful” skill for thermosteric SSH now show skill in low variability regions, suggesting that first, the amplitude of individual contributions is smaller than that of thermosteric signal and second, the contributions add together rather than cancel each other. Hence, “useful” skill is shown by FSI-HIND and FC-HIND for the mixed-layer term over the subtropical North Atlantic, by all the hindcasts for the thermosteric heave term within the latitudinal bands 30°N – 40°N and 30°S – 40°S in the Pacific Ocean and for the spice term over the subpolar North Atlantic. In addition, the thermosteric heave term from AI-HIND shows high correlation over the eastern equatorial Pacific.

The skill for the mixed-layer term is evident in the Southern Ocean and the subtropical Atlantic (up to yr6-9). We expected that the mixed-layer term would have skill due to persistence of heat content anomalies in the winter mixed layer in first lead years (this mechanism has a connection to SST re-emergence mechanism described in Section 3.3.3; Fig. 4.9). Thermal anomalies from previous winter persist below the shallow summer mixed layer and re-entrain into the mixed layer during the following winter. The annual mixed-layer term should incorporate the density changes due to these thermal anomalies. This mechanism is related to persistence skill in the Southern Ocean and North Atlantic. Hall and Manabe (1997) found an indication for persistence of salinity anomalies in the winter mixed layer. An indication of this mechanism was investigated by Alexander and Deser (1994) using observations in the North Atlantic and North Pacific, Deser et al. (2003) using model and observations, Ciasto and Thompson (2009) using observations in the western extratropical South Pacific and others.

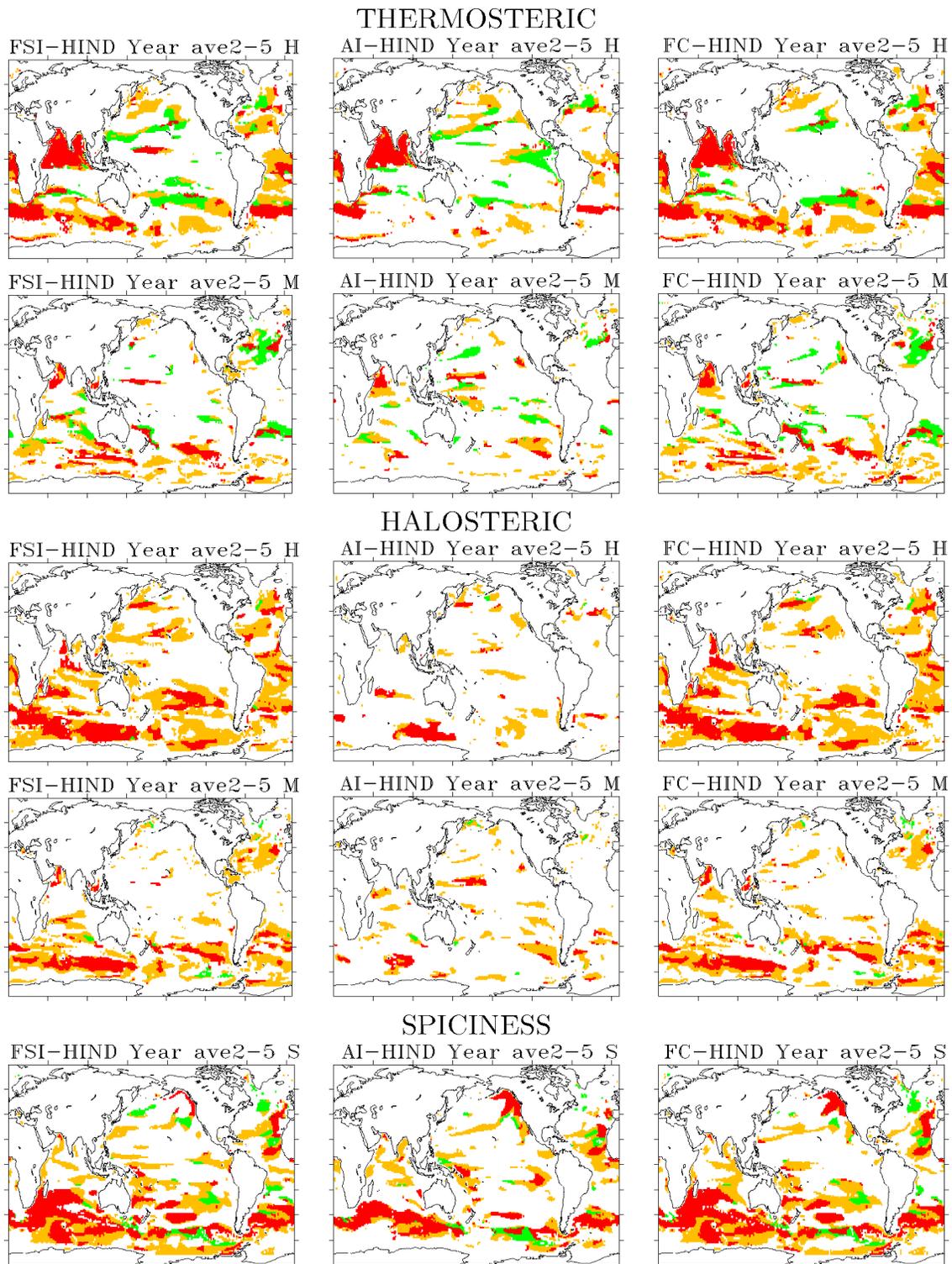


Figure 4.7— Significant COR skill for different thermosteric and halosteric SSH terms computed between detrended GECCO and initialized hindcasts: FSI-HIND (left), AI-HIND (middle) and FC-HIND (right) for lead time yr2-5. Orange area stands for significant COR, green for significant COR over the regions with $STD > 1.5$ cm, and red for significant COR from persistence. M stands for the mixed-layer term, S spice term and H heave term.

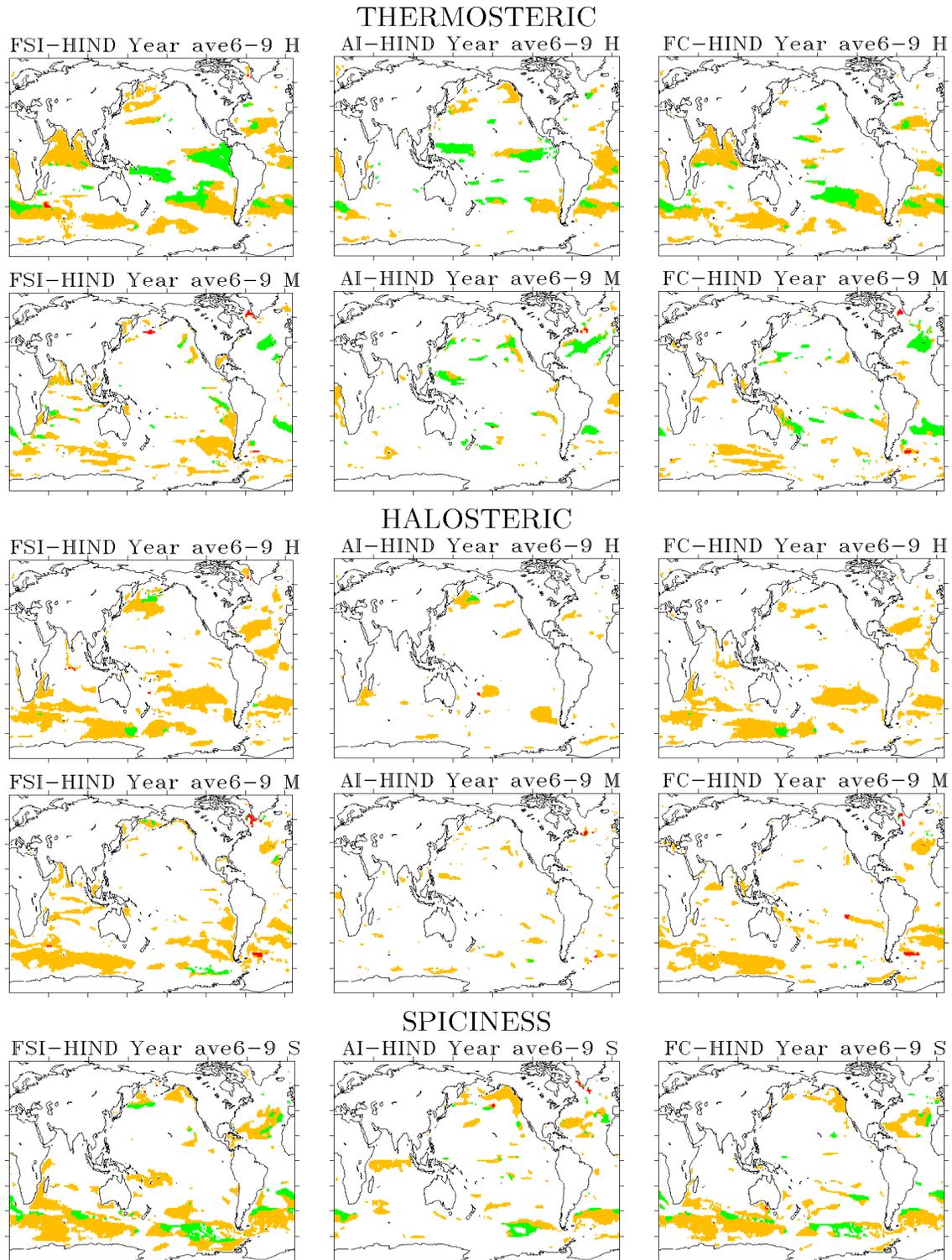


Figure 4.8— Significant COR skill for different thermosteric and halosteric SSH terms computed between detrended GECCO and initialized hindcasts: FSI-HIND (left), AI-HIND (middle) and FC-HIND (right) for lead time yr6-9. Orange area stands for significant COR, green for significant COR over the regions with $STD > 1.5$ cm, and red for significant COR from persistence. M stands for the mixed-layer term, S spice term and H heave term.

Whereas in the Southern Ocean the pattern of skill for thermosteric mixed-layer term is more noisy than for halosteric, in the North Atlantic both halosteric and thermosteric SSH changes have quite similar patterns. The correlation patterns indicate that persistence of heat content anomalies is not the only contribution to SSH skill in the subtropical Atlantic (small regions with persistence skill and large regions with “useful” skill).

Over the North Atlantic, in addition to mixed-layer term, the spice and heave terms show “useful” correlation with GECCO (up to yr2-5). The patterns of high skill for spice terms, in different initialization experiments, are evident over the eastern North Atlantic, subtropical South Atlantic, eastern North Pacific, and Southern Ocean. For AI-HIND, the spice and thermosteric heave terms appear to be the largest contributors to SSH skill.

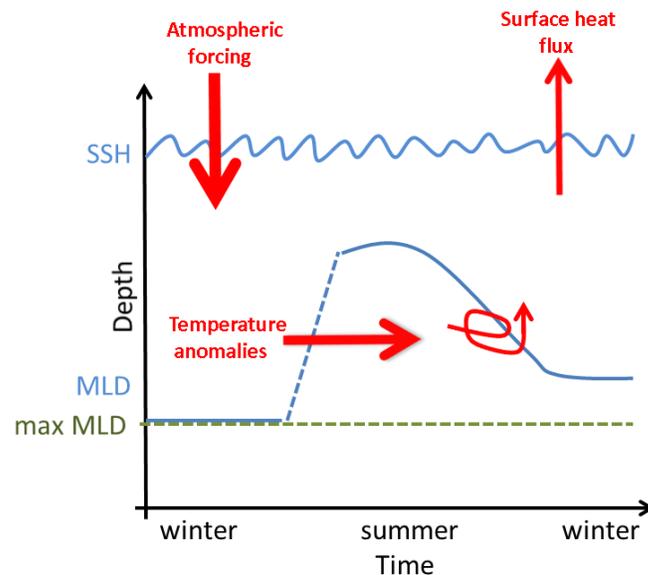


Figure 4.9— Scheme of SST re-emergence mechanism (adapted from Deser et al., 2003). Temperature anomalies in the ocean mixed layer result from atmospheric forcing and damp back to the atmosphere; temperature anomalies created in winter persist below the summer MLD and re-entrain into the MLD the following winter. The mixed layer depth (MLD) is a function of seasonal cycle, and the maximum MLD is used in the decomposition analysis (Sec. 4.2).

The skill due to persistence is present in the thermosteric heave term (in all the initialized experiments) over the Indian Ocean, subtropical South Atlantic and in some regions over the extratropical Southern Hemisphere. Surprisingly, there appears to be almost no significant skill beyond lead time yr2-5 due to persistence. The persistence skill in the extratropical Southern Hemisphere is apparent for all the terms related to halosteric SSH contribution. Now we review the ocean basins which show high predictive skill and possible mechanisms in those regions for future analysis.

4.5.2 Indication of mechanisms for predictability

Atlantic Ocean

The pattern of SSH skill in the Atlantic Ocean (Fig. 4.7, 4.8) is rather complex and represents a combination of different signals. In the North Atlantic north of 20°N the skill (beyond persistence) is dominated by signal from the mixed-layer and heave terms, while in the tropical North Atlantic the skill from spice term is evident. The later term contributes to the thermosteric and halosteric skill but cannot contribute to the skill from total steric SSH changes (Fig. 4.4). In the South Atlantic, predictive skill comes mostly from the mixed-layer term and persistence skill is evident in the heave term of thermosteric and halosteric SSH changes in the tropical Atlantic and spice in latitudinal band $30^{\circ}\text{S} - 50^{\circ}\text{S}$.

The large-scale patterns of skill are attributed to the subtropical mode water regions identified by Talley (1999). As was mentioned in Section 4.1, some fraction of the SSH skill in the North Atlantic might be attributed to the seasonal variations of thermal anomalies stored in the deep winter mixed layer (Fig. 4.9), though such a pronounced seasonal feature, as was present in the skill for the NA SST, was not observed in the skill for the dynamic NA SSH signal. Also in the study by Deser et al. (2003), where the re-emergence mechanism was studied with respect to SST and OHC changes, the skill for OHC demonstrated rather monotonic decay. This is expected, as OHC represents integrated subsurface temperature changes.

The skill for the decomposed terms in the earlier reviewed region over the North Atlantic (Fig. 3.11, where FC-HIND demonstrated the longest duration of predictive skill for SSH changes) is shown in Fig. 4.10. The FSI-HIND skill bears similarities to that of FC-HIND. The thermosteric and halosteric heave terms appear to be skillful up to 7 years. In the FC-HIND, the thermosteric mixed-layer term remains skillful up to 8 years, while the halosteric mixed-layer term contributes to the skill in the first four years. The spice term (advection on isopycnals) in both FSI-HIND and FC-HIND is skillful up to 4 years and comes from the persistence. In the persistence forecast, the skill is decreasing more rapidly for the heave terms and the halosteric mixed-layer term, while more gradually for the spice term and thermosteric mixed-layer term. AI-HIND shows skill for the thermosteric mixed-layer term, whereas the skill for the advection term becomes significant in later years (6-9yrs).

To summarize, for the North Atlantic predictability of steric SSH changes and underlying mechanisms are related to

- the thermosteric and halosteric mixed-layer terms. In the first lead years, predictive skill is associated with persistence of heat content anomalies in winter mixed layer. The other mechanism might be advection of heat toward regions of high predictive skill by the Atlantic meridional overturning circulation (AMOC). High predictive skill for AMOC up to a decade in the FSI and FC experiments supports this idea

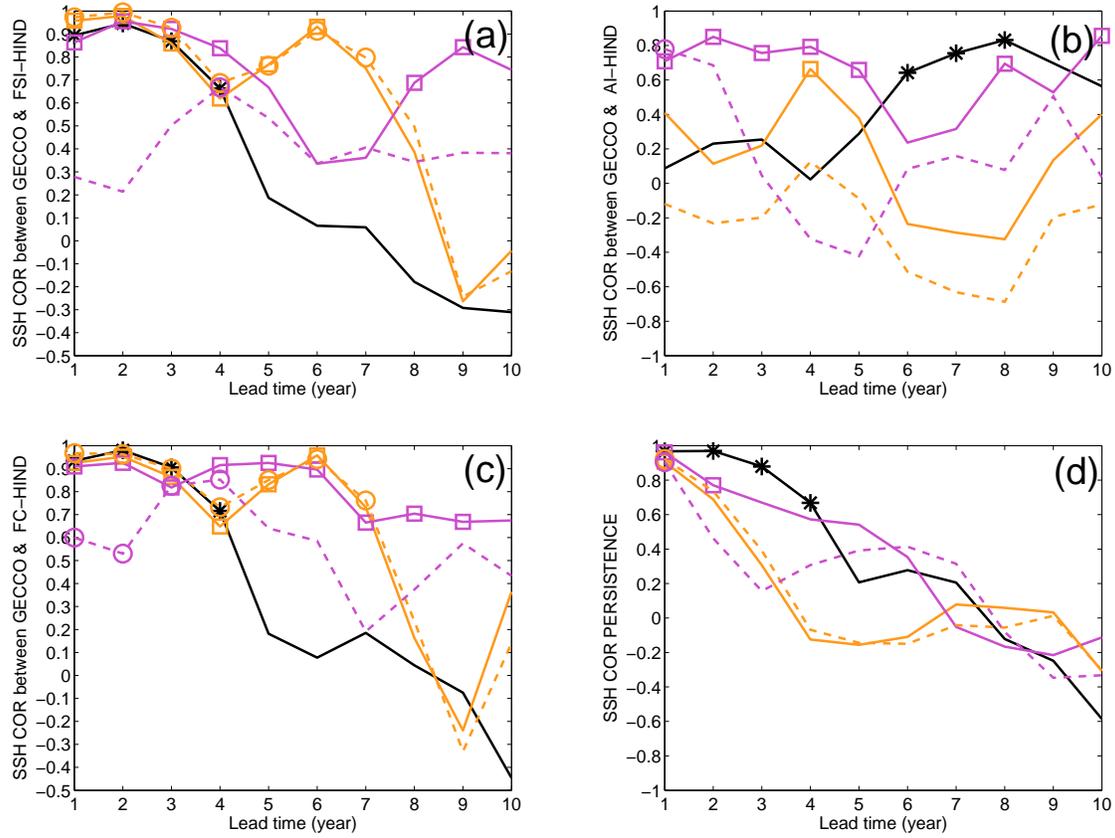


Figure 4.10— COR between different contributions to the NA SSH changes from GECCO and initialized hindcasts: FSI-HIND (a), AI-HIND (b), FC-HIND (c) and persistence (d). The spice term is shown in black, heave term in orange and mixed-layer term in magenta. Solid lines represent thermosteric terms and dashed lines – halosteric terms. Symbols imply significant correlation values at 95% level (Goddard et al., 2013). The COR is computed for detrended data.

(Fig. 3.13 and 3.14). The mechanisms for decadal predictability that involve AMOC were reviewed by e.g., Latif and Keenlyside (2011).

- the thermosteric and halosteric heave terms (heave anomalies induced by persistent wind stress forcing, propagate from the place of origin as baroclinic Rossby waves). Previous studies on processes contributing to SSH changes suggest that the heave term is mostly occur as adjustment to wind forcing through the Rossby waves. As in high latitudes Rossby waves are much slower than in low latitudes (Yang et al., 2003), these propagating features might provide predictability for SSH at decadal timescale. The mechanism of low-frequency Rossby waves forced by wind was earlier investigated by Sturges et al. (1998) for the subtropical North Atlantic. Also Cabanes et al. (2006), investigating mechanisms for interannual variability of SSH changes in the Atlantic Ocean from altimeter data, identified the following processes: local steric SSH changes due to surface buoyancy forcing; local response to

wind stress through Ekman pumping; baroclinic and barotropic oceanic adjustment via propagating Rossby waves and Sverdrup balance.

- the spice term is often generated from the mixed layer, therefore might share the same mechanism as the mixed-layer term. Advection of temperature and salinity anomalies by the mean currents in the North Atlantic subtropical gyre was earlier studied by Laurian et al. (2009).

Understanding whether, in the current experimental setup, the candidate mechanisms of heat advection and Rossby waves provide predictability for steric SSH changes is left for future work. Simple models can be used to isolate some of the physical mechanisms, e.g., wind forcing model as applied by Sturges et al. (1998), Qiu and Chen (2006) or Cabanes et al. (2006), and to analyze their impact on spatial and temporal skill distribution. Correlation skill is computed as a function of lead time, but in order to analyze propagating structures, like Rossby waves, other diagnostics (e.g., complex empirical orthogonal functions, Hannachi et al., 2007) applied to individual starting dates is needed.

Pacific Ocean

In the North Pacific, SSH skill is dominated by two contributions, the spice term (persistence skill) along the western coast of the North America and the heave term (“useful” skill) in the central (30°N–40°N) North Pacific. In the eastern equatorial Pacific the skill is dominated by the thermosteric heave term (for AI-HIND).

The study by Latif and Barnett (1996) indicate that in the North Pacific thermal anomaly variability is attributed partly to Rossby wave adjustment and advection from subtropical mixed layer around the subtropical gyre. Weijer et al. (2013) also proposed that decadal variability in the subpolar North Pacific might result from the pressure (density) adjustment to Rossby wave fronts coming from the western boundary. Capotondi and Alexander (2001) found that for a latitudinal band 10°N–15°N long baroclinic Rossby waves with propagation time about 7–10yr might be responsible for decadal variability.

Example of propagating SSH anomalies from the GECCO dataset and FC-HIND along 40°S in the South Pacific are shown in Fig. 4.11. The SSH anomalies propagate to the west of the basin and can be damped or exited by wind forcing. FC-HIND show resemblance of these propagating features with GECCO in the first lead years. As the wind forcing differs in GECCO and FC-HIND the propagating feature can develop differently. Further analysis is needed to test Rossby-wave mechanism in the current experimental setup.

The Southern Ocean

Additionally to the mixed-layer term contribution, some fraction of the skill also comes from the halosteric spice and heave terms. Vivier et al. (2005) suggest that the dominant dynamical response in the Southern Ocean is barotropic, which results from wind forcing

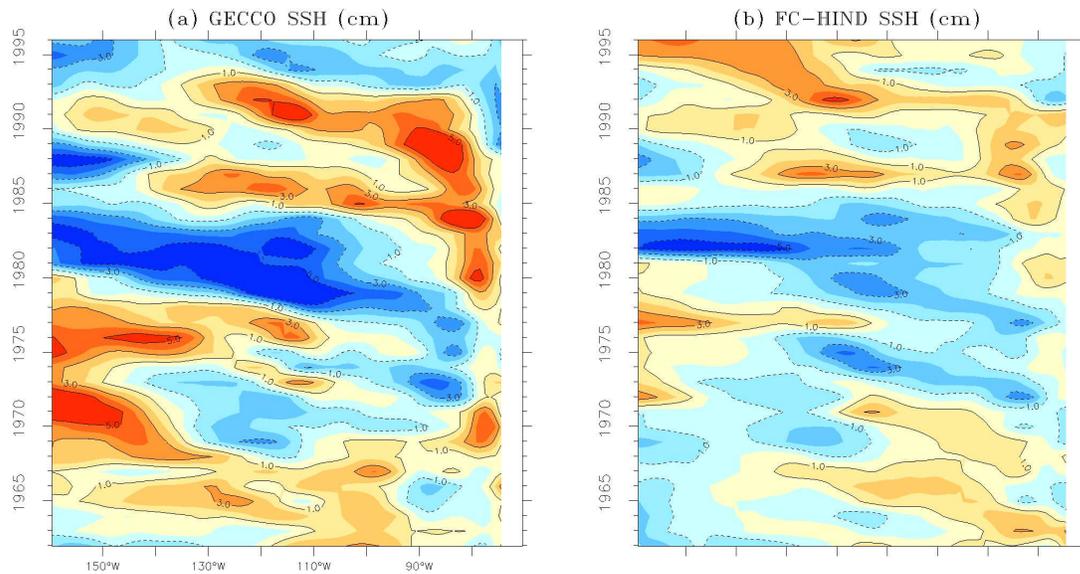


Figure 4.11— Time-longitude (Hovmöller) diagram for the annual SSH anomalies from GECCO (left) and FC-HIND (right) over 1961-1995 at latitude 40°S . For the FC-HIND diagram the first five-year periods of the initialized hindcasts were merged, i.e., 1961-1965, 1966-1970, ..., 1991-1995.

adjustment. The skill over the Southern Ocean in the initialized hindcasts comes to a great extent from the persistence. The planetary wave mechanism suggests that baroclinic Rossby waves become progressively slow toward higher latitudes (about 50yrs, estimated by Qiu and Chen, 2006), this also supports persistence skill at decadal timescale in the Southern Ocean.

Indian Ocean

Patterns of high predictive skill for SSH changes (Fig. 3.9) up to yr2-5 are clearly related to the steric SSH contribution (Fig. 4.4). The analysis of decomposed terms revealed that the skill in the band of latitudes from 20°N to 10°S is mostly dominated by the thermosteric heave term and comes from the persistence. In the Arabian Sea, skill for mixed-layer term is evident (mostly persistence), in addition, the equatorial region shows some skill due to temperature and salinity compensation processes.

4.6 Concluding remarks

The analysis performed in this Chapter aims to detect indications for mechanisms of predictability at decadal timescale and, in particular, addresses the question on the level of predictability of steric (density-related) SSH changes and its components. The contributions to SSH changes, which are caused by temperature- and salinity-induced density changes, can result from vertical displacement of isopycnals caused by dynamical processes such as the response to variations in wind stress curl and through planetary waves; from

mixing processes as the response to wind and buoyancy forcing and from density compensating temperature-salinity anomalies. This may strongly affect dynamic SSH changes in the model owing to the fact that a large fraction of SSH variability is related to steric SSH contribution. Thermosteric and halosteric components of SSH changes that are compensated with no change in total density structure are not detected in total steric SSH changes. In this case temperature and salinity anomalies are advected from the place of origin along isopycnals as passive tracers. Advection and mixing processes are responsible for heat and salt transport into the deep-ocean. Spice anomalies can also be important for teleconnections. More specifically, propagating along mean geostrophic streamlines and appearing at the ocean surface, spice anomalies may interact with atmosphere and not being density-compensated anymore, but thereby cause remote density anomalies (Laurian et al., 2009). In this way they can contribute to variability of ocean circulation. Köhl (accepted, 2014) suggests that advection on isopycnals plays an important role in the variability of SSH in the subpolar regions, also the halosteric SSH changes appear to have the largest impact from this term, while the mixed-layer term is important in mode water regions. When temperature or salinity changes are not compensated, this creates density anomalies and leads to exchange of properties in the ocean. Related to this thermosteric heave component is believed to be a dominant part of the tropical SSH variability (Doney et al., 2007; Piecuch and Ponte, 2011; Köhl, accepted, 2014).

We analyzed the patterns of GECCO SSH interannual variability and found out that the thermosteric SSH variability pattern is dominated by the heave term in the tropical Pacific, Indian Ocean and Atlantic subtropical gyres; by the mixed-layer term in the subpolar and subtropical mode water regions, and along the equator in the Pacific Ocean. The halosteric SSH variability pattern is dominated to a great extent by the spice term. Also there is a contribution from the mixed-layer term in the subpolar region over the North Atlantic and Southern Ocean; and heave term in the subpolar North Pacific.

We then analyzed SSH skill associated with thermosteric and halosteric SSH changes. The term “useful” skill was introduced to study whether the skill of the initialized hindcasts have significant values in the regions of high interannual variability. A lot of similarities in the “useful” skill for steric SSH changes are related to the thermosteric term, also halosteric term appeared to have “useful” predictive skill in the Atlantic Ocean. We continued the analysis by decomposing the thermosteric and halosteric SSH changes into the heave, spice and mixed-layer contributions. In general, FSI-HIND and FC-HIND showed quite similar distribution of skill. The correlation values in the tropics, in particular in the Indian Ocean (for all three initialized experiments) and in the Pacific Ocean (for AI-HIND), comes mostly from the thermosteric heave term, which is also evident in the persistence forecast. In the extratropical Southern Hemisphere, the skill for the halosteric terms is due to persistence as well.

Predictive skill analysis for the decomposed terms over the North Atlantic showed that the mixed-layer term is one of the dominant contributions to predictive skill for SSH changes and appears to be better captured by FC-HIND. Predictability of this term is

associated with persistence of heat content anomalies in the winter mixed layer up to yr2-5. While results showed that this is not the only mechanism in the North Atlantic region and advection of heat toward regions of high predictive skill by AMOC is hypothesized. In AI-HIND, the mixed-layer term in the North Atlantic seems to be better represented signal than other contributions. Robson (2010) found out that anomaly initialization may suffer from the problem of imbalanced initial conditions as this approach assumes that using observed anomalies of temperature and salinity with model climatology will produce the same anomalies of density in the model. Robson demonstrated that this linearity assumption caused errors especially in the regions sensitive to density anomalies. This might be the reason for low skill of the steric SSH changes in the AI-HIND here. Though “useful” predictive skill comes from the mixed-layer term to a large extent, the North Atlantic is dominated by a combination of different processes. Spice term shows high skill in the first four years but not beyond the persistence skill. The heave term is the second long-lasting contribution to SSH skill in FSI-HIND and FC-HIND, after the mixed layer term.

Previous studies associated the heave term with wind forcing, and due to white spectrum of wind stress, the skill for the heave term is expected to vanish eventually. The correlation analysis showed that this term in some regions possess predictability at long time scales, e.g. around 30°S–40°S and 30°N–40°N in the Pacific Ocean up to yr2-5. The candidate mechanism for predictability in these regions is hypothesized to baroclinic Rossby waves. Though, additional analysis would be needed to better understand the mechanisms for the skill of the heave term.

Chapter 5

Conclusions and outlook

The primary objective of this study was to investigate the performance of three different initialization strategies for decadal climate predictions using the UCLA/MITgcm coupled ocean-atmosphere model. We mainly focused on estimating predictive skill for sea surface temperature (SST), sea surface height (SSH) and Atlantic meridional overturning circulation (AMOC). We attempted to gain insight into how predictive skill is distributed in time and space, in order to identify the regions where initialization is able to improve predictions and to further investigate possible mechanisms giving rise to high predictive skill.

In the following we will provide an overview of the main results and the main conclusions in Section 5.1. This section also includes the evaluation of sensitivity of the results to different metrics, bias and trend removal approaches. Recommendations for further work are given in Section 5.2.

5.1 Thesis overview and main conclusions

5.1.1 Improving decadal predictions through different initialization procedures

In this thesis, we test three initialization schemes for decadal predictions, namely full state initialization, anomaly initialization and flux correction. Full state initialization and anomaly initialization schemes were originally designed for seasonal-to-interannual forecasts (e.g. Stockdale, 1997; Schneider et al., 1999). In terms of predictive skill for SST and SSH in current experimental setup, the former one demonstrates more regionally extended and persistent skill, by contrast the later demonstrates high skill in the equatorial region. The skill for full state initialized AMOC hindcasts appears to be more long-

sustained than for anomaly initialized ones.

The flux correction schemes, which were earlier disqualified from transient climate change simulations due to possible feedbacks distortion (Neelin and Dijkstra, 1995; Marotzke and Stone, 1995), is nowadays reconsidered for initializing decadal climate predictions (Magnusson et al., 2012a,b). It is expected that this approach has a potential to improve predictions of internal climate variability. Moreover, because the response to anthropogenic forcing seems to play a minor role in the first years of predictions, flux correction is believed not to contradict with external forcing response. Our flux corrected hindcasts obtain the highest and more persistent skill for the analyzed climate variables. Also for AMOC and SSH the skill is almost always slightly better than just full state initialization.

The highest skill for SST and SSH is achieved in the deep mixed-layer regions in the North Atlantic and Southern Ocean in full state initialized and flux corrected hindcasts. The analysis of skill for monthly mean SST shows that correlation values are high in winter time and low in summer and the root mean squared error is low in winter-time and high in summer. The re-emergence mechanism, which was previously investigated by Frankignoul and Hasselmann (1977), Alexander and Deser (1994) and Deser et al. (2003), is associated with high skill in these regions. The SST re-emergence mechanism provides predictive skill for the North Atlantic SST for up to 2 to 4 lead years, in differently initialized experiments, but the spatial distribution of skill suggests predictability up to lead time yr6-9. Therefore, the second part of this study is dedicated to find other candidate processes that might work in conjunction and provide high predictability for these regions.

5.1.2 Sensitivity of the results

When estimating predictive skill of initialized hindcasts we followed recommendations of the verification framework for decadal predictions (Goddard et al., 2013). The results of this thesis are based on every-five year initialization scheme (following the recommendations of the CMIP5 experimental design for the initialized hindcasts; Taylor et al., 2009), on the detrended (with the GECCO trend removed from the initialized hindcasts), bias corrected hindcasts (with a posteriori mean bias correction scheme applied). In recent years a progress has been made in improving the experimental setup for the decadal predictions using denser initialization periods (every year), improving bias correction procedures and hence initialization techniques. It is now known that the sparse start dates are not good enough to obtain a robust skill estimate (e.g., van Oldenborgh et al., 2012). In addition to deterministic skill scores, which are mainly used to estimate predictive skill of decadal predictions, studies by Goddard et al. (2013), Corti et al. (2012) and Hazeleger et al. (2013a) proposed to assess probability skill scores, which we didn't calculate because of the relatively small ensemble size.

An analysis of the evolution of the spatial averaged North Atlantic (NA) SST, NA SSH and AMOC at 26.5°N from differently initialized hindcasts showed that the coupled model is undergoing some drift in the full state initialization scheme. A drift is also evident in the flux corrected hindcasts. Therefore, we applied the bias correction to both of these schemes as well as to the anomaly initialized experiments to treat all the hindcasts the same way.

In order to compare the results from this thesis with other modeling efforts, one should keep in mind which bias and trend adjustment methods are applied, because different methods may lead to some differences in predictive skill estimate. Some of the decadal prediction studies on statistical skill estimation apply cross-validation scheme rather than mean bias correction scheme, trend adjustment rather than subtraction of the observed trend from the initialized hindcasts (Kharin et al., 2012; Smith et al., 2013). We provide here examples of predictive skill estimates for the North Atlantic SST, SSH and AMOC at 26.5°N (Fig. 5.1), which demonstrate how sensitive are the results to different bias and trend correction methods. The correlation skill (COR) and the root-mean-squared-error skill score (RMSS) estimated for the NA SST, NA SSH and AMOC at 26.5°N show that the cross-validated hindcasts have somewhat reduced skill in comparison to the bias-corrected hindcasts. Whereas the skill obtained by the trend adjustment method (where the lead years are detrended; Kharin et al., 2012), beats the skill for the NA SST and NA SSH computed using the first two methods, with the exception for the AMOC RMSS skill. The reason for lower skill when the first two methods are used might be that the skill is computed between data that do not have trend (GECCO) and those that still have some remaining trend (the initialized hindcasts). This suggests that the trend adjusting technique (Kharin et al., 2012) might be more appropriate approach to deal with hindcast's trend, especially if the observational and modeled long-term trends are not the same.

Though predictive skill does show differences when different metrics, bias and trend adjustment schemes are applied, the qualitative conclusion we make out of these results is not changed and is that in the current experimental setup the flux correction procedure does show an advantage over full state initialization and anomaly initialization procedures.

5.1.3 Processes giving rise to predictive skill

We analyzed possible mechanisms giving rise to high predictive skill of the initialized hindcasts using the SSH decomposition approach proposed by Köhl (accepted, 2014). This approach allows to investigate the density-related (steric) SSH changes and its contributions, namely the mixed-layer term that results from temperature and salinity anomalies in the mixed layer, the heave term that results from temperature and salinity anomalies below the mixed layer and the spice term that results from density compensating temperature-salinity anomalies.

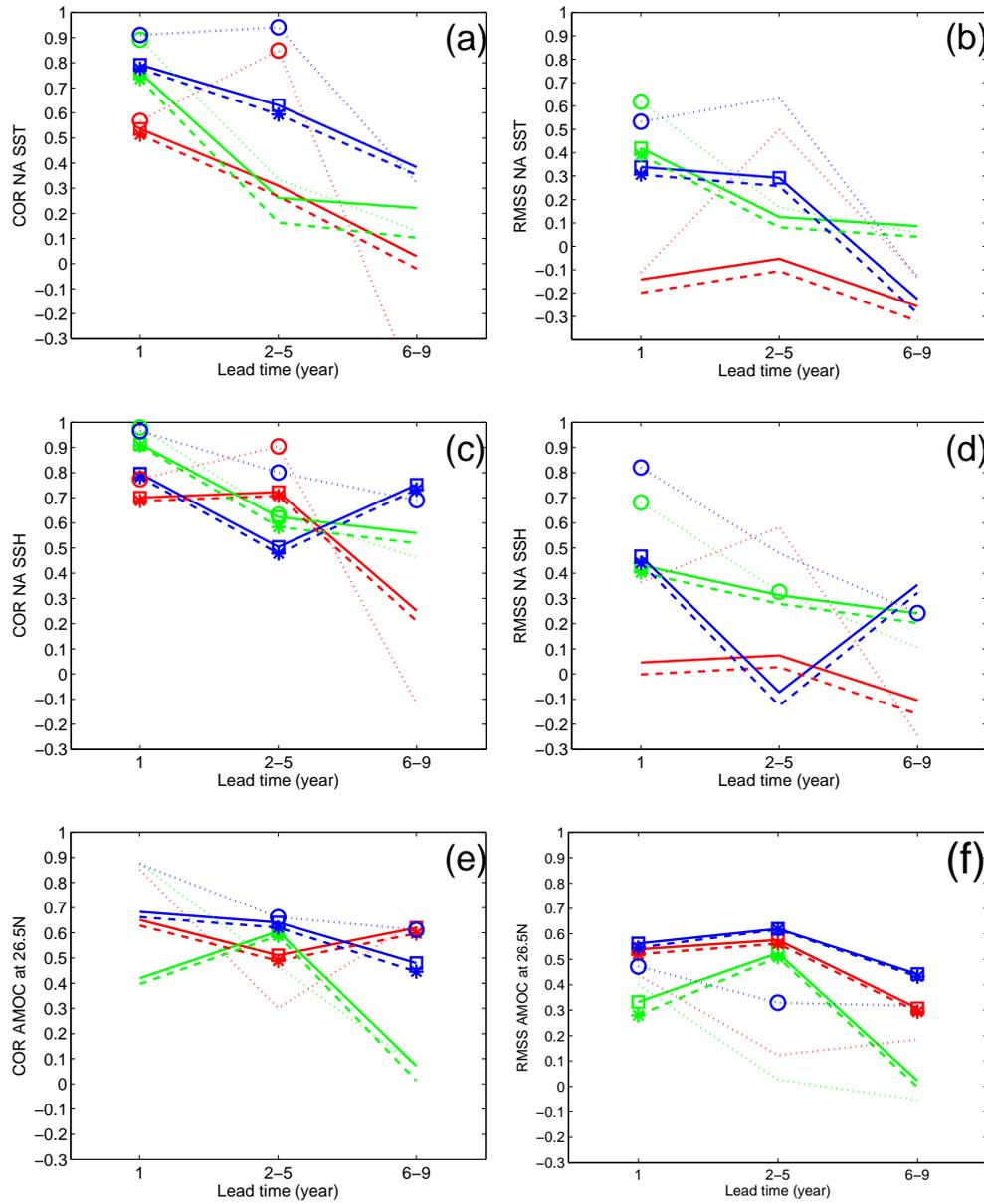


Figure 5.1— Anomaly correlation coefficient (left) and root mean square error skill score (right), both are computed with respect to GECCO, for the North Atlantic SST (a, b) and SSH (c, d) averaged over the region 50°W - 10°W , 20°N - 60°N , and AMOC at 26.5°N (e, f) from FSI-HIND (red), AI-HIND (green), FC-HIND (blue). Marks imply significant correlation at the 90% confidence interval estimated with a bootstrap procedure. Solid lines imply bias corrected hindcasts with the GECCO trend taken out, dashed lines – cross validated hindcasts with the GECCO trend taken out, dotted – trend adjustment.

The decomposition analysis showed that “useful” predictive skill is associated with mixed-layer term in mode water regions (in the Atlantic Ocean). We refer to predictive skill of initialized hindcasts as “useful”, if it satisfies the conditions: significant predictive skill beyond the persistence in the regions of large interannual variability. Also the density compensation (spice) term showed “useful” skill in the North Atlantic and Southern Ocean up to lead time yr2-5 and the heave term showed skill in the subtropical Pacific up to lead time yr2-5. The large-scale patterns of skill for steric SSH changes (including heave, spice and mixed layer terms) in the extratropical Southern Hemisphere are associated with persistence. Whereas the spice term represents the processes related to advection on isopycnals, the heave term is usually associated with the ocean response to the wind forcing (Doney et al., 2007). We expected that there will be less skill from the later term, because the directly forced by the atmosphere Ekman transport provides the response over the relatively short time scales (about days; Visbeck et al., 2003). In addition, some studies (e.g., Hirschi et al., 2007) speculate that for AMOC variability the Ekman transport variability is limited to subannual and seasonal timescales. On the other hand, studies by Capotondi and Alexander (2001) and Cessi and Louazel (2001) have shown some evidence for decadal variability related to wind-driven response.

In terms of the performance of differently initialized hindcasts in the decomposition analysis, the full state initialized and flux corrected hindcasts showed very similar skill distribution and skill duration for different contributions to steric SSH changes, and in general outperformed the skill from the anomaly initialized hindcasts.

5.2 Recommendations and future work

This study provides new insight into the processes that potentially contribute to predictive skill for decadal predictions but also raises new questions. The next steps that need to be undertaken for a continuation of the present work are:

- The studies by Magnusson et al. (2012a,b) and this thesis in particular showed the potential for flux correction to improve predictive skill on decadal timescales. The flux corrected hindcasts demonstrated somewhat better results than full state initialization and anomaly initialization. Yet, this technique is still a work-around strategy until model biases are eliminated. We found out that while significantly reducing the error for SST, heat and freshwater flux correction seems to have very little effect on the bias of quantities that represent the subsurface of the ocean. In this respect, the SSH and AMOC initialized hindcasts required additional a posteriori bias correction. Therefore, more investigation is required to understand what causes model biases and possibly eliminate their effect.

- Modeling efforts in estimation of predictive skill for decadal predictions usually follow certain requirements for experimental design and verification metrics; however, there are still many differences in experimental setup and how the skill is assessed. The verification framework developed by Goddard et al. (2013) consists of many recommendations, but still does not cover many aspects related to standardization in predictive skill evaluation. For instance, the verification framework does not consider flux correction scheme as an alternative initialization method. Hence, there are no requirements to what the relaxation timescales and to the duration of additional runs, from which the relaxation terms are constructed, should be. The statistical significance estimation is also an important issue for decadal predictive skill because we draw conclusions based on this value. It tells us about the timescale beyond which the climate variable can not be predicted. For instance, the lowest significant correlation value based on one-tailed t-test would be about 0.6, while the nearest-neighbor bootstrap method might give the value of 0.4.
- Given that uncertainties in climate change predictions can be reduced by proper initialization technique, it is important to understand the reasons why some initialization methods lead to higher skill than the others. Some efforts have been done in this direction. For instance, the study by Robson (2010) investigates the possible issues related to the representation of density fields in the anomaly initialization scheme. The issues of the anomaly initialization scheme in this thesis, such as the amplification of the AMOC trend at 26.5°N and low skill of the density-related SSH changes, might be due to the problem raised by Robson (2010). On the other hand, full state initialization allows large error growth in the system which seems not to interfere with internal variability in the decadal prediction system. Magnusson et al. (2012a,b) investigated the bias development in differently initialized experiments. They underlined the possible non-linear effects of the model drift in the full state initialized experiments called “over-shooting”, which means that the bias in the initialized hindcasts can become larger than the bias of the model climate during a transient period. We did not investigate the bias development in different initialization strategies in details. Though, the high performance of full state initialization technique in comparison with anomaly initialized should be further investigated.
- The North Atlantic and Southern Oceans appear to provide high predictive skill for SST and SSH on decadal timescale when the skill is estimated with respect to the GECCO ocean synthesis, which is also used as a source of initial conditions in this thesis. When comparing the SST hindcasts to other data, like HadISST, the correlation analysis demonstrates rather poor skill in the Southern Ocean. Though both of the datasets represent the ocean reconstructed state and are based on actual observations, the difference in this region is remarkable. In this thesis we used GECCO and HadISST as the verification dataset for SST skill, and only GECCO for AMOC and SSH skill owing to the relatively short data records for the later two

climate variables. The study by Boer et al. (2013) demonstrates that only the skill which is based on real observations gives the information about actual forecast skill of the decadal prediction system.

- For getting an insight into possible physical and dynamic processes that give rise to predictive skill, a decomposition analysis of different contributions to steric SSH and associated skill is to be further explored.
- In this study the MITgcm ocean model was used for producing initial conditions and initialized hindcasts. It is expected that initializing the model with initial conditions obtained from the same coupled system can minimize the initial shock and model drift. In this study we initialized only the ocean component of the climate system and based the analysis of predictive skill on the ocean variables. Because of the large thermal capacity and long equilibration time, the ocean is considered to possess a long-term memory of the climate system and should be primarily initialized. However, the role of land and ice contribution to predictive skill should not be discarded. For instance, Koster and Suarez (2003) showed that land initialization provides improvement of seasonal predictions of precipitation and surface temperature. For future work, understanding whether initialization of land and ice climate components provides predictability on decadal timescale is needed.

Acronyms

AI anomaly initialization

AI-ASSIM assimilation run with anomaly initialization

AI-HIND anomaly initialized hindcasts

AMOC Atlantic Meridional Overturning Circulation

AVISO Archiving, Validation and Interpretation of Satellite Oceanographic Data project

COR correlation coefficient

ERA-Interim atmospheric reanalysis produced by the European Centre for Medium Range Weather Forecasts

FC flux correction

FC-ASSIM assimilation run with flux correction

FC-HIND flux corrected hindcasts

FSI full state initialization

FSI-ASSIM assimilation run with full state initialization

FSI-HIND full state initialized hindcasts

GECCO German contribution to Estimating the Circulation and Climate Ocean project

GCM, AGCM, CGCM, OGCM general circulation model, atmospheric GCM, coupled GCM, ocean GCM

H heave term

HadISST Met Office Hadley Centre's sea surface temperature

KPP "k profile" parameterization

M mixed-layer related

MIT Massachusetts Institute of Technology

OHC ocean heat content

P-E+R freshwater flux: precipitation – evaporation + river runoff

RMSE root mean square error

RMSS root mean square error skill score

S spice term

SSH sea surface height

SST sea surface temperature

MLD mixed layer depth

UCLA University of California, Los Angeles

Bibliography

- Adcroft, A., Campin, J., Dutkiewicz, S., Evangelinos, C., Ferreira, D., Forget, G., Fox-Kemper, B., Heimbach, P., Hill, C., Hill, E., et al.: MITgcm user manual, MIT Department of EAPS, p. 470, URL http://mitgcm.org/public/r2_manual/latest/online_documents/, 2008.
- Alexander, M. and Deser, C.: A Mechanism for the recurrence of wintertime midlatitude SST anomalies, *Journal of Physical Oceanography*, 25, 122–137, 1994.
- Anderson, D. L. and Willebrand, J.: Decadal Climate Variability: Dynamics and Predictability, vol. 44, pp 496, Springer, 1996.
- Antonov, J. I., Levitus, S., and Boyer, T. P.: Steric sea level variations during 1957–1994: Importance of salinity, *Journal of Geophysical Research*, 107, 14–1–14–8, 2002.
- Arakawa, A. and Schubert, W.: Interaction of a cumulus cloud ensemble with the large-scale environment, Part I, *Journal of the Atmospheric Sciences*, 31, 674–701, 1974.
- Balmaseda, M. and Anderson, D.: Impact of initialization strategies and observations on seasonal forecast skill, *Geophysical Research Letters*, 36, L01701, doi:10.1029/2008GL035561, 2009.
- Banks, H. T. and Gregory, J. M.: Mechanisms of ocean heat uptake in a coupled climate model and the implications for tracer based predictions of ocean heat uptake, *Geophysical Research Letters*, 33, L07608, doi:10.1029/2005GL025352, 2006.
- Barthelet, P., Terray, L., and Valcke, S.: Transient CO₂ experiment using the ARPEGE/OPAICE non-flux corrected coupled model, *Geophysical Research Letters*, 25, 2277–2280, 1998.
- Bindoff, N. L., Willebrand, J., Artale, V., Cazenave, A., Gregory, J. M., Gulev, S., Hanawa, K., Le Quere, C., Levitus, S., Nojiri, Y., et al.: Observations: oceanic climate change and sea level, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, pp. 385–432, 2007.
- Bloom, S., Takacs, L., Da Silva, A., and Ledvina, D.: Data assimilation using incremental analysis updates, *Monthly Weather Review*, 124, 1256–1271, 1996.

- Boer, G.: A study of atmosphere-ocean predictability on long time scales, *Climate Dynamics*, 16, 469–477, 2000.
- Boer, G.: Long time-scale potential predictability in an ensemble of coupled climate models, *Climate Dynamics*, 23, 29–44, 2004.
- Boer, G. and Lambert, S.: Multi-model decadal potential predictability of precipitation and temperature, *Geophysical Research Letters*, 35, L05 706, doi:10.1029/2008GL033234, 2008.
- Boer, G., Kharin, V., and Merryfield, W.: Decadal predictability and forecast skill, *Climate Dynamics*, 41, 1817–1833, 2013.
- Boer, G. J.: Decadal potential predictability of twenty-first century climate, *Climate Dynamics*, 36, 1119–1133, 2011.
- Branstator, G. and Teng, H.: Two limits of initial-value decadal predictability in a CGCM, *Journal of Climate*, 23, 6292–6311, 2010.
- Cabanes, C., Huck, T., and Colin de Verdière, A.: Contributions of wind forcing and surface heating to interannual sea level variations in the Atlantic Ocean, *Journal of Physical Oceanography*, 36, 1739–1750, 2006.
- Capotondi, A. and Alexander, M.: Rossby waves in the tropical North Pacific and their role in decadal thermocline variability, *Journal of Physical Oceanography*, 31, 3496–3515, 2001.
- Carton, J. and Santorelli, A.: Global decadal upper-ocean heat content as viewed in nine analyses, *Journal of Climate*, 21, 6015–6035, 2008.
- Cazes-Boezio, G., Menemenlis, D., and Mechoso, C.: Impact of ECCO ocean-state estimates on the initialization of seasonal climate forecasts, *Journal of Climate*, 21, 1929–1947, 2008.
- Cessi, P. and Louazel, S.: Decadal oceanic response to stochastic wind forcing, *Journal of Physical Oceanography*, 31, 3020–3029, 2001.
- Church, J. A., Woodworth, P. L., and Aarup, Thorkild and. Wilson, W. S.: *Understanding Sea-level Rise and Variability*, pp 428, Wiley-Blackwell, 2010.
- Church, J. A., White, N. J., Konikow, L. F., Domingues, C. M., Cogley, J. G., Rignot, E., Gregory, J. M., van den Broeke, M. R., Monaghan, A. J., and Velicogna, I.: Revisiting the Earth’s sea-level and energy budgets from 1961 to 2008, *Geophysical Research Letters*, 38, L18 601, doi:10.1029/2011GL048794, 2011.
- Ciasto, L. and Thompson, D.: Observational evidence of reemergence in the extratropical Southern Hemisphere, *Journal of Climate*, 22, 1446–1453, 2009.

- Collins, M.: Climate predictability on interannual to decadal time scales: the initial value problem, *Climate Dynamics*, 19, 671–692, 2002.
- Collins, M., Botzet, M., Carril, A., Drange, H., Jouzeau, A., Latif, M., Masina, S., Otteraa, O., Pohlmann, H., Sorteberg, A., et al.: Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study, *Journal of Climate*, 19, 1195–1203, 2006.
- Corti, S., Weisheimer, A., Palmer, T., Doblas-Reyes, F., and Magnusson, L.: Reliability of decadal predictions, *Geophysical Research Letters*, 39, L21 712, doi:10.1029/2012GL053354, 2012.
- Deardorff, J.: Parameterization of the planetary boundary layer for use in general circulation models 1, *Monthly Weather Review*, 100, 93–106, 1972.
- Delworth, T. L. and Knutson, T. R.: Simulation of early 20th century global warming, *Science*, 287, 2246–2250, 2000.
- Delworth, T. L., Ramaswamy, V., and Stenchikov, G. L.: The impact of aerosols on simulated ocean temperature and heat content in the 20th century, *Geophysical Research Letters*, 32, L24 709, doi:10.1029/2005GL024457, 2005.
- Deser, C., Alexander, M. A., and Timlin, M. S.: Understanding the persistence of sea surface temperature anomalies in midlatitudes, *Journal of Climate*, 16, 57–72, 2003.
- Dijkstra, H. A. and Ghil, M.: Low-frequency variability of the large-scale ocean circulation: A dynamical systems approach, *Reviews of Geophysics*, 43, RG3002, doi:10.1029/2002RG000122, 2005.
- Doblas-Reyes, F., Balmaseda, M., Weisheimer, A., and Palmer, T.: Decadal climate prediction with the European Centre for Medium-Range Weather Forecasts coupled forecast system: Impact of ocean observations, *Journal of Geophysical Research*, 116, D19 111, doi:10.1029/2010JD015394, 2011a.
- Doblas-Reyes, F., van Oldenborgh, G., García-Serrano, J., Pohlmann, H., Scaife, A., and Smith, D.: CMIP5 near-term climate prediction, WCRP Coupled Model Intercomparison Project – Phase 5 (CMIP5), p. 8, 2011b.
- Doblas-Reyes, F., Andreu-Burillo, I., Chikamoto, Y., García-Serrano, J., Guemas, V., Kimoto, M., Mochizuki, T., Rodrigues, L., and van Oldenborgh, G.: Initialized near-term regional climate change prediction, *Nature Communications*, 4, 1–9, doi:10.1038/ncomms2704, 2013.
- Doney, S. C., Yeager, S., Danabasoglu, G., Large, W. G., and McWilliams, J. C.: Mechanisms governing interannual variability of upper-ocean temperature in a global ocean hindcast simulation, *Journal of Physical Oceanography*, 37, 1918–1938, 2007.

- Dorman, J. and Sellers, P. J.: A global climatology of albedo, roughness length and stomatal resistance for atmospheric general circulation models as represented by the simple biosphere model (SiB), *Journal of Applied Meteorology*, 28, 833–855, 1989.
- Dukowicz, J. K. and Smith, R. D.: Implicit free-surface method for the Bryan-Cox-Semtner ocean model, *Journal of Geophysical Research*, 99, 7991–8014, 1994.
- Frankignoul, C. and Hasselmann, K.: Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability, *Tellus*, 29, 289–305, 1977.
- Gent, P. and McWilliams, J.: Isopycnal mixing in ocean circulation models, *Journal of Physical Oceanography*, 20, 150–155, 1990.
- Gill, A. and Niller, P.: The theory of the seasonal variability in the ocean, *Deep Sea Research and Oceanographic Abstracts*, 20, 141–177, 1973.
- Glickman, T. S.: *Glossary of Meteorology*, vol. 2, pp 850, American Meteorological Society, 2000.
- Goddard, L., Hurrell, J. W., Kirtman, B. P., Murphy, J., Stockdale, T., and C., V.: Two time scales for the price of one (almost), *Bulletin of the American Meteorological Society*, 93, 621–629, 2012.
- Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S. J., et al.: A verification framework for interannual-to-decadal predictions experiments, *Climate Dynamics*, 40, 245–272, 2013.
- Greatbatch, R. J.: A note on the representation of steric sea level in models that conserve volume rather than mass, *Journal of Geophysical Research: Oceans (1978–2012)*, 99, 12 767–12 771, 1994.
- Gregory, J., Dixon, K., Stouffer, R., Weaver, A., Driesschaert, E., Eby, M., Fichefet, T., Hasumi, H., Hu, A., Jungclaus, J., et al.: A model intercomparison of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO₂ concentration, *Geophysical Research Letters*, 32, L12 703, doi:10.1029/2005GL023209, 2005.
- Griffies, S. and Bryan, K.: Predictability of North Atlantic multidecadal climate variability, *Science*, 275, 181–184, 1997.
- Guemas, V., Doblas-Reyes, F. J., Andreu-Burillo, I., and Asif, M.: Retrospective prediction of the global warming slowdown in the past decade, *Nature Climate Change*, 3, 649–653, 2013.
- Hall, A. and Manabe, S.: Can local linear stochastic theory explain sea surface temperature and salinity variability?, *Climate Dynamics*, 13, 167–180, 1997.

- Hannachi, A., Jolliffe, I., and Stephenson, D.: Empirical orthogonal functions and related techniques in atmospheric science: A review, *International Journal of Climatology*, 27, 1119–1152, 2007.
- Harshvardhan, R., Randall, D., Corsetti, T., et al.: A fast radiation parameterization for atmospheric circulation models, *Journal of Geophysical Research*, 92, 1009–1016, 1987.
- Harshvardhan, R., Randall, D., and Corsetti, T.: Earth radiation budget and cloudiness simulations with a general circulation model., *Journal of Atmospheric Sciences*, 46, 1922–1942, 1989.
- Hawkins, E. and Sutton, R.: The potential to narrow uncertainty in regional climate predictions, *Bulletin of the American Meteorological Society*, 90, 1095–1107, 2009.
- Hawkins, E., Robson, J., Sutton, R., Smith, D., and Keenlyside, N.: Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach, *Climate Dynamics*, 37, 2495–2509, 2011.
- Hazeleger, W., Guemas, V., Wouters, B., Corti, S., Andreu-Burillo, I., Doblas-Reyes, F., Wyser, K., and Caian, M.: Multiyear climate predictions using two initialisation Strategies, *Geophysical Research Letters*, 40, 1794–1798, 2013a.
- Hazeleger, W., Wouters, B., Oldenborgh, G., Corti, S., Palmer, T., Smith, D., Dunstone, N., Kröger, J., Pohlmann, H., and Storch, J.-S.: Predicting multiyear North Atlantic Ocean variability, *Journal of Geophysical Research: Oceans*, 118, 1087–1098, doi:10.1002/jgrc.20117, 2013b.
- Hirschi, J. J., Killworth, P. D., and Blundell, J. R.: Subannual, seasonal, and interannual variability of the North Atlantic meridional overturning circulation, *Journal of Physical Oceanography*, 37, 1246–1265, 2007.
- Huber, M. and Knutti, R.: Anthropogenic and natural warming inferred from changes in Earth's energy balance, *Nature Geoscience*, 5, 31–36, 2011.
- Hurrell, J. W. and Van Loon, H.: Decadal variations in climate associated with the North Atlantic Oscillation, *Climatic Change*, 36, 301–326, 1997.
- Hurrell, J. W., Visbeck, M., Busalacchi, A., Clarke, R., Delworth, T., Dickson, R., Johns, W. E., Koltermann, K., Kushnir, Y., Marshall, D., et al.: Atlantic climate variability and predictability: A CLIVAR perspective, *Journal of Climate*, 19, 5100–5121, 2006.
- IPCC: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2013.

- Ishii, M., Kimoto, M., Sakamoto, K., and Iwasaki, S.-I.: Steric sea level changes estimated from historical ocean subsurface temperature and salinity analyses, *Journal of Oceanography*, 62, 155–170, 2006.
- Keenlyside, N., Latif, M., JungCLAUS, J., Kornbluh, L., and Roeckner, E.: Advancing decadal-scale climate prediction in the North Atlantic sector, *Nature*, 453, 84–88, 2008.
- Kerr, R. A.: A North Atlantic climate pacemaker for the centuries, *Science*, 288, 1984–1985, 2000.
- Kharin, V., Boer, G., Merryfield, W., Scinocca, J., and Lee, W.-S.: Statistical adjustment of decadal predictions in a changing climate, *Geophysical Research Letters*, 39, L19 705, doi:10.1029/2012GL052647, 2012.
- Knight, J., Kennedy, J., Folland, C., Harris, G., Jones, G., Palmer, M., Parker, D., Scaife, A., and Stott, P.: Do global temperature trends over the last decade falsify climate predictions?, *Bulletin of the American Meteorological Society*, 90, 22–23, 2009.
- Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., and Mann, M. E.: A signature of persistent natural thermohaline circulation cycles in observed climate, *Geophysical Research Letters*, 32, L20 708, doi:10.1029/2005GL024233, 2005.
- Köhl, A.: Detecting the origin of interannual steric sea level changes, *Journal of Climate*, doi:10.1175/JCLI-D-13-00412.1, accepted, 2014.
- Köhl, A. and Stammer, D.: Decadal sea level changes in the 50-year GECCO ocean synthesis, *Journal of Climate*, 21, 1876–1890, 2007.
- Köhl, A. and Stammer, D.: Variability of the meridional overturning in the North Atlantic from the 50-year GECCO state estimation, *Journal of Physical Oceanography*, 38, 1913–1930, 2008.
- Köhl, A., Dommenges, D., Ueyoshi, K., and Stammer, D.: The global ECCO 1952 to 2001 ocean synthesis, *ECCO Report*, 40, 1–43, 2006.
- Köhl, A., Stammer, D., and Cornuelle, B.: Interannual to decadal changes in the ECCO global synthesis, *Journal of Physical Oceanography*, 37, 313–337, 2007.
- Köhler, M.: Explicit prediction of ice clouds in general circulation models, Ph.D. Dissertation, Department of Atmospheric Sciences, University of California, Los Angeles, 167pp, 1999.
- Konikow, L. F.: Contribution of global groundwater depletion since 1900 to sea-level rise, *Geophysical Research Letters*, 38, L17 401, doi:10.1029/2011GL048604, 2011.
- Konor, C., Boezio, G., Mechoso, C., and Arakawa, A.: Parameterization of PBL processes in an atmospheric general circulation model: Description and preliminary assessment, *Monthly Weather Review*, 137, 1061–1082, 2009.

- Koster, R. D. and Suarez, M. J.: Impact of land surface initialization on seasonal precipitation and temperature prediction, *Journal of Hydrometeorology*, 4, 408–423, 2003.
- Kröger, J., Müller, W. A., and von Storch, J.-S.: Impact of different ocean reanalyses on decadal climate prediction, *Climate Dynamics*, 39, 795–810, 2012.
- Kumar, A., Chen, M., Zhang, L., Wang, W., Xue, Y., Wen, C., Marx, L., and Huang, B.: An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2, *Monthly Weather Review*, 140, 3003–3016, 2012.
- Landerer, F. W., Jungclaus, J. H., and Marotzke, J.: Regional dynamic and steric sea level change in response to the IPCC-A1B scenario, *Journal of Physical Oceanography*, 37, 296–312, 2007.
- Large, W., McWilliams, J., and Doney, S.: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization, *Reviews of Geophysics*, 32, 363–403, 1994.
- Large, W., Danabasoglu, G., Doney, S., and McWilliams, J.: Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology, *Journal of Physical Oceanography*, 27, 2418–2447, 1997.
- Latif, M. and Barnett, T. P.: Decadal climate variability over the North Pacific and North America: Dynamics and predictability, *Journal of Climate*, 9, 2407–2423, 1996.
- Latif, M. and Keenlyside, N. S.: A perspective on decadal climate variability and predictability, *Deep Sea Research Part II: Topical Studies in Oceanography*, 58, 1880–1894, 2011.
- Laurian, A., Lazar, A., and Reverdin, G.: Generation mechanism of spiciness anomalies: An OGCM analysis in the North Atlantic subtropical gyre., *Journal of Physical Oceanography*, 39, 1003–1018, 2009.
- Lee, T., Fukumori, I., Menemenlis, D., Xing, Z., and Fu, L.-L.: Effects of the Indonesian throughflow on the Pacific and Indian Oceans, *Journal of Physical Oceanography*, 32, 1404–1429, 2002.
- Lee, T., Awaji, T., Balmaseda, M., Ferry, N., Fujii, Y., Fukumori, I., Giese, B., Heimbach, P., Köhl, A., Masina, S., et al.: Consistency and fidelity of Indonesian-throughflow total volume transport estimated by 14 ocean data assimilation products, *Dynamics of Atmospheres and Oceans*, 50, 201–223, 2010.
- Leeuwenburgh, O. and Stammer, D.: The effect of ocean currents on sea surface temperature anomalies, *Journal of Physical Oceanography*, 31, 2340–2358, 2001.

- Leuliette, E. W. and Wahr, J. M.: Coupled pattern analysis of sea surface temperature and TOPEX/Poseidon sea surface height, *Journal of Physical Oceanography*, 29, 599–611, 1999.
- Levitus, S. and Boyer, T.: World Ocean Atlas 1994, vol. 4, Temperature, NOAA Atlas NESDIS 4, US Department of Commerce, Washington, DC, 117pp, 1994.
- Levitus, S., Burgett, R., and Boyer, T.: World Ocean Atlas 1994, vol. 3, Salinity, NOAA Atlas NESDIS 3, US Gov. Print. Off., Washington, DC, 1994.
- Li, D. and Shine, K.: A 4-dimensional ozone climatology for UGAMP models, UGAMP Internal Report, 35, 35pp, 1995.
- Li, J., Köhler, M., Farrara, J., and Mechoso, C.: The impact of stratocumulus cloud radiative properties on surface heat fluxes simulated with a general circulation model, *Monthly Weather Review*, 130, 1433–1441, 2002.
- Lo, J. C.-F., Yang, Z.-L., and Pielke Sr, R. A.: Assessment of three dynamical climate downscaling methods using the Weather Research and Forecasting (WRF) model, *Journal of Geophysical Research*, 113, D09 112, doi:10.1029/2007JD009216, 2008.
- Lowe, J. A. and Gregory, J. M.: Understanding projections of sea level rise in a Hadley Centre coupled climate model, *Journal of Geophysical Research: Oceans (1978–2012)*, 111, doi:10.1029/2005JC003421, 2006.
- Lynn, R. J. and Reid, J. L.: Characteristics and circulation of deep and abyssal waters, *Deep Sea Research and Oceanographic Abstracts*, 15, 577–598, 1968.
- Ma, H., Mechoso, C., Xue, Y., Xiao, H., Wu, C., Li, J., and De Sales, F.: Impact of land surface processes on the South American warm season climate, *Climate Dynamics*, 37, 187–203, 2010.
- Magnusson, L., Alonso-Balmaseda, M., Corti, S., Molteni, F., and Stockdale, T.: Evaluation of forecast strategies for seasonal and decadal forecasts in presence of systematic model errors, *Climate Dynamics*, 41, 2393–2409, doi:10.1007/s00382-012-1599-2, 2012a.
- Magnusson, L., Alonso-Balmaseda, M., and Molteni, F.: On the dependence of ENSO simulation on the coupled model mean state, *Climate Dynamics*, 41, 1509–1525, doi: 10.1007/s00382-012-1574-y, 2012b.
- Manganello, J. V. and Huang, B.: The influence of systematic errors in the Southeast Pacific on ENSO variability and prediction in a coupled GCM, *Climate Dynamics*, 32, 1015–1034, 2009.
- Mantua, N. J. and Hare, S. R.: The Pacific decadal oscillation, *Journal of Oceanography*, 58, 35–44, 2002.

- Marotzke, J. and Stone, P. H.: Atmospheric transports, the thermohaline circulation, and flux adjustments in a simple coupled model, *Journal of Physical Oceanography*, 25, 1350–1364, 1995.
- Marshall, J., Nurser, A., and Williams, R.: Inferring the subduction rate and period over the North Atlantic, *Journal of Physical Oceanography*, 23, 1315–1315, 1993.
- Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R., and Visbeck, M.: North Atlantic climate variability: phenomena, impacts and mechanisms, *International Journal of Climatology*, 21, 1863–1898, 2001.
- Matei, D., Pohlmann, H., Jungclauss, J., Müller, W., Haak, H., and Marotzke, J.: Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model, *Journal of Climate*, 25, 8502–8523, 2012.
- McDougall, T. J.: Neutral surfaces, *Journal of Physical Oceanography*, 17, 1950–1964, 1987.
- Mechoso, C. R., Robertson, A. W., Barth, N., Davey, M., Delecluse, P., Gent, P., Ineson, S., Kirtman, B., Latif, M., Le Treut, H., et al.: The seasonal cycle over the tropical Pacific in coupled ocean–atmosphere general circulation models, *Monthly Weather Review*, 123, 2825–2838, 1995.
- Meehl, G., Goddard, L., Murphy, J., Stouffer, R., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M., Greene, A., Hawkins, E., et al.: Decadal prediction, *Bulletin of the American Meteorological Society*, 90, 1467–1485, 2009.
- Meehl, G. A., Arblaster, J. M., Fasullo, J. T., Hu, A., and Trenberth, K. E.: Model-based evidence of deep-ocean heat uptake during surface-temperature hiatus periods, *Nature Climate Change*, 1, 360–364, 2011.
- Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Reyes, F., Hawkins, E., et al.: Decadal climate prediction: an update from the trenches, *Bulletin of the American Meteorological Society*, doi: 10.1175/BAMS-D-12-00241.1, accepted, 2013.
- Menary, M. B., Roberts, C. D., Palmer, M. D., Halloran, P. R., Jackson, L., Wood, R. A., Müller, W. A., Matei, D., and Lee, S.-K.: Mechanisms of aerosol-forced AMOC variability in a state of the art climate model, *Journal of Geophysical Research: Oceans*, 118, 2087–2096, 2013.
- Met Office Hadley Centre: The recent pause in global warming (2): What are the potential causes?, URL <http://www.metoffice.gov.uk/research/news/recent-pause-in-war%ming>, 2013.

- Mochizuki, T., Ishii, M., Kimoto, M., Chikamoto, Y., Watanabe, M., Nozawa, T., Sakamoto, T., Shiogama, H., Awaji, T., Sugiura, N., et al.: Pacific Decadal Oscillation hindcasts relevant to near-term climate prediction, *Proceedings of the National Academy of Sciences*, 107, 1833–1837, 2010.
- Moulin, S., Bondeau, A., and Delecolle, R.: Combining agricultural crop models and satellite observations: from field to regional scales, *International Journal of Remote Sensing*, 19, 1021–1036, 1998.
- Munk, W.: Internal waves and small-scale processes, *Evolution of physical oceanography*, pp. 264–291, 1981.
- Munoz, E., Kirtman, B., and Weijer, W.: Varied representation of the Atlantic Meridional Overturning across multidecadal ocean reanalyses, *Deep Sea Research Part II: Topical Studies in Oceanography*, 58, 1848–1857, 2011.
- Murphy, J., Kattsov, V., Keenlyside, N., Kimoto, M., Meehl, G., Mehta, V., Pohlmann, H., Scaife, A., and Smith, D.: Towards prediction of decadal climate variability and change, *Procedia Environmental Sciences*, 1, 287–304, 2010.
- Neelin, J. D. and Dijkstra, H. A.: Ocean-atmosphere interaction and the tropical climatology. Part I: the dangers of flux correction, *Journal of climate*, 8, 1325–1342, 1995.
- Pan, D. and Randall, D.: A cumulus parameterization with a prognostic closure, *Quarterly Journal of the Royal Meteorological Society*, 124, 949–981, 1998.
- Parker, D., Folland, C., Scaife, A., Knight, J., Colman, A., Baines, P., and Dong, B.: Decadal to multidecadal variability and the climate change background, *Journal of Geophysical Research: Atmospheres* (1984–2012), 112, D18 115, doi:10.1029/2007JD008411, 2007.
- Piecuch, C. and Ponte, R.: Mechanisms of interannual steric sea level variability, *Geophysical Research Letters*, 38, L15 605, doi:10.1029/2011GL048440, 2011.
- Pierce, D., Barnett, T., Tokmakian, R., Semtner, A., Maltrud, M., Lysne, J., and Craig, A.: The ACPI project, element 1: Initializing a coupled climate model from observed conditions, *Climatic Change*, 62, 13–28, 2004.
- Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., and Tschuck, P.: Estimating the decadal predictability of a coupled AOGCM, *Journal of Climate*, 17, 4463–4472, 2004.
- Pohlmann, H., Jungclaus, J., Köhl, A., Stammer, D., and Marotzke, J.: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic, *Journal of Climate*, 22, 3926–3938, 2009.

- Pohlmann, H., Smith, D. M., Balmaseda, M. A., Keenlyside, N. S., Masina, S., Matei, D., Müller, W. A., and Rogel, P.: Predictability of the mid-latitude Atlantic meridional overturning circulation in a multi-model system, *Climate Dynamics*, 41, 775–785, 2012.
- Ponte, R. M.: A preliminary model study of the large-scale seasonal cycle in bottom pressure over the global ocean, *Journal of Geophysical Research*, 104, 1289–1300, 1999.
- Power, S., Haylock, M., Colman, R., and Wang, X.: The predictability of interdecadal changes in ENSO activity and ENSO teleconnections, *Journal of Climate*, 19, 4755–4771, 2006.
- Primeau, F.: Long Rossby wave basin-crossing time and the resonance of low-frequency basin modes, *Journal of Physical Oceanography*, 32, 2652–2665, 2002.
- Qiu, B. and Chen, S.: Decadal variability in the large-scale sea surface height field of the South Pacific Ocean: Observations and causes, *Journal of Physical Oceanography*, 36, 1751–1762, 2006.
- Qiu, B. and Huang, R.: Ventilation of the North Atlantic and North Pacific: subduction versus obduction, *Journal of Physical Oceanography*, 25, 2374–2390, 1995.
- Rayner, N., Parker, D., Horton, E., Folland, C., Alexander, L., Rowell, D., Kent, E., and Kaplan, A.: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, *Journal of Geophysical Research*, 108, 2–1–2–22, doi:10.1029/2002JD002670, 2003.
- Redi, M.: Oceanic isopycnal mixing by coordinate rotation, *Journal of Physical Oceanography*, 12, 1154–1158, 1982.
- Robertson, A., Overpeck, J., Rind, D., Mosley-Thompson, E., Zielinski, G., Lean, J., Koch, D., Penner, J., Tegen, I., and Healy, R.: Hypothesized climate forcing time series for the last 500 years, *Journal of Geophysical Research*, 106, 14 783–14 803, 2001.
- Robson, J. I.: Understanding the performance of a decadal prediction system, Ph.D. thesis, The University of Reading, 2010.
- Rosati, A., Miyakoda, K., and Gudgel, R.: The impact of ocean initial conditions on ENSO forecasting with a coupled model, *Monthly Weather Review*, 125, 754–772, 1997.
- Santer, B., Mears, C., Doutriaux, C., Caldwell, P., Gleckler, P., Wigley, T., Solomon, S., Gillett, N., Ivanova, D., Karl, T., et al.: Separating signal and noise in atmospheric temperature changes: The importance of timescale, *Journal of Geophysical Research*, 116, D22 105, doi:10.1029/2011JD016263, 2011.
- Sausen, R., Barthel, K., and Hasselmann, K.: Coupled ocean-atmosphere models with flux correction, *Climate Dynamics*, 2, 145–163, 1988.

- Schneider, E., Huang, B., Zhu, Z., DeWitt, D., Kinter III, J., Kirtman, B., and Shukla, J.: Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM, *Monthly Weather Review*, 127, 1187–1207, 1999.
- Schneider, N.: A decadal spiciness mode in the tropics, *Geophysical Research Letters*, 27, 257–260, 2000.
- Shackley, S., Risbey, J., Stone, P., and Wynne, B.: Adjusting to policy expectations in climate change modeling, *Climatic Change*, 43, 413–454, 1999.
- Simmons, A., Uppala, S., Dee, D., and Kobayashi, S.: ERA-Interim: New ECMWF reanalysis products from 1989 onwards, *ECMWF newsletter*, 110, 25–35, 2007.
- Smith, D., Cusack, S., Colman, A., Folland, C., Harris, G., and Murphy, J.: Improved surface temperature prediction for the coming decade from a global climate model, *Science*, 317, 796–799, 2007.
- Smith, D. M., Eade, R., and Pohlmann, H.: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction, *Climate Dynamics*, 41, 3325–3338, doi:10.1007/s00382-013-1683-2, 2013.
- Spencer, H., Sutton, R., and Slingo, J. M.: El Niño in a coupled climate model: sensitivity to changes in mean state induced by heat flux and wind stress corrections, *Journal of Climate*, 20, 2273–2298, 2007.
- Stammer, D., Kohl, A., Awaji, T., Balmaseda, M., ECMWF, S., Reading, R., Behringer, D., Center, C., Carton, J., Ferry, N., et al.: Ocean information provided through ensemble ocean syntheses, *Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society*, Venice, Italy, 21-25 September 2009, 2, 2009.
- Stammer, D., Agarwal, N., Herrmann, P., Köhl, A., and Mechoso, C.: Response of a coupled ocean–atmosphere model to greenland ice melting, *Surveys in Geophysics*, 32, 621–642, 2011.
- Stammer, D., Cazenave, A., Ponte, R. M., and Tamisiea, M. E.: Causes for contemporary regional sea level changes, *Annual Review of Marine Science*, 5, 21–46, 2013.
- Stockdale, T.: Coupled ocean-atmosphere forecasts in the presence of climate drift, *Monthly Weather Review*, 125, 809–818, 1997.
- Sturges, W. and Hong, B.: Decadal variability of sea level, *International Geophysics*, 75, 165–180, 2001.
- Sturges, W., Hong, B., and Clarke, A. J.: Decadal wind forcing of the North Atlantic subtropical gyre, *Journal of Physical Oceanography*, 28, 659–668, 1998.

- Suarez, M., Arakawa, A., and Randall, D.: The parameterization of the planetary boundary layer in the UCLA general circulation model- Formulation and results, *Monthly Weather Review*, 111, 2224–2243, 1983.
- Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H., and Ishikawa, Y.: Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations, *Journal of Geophysical Research: Oceans (1978–2012)*, 113, doi:10.1029/2008JC004741, 2008.
- Talley, L. D.: Some aspects of ocean heat transport by the shallow, intermediate and deep overturning circulations, *Mechanisms of global climate change at millennial time scales: Geophysical Monograph*, 112, 1–22, 1999.
- Tans, P. and Keeling, R.: Mauna Loa CO₂ annual mean data, URL www.esrl.noaa.gov/gmd/ccgg/trends/, NOAA/ESRL, Scripps Institution of Oceanography, 2011.
- Taylor, K., Stouffer, R., and Meehl, G.: A Summary of the CMIP5 Experiment Design, available at: <http://cmip-pcmdi.llnl.gov/cmip5/>, 2009.
- Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An overview of CMIP5 and the experiment design, *Bulletin of the American Meteorological Society*, 93, 485–498, 2012.
- Tebaldi, C. and Knutti, R.: The use of the multi-model ensemble in probabilistic climate projections, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365, 2053–2075, 2007.
- Teng, H., Branstator, G., and Meehl, G.: Predictability of the Atlantic Overturning Circulation and Associated Surface Patterns in Two CCSM3 Climate Change Ensemble Experiments, *Journal of Climate*, 24, 6054–6076, 2011.
- Troccoli, A. and Palmer, T.: Ensemble decadal predictions from analysed initial conditions, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 365, 2179–2191, 2007.
- van Oldenborgh, G. J., Doblas-Reyes, F. J., Wouters, B., and Hazeleger, W.: Decadal prediction skill in a multi-model ensemble, *Climate Dynamics*, 38, 1263–1280, 2012.
- Vinogradova, N., Ponte, R., and Stammer, D.: Relation between sea level and bottom pressure and the vertical dependence of oceanic variability, *Geophysical Research Letters*, 34, L03 608, doi:10.1029/2006GL028588, 2007.
- Visbeck, M., Chassignet, E. P., Curry, R. G., Delworth, T. L., Dickson, R. R., and Krahnemann, G.: The ocean’s response to North Atlantic Oscillation variability, *The North Atlantic Oscillation: climatic significance and environmental impact*, 134, 113–145, 2003.

- Vivier, F., Kelly, K. A., and Harismendy, M.: Causes of large-scale sea level variations in the Southern Ocean: Analyses of sea level and a barotropic model, *Journal of Geophysical Research*, 110, C09 014, doi:10.1029/2004JC002773, 2005.
- von Storch, H. and Navarra, A.: Analysis of climate variability: applications of statistical techniques, vol. 2, updated a. extended ed., pp 342, Springer, 1999.
- Weaver, A. J. and Sarachik, E.: Evidence for decadal variability in an ocean general circulation model: An advective mechanism 1, *Atmosphere-Ocean*, 29, 197–231, 1991.
- Weijer, W., Muñoz, E., Schneider, N., and Primeau, F.: Pacific decadal variability: Paced by Rossby basin modes?, *Journal of Climate*, 26, 1445–1456, 2013.
- Xue, Y., Sellers, P., Kinter, J., and Shukla, J.: A simplified biosphere model for global climate studies, *Journal of Climate*, 4, 345–364, 1991.
- Xue, Y., Fennessy, M., and Sellers, P.: Impact of vegetation properties on US summer weather prediction, *Journal of Geophysical Research. D. Atmospheres*, 101, 7419–7430, 1996.
- Yang, H., Hugh, A., and Liu, Z.: Basin Modes in a Tropical–Extratropical Basin., *Journal of Physical Oceanography*, 33, 2751–2763, 2003.
- Yeager, S. G. and Large, W. G.: Late-winter generation of spiciness on subducted isopycnals, *Journal of Physical Oceanography*, 34, 1528–1547, 2004.
- Yu, J.-Y. and Mechoso, C. R.: A discussion on the errors in the surface heat fluxes simulated by a coupled GCM, *Journal of Climate*, 12, 416–426, 1999.
- Zhang, X. and Church, J. A.: Sea level trends, interannual and decadal variability in the Pacific Ocean, *Geophysical Research Letters*, 39, L21 701, doi:10.1029/2012GL053240, 2012.

Declaration

I hereby declare, on oath, that I have written the present dissertation by my own and have not used other than the acknowledged resources and aids.

Hamburg

Iuliia Polkova

