Impact of Consistency between the Climate Model and its Initial Conditions on Decadal Climate Prediction

Dissertation zur Erlangung des Doktorgrades

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Geowissenschaften

der Universität Hamburg

vorgelegt von Xueyuan Liu

Hamburg, 2014

Tag der Disputation:

23. Januar 2015

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Detlef Stammer

Dr. Armin Köhl

Contents

Abstract

In my thesis I investigated the influence of dynamical consistency of initial conditions with the model used to perform forecasts starting from this initial condition, on the predictive skill of climate predictions on decadal time scales. The investigation builds on the coupled global model "Coupled GCM for Earth Simulator" (CFES) developed by Japan Agency for Marine-Earth Science and Technology (JAMSTEC). It is initialized in the ocean using a full state initialization strategy with two different initial fields: (1) oceanic initial conditions obtained through the same CFES coupled model using a 4-DVAR assimilation scheme (CDA: Coupled Data Assimilation) as provided by JAMSTEC; and (2) interpolated oceanic initial conditions obtained from the ocean-only GECCO2 Synthesis, while in the atmosphere initialized with bulk parameters controlling air-sea fluxes assimilated through CFES. CDA initial conditions are more dynamically consistent with the CFES model, because it is assimilated with the same CFES model. The relative skill of initialized hindcasts is then evaluated by comparing them with the persistence forecast, and with the uninitialized hindcasts (i.e. the twentieth-century simulation), respectively. To this end the spatial distribution of respective root mean square skill score (RMSS) in terms of SST is analyzed. Estimations on the performances of two differently initialized forecasts, i.e. CDA initialized hindcasts (CIH) and GECCO2 initialized hindcasts (GIH), are also done through anomaly correlation coefficient (ACC) of SST.

The spatial distribution of ACC for SST reveals that the forecast skill is better when initializing the CFES model with CDA initial conditions rather than with GECCO2, with the most significant improvement of skill observed at the first lead year. For this specific model, improvements in SST predictive skill are especially obvious over the tropical Pacific, suggesting that the dynamical consistency of initial conditions can indeed improve the predictive skill of climate predictions at least in the first lead year. For longer lead times of 4-yr averages, a large decrease of SST predictive skill is observed almost everywhere for both CIH and GIH. This holds especially over the North Atlantic (NA) where previous studies indicate that predictive skill of SST is associated with the predictability of Atlantic Meridional Overturning Circulation (AMOC), with the latter leading the former several years. Our results indicate that significant predictive skill in terms of AMOC is obtained at the first lead year for both CIH and GIH. However, significant AMOC predictive skill in the subpolar region of North Atlantic is observed at lead yr 2-5 in CIH, but not in GIH. The poor predictive skill of NA SST is consistent with poor predictability of AMOC in our solutions.

To investigate to what extent the reduced predictive skill in the North Atlantic in our solution may result from relatively short hindcast runs, the performances of CIH and GIH are also evaluated against previous results from the Mittelfristige Klimaprognosen (MiKlip hereafter) project with Max-Planck-Institute Earth System Model (MPI-ESM), which are initialized with assimilated data through nudging MPI-ESM towards GECCO2. The comparison reveals that the MiKlip hindcasts of 1980-2006 outperforms GIH, while its performance is compatible with that of CIH. Insights into how the number of initial date impacts predictive skill are provided through comparing the performances of MiKlip hindcasts of 1980-2006 and that of 1961-2009. It is shown that a larger number of initialization dates reduces the threshold of significance level of predictive skill, due to larger verification period. Therefore, more areas with significant skill are likely to occur for the latter.

Looking into processes that lead to improved predictive skill in the tropical Pacific, the significant, and high SST predictive skill over the tropical Pacific in CIH indicates a good reproduction of El Niño events at lead year one. In contrast, GIH produces additional erroneous El Niño events, which may contribute to the poor predictive skill of SST over the tropical Pacific. During balanced states, a zonal momentum balance between the wind stress and pressure gradient force of the upper equatorial Pacific exists. GIH is initialized with interpolated GECCO2 ocean estimations and atmospheric conditions that result from the same coupled model as the model used in the forecasts. The differences between the initialized ocean and atmosphere modes cause incompatibilities to the coupled model, and lead to imbalance between the zonal wind stress and pressure gradient force over the equatorial Pacific. The reduced predictive skill of GIH in the Pacific equatorial region is therefore found to mainly related to dynamical imbalance between zonal wind stress and pressure gradient over the central Pacific. These imbalances in the central Pacific subsequently propagate eastwards as Kelvin waves. Further imbalance may result from the differences in topography between GECCO2 and the model system, and perturbations may give rise to propagating waves. Jointly, these imbalances are highly likely to lead to additional pseudo El Niño events in GIH, as well as the poor predictive skill over the tropical Pacific. Our results underpin the requirement of a momentum balance between zonal wind stress and pressure gradient force along the equatorial Pacific when initializing the model from any oceanic state. Initializing a coupled model with self-consistent initial conditions therefore improves the skill of decadal climate prediction in the tropical Pacific.

Zusammenfassung

In meiner Diplomarbeit habe ich den Einfluss, auf die prädiktive Fähigkeit der Klimavorhersagen auf dekadischen Zeitskalen, dynamischer Konsistenz der Anfangsbedingungen mit dem verwendet werden, um tatsächlich Prognosen des gekoppelten Klimasystems ausgehend von diesem Anfangszustand durchzuführen Modell. Die Untersuchung basiert auf der vollständig gekoppelten globalen Modell "Coupled GCM für Earth Simulator" (CFES) von Japan Agency for Marine Earth Science and Technology (JAMSTEC), die in den Ozean mit einem vollen Zustand Initialisierung Strategie mit zwei verschiedenen initialisiert wird entwickelt Anfangsfelder: (1) ozeanischen durch die gleiche CFES gekoppelt Modell mit einem 4-DBVARIANZ Assimilationsschema erhalten Anfangsbedingungen (CDA: gekoppelt Datenassimilation), wie von JAMSTEC vorgesehen ist; und (2) interpoliert ozeanischen vom Meer nur GECCO2 Synthese erhalten Anfangsbedingungen sowie Groß Parameter, die Luft und Meer Flüsse durch CFES als atmosphärischen Anfangsbedingungen für beide assimiliert. CDA Anfangsbedingungen sind dynamisch in Übereinstimmung mit der CFES Modell, weil es mit der gleichen CFES Modell assimiliert. Die räumliche Verteilung der ACC für SST zeigt, dass die Vorhersagegüte ist besser bei der Initialisierung des CFES Modell mit CDA Anfangsbedingungen nicht mit GECCO2, mit den meisten signifikante Verbesserung der Fähigkeiten in der ersten Leitung Jahr beobachtet. Aus diesem bestimmten Modell sind Verbesserungen in SST prädiktive Fähigkeit insbesondere über den tropischen Pacific offensichtlich, was darauf hindeutet, dass die dynamische Kohärenz von Anfangsbedingungen kann in der Tat die prädiktive Fähigkeit der Klimavorhersagen zumindest in der ersten Leitung Jahr verbessern. Für längere Laufzeiten von 4-Jahresdurchschnitte wird ein großer Rückgang von SST prädiktiven Fähigkeiten fast überall sowohl für CIH und GIH beobachtet. Dies gilt vor allem über dem Nordatlantik (NA), wo frühere Studien zeigen, dass vorausschauende Fähigkeiten des SST mit der Vorhersagbarkeit der Ozeanzirkulation im Atlantik (AMOC) verbunden sind, wobei letztere führt die ehemaligen mehreren Jahren. Unsere Ergebnisse zeigen, dass signifikante prädiktive Fähigkeiten in Bezug auf AMOC an der ersten Führung Jahr sowohl für CIH und GIH erhalten, während bei längeren Liefer Jahren schlechte Vorhersagemann subpolaren Region Nordatlantik beobachtet, was darauf hindeutet, dass die Armen prädiktive Fähigkeit des NA SST ist teilweise auf schlechte Vorhersagbarkeit AMOC in unsere Lösungen.

Die Leistungen der CIH und GIH auch gegen frühere Ergebnisse aus dem Projekt Mittelfristige Klimaprognosen (MiKlip unten) mit Max-Planck-Institut Erdsystemmodells (MPI-ESM), der mit assimilierten Daten durch Antippen MPI-ESM gegen GECCO2 initialisiert werden ausgewertet. Der Vergleich zeigt, dass die MiKlip hindcasts von 1980-2006 übertrifft GIH, während seine Leistung mit der des CIH-kompatibel. Weitere Erkenntnisse bei jedem Startdatum ergeben sich Vorhersagen Geschick durch die Verwendung von Ensemble bedeuten, anstatt individuelle Umsetzung.

Die deutliche und hoch SST prädiktiven Fähigkeiten über den tropischen Pazifik in CIH zeigt eine gute Wiedergabe von El-Niño-Ereignissen an Blei Jahr ein. Im Gegensatz dazu produziert GIH weitere fehlerhafte El-Niño-Ereignisse, die für die Armen prädiktive Fähigkeit der SST über den tropischen Pazifik beitragen können. Während ausgewogene

Staaten besteht eine zonale Impulsbilanz zwischen Wind Stress und Druckgradientkraft der oberen äquatorialen Pazifik. GIH mit interpolierten GECCO2 Ozean Schätzungen und atmosphärischen Bedingungen, die von derselben gekoppelt Modell als in den Prognosen verwendete Modell führen initialisiert. Die Unterschiede zwischen den initialisiert Ozean und Atmosphäre Modi Inkompatibilitäten mit dem gekoppelten Modell, und führen zu Ungleichgewicht zwischen der zonalen Wind Stress und Druckgradientkraft über den äquatorialen Pazifik. Die reduzierte prädiktive Fähigkeit des GIH im Pazifik Äquatorregion wird daher festgestellt, dass vor allem auf dynamische Ungleichgewicht zwischen Zonalwind Stress und Druckgefälle über dem zentralen Pazifischen zusammen. Diese Ungleichgewichte in der zentralen Pazifik anschließend nach Osten ausbreiten als Kelvin-Wellen. Weitere Ungleichgewicht kann sich aus den Unterschieden in der Topographie zwischen GECCO2 und des Modellsystems zur Folge haben, und Störungen können zu ausbreitende Wellen geben. Gemeinsam sind diese Ungleichgewichte sehr wahrscheinlich zusätzliche Pseudo El-Niño-Ereignisse in GIH sowie die Armen prädiktiven Fähigkeiten über den tropischen Pazifik führen. Unsere Ergebnisse untermauern die Forderung nach einer Impulsbilanz zwischen Zonalwind Stress und Druckgradientkraft entlang der äquatorialen Pazifik bei der Initialisierung des Modells von jedem ozeanischen Staat. Initialisieren einer gekoppelten Modell mit selbstkonsistenten Ausgangsbedingungen verbessert daher die Fähigkeit des dekadischen Klimavorhersage im tropischen Pazifik.

Acknowledgement

I would sincerely thank my scientific supervisor Prof. Detlef Stammer for his support during my PhD study, for motivating my research and encouraging me on every scientific outcome. I am also thankful to my co-advisor Dr. Armin Köhl for all the advices and useful discussions on the model setup, brilliant ideas about the analysis and plenty of comments on the thesis. I am really lucky to have both as advisors for my PhD research, and I want to thank you again just for all. I also want to thank Prof. Jörn Behrens as chairman of my advisory panel, for keeping me under track of the PhD procedure, for useful discussions and plenty of advices on future career.

I want to thank Japan Agency for Marine-Earth Science and Technology for letting me use their couple model and the assimilated data as initial conditions in my experimental runs. Many thanks for their support in the model set up, and for the useful answers on all my questions about the experimental design. I also want to thank Mittelfristige Klimaprognosen (MiKlip hereafter) project group of Max Plank Institute for Meteorology for allowance of analyzing their outputs.

I am really happy to work with the colleagues of Remote Sensing and Assimilation Group of the Institute of Oceanography. They are all nice and warm-hearted, and I am really grateful to get so much help and advices from them, either on problems with codes or on scientific analysis. I would especially thank Iullia Polkova and Nicolay Koldunov for their help when I started my PhD here in Hamburg, for helping me use the ZMAW Linux system and get the first idea of PhD life. Very special thanks to Chuanyu Liu for his willingness to offer help and advices anytime, and his inspiration of finishing my work. I would also thank Ms. Karin Ziemann-Cohrs for preparing all the documents on my arrival, helping me get enrolled, and a lot of help on theses important issues with the University. I wish to thank Nuno Serra for his willingness to help and useful advices with the Ferret codes. Special thanks to Guokun for his patience in many discussions, and questions about my model setups. I would also like to thank Mark for proof reading the third chapter of my thesis, and for bringing me a lot of good ideas. Thanks shall also go to Xin just for everything. I would like to thank them all for so many suggestions and being patient in preparing the presentation for the defense.

I would also like to thank IT support group of CIS. All of my experimental runs are made on the server of the Deutsches Klimarechenzentrum (DKRZ). Many thanks for the support from them during the model simulations and analysis. I want to thank the Office of the graduate school SICSS for offering me opportunities getting to know other PhD students, organizing activities, financing German course attendance, and offering many helpful courses and retreats.

I am especially thankful to my dear friend Ursula Ullmann for her inspiration and motivation of starting my PhD study under Prof. Stammer in Hamburg, and for her endless advices and ideas about life in Hamburg. I would also like to thank my former supervisor Prof. Yuguang Liu, for his support and motivation in the application, and later on encouragement during my PhD study, and the comments on my thesis.

Many, many thanks to Yuan for everything: for every up and down that we went through, for every sorrow and delight that we shared, and most of all, for the support from him. Many thanks also go to my family back in China. I am especially thankful to my mother, for always having confidence in me and for her love and support. I would also like to thank my friends that I have shared my PhD life with: Xin, Feifei, Jian, Zhuhua, Fangxing, Huan, Hongmei, Chao and Mark. Special thanks to my former office mates Iullia and Reema. I wish also to thank my former colleague Eifu, for his kindness and a lot of useful information about visa application after graduation and possible careers.

Chapter 1

Introduction

1.1 Background

Large amount of climate variations observed in the past (Seager et al., 2008; Mehta and Lau, 1997; Barnett et al., 2008; …) and extreme events have made the information about near-term changes of climate variables a key factor to policy and decision makers. Under the requirement of adapting to time-evolving climate change and global warming, a new field referred to as "decadal prediction" was brought up by climate researchers in the early 2000s (Vera et al., 2009; Meehl et al., 2009; Taylor et al., 2008). It was meant to predict climate fluctuations for the nearest future (out to 10-30-yr, a time scale that decision makers are interested in, e.g. Pulwarty 2003, Barsugli et al. 2009; Means et al., 2010). Hence, this new field will bridge the gap between seasonal-to-interannual (SI) forecasting and the future climate change (long-term) projections. The importance of decadal prediction in governmental decision making, business planning (Vera et al., 2009) was recognized by the Intergovernmental Panel on Climate Change (IPCC) assessments. Much effort has been devoted by climate researchers in several countries and under partnership between organizations, aiming at providing reliable information of climate change in the coming decade or longer time scale (Meehl et al., 2009b).

In daily weather forecasts and shorter-term Seasonal to Interannual (SI) climate prediction, the impact from external forcing is negligible. However, anthropogenically-forced and naturally-forced climate changes shall not be ignored (Murphy et al., 2010) for climate prediction on decadal time scales, although the signal is often weaker than or comparable to the internal variability (Meehl et al., 2009; Murphy et al., 2010). Studies show that both the internal variability of climate change and its response to anthropogenic forcing offer potential for decadal predictability (Hwakins and Sutton, 2009a), although their importance is region-dependent. For example, Boer (2011) reported that anthropogenically-forced change is typically larger than internal variability over tropical oceans (except the equatorial Pacific), while over the mid- and high-latitude oceans the internal variability is more important since overall variability is high there. For a skillful decadal prediction, predictable signal in the initial state and forced climate change signal shall all be considered. Climate projections depending on the response to anthropogenic forcing from green-house gases and aerosols and external forcing from major volcanic eruptions are referred to as the "Boundary condition problems" (Meehl et al., 2009, Fig 1.1), while building the starting state of climate predictions based on the current knowledge of observed ocean/atmosphere variables can be addressed as "initial value problems". Hence decadal climate prediction is a joint initial-boundary problem.

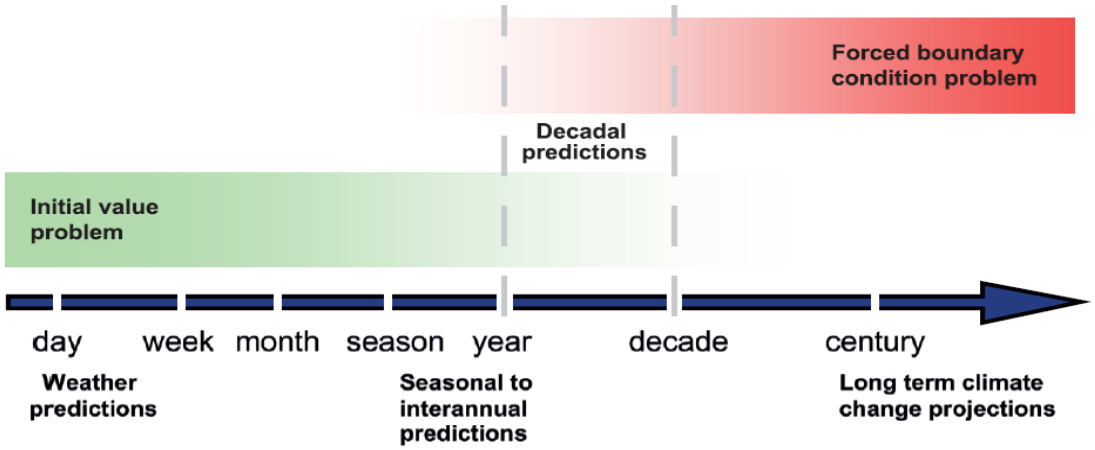

Fig 1.1 Schematic illustrating different portion of impact from initial value and boundary conditions on climate forecasts at different scale from IPCC (2013, based on Meehl et al. 2009). This figure shows us the progression with daily weather forecasts at one end (left), seasonal and decadal prediction in between, and multi-decadal to century projections as a forced boundary condition problem at the other end.

However, the importance of external forcing for decadal prediction was recognized only recently. At the early stage of decadal prediction, researchers treated decadal prediction as an initial value problem, following the experience from seasonal-to-interannual (SI) prediction investigation. These studies reported potential predictability of some aspects of the internal variability for a decade or longer, such as in the North Atlantic and North Pacific (e.g. Latif et al., 2006). Emerging evidence about possibility of skill achieved in surface temperature is also provided by decadal projections that initialize global coupled ocean-atmosphere climate models (Smith et al., 2007; Keenlyside et al., 2008). On the other hand, the skill of predictability on decadal scale in both regional and global surface temperature is improved through radiative forcing changes during the 20th century by study of Lee et al., 2006; Laepple et al., 2008. The new Coupled Model Intercomparison Project phase 5 (CMIP5) protocol extended the focus of their experiments to predictability /prediction over decadal scales which is significantly influenced by anthropogenic external forcing from changing concentrations of greenhouse gases and aerosols (Meehl et al., 2009b; Taylor et al., 2008). Up to date there are more and more studies follow the joint initial-boundary value approach in decadal prediction exploration (Pierce et al., 2004; Troccoli and Palmer, 2007; Smith et al., 2007…). Since very little skill is known about natural external forcing and no future information will be provided, the target of providing the best information on time-evolving climate prediction is basically directed to intrinsic climate variability and anthropogenically forced climate changes, which still remain a major challenge in climate science.

Attempting to provide reliable decadal climate forecasts (e.g. Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009), the elements affecting near-term predictions are highlighted (Hurrell et al., 2009). These particular elements can be summarized as the dynamical models to be used, and the initialization problems including the initialization strategy and the reanalyses constraining the initial state of the model (Bellucci, et al., 2012). By applying each of these elements in the prediction system, large amount of uncertainties are potentially introduced. Take the prediction of global-mean surface air temperature for example, the relative uncertainty changed with lead time from “initial state dependence” at

the first decade to "the forced response" out to around 40 years (Meehl et al., 2009, Fig 1.2). Meehl et al. (2009) reviewed some of the first results about decadal climate prediction published by that time, and summarized the highlights involved with decadal prediction as:

1) how to initialize decadal predictions and

2) how to evaluate the predictability and forecast skill.

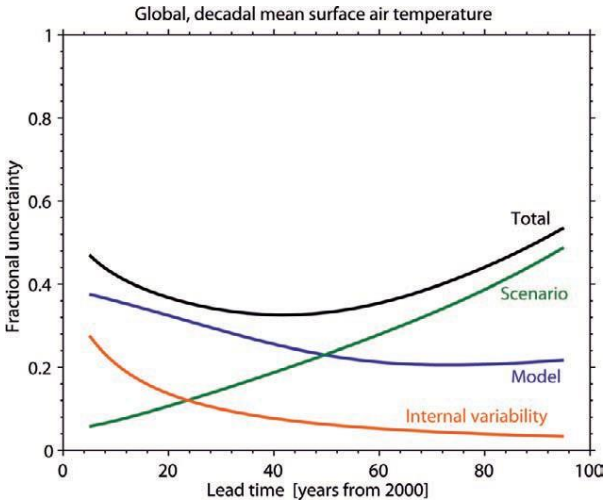

Fig 1.2 Schematic showing relative importance of uncertainty from different sources in decadal mean global surface air temperature by the fractional uncertainty (fractional uncertainty $= \frac{prediction\ uncertainty}{expected\ mean\ change}$) from Hawkins and Sutton (2009). Model uncertainty acts as the main source for uncertainty up to lead time 50yr. Internal variability is another important source of uncertainty for the first decade, while scenario uncertainty being important for multi-decadal lead time.

Afterwards, coordinated climate modeling exercises (e.g. ENSEMBLES; van der Linden and Mitchell, 2009; van Oldenborgh et al., 2012; García-Serrano and Doblas-Reyes, 2012) and considerable researches are carried out to advance the science of decadal climate prediction, making this field evolve rapidly. Much information about time-evolving regional climate predictions based on new results from initializing climate models (e.g. CMPI5; Taylor et al., 2012) is provided.

With the full-field initialization techniques, the model is started from the observational initial conditions. However, the model can never represent neither the oceanic nor the atmospheric physical process perfectly. There are always systematic errors within a climate model, and the model preferred climate is different from the observed climate. When initializing the model from the observations, the model state is brought close to the observations at the beginning of the integration. The imbalance between the model dynamics and the observed initial conditions may lead to errors, which is referred to as "initial shock" (Balmaseda and Anderson, 2009). As the simulation evolves, the model always drifts towards its own climatology (Pierce et al., 2004), leading to bias in the forecasts. From the experience of SI forecasts, a posteriori bias correction procedure is needed to remove such bias (Stockdale, 1997). However this practice can be very expensive because a large set of

ensembles is required in order to estimate the mean drift. Besides, a bias correction might be inappropriate due to the small magnitude of the predictable signal and the growth of nonlinearities (Meehl et al., 2009).

An alternative strategy, initializing the model with the observed anomalies added to the model climatology, was introduced and explored (Barnett et al., 2004; Pierce et al., 2004; Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009, Polkova et al., 2013). The so called "anomaly initialization" (Schneider et al., 1999) avoids the model drift. Hence the bias correction is less expensive compared with the time-dependent bias correction due to model drift. Instead, an estimation of model climatology is needed to build the initial conditions as well as an estimation of the bias. However, this strategy could not assure the reduction of initial shock at the beginning of the prediction due to the possibility of inconsistency between the model mean state and the observed anomaly (Magnusson et al., 2012a; Goddard et al., 2013), e.g. mismatch between the observational anomalies and the model climatology in sharp Gulf Stream gradient locations (Meehl et al., 2014). Besides, large errors introduced in the mean climate state affect internal variability (Schneider et al., 1999) and potential drifts are expected due to incorrect model response to forcing (Goddard et al., 2013).

To deal with the model drift, another strategy based on empirical corrections of heat, freshwater and momentum fluxes is tried by researchers (e.g. Rosati et al., 1997; Manganello and Huang, 2009; Polkova et al., 2014). This strategy named "flux correction" is to initialize a coupled model with full state observations, and additionally corrected with heat, freshwater or momentum flux correction during the model integration, i.e. full state initialization employing flux correction (Polkova et al., 2014). By starting the model close to the observed climatology and constraining the model with empirical corrections of fluxes, this strategy offers another possibility to deal with model drift and solve the model systematic errors. However, disadvantages still emerge with this strategy such as the difficult construction of flux correction for future climate. The initialization technique of the modeled ocean atmosphere models still remains a chief challenge and further enhancement is highly required (Meehl et al., 2009b, 2014; Murphy et al., 2007). Magnusson et al. (2012a) compared two methods of full-state and anomaly initialization, and showed that higher predictive skill on seasonal time scale was provided with the former, while study by Smith et al. (2012a) derived more regional skillful predictions in hindcasts with the latter. A study by Hazeleger et al. (2013b) revealed that there was no significant difference in decadal predictability between the two different methods. Polkova et al. (2014) compared these three approaches in decadal forecasts with a coupled ocean-atmosphere model. The highest predictive skill for SST is obtained for flux corrected hindcasts in regions of a deep mixed layer because of their smallest bias. But this strategy hasn't been used by any of the atmosphere-ocean general circulation models (AOGCMs) in CMPI5 (Meehl et al., 2014). Right now most of modeling groups tend to prefer full-state initialization. To find out the best initialization strategy, further evaluations are needed with more models (Meehl et al., 2014).

Besides the strategy of initialization, the initial conditions are also important in decadal prediction, which are used to constrain the initial state of the model with a realistic representation of the climatic system (Bellucci et al., 2012). So far the most commonly used source of initial conditions is estimations of climate components (e.g. the ocean, atmosphere) incorporating observational information, which are derived from atmospheric or oceanic

general circulation models. Predictability due to initial conditions mainly arises in the slow components of the climate system, most notably the ocean, despite the necessary but not sufficient low-frequency component of climate variability (Boer 2004; Pohlmann et al, 2004). Hence a lot of studies only initialize the ocean component (Keenlyside et al, 2008; Pohlmann et al., 2009; Matei et al., 2012c; Carton and Santorelli, 2008; Stammer et al., 2009). Modeling groups participating in CMPI5 initialize the models with estimation of assimilated ocean and/or atmospheric observations with partial/coupled models (Meehl et al., 2014). However, to offer the most forecast skill, all the elements of the climate system shall be considered, including the atmosphere, the ocean, sea ice and so on (Doblas-Reyes et al., 2011a; Matei et al., 2012c; Smith et al., 2007; Troccoli and Palmer, 2007).

The study by Sugiura et al. (2008) proposed that by assimilating the historical observations (ocean, atmosphere, ice and land surface) through a coupled ocean-atmosphere model, better initial conditions can be derived, which may contribute to reduced initial shock as mentioned above. The coupled data assimilation schemes (four-dimensional variational (4D-VAR) data assimilation system) takes advantage of covariance between ocean and atmosphere variables, and may offer the most forecast skill in decadal prediction (Meehl et al., 2009). This may offer potential possibility in improving the predictability of climate prediction at decadal time scale.

Another important issue remained unresolved about decadal climate prediction summarized by Meehl et al. (2009) is "how to evaluate decadal prediction skill". The evaluation of prediction skill in hindcasts is necessary either for analyzing the spatial-temporal quality of decadal predictions or for a lower bound for the predictability of the system (e.g. Hawkins et al., 2011; Goddard et al., 2012a; Wang et al., 2012). In the "perfect model experiments", the ability of the coupled model to reproduce itself given the uncertainties in initial conditions is usually assessed, rather than assessing the predictability of the model system by comparing the model simulations to observations (Boer, 2000; Branstator and Teng, 2010; Teng et al., 2011). The "diagnostic" potential predictability approach distinguishes regions with potential predictability through variance analysis (Boer, 2000, 2004, 2011; Pohlmann et al., 2004; Boer and Lambert, 2008). Besides the above-mentioned approaches of predictability assessment, one of the widely accepted approaches is to qualify the initialized forecasts through deterministic metrics such as the root mean square error (RMSE) and the anomaly correlation coefficient (Smith et al., 2007; Matei et al., 2012c; Goddard et al., 2012a). As is mentioned above, by introducing each element to improve near-term predictability, uncertainties are brought in the prediction. Multiple diagnostics of performance that includes uncertainty identification, possible correlation and limitations in the model's performance works as a better verification approach (Tebaldi and Knutti, 2007). Probabilistic metrics that measures improvement of probabilistic forecast relative to reference forecast proposed by Goddard et al. (2013) and the "Bier skill score" used by Corti et al. (2012) and Hazeleger et al. (2013a) are alternative framework for decadal prediction. The multi-model approach, which shows advantage for the seasonal to interannual predictions and long-term climate projections (Palmer et al., 2004; Tebaldi and Knutti, 2007), offers possibility of narrowing down the uncertainties in decadal prediction through sampling structural differences among climate prediction systems (Bellucci et al., 2012). Its influence on decadal prediction has been explored by climate projections (ENSEMBLES, THOR, CMIP5 coordinated effort) and is recommended for reducing the uncertainty of the near-term climate predictions.

To sum up, much effort has been devoted by climate researchers to improve the skill of decadal climate prediction and to estimate the predictability of the prediction system. So far the initialization is thought to offer potential decadal predictability for important climate variables due to the internal natural variability, including initial conditions and initialization strategy. It is highly possible to improve predictive skill through initialization of more climate components such as sea-ice, land surface. The application of coupled data assimilation scheme is supposed to offer skill in decadal prediction, due to the generation of optimal estimation of climate system (Meehl et al., 2009). With the coupled ocean-atmosphere data assimilation system, remarkable skill into seasonal to interannual process was found by Sugiura et al. (2008). Considering the advantage in SI climate prediction, it's interesting to explore the performance of this system in decadal prediction. In this paper, the coupled model will be initialized with full-state initial conditions resulting from the assimilation of observations into the same coupled climate model. Since the estimation of initial conditions of both the atmosphere and the ocean is derived from the model used for the hindcasts, the systematic errors and model drift due to different dynamics and climatology between the initial states and the model are reduced. Hence, improved predictive skill is expected. This will be an important improvement in decadal prediction considering the possible relatively large magnitude of the bias adjustment compared with the signals predicted (Kharin et al., 2012; Kim et al., 2012).

In climate predictions, it is agreed among researchers that self-consistent initial conditions will improve predictive skill. However, a self-consistent initial condition has never really been used in decadal climate prediction, and the influence of model consistency has never been explored. To better investigate this, the coupled climate system will also be initialized with different initial conditions that are not consistent with the model as a comparison. Possible mechanism responsible for predictive skill will also be investigated.

1.2 Goals

In this thesis we aim to investigate the influence of the model consistency on the predictive skill of decadal prediction with JAMSTEC's updated version of the coupled CFES climate system. Full states 3-dimensional oceanic variables from two different datasets will be used to initialize the coupled model: the CDA (Coupled Assimilation Data) that is estimated from the same coupled CFES model, and the GECCO2 Synthesis. An ensemble of twentieth-century simulation (un-initialized hindcast) will also be done in order to provide insight into the influence of initialization. Based on the two sets of hindcasts, the following work will be done:

1) Assessment on predictive skill due to initialization will be carried out by comparing the performances of the initialized hindcasts with that of the un-initialized forecast (i.e. twentieth-century simulation) in terms of sea surface temperature (SST), as well as with that of the low-skill persistence forecast. The root mean squared error skill scores (RMSS) and the anomaly correlation skill will be used for this purpose. The comparison between the performances of the two sets of differently initialized hindcasts will provide us insight into whether model consistency gives rise to predictive skill and if so, which region resides the highest improvement of predictive skill and what is the duration? The investigation of the skill will be carried out with respect to CDA/GECCO2 and observations respectively. AMOC predictive skill will

also be evaluated, which is associated with SST predictive skill in the North Atlantic, with the former leading the latter several years as is indicated by former studies.

2) To evaluate the predictive skill of the system in our solution with previous results, the performance of our hindcasts will be compared with the MiKlip hindcasts derived by initializing a different model (MPI-ESM) with GECCO2 Synthesis, in terms of anomaly correlation skill for SST/AMOC. Restricted by the available CDA Synthesis from 1980-2006, the evaluation of skill is limited to 27 realizations, whereas previous studies typically use 50-years period. Therefore, the impact of the number of initial dates on decadal prediction will be investigated first. For this, ensemble mean of MiKlip hindcasts of the period 1980-2006 (which is the same as the period available from CDA) will all be compared with that of 1961-2009, given that the MiKlip hindcasts have larger number of initial dates, and each initial date with three realizations. The comparison between performances of ensemble mean of MiKlip hindcasts of 1980-2006 and that of CIH/GIH will provide information of impact from more realizations at each initial date on decadal climate predictions.

3) To provide useful information of decadal climate predictions, further exploration of the possible improvement/decrease in SST/AMOC predictive skill will also be carried out. The possible underlying mechanism giving rise/decrease in predictive skill will be investigated based on the two differently initialized hindcasts. If a decrease of predictive skill is observed in the initialized hindcasts, possible solutions will be discussed.

1.3 Outline of the thesis

The following part of this thesis is divided into four chapters:

Chapter 2 provides information about the methodology for exploring the influence of model consistency on decadal climate prediction. The coupled model CFES, CDA and GECCO2 Synthesis used as oceanic initial conditions, application of trend removal, bias-correction, predictive skill verification and significant test are all described in this chapter. The model description and experimental setup of hindcasts from MiKlip are also included.

Chapter 3 explores the influence of model consistency on decadal climate prediction by comparing the performances among differently initialized hindcasts, the un-initialized hindcast (20C simulation) and the low-skill persistence forecast. Predictive skill of SST and AMOC in terms of root mean square error score and anomaly correlation skill is evaluated in three time scales: lead year 1, average over lead year 2-5 and 6-9. Regional and duration of SST/AMOC predictive skill are also investigated. Predictive skill of the coupled system in our solution is also evaluated against previous results with relative long hindcast runs. The influence of the number of initial dates is studied through comparison of SST/AMOC predictive skill among our initialized hindcasts and MiKlip hindcasts in this chapter. Comparison between ensemble mean of MiKlip hindcasts with different length of hindcast runs gives us insights into how more realizations impact predictive skill.

Chapter 4 investigates the underlying mechanism of possible improvement or decrease of predictive skill from hindcasts initialized with self-consistent or non-self-consistent initial conditions. For this purpose, the climate properties that are most possibly related with SST or AMOC predictability on decadal time scale are analyzed. Solutions to the possible decrease of predictive skill are also discussed in this chapter.

Chapter 5 summarizes the main findings and outlines the future work.

Chapter 3 and 4 of this thesis will be rewritten into a paper. This will be done after the defense.

Chapter 2

Methodology for exploring influence of model consistency on decadal prediction

The idea is to initialize the coupled climate system with 1) the optimized initial conditions (both atmosphere and ocean), which were obtained by assimilating data into the same coupled assimilation system with the 4 dimensional variational method referred to as the Coupled Data Assimilation (CDA) Synthesis, 2) the re-gridded ocean synthesis from the project "German contribution to Estimating the Circulation and Climate of the Ocean" (GECCO2, Köhl and Stammer, 2008; Köhl, 2014), as well as an un-initialized run (i.e. the twentieth-century simulation). The hypothesis is that with the dynamically self-consistent initial conditions and control variables, the model will suffer less from initial shocks and artificial model drift and produce more skillful predictions than initialized with GECCO2. Hence, assessment of the region and duration with skill through variables such as Sea Surface Temperature (SST) and Atlantic Meridional Overturning Circulation (AMOC hereafter) of different experiments will be offered at first. Comparison against the control run will provide an evaluation of the improvement of predictive skill through initialization.

A second step is to compare the performance of different initializations in order to find out which has advantage over the other and the reasons. As a necessary requirement before the influence of different initial conditions can be investigated, the skill of the system needs to be evaluated against previous results. Since the CDA Synthesis provides data only for the period 1980-2006, the evaluation of skill is restricted to little more than 20 realizations, which is only half the size of the 50-years period that typically has been used in previous studies for evaluating the skill of hindcasts. In order to be able to compare with previous skill, estimating the impact of the length of the period of the numerical runs that provide initial conditions and that can be used for evaluation of the hindcasts on decadal prediction has to be investigated first. For this, full-state initialized hindcasts performed as part of the Mittelfristige Klimaprognosen (MiKlip hereafter) project based on the Max-Planck-Institute Earth System Model (MPI-ESM) and initialized from the GECCO2 Synthesis are compared to the hindcasts presented here. Hindcasts from the MPI-ESM cover the period 1961 to 2009. Hence, a comparison over the same period and the investigation of the influence of the length of the period is feasible. With the more realizations of hindcasts, better statistics are expected for MiKlip hindcasts. Besides, better predictable skill is also seen for the persistence forecasts due to long-term variability. These two hypotheses will be assessed through comparison between the performances of initialized hindcasts.

The following part is about the details of different methods applied in exploring the questions addressed above and the models involved. Note the model runs described in 2.1.2 are developed by MPI-M and the hindcasted data are only used in this thesis under permission. The outline of the methodology involved includes: 1) a coupled climate model CFES developed by JAMSTEC; 2) a set of full-field initialized hindcasts by MPI-M in comparison; 3) verification metrics for decadal predictability

2.1 Description of the climate models

2.1.1 Model used in decadal prediction experiments

To explore the skill of decadal prediction, the coupled model CFES (Coupled GCM of Earth Simulator) developed by Japan Agency for Marine-Earth Science and Technology (JAMSTEC) is used in this study. It is composed of the Atmospheric GCM for the Earth Simulator (AFES) and the Ocean-Sea Ice GCM for the Earth Simulator (OIFES) (Ohfuchi et al., 2004; Masuda et al., 2006; Sugiura et al., 2008). The AFES component is the atmospheric GCM developed by the Center for Climate System Research of the University of Tokyo/Japanese National Institute for Environmental Studies (CCSR/NIES), with the radiation code updated using MstrnX (Nakajima et al., 2000) and a simple diagnostic calculation of marine stratocumulus cloud cover (Mochizuki et al., 2007b) implemented together with a land surface model called MATSIRO (Takata et al., 2003). With an improved turbulent closure ocean mixed layer scheme developed by Noh (2004) embedded in, the OIFES component was developed from version 3 of the Modular Ocean Model (MOM3) (Pacanowski and Griffies, 1999) and the sea ice model of the International Arctic Research Center (Hibler, 1980). The adjoint codes of the AFES and OIFES component models were obtained using the Tangent linear and Adjoint Model Compiler (TAMC) (Giering and Kaminski, 1998) and the Transformation of Algorithms in Fortran (TAF) (Giering and Kaminski, 2003). The resolution of the AFES component is horizontally the same as the commonly-used T42 spectral model and vertically 24 layers in σ coordinates and that of the OIFES component is 1 degree in both latitude and longitude, with 45 vertical layers (Sugiura et al., 2008).

In the system, the bulk adjustment factors of the latent heat, sensible heat, and momentum fluxes were chosen as the control variables determining air-sea-land surface fluxes, together with the oceanic initial conditions of the model variables, inspired by study of Fedorov and Philander (2000) on the significant role of subsurface temperature in the upper ocean that regulates the basic state of a coupled ocean-atmosphere field. This model was used for the assimilation in order to get the optimization of bulk adjustment factors and oceanic initial conditions (Sugiura et al., 2008).

OIFES (the Ocean-Sea Ice GCM for the Earth Simulator)

Together with the sea ice model of the International Arctic Research Center (Hibler, 1980), an updated version of the GFDL Modular Ocean Model (MOM3) is used for the coupled system here (Pacanowski and Griffies, 1999; Masuda et al., 2003, 2006). Designed and developed by researchers at the Geophysical Fluid Dynamics Laboratory (GFDL/NOAA Department of Commerce), the state of the art numerical ocean model MOM 3 is a widely used tool in studying spatial and temporal ocean. It is a finite difference model that realizes solutions of the ocean primitive equations with numerical methods, which govern much of the large-scale ocean circulation. The model physics is implemented with several sophisticated schemes, such as the nonlocal K Profile Parameterization for mixed layer physics, the Gent and McWilliams scheme for isopycnal mixing and the Held and Larichev and Visbeck, Marshall, Haine and Spall closures for the tracer diffusivities, as well as an improved turbulent closure ocean mixed layer scheme developed by Noh (2004). Embedded with a rectangular grid system, the model has a global domain with horizontal resolution of 1° in

both latitude and longitude, and 45 vertical levels spaced from 5m near the sea surface to 5261m at the bottom (vertical layers with thickness of 10m until a depth of 205m, then gradually increased to around 470m of bottom). With the sophisticated model and finer resolution than previous version, this model is supposed to form a better platform suitable for the 4D-VAR adjoint model due to its capability of reproducing ocean circulation processes. For more detail of the MOM3 model, please refer to the manual by Pacanowski and Griffies (1999).

-----AFES (Atmospheric GCM for the Earth Simulator)

The AFES is version 5.4.02 of an Atmospheric General Circulation Model (AGCM) developed under cooperation between the Center for Climate System Research (CCSR) of the University of Tokyo and the Japanese National Institute for Environmental Sciences (NIES) to run on the Earth Simulator (ES) manufactured by the Nippon Electronic Company (NEC) (Ohfuchi et al., 2004). It is a σ-coordinate, 3-dimensional global model that integrates the primitive equation (under the hydrostatic approximation,) with a semi-implicit method temporally. The vorticity-divergence form with the spectral Eulerian advection scheme is employed in the governing equations (Krishnamurti et al., 1998). The model's discretization is taking advantage of the spectral transform method horizontally (Orszag, 1970) and the Lorenz differenced vertical σ –coordinate (Phillips, 1957). An accurate radiative transfer scheme developed by Nakajima and Tanaka (1986) is used for the physics parameterization. The model employs a simplified Arakawa-Schubert convective scheme (1974) and Mellor-Yamada level-2 vertical diffusion scheme (1974). The resolution of the model is a triangular truncation at wavenumber 42 horizontally (T42, or about 2.8° by 2.8°) and 24 levels with the top level placed at about σ = 0.003 (about 3 hPa) vertically.

-----An updated version of JAMSTEC coupled climate model for decadal prediction experiments

For the decadal hindcasts, an updated model system forced with historical radiative forcing conditions including greenhouse gas (GHG), aerosol and volcano, is used, based on the radiation code MstrnX (Nakajima et al., 2000). Zonal averaged 2D-xy grid annual-mean GHGs (CH4, CO2, N2O) and monthly volcano are taken from historical and RCP4.5 scenario-based data in CMPI5, both vertically averaged. As is commonly used in global warming simulations, radiative effects of volcanic ash (historical major volcanic eruptions) are taken into account. Instead of directly using concentration of volcanic ash, additional optical thickness for a specific band of radiation spectrum at the lowest level of the model stratosphere is used. Aerosols (black and organic carbon, dust, sulfur) used here are column integrated monthly-mean historical and RCP4.5 scenario-based simulations in CMIP5. Since the version of MstrnX used does not calculate chlorofluorocarbon (CFC) effect, the CFC values are set to zeros and ozone is set to 3-D monthly climatology.

2.1.2 Model used for the MiKlip experiments

In order to evaluate the performance of the experiments from the CFES climate system, full-state initialized hindcasts performed as part of the Mittelfristige Klimaprognosen (MiKlip

hereafter; Müller et al., 2012; Pohlmann et al., 2013) project based on the Max-Planck-Institute Earth System Model (MPI-ESM; Giorgetta et al., 2013; Stevens et al., 2013; Jungclaus et al., 2013) are also used for further analysis considering different initialization and length of the covering period of the numerical runs. The coupled model is composed of the Max Planck Institute ocean-sea ice model (MPI-OM) for the ocean (Marsland et al., 2003) and European Centre-Hamburg model version 6 (ECHAM6) for the atmosphere. Observational estimate of the ocean state derived through nudging the MPI-OM to the oceanic 3-D temperature and salinity from GECCO2 Synthesis for the period of 1960-2009 is assigned as the oceanic initial condition of the coupled model at the start of each simulation. The atmospheric component is initialized with full-field 3-D spectral temperature, vorticity, divergence, and surface pressure fields (Roeckner et al., 2003), represented by a truncated series of spherical harmonics with triangular truncation 63 (T63), with the data from ECMWF Re-Analysis (ERA)-40 (Uppala et al., 2005) for the period 1960-1989 and ERA-Interim (Dee et al., 2011) for the period 1990-2012 respectively. The coupled model MPI-ESM is first nudged towards GECCO2 to generate initial conditions for the hindcasts. With the full-field initialization technique, these fields are used to initialize the decadal hindcasts, starting yearly around 1 January over the period 1960-2009, each with an ensemble simulation of three experiments. The numerical run are realized with lagged 1-day initialization, where again restart files from consecutive days are used for the initial conditions. The MPI-ESM in low resolution (LR, atmosphere: T63L47, ocean: 1.5 degrees, 40 vertical layers) is employed for this set of simulations (Pohlmann et al., 2013).

2.2 Experiments

The CFES climate system was originally designed to run with the JAMSTEC NEC vector machine (Earth Simulator 2). For all the experiments done in this thesis with the DKRZ scalar machine, an updated version of CFES embedded with external forcing is used. The configuration is modified to fit the DKRZ parallel system. Outlines of this chapter are listed as below:

- The twentieth-century integrations (20C) with the updated coupled climate system CFE (externally forced) as reference for robust evaluation of initialized hindcasts

- Two sets of differently initialized hindcasts with the updated coupled climate system CFES (externally forced) and comparison ensembles of initialized hindcasts with MPI-ESM from Max Planck Institute for Meteorology

Table 2.1 Summary of the experiments

Experiments	Initialization	Forecast period	Initial condition	Forcing	realization
20C*	Jan of 1946	1946-2007	1980 CDA*	GHG, Aerosol, volcano	1
CIH[a]	Jan of each year	1980-2007	CDA, full state	GHG, Aerosol, volcano	1
GIH[b]	Jan of each year	1980-2007	GECCO2, full state	GHG, Aerosol, volcano	1
MiKlip[c]	Jan 1st of each year	1961-2009	GECCO2, full state	GHG, Ozone, volcano	3

a) CIH: CDA initialized hindcasts
b) GIH: GECCO2 initialized hindcasts
c) MiKlip: GECCO2 initialized hindcasts provided by Max Planck Institute for Meteorology

* Due to technique problems, the 20C run is initialized with initial conditions from CDA of year 1980. The first 33 years are treated as spin-up run to get rid of the effects from initialization.

2.2.1 Twentieth-Century Simulation

An integration of the 20C simulation is performed with the updated version of JAMSTEC's CFES system, which includes external forcing (including GHG, aerosol concentrations from historical volcanic eruptions; GHG shown in Fig 2.2.1). This uninitialized climate forecast intends to capture only the forced response to changing atmospheric components. Hence, the 20C can be used to verify whether initial conditions are playing a more significant role in improvement of predictive skill. For this experiment, all the control variables are set to unity, and the 1980's oceanic restarts are taken as the oceanic initial conditions due to technique problems with the model. Therefore, the first 30-years are treated as a spin-up phase to allow the model to get used to the parameter changes and get rid of the impact of oceanic initial conditions. The modelled fields of the period after 1980 (1980-2006 in this thesis) are used to investigate predictive skill resulting from the radiative forcing. Since the initialized forecasts are intended to capture both the forced response to radiative forcing and the internally generated climate variability, we can investigate whether initialization provides useful information on the time evolution of internally generated climate variability, by comparing the 20C integration with the initialized hindcasts (Meehl et al., 2014).

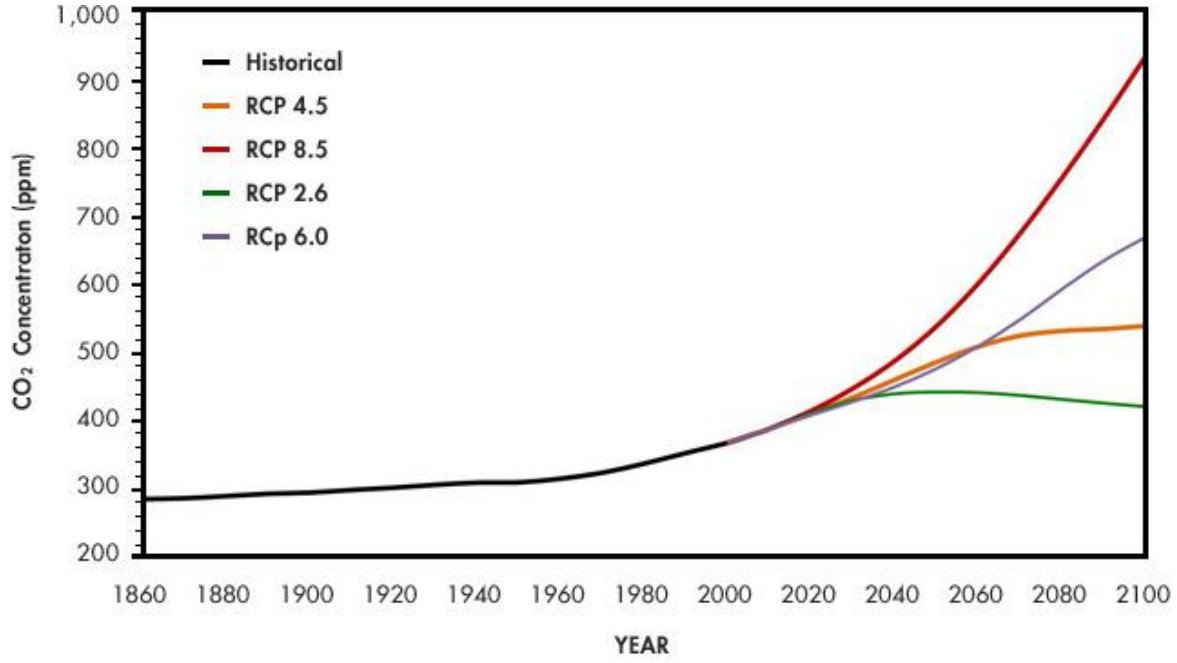

Fig 2.2.1 The time evolution of global CO2 concentrations (ppm) from HadGEM2 Model Development Team (2011) supported by CMIP5 for historical period (1860-2005) and for four RCPs (2006-2100). In our hindcasts (CIH, GIH and 20C), the historical (black) and RCP 4.5 (orange) CO2 concentrations from 1946 to 2045 are used. (Figure available from website: http://cordex-ea.climate.go.kr/main/modelsPage.do)

2.2.2 Initialized hindcasts

Two sets of hindcasts are initialized in January of each year for the period of 1980-2006, with the full state initialization strategy. By comparison, insights regarding the influence of initial conditions (or model-consistency) can be derived. One set is initialized with oceanic initial conditions from the coupled data assimilation synthesis (CDA hereafter) obtained by JAMSTEC through the 4-dimensional variational method. Another set of hindcasts is initialized with the 3-dimentional temperature, salinity, velocity, and sea surface height from GECCO2 Synthesis for the ocean of the same period. Since the GECCO2 Synthesis has a different grids and topography, the data are interpolated to the CDA grid. The atmosphere is controlled by air-sea surface fluxes through the optimal estimation of bulk adjustment factors (hence control variables of sensible heat, latent heat, and momentum fluxes), which are also assimilated through the same CFES model by JAMSTEC. The control variables are set to climatological values of the target periods for both sets of hindcasts, i.e., the two sets of hindcasts use the same atmosphere fields. According to the core experiments of CMIP5 regarding decadal predictability and prediction, a series of 10-year hindcasts with initial observed climate state every five years (starting near 1960) is required in order to estimate the theoretical limits and evaluate the ability in decadal prediction. However in this study, limited by the length of CDA analysis dataset and computational time of the DKRZ machine, all the numerical experiments are initialized every year on January from 1980 to 2006. They all have one ensemble experiment lasting for 9 years.

2.3 Initialization

As a key point in decadal prediction, initialization is believed to offer potential predictability of climate change on decadal time scale. How to initialize a climate model is one of the highlights in decadal prediction. Modelling groups have been exploring different methodologies and techniques to initialize decadal climate predictions. Up to date, there are three popular initialization strategies: the "full-state initialization", the "anomaly initialization" and full-state "fluxes correction" (full-state employing flux correction). These strategies are applied to deal with the model's drifting away from the observed initial state. However, none of these strategies have succeeded in minimizing the drift, as is discussed in Section 1.1. Strategies aiming at minimizing the bias and overcoming the shortage of observations have also been made by perturbing initial conditions to generate small ensembles of hindcasts (Smith et al., 2007; Keenlyside et al., 2008). Such strategies are introduced because climate predictions are sensitive to small perturbations in the initial state (e.g. Du et al., 2012). These efforts include perturbations with breeding vectors (Toth and Kalnay, 1997; Vikhliaev et al., 2007; Ham et al., 2014) or singular vectors (Molteni et al., 1996), initialized the forecasts with different start days of the atmosphere state (e.g. Yeager et al., 2012). Another alternative is the ensemble assimilation method developed to evaluate the initial state errors and their effects on the growth of forecast uncertainties (e.g. Evensen, 1994; Karspeck et al., 2013). The multi-model approach (Hagendorn et al., 2005; Zhang et al., 2007a; Tebaldi and Knutti, 2007; Meehl et al., 2007; Murphy et al., 2010), perturbed physics approach (e.g. Stainforth et al., 2005; Collins et al., 2006; Murphyet al., 2007) and stochastic physics approach (e.g. Berner et al., 2008) have also been used to construct ensembles and they showed similar levels of skill (Doblas-Reyes et al., 2009). Despite this diversity of approaches aiming at the initialization and ensemble problem, the full-state initialization still remains the most popular strategy among the majority of modeling groups. Studies show that using the full-state of the ocean initial conditions from an assimilation run with the same coupled model is another possibility to offer high skill in decadal prediction (Smith et al., 2007; Keenlyside et al., 2008; Pohlmann et al., 2009; Mochizuki et al., 2010; Chikamoto et al., 2012). In this thesis, we choose the full-state technique of initialization for the predictions. Details about the initialization will be described below.

2.3.1 Initial conditions

- **CDA (coupled data assimilation synthesis) from JAMSTEC**

In this thesis, the optimal estimations of climatological control variables and oceanic initial conditions, derived from the coupled ocean-atmosphere model (Sugiura et al., 2008) CFES by JAMSTEC, which is referred to as CDA, are used as initial conditions for the set of hindcasts CIH. CDA was assimilated through 4D-VAR assimilation run of the coupled model. Aiming at improve the predictability of ENSO events, the assimilation was done through a 9-month assimilation window cored from January to September and July to the following March, with 1.5-month margins set at both ends of each window. The first-guess fields for all the 9-months

assimilation experiments were derived by initializing the model with a 30-year-long spin-up run, for which the incremental analysis updates (IAU) method incorporating reanalysis data of temperature and salinity by Masuda et al. (2003) was applied. Based on the first-guess fields and a series of 10-day mean observational data (temperature and salinity from the Fleet Numerical Meteorology and Oceanography Center Data set, OISST values, and TOPEX/Poseidon altimetry sea surface dynamic height anomaly, monthly temperature and salinity from World Ocean Database 2001), the control variables of the air-sea surface fluxes and oceanic initial conditions were adjusted in each time window via an optimal synthesis. The conjugate gradient method was used to solve the 4D-VAR optimization problem by iteration. For higher latitudes north of 60°N and south of 60°S CDA didn't assimilate the observation. The constraint of assimilation is stronger in the tropics than in the extra tropics.

Study based on CDA reported improvement on seasonal to interannual climate predictability and possibility of corrected initial location of the model climate attractor on the basis of observational data (Sugiura et al., 2008). Further analysis showed improvement in prediction of the 1997-1998 El Niño event. Another reason of taking CDA as the initial condition is that the initial condition is highly self-consistent with the model since the assimilation used the same coupled atmosphere and ocean model. By doing this, the initial shock which usually arises from the difference between observational and model-preferred climatology, is possibly to be reduced. Hence, the optimal estimation of oceanic initial states and climatological control variables are used to initialize the same coupled-model. Note, the assimilation run that produced CDA by JAMSTEC was done with the vector machine Earth Simulator 2 under the original set-ups, without external forcing included. For the hindcasts in this thesis, an updated version of the coupled climate model embedded with prescribed historical external forcing conditions (e.g., greenhouse gases, aerosols, and major volcanic eruptions) is used. Hence, the impact of forcing on decadal climate prediction is considered. This strategy is supposed to offer potential possibility in improving climate prediction of past case on decadal time scale. The data has a resolution of 1° by 1° horizontally and 45 layers vertically for the ocean, and covers the period of 1980-2006.

- **GECCO2**

In this thesis the GECCO2 ocean state estimation is also used as the ocean initial conditions for another set of hindcasts with the JAMSTEC coupled system CFES. As a new version of ocean synthesis available from the German contribution to Estimating the Circulation and Climate of the Ocean project (Köhl and Stammer, 2007, 2008; Köhl, 2014), GECCO2 covers the period of 1948-2013. The variables used for the ocean initial condition are 3-D temperature, salinity, velocity (zonal and meridional), and SSH. GECCO2 uses the MIT GCM (Massachusetts Institute of Technology general circulation model, Marshall et al., 1997) model, and has a horizontal resolution of 1° by 1° with 50 vertical layers. A higher meridional resolution of 1/3° is used at the Equator, as well as approximately isotropic cells between 20° and 60° (Köhl, 2012). Different from GECCO synthesis (GECCO doesn't include a sea ice model and therefore ends at 80°N), in GECCO2 the Arctic Ocean with roughly 40 km resolution is included and a dynamic/thermodynamic sea ice model is used. It's forced by the NCEP RA1 atmosphere fields and bulk formulae. Through a 4D-VAR (adjoint) assimilation scheme, the model is brought into consistency with the observations and available satellite

data as well as estimation of prior surface fluxes momentum within 28 iterations for 1948-2011 and 37 iterations for 2008-2013. GECCO2 combines most of the World Ocean Circulation Experiment observations available for the synthesis period and provides better consistency with the data than the control run. The Denmark Strait overflow in GECCO is replaced by water mass transformation in the subpolar gyre (Köhl, 2014). Analysis on GECCO synthesis already showed improvement of decadal predictability in the AMOC at 25°N considering the long-time changes (Köhl and Stammer, 2007, 2008). The enhanced version of GECCO2 will also work as a good source of initial conditions for decadal predictability, especially for improvement in AMOC and NA SST prediction skill (Köhl et al., 2009; Pohlmann et al., 2013; Polkova et al., 2014).

2.3.2 Full state initialization

Inspired by experiments from SI climate prediction, the state-of-art technique that starts the model with an ocean state, which is close to the observations, is an important initialization strategy. The advantage of this strategy is that it starts the model from a state closer to actual observed climate state (Troccoli and Palmer, 2007; Magnusson et al., 2012a, b; Smith et al., 2013). However, at the beginning of the numerical run, initial shock due to the imbalance between initial conditions and the model dynamic may lead to errors in the forecast (Balmaseda and Anderson, 2009). As time evolves, systematic errors due to quantification of oceanic and atmospheric physics develop and forecast loses predictability (Pierce et al., 2004). As a result, the forecasts drift way towards the model's climate. Besides, computational limits, lack of long period observations also lead to unavoidable biases in the forecasts at decadal time scale. Hence, with the application of full state initialization, reducing the initial shock and systematic errors shall give rise to predictability.

In this study, all the decadal hindcasts are initialized with full-field monthly mean ocean state from CDA/GECCO2 that has a realistic representation of the ocean states. The properties used as the oceanic initial conditions are 3-D monthly mean temperature, salinity, velocity (both zonal and meridional), and sea surface height. Since the GECCO2 uses a different grid (vertical, zonal and meridional near the equator) and topography from the coupled CFES system, the GECCO2 fields are expanded with the nearest grid value towards the continent and bottom of ocean first. Afterwards linear (over vast ocean) and nearest (along the coast lines) interpolations are used to put the GECCO2 data into the model grid. In order to get rid of the dynamical adjustment to differences in topography, the ocean state after a 3-day forward run is taken as the initial condition for the experiments.

Through initializing the model with the optimal estimation of ocean states and atmospheric control variables assimilated from the same coupled model, the initial shock and systematic error are supposed to be reduced. The consistency between the model used for estimating the oceanic initial conditions and that used for model simulations makes the representation of the oceanic and atmospheric processes more dynamically self-consistent. Considering these advantages of the CDA initialized hindcasts (CIH), higher predictive skill on decadal time scale is supposed than forecast experiments initialized with GECCO2 Synthesis that is derived from a different coupled model. Evaluation of the predictive skill of the two sets of

hindcasts will be carried out against observational data, uninitialized 20C hindcasts and the low-skill persistence forecast.

2.4 Assessment skill on decadal predictability

Primary processing of hindcasts

Due to the configuration of the JAMSTEC's CFES coupled climate system, all the oceanic outputs of hindcast are 10-day mean data. The output of each month contains three 10-day mean values: 1) the first/second one averaged over the first/second ten days, 2) the third one averaged over the residual of the month. A regular calendar is used within the system (without leap year). For example, the three outputs of February are averages over 1^{st}-10^{th}, 11^{th}-20^{th} and 21^{st}-28^{th}. The hindcasted ocean properties are converted to monthly mean data of corresponding period according to the implemented calendar in the model.

2.4.1 Verification metrics

In order to verify whether the initialization provides greater skill in decadal prediction than the uninitialized 20C run, certain metrics shall be chosen to offer assessment of the hindcasts. The primary verification metric used here is Root Mean Square Error (RMSE). RMSE between the paired ensemble mean hindcasts H_j (both raw and bias-corrected) and corresponding observations O_j is given for a set of commonly chosen lead times: lead year 1, average of lead years 2-5 and average of lead years 6-9 (Wilks, 2006), in equation (Goddard et al., 2013):

$$RMSE = \sqrt[2]{\frac{1}{n}\sum_{j=1}^{n}(H_j - O_j)^2}$$

By comparing the hindcasted variables with the observations, primary assessment on prediction skill can be achieved (Tebaldi and Knutti, 2007; Goddard et al., 2013). As is indicated in previous section, model drifts can be reduced for the consistent hindcast CIH by initializing the coupled model CFES with dynamically consistent initial conditions CDA. In order to verify whether this is true, RMSE of raw hindcast H_j and bias-corrected hindcast H_j^{b-c} are all analyzed.

However, RMSE only is not sufficient to assess prediction skill due to its sensitivity to outliers that leads to a possibility of large error (Collins, 2002). Studies show that multiple diagnostics work as a better estimation on the model performance. Another measurement used here is the anomaly correlation coefficient (ACC), which is a scale-invariant measure of the linear associations between the two sets of forecasts. The anomaly correlation coefficient measures only the phase difference between the observations and forecasts experiment (Matei,

et al., 2012). Since ACC is quite sensitive to the trend, all the variables are detrended first in order to avoid overestimation in the performance of hindcasts. The ACC is also calculated as a function of lead time. The correlation coefficient of hindcasts with their own initialization (CDA and GECCO2 respectively) is computed as below:

$$COR_t = \frac{\sum H_{at}^D O_{at}^D}{\sqrt{\sum {H_{at}^D}^2 {O_{at}^D}^2}}$$

where H_{at}^D is the predicted anomalies at lead time t with its linear trend removed, O_{at}^D is the detrended anomalies of observations or the data that serve as initial conditions of the hindcasts (i.e. CDA/GECCO2). The observations for verification of sea surface temperature are taken from the UK Met Office Hadley Center dataset (HadISST; Rayner et al., 2003), with the corresponding cover period. Due to the lack of AMOC observations (e.g. Cunningham et al., 2007; Kanzow et al., 2007), we take AMOC from the GECCO2 Synthesis for evaluation of predictive skill, also with the corresponding cover period.

To compare and assess the relative skill in different forecast experiments, the root mean square error skill score (i.e. RMSS; Jolliffe and Stephenson, 2003) based on a set of RMSE from paired hindcasts (i.e. the reference prediction and the prediction needed to be evaluated) is used in this study. The skill score is defined as:

$$RMSS = 1 - \frac{RMSE_{hind}}{RMSE_{20c}},$$

It is supposed to give the accuracy skill of the test prediction (hindcasts initialized with CDA and GECCO2) against the reference prediction (20C). Positive values denote that the errors in the initialized hindcasts are smaller than the un-initialized hindcasts. The verification metric is computed in terms of annual mean SST.

The lead time mentioned here is defined as the length of time from the starting date of the experiment to the hindcasted year respectively. Temporal smoothing can typically reduce higher frequency noise and increase skill. In this thesis, assessment of SST predictive skill will be performed at three different time scales: lead year 1, averages over lead years 2-5 and 6-9 (Wilks, 2006). The three different scales chosen work as a small set of cases to be illustrated to qualify the information for different lead time and temporal averaging (e.g. Smith et al., 2007; Goddard et al., 2012a). Resulting from the proximity to the observed initial conditions, the year 1 of the prediction shall give the best predictability; it overlaps with currently available seasonal-to-interannual predictions. As the model evolves, it discards the strong imprint of initial conditions and tends to its own climatology. On the contrary, the predictive information from past and projected changes in anthropogenic external forcing starts to increase (Hawkins and Sutton, 2009a; Branstator and Teng, 2010, 2012; Fig 2.4.1 from Branstator and Teng, 2012). Although the skill arises from these two factors are not always possible to distinguish (Solomon et al., 2011), it is widely accepted that after about the first 5 years of a decadal prediction, skill arises due to response to external forcing. Recent estimations and measurements of predictability and predictive skill proved that greater forecast skill arises from external forcing after about 8 years of the forecasts (Boer 2011; Boer et al., 2013; Fig 2.4.1). Overall, the 4-yr average forecasts (year 2-5 and 6-9) within the

decade are likely to have lower skill compared to lead year 1, and there are potential differences between them. The year 2-5 average still represents the interannual time scale; it is likely still to be dominated by year-to-year variability and less by the climate change signal. Comparison between 2-5 and 6-9 year average predictions can provide information of the dependence of skill on lead time.

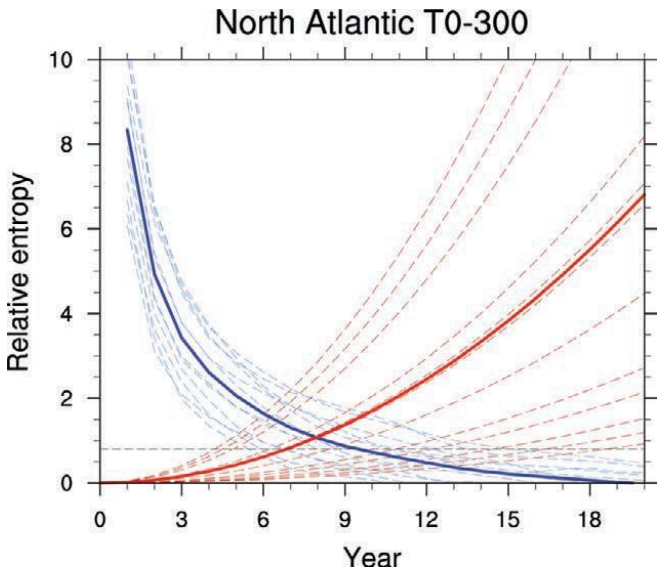

Fig 2.4.1 Predictability of upper 300-m temperature of the North Atlantic for 12 CMIP5 models resulting from initialization (dashed blue lines) and the response to RCP4.5 forcing (dashed red lines) from Branstator and Teng (2012). The solid lines are multi-model averages. Predictability is measured by relative entropy. The crossover point near year 8 for the multi-model averages indicates that after 8 years of the hindcasts, information resulting from external forcing (red line) becomes more important than information originating in the initial conditions (blue line).

2.4.2 Trend removal

Technically speaking, a "trend" in a time series is a slow, gradual change in some property of the series over the whole interval under investigation, which can be a long term change in the mean, or change in other statistical properties. Being one of the most critical quantities, the trend is involved in many applications of climate analysis, such as correlation computation and spectral analysis (Wu et al., 2007; Polkova, 2014). The statistical or mathematical operation "detrending" is often applied to remove the trend that represents a feature thought to distort or obscure the relationships of interest. It is a necessary step in order to avoid the analysis being overwhelmed by the nonzero mean and the trend terms.

For climate prediction at decadal time scale, which is a joint initial-boundary problem, it's always important to include information about external forcing within the observed state of the climate system (usually through the prescribed historical and RCP 4.5 radiative forcing). Hence, the evaluation of total skill score of SST indicates the predictive skill associated with both the external forcing and internally generated natural climate variability. Many studies have indicated improvement of skill in global SST predictability along with the warming

trend associated with external forcing (Smith et al., 2010; van Oldenborgh et al., 2012), as is the case for North Atlantic and Western Pacific shown in Fig 2.4.2.

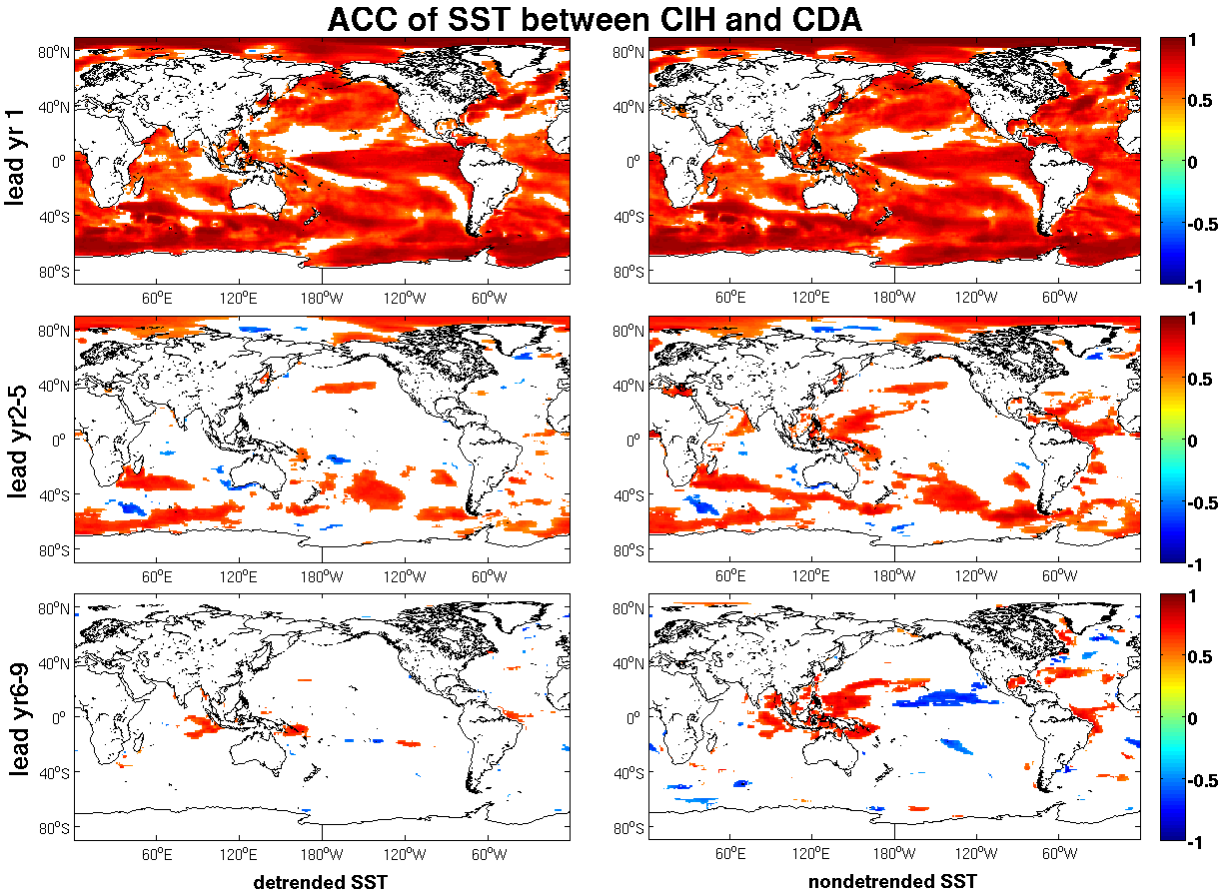

Fig 2.4.2 Spatial distribution of anomaly correlation coefficient (ACC) between hindcasted SST (hindcasts initialized with CDA, short as "CIH") and CDA (estimations derived by assimilation through the coupled climate model CFES) at different lead years: 1) lead year 1 (top panels); 2) averages of lead years 2-5 (middle panels); and 3) years 6-9 (bottom panels). The SST data shown in left panels are all detrended and these in the right panels are not. Only the significant ACC (at 95% level) are shown here. The trend gives rise to predictability on decadal time scale regionally from comparison of the two columns of figures. For the first lead year, larger areas with predictive skill are observed over the North Atlantic for the non-detrended SST (upper panels). For the 4-yr average, the trend gives rise to predictability of SST over part of the North Atlantic and western Pacific (middle and bottom panels).

For the initialized hindcasts CIH, larger areas with predictive skill are observed over the North Atlantic for the non-detrended SST at the first lead year (upper panels of Fig 2.4.2). For the 4-yr average, the trend gives rise to predictability of SST over part of the North Atlantic, the Indian Ocean and western Pacific (middle and bottom panels). In fact, a great fraction of predictive skill for the global SST is due to external forcing. The differences between spatial distribution of RMSE skill score of detrended and non-detrended SST from the initialized hindcasts (CIH here) against un-initialized forecast reveals relative predictability due to internal variability and externally forced climate change (Fig 2.4.3), and therefore indicates improvement of predictive skill due to external forcing. As is indicated by Fig 2.4.3, the smaller errors when SST from CIH is not detrended than that is detrended prove that

predictive skill arises due to external at different lead years, in the North Atlantic and Pacific. Hence the "detrending" realized by the removal of the influence of greenhouse gas-induced global warming from the analysis (i.e. including only the predictable signal related with natural climatic fluctuations) is important for the investigation on the skill of decadal prediction associated with internal variability through the initialization procedure.

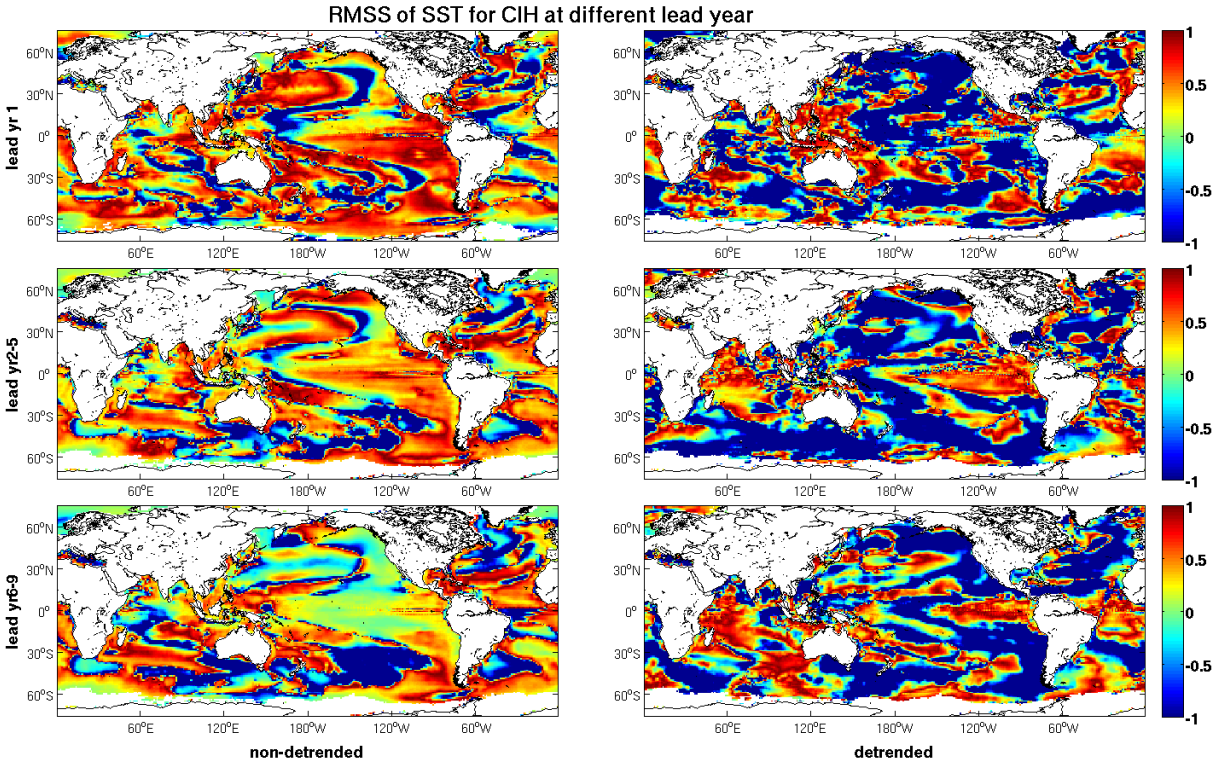

Fig 2.4.3 Spatial distribution of SST RMSS (RMSE skill score) for CIH (CDA initialized hindcast) against 20C (un-initialized hindcast) at different lead years: 1) lead year 1 (upper panels); 2) averages of yr2-5 (middle panels); and 3) yr6-9 (bottom panels). SST from HadISST is used as the observation. The SST data used in right panels are all detrended and these in the left panels are not. Red area in the figure indicates improvement due to initialization. Improvement in predictive skill for initialized hindcasts is mainly due to the internal variability and externally forced climate changes (left panels). With an application of trend removal, improved predictability for initialized hindcasts is only about the internal variability, as is shown in right panels. Comparison between the left panels and right panels indicates that the trend plays an important role in improving predictive skill on decadal prediction.

Several methods have been proposed to remove the global trend and ENSO influence over global SST (eg. the least-squares-fit straight line for linear trend). In this thesis, the strategy for detrending time series that contain fields at particular lead years proposed by Kharin et al. (2012) is applied for the anomaly correlation analyses of distribution for global SST at each grid point. This methodology removes the period dependent "mean bias" from the predictions that is different from the observed. Compared with other approaches such as removing the observational trend from both the hindcasts and observations, this strategy could be more appropriate to deal with hindcasts' trend, especially for the case with different observational and hindcasted trends (Polkova, 2014). In detail, the long-term linear trend of SST is calculated from hindcasted fields over the initialization period of 1980-2006 at different lead year for each grid point, employing the least square statistical method. Afterwards, the trend is

subtracted from the hindcasts over lead years for each grid point. The equation follows like this:

$$O_{jl} = \mu_l^o + s_l^o j' + e_{jl}^o,$$

$$H_{jl} = \mu_l^H + s_l^H j' + e_{jl}^H,$$

$O(t_j, l) \equiv O_{jl}$ represents the observations corresponding to the simulated variable, $H(t_j, l) \equiv H_{jl}$ denotes the hindcasted variable initialized at year j. l represents the lead year of the model prediction between the forecast initialized date and the forecast preceeding period considered. Here in this paper, $t_j = 1980,1981,1982, \ldots,2006$ and $l = 1,2, \ldots,9$; μ_l^o and μ_l^H are the long-term means of the observations and hindcasts respectively at lead year l, $s_l^o j'$ and $s_l^H j'$ are the corresponding long-term linear components. The residual e_{jl}^o and e_{jl}^H are predicted and observed deviation (e.g. Fig 2.4.4) that will be used for further analysis such as anomaly correlation coefficient here.

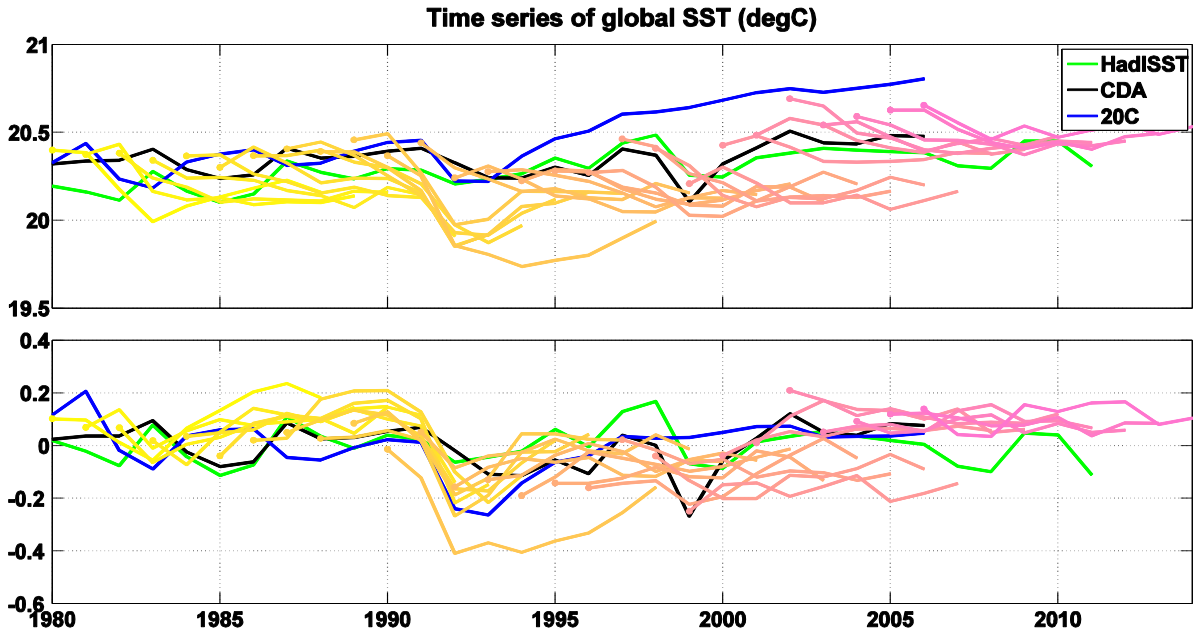

Fig 2.4.4 Time series of global-averaged annual mean SST in the HadISST dataset (green), initialized (colours, CIH) and un-initialized (blue) predictions and CDA (black). The upper panel shows the raw initialized hindcasts before detrending. The bottom panel shows the evolution of global mean SST with the linear long-term trend removed.

The above-mentioned trend removal is applied to SST for further evaluation of predictability on decadal scale. However, study has shown that for regional climate parameters such as the North Atlantic SST, different portions of impact from external forcing relative to internally generated variability are reported at the end of the twentieth century (Ting et al., 2009). Branstator and Teng (2010) point out that North Atlantic SST is likely to be dominated by internal natural variability for the climate system and the impact of the radiative forcing will be small at least in the first few years of decadal predictions. As is agreed by researchers, the skillful predictions in AMOC is related with improved predictive

skill of SST in the North Atlantic (Delworth et al., 2007; Knight et al., 2005; Swingedouw et al., 2012), possibly due to a connection between them associated with the Atlantic multi-decadal oscillation (i.e. AMO; Meehl et al., 2014). Considering the relatively small impact from external forcing of SST in the North Atlantic, both the detrended AMOC and non-detrended AMOC hindcasts will be analyzed in this thesis.

2.4.3 Bias adjustment

The bias adjustment procedure is commonly used in seasonal forecasting (Stockdale, 1997). In decadal climate predictions, the characteristics of bias growth may help understand the physical processes that lead to prediction errors (Meehl et al., 2014). Nevertheless, to provide useful information that leads to predictability, biases that arise in decadal prediction must be removed. A posterior bias correction is specially needed for prediction with full-state initialization strategies. This is because with full-state initialization strategy, the model is constrained to be close to the observations by initialization with assimilated observations of the current climate state. However, as the prediction evolves, the model drifts back towards its preferred model climatology since it can no longer be constrained with observations (Smith et al., 2013). Such drifts due to difference between the model climatology and observations will introduce errors in the forecasts and may affect predictive skill on all time scales. Therefore, a bias correction is usually needed for climate predictions.

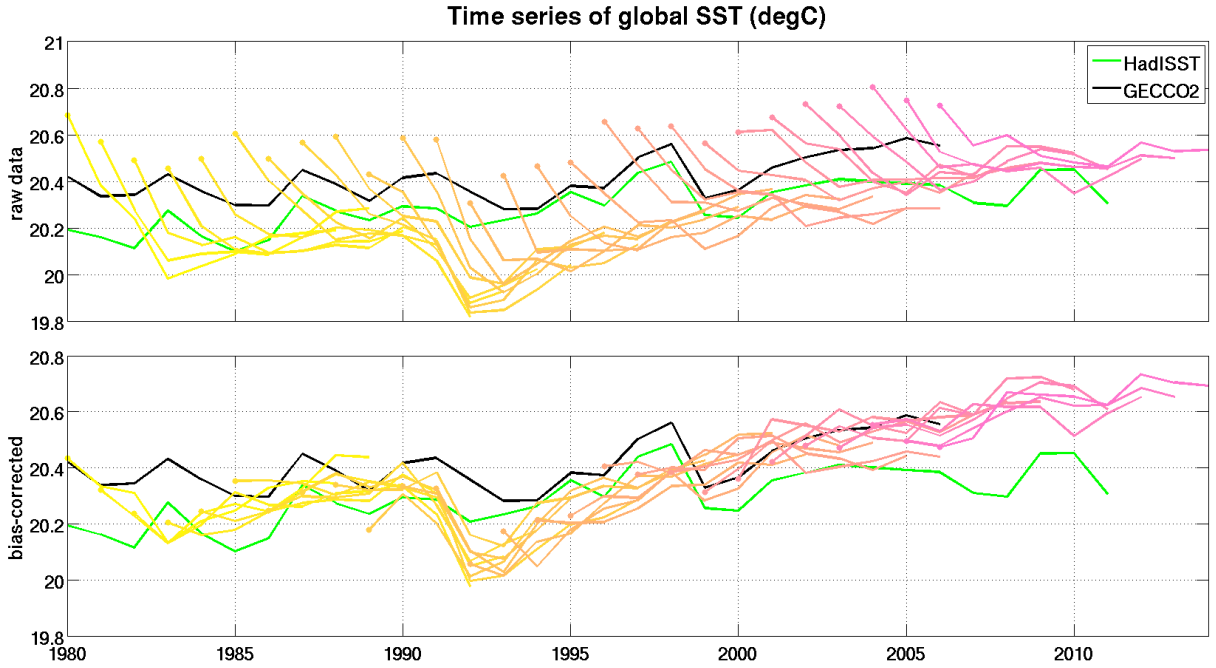

Fig 2.4.5 Time series of global-averaged annual mean SST in the HadISST dataset (green), initialized (colours, GIH) and GECCO2 (black). The upper panel shows the raw initialized hindcasts before bias correction. The bottom panel shows the evolution of global mean SST with mean bias removed.

Bias adjustment can be achieved by subtracting the climatological averages of a series of initialized hindcast over all starting dates (e.g. Smith et al., 2013; ICPO, 2011), or by

removing the model climatological average for each time period from the hindcasts/observations from the corresponding period (e.g. García-Serrano and Doblas-Reyes, 2012), or by a correction that varies with the period (i.e. the "average" model drift with a trend, van Olderborgh et al., 2012). In this thesis, the lead-time dependent mean bias is removed from the predictions by subtracting the difference between the climatological average of hindcasts and that of observations on each starting date at different lead years from the hindcasts of the same period. Hence hindcasts after mean bias correction are used for further assessment. As is shown in top panel of Fig 2.4.5, for the raw hindcasts GIH (colors), global SST drifts quickly after the initialization, with an average value up to 0.7°C. The application of mean bias correction results in a better resemble of hindcasts to observations (HadISST, green).

The mean bias at different lead times is calculated as below:

$$\widehat{h_{it}} = h_{it} - \sum_{i=1}^{N}(h_{it} - O_{it})/N \ ,$$

where *i* represents a particular starting date of the initialized hindcast, *t* represents a particular hindcast year or an average over lead years. Hence, $\widehat{h_{it}}$ are the bias-adjusted values for one specific initialized hindcast at lead year/averaged over lead year *t*, O_{it} are the observed values corresponding to raw hindcasts h_{it}. Since the initial conditions from CDA goes from 1980 only until 2006, the number of integrations accounted in the bias adjustment are $N = 27$. In this thesis, the trend removal is applied to the observational SST and bias-adjusted values of hindcasted SST as is above-mentioned.

2.4.4 Persistence

Because of inertial or carryover processes in the physical system, geophysical time series are often auto-correlated. In a climate system, the system may remain in the same state from one observation to the following one, which can be associated with positive autocorrelation. The persistence is used as a reference forecast to verify whether the usage of the coupled climate model CFES helps in improving the skill on decadal predictability through comparison with hindcasts derived from initializing the model. As a commonly used statistical forecast, the persistence forecast assumes that future conditions will be the same as past conditions (Matei et al., 2012). In this paper, annual mean data of GECCO2/CDA fields are used to compute the persistence forecasts respectively, started 1 year preceding the initialization year. Please note that CDA reanalysis covers only the period of 1980-2006. Hence for the first lead year, the period available for persistence is 1980-2005, corresponding to the hindcasts from 1981-2006. When the skill is computed for the averages over lead year 2-5 and year 6-9, the persistence forecasts are calculated from the averages over 4 years at lag-5, which precede the initialization date. For instance, for the start date 1983, the hindcasts at lead time yr 2-5 (1984-1987) and the hindcasts at yr 6-9 (1988-1991) are compared with the persistence forecast of 4-yr average over the preceding years, i.e. 1980-1983. Another option using the damped persistence computed with first-order auto-regression following the equation given by von Storch and Navarra (1999) is also popular for comparing performance of initialized hindcasts and persistence at every lead year (e.g. Polkova et al., 2014).

2.4.5 Determination of statistical significance level

The estimation of significant levels for the predictive skill of spatial ACC and RMSS is based on a Student's *t*-test introduced by William Sealy Gosset (Mankiewicz, R., 2004). This commonly applied statistical test is a parametric test of the null hypothesis that two univariate random variables have equal means (von Storch and Zwiers, 1999). In the experiments mentioned above, each initialized hindcast has one realization. Hence all of the realizations occur independently. However, the initial conditions of experiments are derived through assimilated observations. Geophysical time series are generally auto-correlated because of inertial or carryover systems in the physical system. The observed series of oceanic variable is not random in time, and the information within the observation is not independent. If a time series is auto-correlated, less information of the population mean will be provided, compared with a non-auto-correlated time series with the same sample size. Under this circumstance, the denominators of *t* statistics under independence assumption will underestimate the variability of variables and therefore results in a too large absolute value of *t*.

To resolve this problem, adjustment to the sample size of autocorrelation is introduced. A factor related to the first-order auto-correlation scales the original sample size to reach the "effective sample size" or "equivalent number of independent samples", and hence an estimation of degrees of freedom (von Storch and Zwiers, 1999) follows:

$$df = N\frac{1-r_1}{1+r_1} - 1$$

where N represents the sample size of the hindcasts and r_1 denotes its lag-1 auto-correlation coefficient.

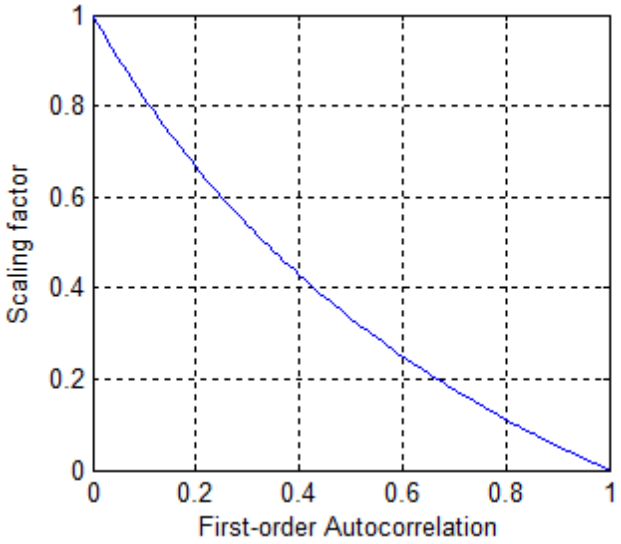

Fig 2.4.6 Scaling factor applied in computation of effective sample size from original sample size of auto-correlated time series.

The adjustment to effective sample size becomes less important as the auto-correlation becomes lower (Fig 2.4.6 from Notes_3, GEOS 585A by Meko). Note the definition of the equivalent sample size depends upon the parameter that is being tested and the way in which information is measured. It's applicable only when the time series are persistent. It is possible to have larger effective sample size than original sample size when adjacent anomalies tend to have opposite sign, that is, a negative lag-1 auto-correlation coefficient (von Storch and Zwiers, 1999). In this paper, the "effective sample size" is applied to calculate the degrees of freedom for the reference run involved in the assessment of predictability. The ensembles' degrees of freedom are the scaled total sum of the differently initialized ensemble numbers minus 1 at different lead years.

Chapter 3

Evaluation of model consistency on decadal climate prediction

Working as a control parameter for the heat flux exchanges between the atmosphere and the ocean, SST is always a key variable in studies of these two fields. It is easier to observe than other oceanic variables and has long observation records dating back to the early stage of oceanography. Local variations of SST are important for the study of tropical cyclones, Ekman transport, rapid influxes of glacial fresh water, concentrated phytoplankton blooms, and so on. The signature of its regional variations over the Tropical Eastern Pacific Ocean denotes the occurrence of El Niño/La Niña events. So far, SST has been used as a primary assessment of skill from initializations, and improvement in SST skill is observed from initialization over parts of the ocean in recent studies of decadal prediction (Pohlmann et al., 2009; Mochizuki et al., 2010; Matei et al., 2012). Here, the SST is analyzed from multiple initialized hindcasts as well as reference data in order to explore whether the initialized forecasts are skillful, and furthermore the regions and duration with predictive skill (if they are skillful). Insights to the causes leading to predictability will also be studied. By doing this, assessment of the consistency of the initialization with/without assimilation to the model will be provided. Additionally, regional SST (region 3.4) is used to calculate an index of ENSO (El Niño–Southern Oscillation) events to provide an example how different levels of model consistency can affect the predictive skill.

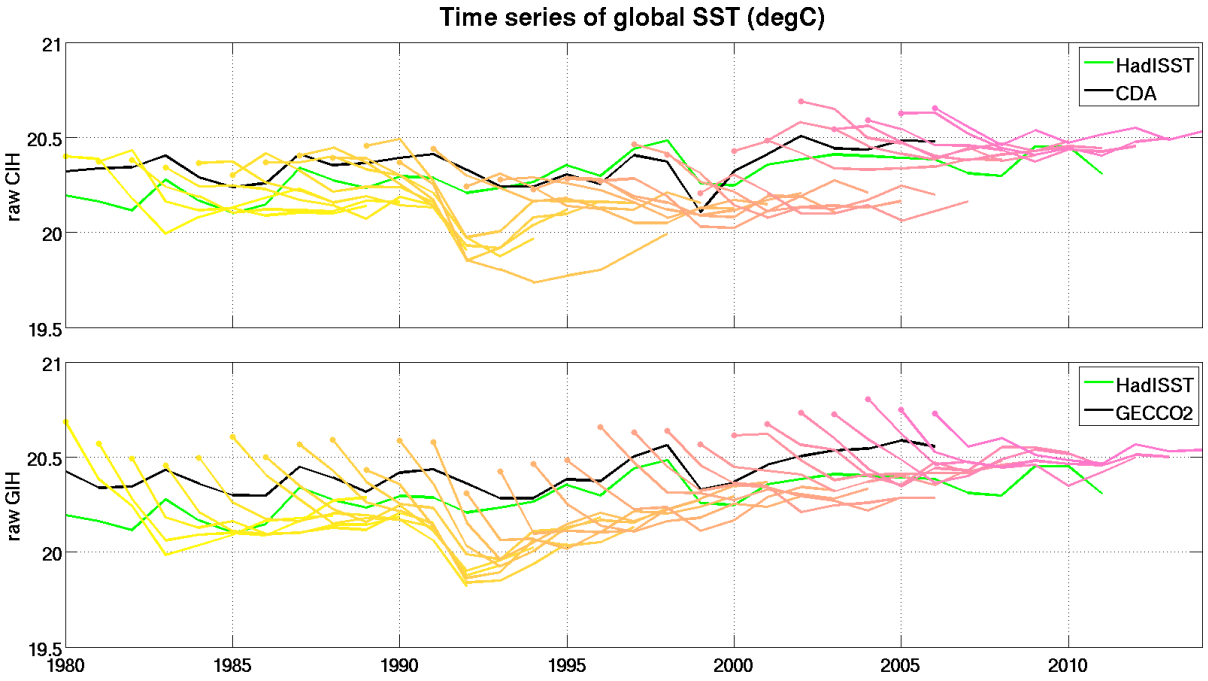

Fig 3.1 Evolution of global-averaged annual mean SST in the HadISST dataset (green), raw initialized hindcasts (colours) and ocean estimation (CDA/GECCO2) from which the hindcasts are initialized (black). The colored lines in top/bottom panel are raw CIH/GIH.

We start our analysis with hindcast SST variations through comparison with the observed and assimilated (CDA in short) SST. However, before this, it's necessary to verify whether the drifts in CIH are smaller than in GIH due to initialization with dynamically-consistent initial conditions, as is expected in the motivation. For this purpose, the global SST is analyzed. The evolution of global averaged SST from initialized hindcasts (colors) is displayed in Fig 3.1. As is indicated by the colored lines in the top panel, there is a smaller drift with an average around 0.4^{o}C in CIH one year after it is initialized with CDA. For GIH, the forecasts drift immediately after initialization, and an average of drifts up to 0.7^{o}C is shown in the bottom panel. For the globally averaged SST, the drifts in CIH are indeed reduced as is expected. Further information can be gained through the spatial distribution of RMSS between initialized hindcasts of SST without bias correction against the un-initialized hindcasts. As is indicated by comparison between left and right columns, errors are much smaller over vast areas of the ocean (e.g. over the Pacific) in CIH than in GIH until lead years 2-5. Therefore, our expectation that dynamically consistent initial conditions help to reduce model drifts is proved to be true here. Flowingly, assessment of SST predictive skill will be provided.

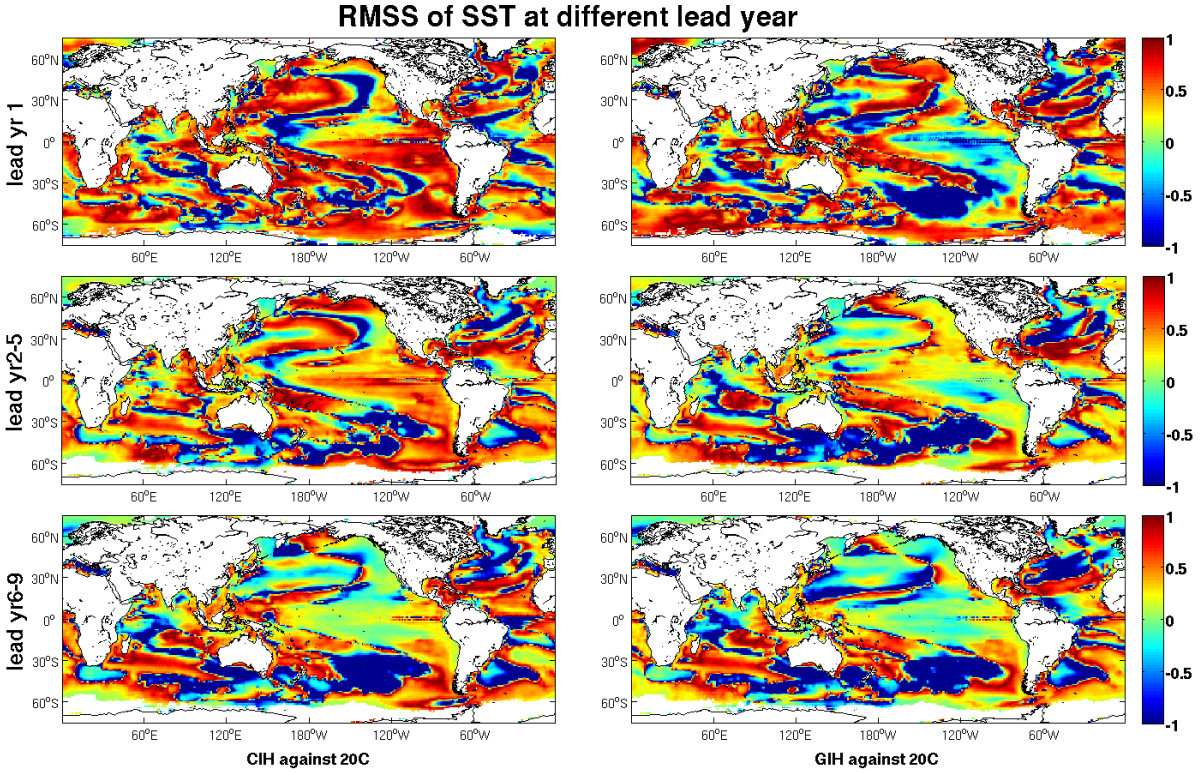

Fig 3.2 Spatial distribution of RMSS for annual mean SST for raw hindcast SST against the 20C at lead year 1(upper panels), averages over yr 2-5 (middle panels), and 6-9 (bottom panels). Left panels are CIH against 20C and right panels are GIH against 20C, without bias correction. Red areas indicate that the initialized hindcasts are more accurate than the 20C and blue indicates the opposite.

3.1 Predictive skill of global and regional SST

3.1.1 Representation of hindcasted global annual SST

The term "global warming" has become a focus in climate research and its significant and harmful effects on the environment, communities, and human's health are widely recognized. Ever since 1971, the sea surface temperature accounts for 90% of the warming (IPCC AR5, 2013). About two-thirds of the warming in the last three decades of the 20th century happened after 1980 (IPCC AR5, 2013). The warming is proven by a wide range of independent scientific analysis of variables, such as sea level rise (e.g. Stammer et al., 2013), snow and ice melting (IPCC, AR4, 2007), increased humidity (Kennedy et al., 2010), and so on. As a general consensus by scientists, increasing concentrations of greenhouse gases produced by human activities play a dominant role in global warming (more than 90% certain, IPCC, AR4, 2007; National Research Council, 2010; IPCC, AR5, 2013;). Decadal forecasts that take the anthropogenic forcing into account to predict the near-term variability of the climate system work as a valuable and scientifically significant application in terms of policy making and investment decisions.

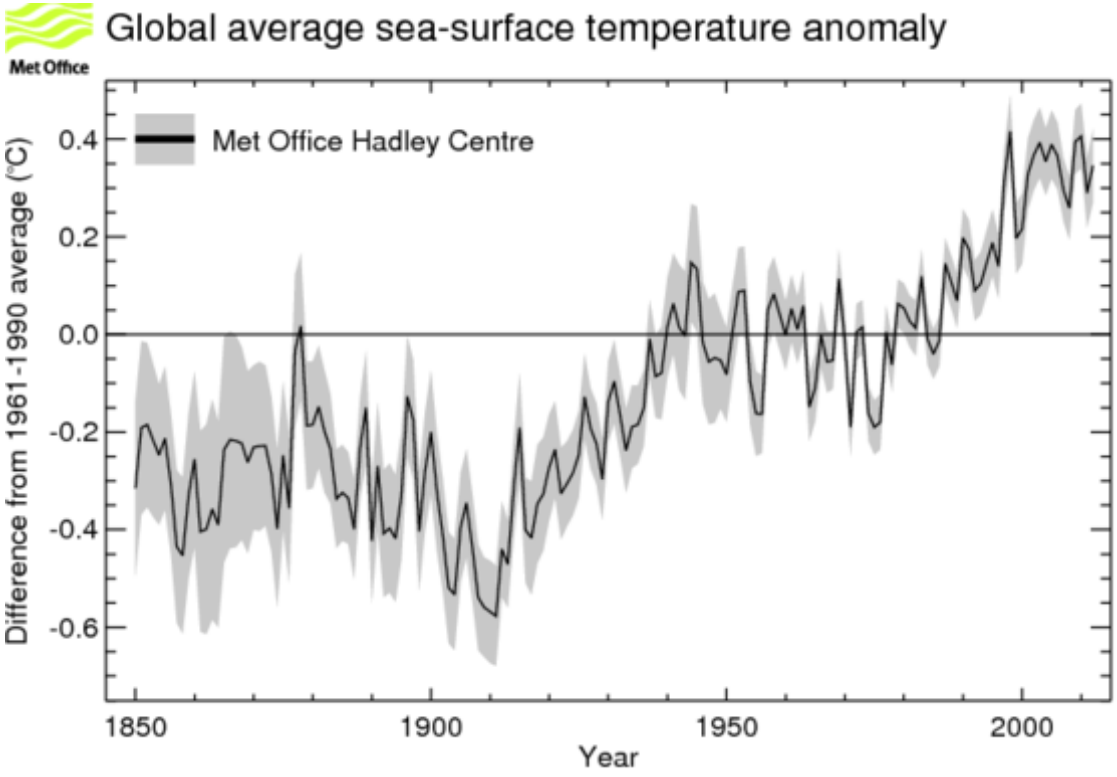

Figure 3.1.1 The time series of annual-mean global average sea surface temperature anomalies relative to 1961-1990 in observations (black line) from Met Office. The calculation is based on the sea surface temperature of dataset HadCRUT4 (abbreviated as HadSST3). The grey shade indicates the ranges that are 95% significant. As is shown in the figure, the SST warming trend is much larger after the 1900s. There is a warm pause from 1950-1980. After 1980, a large warm trend is found in SST again. (figure from Met Office website, available at: http://www.metoffice.gov.uk/research/monitoring/climate/surface-temperature)

The Fifth Assessment Report (AR5) by IPCC (the Intergovernmental Panel on Climate Change, 2013) reported that the global surface temperature will be more than 1°C (but most likely less than 1.5°C) warmer for the period 2016–2035 above the 1850–1900 mean, although the uninitialized simulations in the IPCC AR5 (Kirtman et al., 2013) indicate more near-term (2016-2035) warming than the initialized decadal predictions from CMIP5. Projections based on both present-day atmospheric concentrations of GHGs ('committed warming'; e.g., Meehl et al., 2006) and projected future changes in radiative forcing show a rise for globally averaged sea surface temperatures over the early 21st century (IPCC AR5, 2014). A figure based on the observational dataset HadCRUT4 from Met Office shows consistency in global warming (Fig 3.1.1). The observed SST anomaly relative to 1961-1990 shows that a much larger warming rate happened after 1980, compared with the period of 1960-1980. In this thesis, limited by the available data of CDA, both CIH and GIH cover the period from 1980 to 2006, which is a period with a high rate of global warming. We will compare the hindcast global SST with the observations to check whether the posterior forecasts can capture the global warming trend. This will offer us a general idea of the performance of the hindcasts and help in further analysis.

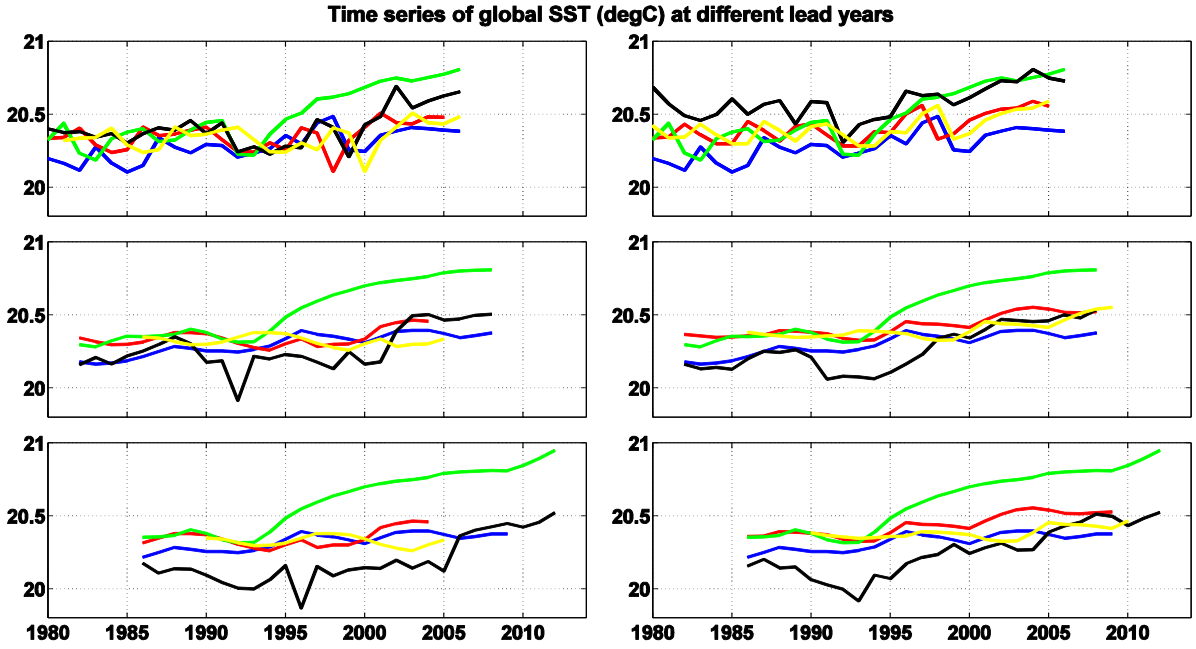

Fig 3.1.2 Time series of annual mean global SST at different lead years in observations (blue, HadISST data set), persistence forecast (yellow), 20C (green), initialized hindcasts (black, left panels CIH, right panels GIH), CDA (left panels, red) and GECCO2 (right panels, red). The top panels are SST at lead year 1. The 4-yr means of lead years 2-5 are shown in middle panels, with the time series plotted centered in year 3. Averages over lead years 6-9 are shown in the bottom panels, with time series plotted centered in year 7.

As is shown in Fig 3.1.2, global SST from CIH follows CDA closely at lead year 1, resembling well the interannual variability and the warming trend. The 4-yr average of CIH becomes cooler and a drift around -0.2°C /-0.4°C appears for lead years 2-5 and 6-9 separately, yet a warming trend is also captured for both. On the contrary, such cooling is not observed for global SST in the uninitialized forecast for the 4-yr average. Actually until 1992, when a major volcanic eruption happened, the 4-yr averaged SST from the 20C run resembles CDA

quite well. At lead year 1, the two major coolings around 1982 and 1992 of uninitialized SST are also highly likely due to volcanic eruptions. For CIH, the predictions are initialized with an estimation of climate fields that result from assimilating data into the same coupled global climate model as the simulation, which is noted here as being self-consistent. The assimilation scheme used reduces the initial shock at the very beginning of the prediction. While for GIH, a warmer bias is shown for lead year 1 over all the starting years (upmost of right panels). For the 4-yr average of lead years 2-5 and 6-9, the uninitialized SST is warmer than CDA/GECCO2 after 1992. Therefore, the response from external forcing shall make the hindcasted SST warmer. However, global SST from 20C is warmer than the initialized hindcast, but with the same trend. Since the initialization is not making the hindcasts cooler (CDA V.S. CIH at the first lead year, up-left of Fig 3.1.2), and they are both externally forced, the warmer mean of the 20C forecast could possibly be due to the bulk coefficients used to control the air-sea surface fluxes in the initialized hindcasts. SST from GIH is even warmer than SST from 20C for most of the time at lead year 1. This could probably due to the different dynamics and climatology between the initial conditions from GECCO2 and the model. Such difference introduces initial shock and systematic errors. Hence, the hindcast SST follow the CDA better at lead year 1 than that of GIH and GECCO2. This indicates the advantage of model consistency in decadal predictions, which is significant at lead year 1 but relatively small for the 4-yr average of lead years 2-5. Note for the 1990s, drops are observed for both averages over lead years 2-5 and 6-9, possibly resulting from the volcanic eruption in early 1990s, as is the case observed for GIH. Another drop is also evident after the mid-1990s for CIH.

Generally, CIH does a better job in reproducing annual mean time series of global SST than GIH, most possibly due to a more self-consistent full-field initialization. Skill is expected in CIH for climate predictability at decadal time scales. For both CIH and GIH, global SST is warmer than observations at lead year 1, but colder at the 4-yr average. Different from lead year 1, initialized global SST is also colder than the uninitialized forecast at the 4-yr average. Different from CIH initialized every year, CDA from JAMSTEC was reinitialized every 9-months, with an overlap of 1.5 months at both margins. This strategy helps to prevent a slow drift in CDA. The difference between models used to derive CDA (CFES without response to radiative forcing) and for CIH (CFES with external forcing included) may also contribute to the drift at longer lead times.

3.1.2 Assessment on spatial distribution of SST predictive skill

We explore the spatial distribution of predictive skill of decadal SST for the different hindcasts separately at three different timescales as is described in section 2.4.2: lead year 1 (yr 1), average over lead years 2 to 5 (yr 2-5), and average over lead years 6 to 9 (yr 6-9). By comparing the performance with that of the persistence forecast and the 20C run, the predictive skill related to persistence and to forced response can be identified.

The most common way to evaluate decadal predictive skill is through comparison with uninitialized simulations. Another alternative approach that attempts to assess the skill of the remaining variability after removing the forced component from the hindcasts and observations has also been explored (van Oldenborgh et al., 2012). However, this strategy has

difficulty in the identification of the forced component and hence much uncertainty is usually introduced. The uninitialized 20C experiment can be used to estimate climate change due to boundary conditions (Pohlmann et al., 2009). To explore the relative skill of different hindcasts, we start the analysis from the spatial distribution of root mean square error skill (RMSS) between two differently initialized hindcasts against the externally-forced, uninitialized hindcast (20C). The RMSS shown in Fig 3.1.3 is based on the root mean square error (RMSE) between each hindcast and HadISST as the observation. With the benchmark skill of the 20C as comparison, identifying the regions and duration with predictability due to initialization is possible.

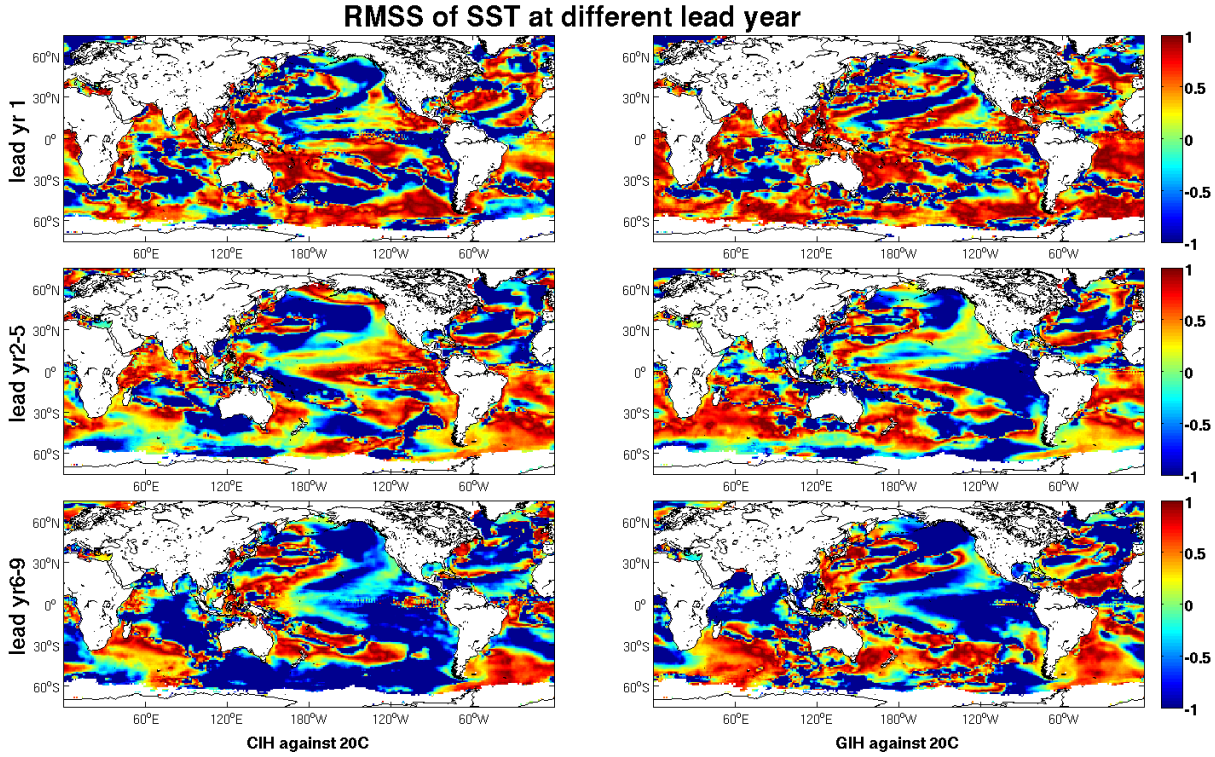

Fig 3.1.3 Spatial distribution of RMSS for hindcast SST against the 20C at lead year 1(upper panels), averages over yr 2-5 (middle panels), and 6-9 (bottom panels). Left panels are CIH against 20C and right panels are GIH against 20C. Red areas indicate that the initialized hindcasts are more accurate than the 20C and blue indicates the opposite. All the hindcast SST involved is bias-corrected

The spatial distribution of RMSS is shown in Fig 3.1.3. Red areas indicate that the initialized hindcasts are more accurate than the 20C hindcast, and blue areas indicate the opposite. On shorter time scales, predictive skill is strongly influenced by initial conditions. For longer time scales, the influence from boundary conditions becomes larger. From the RMSS of initialized hindcasts against 20C, large areas (almost over the whole global ocean) with better skill are observed for CIH. For GIH, compatible red and blue areas are observed between 50°S-50°N, while in the Southern Ocean and the tropical Equatorial Ocean (except the eastern Pacific), large areas with skill are detected. Both sets of initialized hindcasts exhibit enhancement of SST RMSS skill at lead year 1 compared to the non-initialized hindcast. The most significant increase in SST RMSS develops in the tropical Pacific and the Southern Ocean (Fig 3.1.3). The pattern for the 4-yr means at lead years 2-5 for CIH is similar to the first year. Skill over the ENSO area is robust until when it comes to lead years 6-9.

However, the 4-yr averages of GIH show large areas of negative values over the ocean. Moreover, the skill over the Southern Ocean disappears. This indicates that CIH shows more predictability due to initialization until lead years 2-5 than 20C and, above all, CIH is superior to GIH. The strongest increase in RMSS is over the tropical area of Pacific and Atlantic. For the 4-yr average, the portion of predictability that rises from boundary conditions becomes larger, especially for GIH.

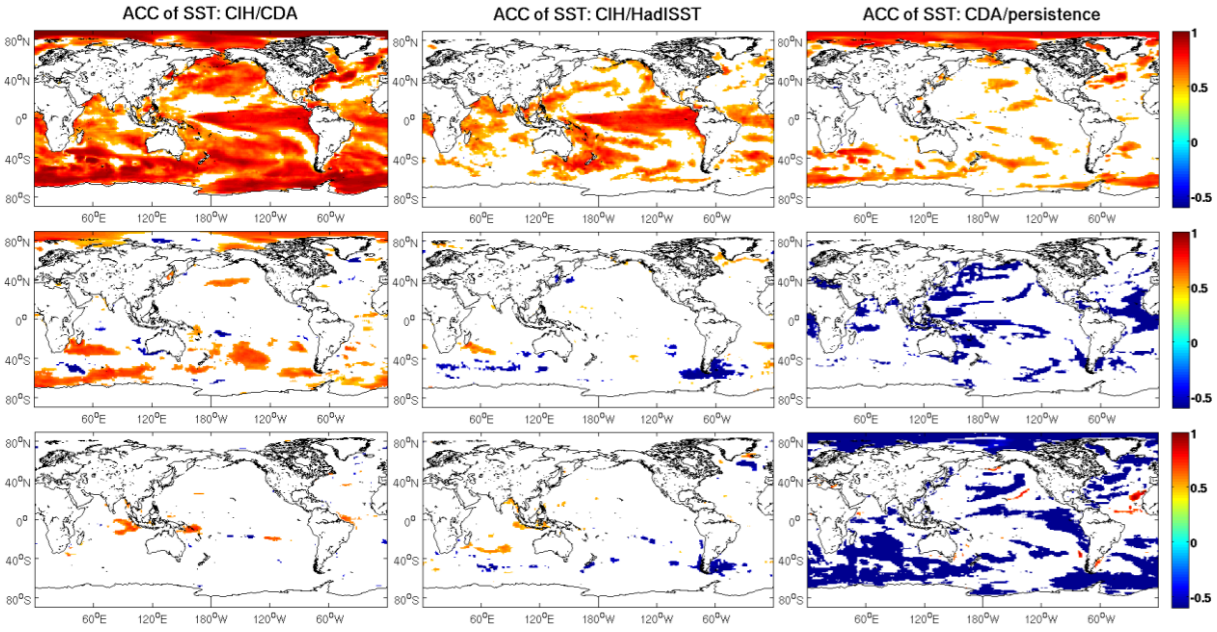

Fig 3.1.4 Spatial distribution of SST anomaly correlation coefficient between CIH and CDA (left), CIH and observed SST (HadISST, middle), and CDA and persistence (right), at lead year 1 (top panels), averages of lead yr 2-5(middle panels) and lead yr 6-9 (bottom panels). Only the significant coefficients (at 95% level) are shown here. A linear trend removal is applied to all the SST involved before calculation of ACC.

We now evaluate the predictability of SST decadal prediction through the spatial distribution of the SST anomaly correlation coefficients (ACC). ACC of SST in Fig 3.1.4 offers us an overview of the regions and time scales with predictability. Before calculation of ACC, all the data used here are linearly detrended in order to remove the influence of linear-in-time greenhouse gas-induced global warming from the analysis. For the lead year 1, wide areas over almost the whole global ocean with significant prediction skill are observed from the distribution of ACC between CIH and CDA. Only one exception offshore near the western Pacific is observed, corresponding to the only reds of upper panel in the third column (ACC of persistence). The small errors as is indicated in red in upper panel of RMSS in Fig 3.1.3 contribute to the high predictive skill. The highest ACC skill is found near the equator and middle of North Atlantic. A significant improvement over the eastern tropical Pacific is observed, where a lot of forecasts failed to capture the variability (e.g. Matei et al., 2012; Pohlmann et al., 2009; Polkova et al., 2014). Predictive skill at the first lead year is mostly achieved due to initialization. The consistent dynamics and compatibility of initial conditions (CDA) with the model used for hindcasts improve the predictive skill of SST from CIH.

However, as the forecast evolves, the influence from initialization decreases, the response to externally forced climate variation starts to influence the forecasts. Meanwhile, the forecast

"drifts" towards the model-preferred climatology that is partially different with the CDA climatology (mostly resulting in the absence of external forcing in deriving CDA). A significant decrease of skill is observed when the lead time of the forecasts increases to lead yr 2-5 and 6-9, especially over the North Pacific and tropical Pacific, where significant skill is observed at lead year 1. Nevertheless, researchers do report less prediction skill over the North Pacific compared with the Indian Ocean (e.g. Goddard et al., 2013; Doblas-Reyes et al., 2013). The sensitivity to uncertainties from the initial state (Branstator et al., 2012; Branstator and Teng, 2012) and internal variability (Meehl et al., 2014) may cause a decrease in predictive skill in the Pacific. A large variety of studies show that initialization is playing a more significant role in predictive skill than the uninitialized simulations in the Pacific (Guemas et al., 2013b; Meehl and Teng, 2012, 2014). A prediction system through initialization shows that the model's ability to reproduce subsurface temperature variability of the North Atlantic is important in improving predictive skill (Mochizuki et al., 2010; Chikamoto et al., 2012a). The small uncertainty in initial conditions in CIH leads to high predictability in the North Atlantic and eastern tropical Pacific at the first lead year, while the uncertainty in the variability leads to a reduction in predictive skill for longer lead times. The difference in the model used for deriving CDA and for hindcasts (with/without external forcing) in CIH possibly leads to a poor performance of the model in capturing the internally generated component of climate change, such as the cold bias and temperature drops observed for initialized hindcasts at longer lead years in Fig 3.1.2, and therefore leads to poor predictive skill shown for the 4-yr averages.

Some areas with predictive skill in the southern Indian Ocean and sub-polar Southern Ocean are still achieved at an average of lead yr 2-5 for CIH/CDA. Predictability is achieved along the coast of South America in the Atlantic (around 5 ^{o}N-16 ^{o}S) until lead years 2-5. Despite this, hardly any skill is observed in North Atlantic. An assessment of predictive skill based on CMIP5 muti-model ensemble mean hindcasts reveals that improved skill in the North Atlantic region is related to skillful predictions of the AMOC due to initialization (Doblas-Reyes et al., 2013; Latif and Keenlyside, 2011; Srokosz et al., 2012; Matei et al., 2012c; Yeager et al., 2012; Robson et al., 2012a,b). The connection between AMOC and surface temperature is possibly related with the Atlantic multidecadal oscillation (AMO; Meehl et al., 2014). By initializing the model with a more accurate initialization of AMOC, higher predictive skill of SST in the North Atlantic can be achieved. CDA from JAMSTEC is primarily directed towards improving the ENSO simulation and prediction. Hence the constraint is possibly weaker in the North Atlantic, and therefore results in less accuracy and possible different variability of the AMOC. Besides, the AMO (Atlantic multi-decadal oscillation) is in a low phase transition with a long-term trend during the hindcast period of 1980-2006. The trend removal applied may also lead to the low skill of SST in the North Atlantic region.

To summarize, initialization gives rise to decadal predictability of SST. The strongest enhancement of skill is achieved globally in the first year. The consistent dynamics and compatibility of initial conditions (CDA) with the model used for hindcasts all lead to a rise in predictability of SST from CIH. A significant improvement over the eastern tropical Pacific is observed. The model consistency between the one used to provide initial conditions and for hindcasts gives rise to predictive skill at lead year 1 in our solution, especially in the tropical Pacific.

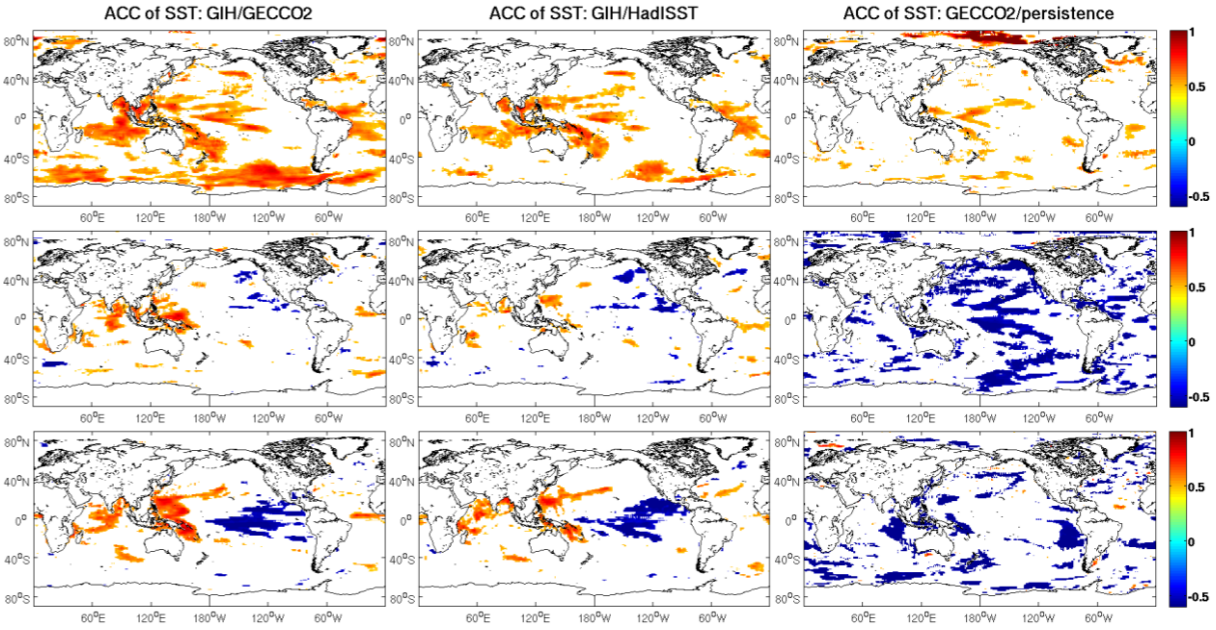

Fig 3.1.5 Spatial distribution of the SST anomaly correlation coefficient between GIH and GECCO2 (left), GIH and observed SST(middle), and GIH and persistence (right), at lead year 1 (top panels), averages of yr2-5(middle panels) and yr6-9(bottom panels). Only the significant correlation coefficients (at 95% level) are shown here.

For the GECCO2 initialized hindcasts (GIH), the patterns in terms of anomaly correlation skill are similar with that of CIH, but with much smaller areas with significant SST predictive skill. The highest predictive skill for GECCO2 initialized SST is observed in the southern tropical Indian Ocean, sub-polar Southern Ocean and western Pacific, pretty similar to a paper by Matei et al. (2012c). Compared to GIH, CIH shows better correlation with the data from which it was initialized (i.e. CDA). For the first year, large areas with SST predictive skill are shown in Fig 3.1.5 over middle and western tropical Pacific, in the Indian Ocean and in the Southern Ocean from GIH. Areas with skill are found in part of the Indian Ocean and around Indonesia for the average of lead years 2-5, and it remains also for lead years 6-9. If compared with GIH, a significant improvement in lead year 1 for CIH is evident over much of the ocean, especially for the eastern tropical Pacific. Poor predictive skill for SST at lead year 1 has also been observed in other studies (e.g. Matei et al., 2012c; Pohlmann et al., 2009) from GECCO initialized hindcasts, mainly due to a less accurate initialization that results from the GECCO configuration. As is found in Section 3.1.1, global SST from GIH is warm-biased at lead year 1 (upmost of right panels in Fig 3.1.2). Regional decomposition of SST (Fig 3.1.6) shows that GECCO2 (red dashed lines) is warmer than CDA (red solid lines) in the tropical Pacific (top panel of Fig 3.1.6), where CIH shows a higher predictive skill than GIH at lead year 1. Such differences between initial conditions are less in the extra-tropical and sub-polar regions (bottom panel of Fig 3.1.6). For GIH, the CFES is initialized with oceanic estimations from interpolated GECCO2. Since GECCO2 has a different topography and dynamic from the model, by initializing the model with a different ocean mode from GECCO2, there will be adjustment through volume transport somewhere in the tropical Pacific (e.g. near Indonesia) at the beginning of the prediction. Such changes in velocity may lead to anomalously warm SST over the tropical Pacific, as shown in Fig 3.1.6. The relative warmer SST due to initialization then possibly leads to the relatively low skill of SST prediction in GIH.

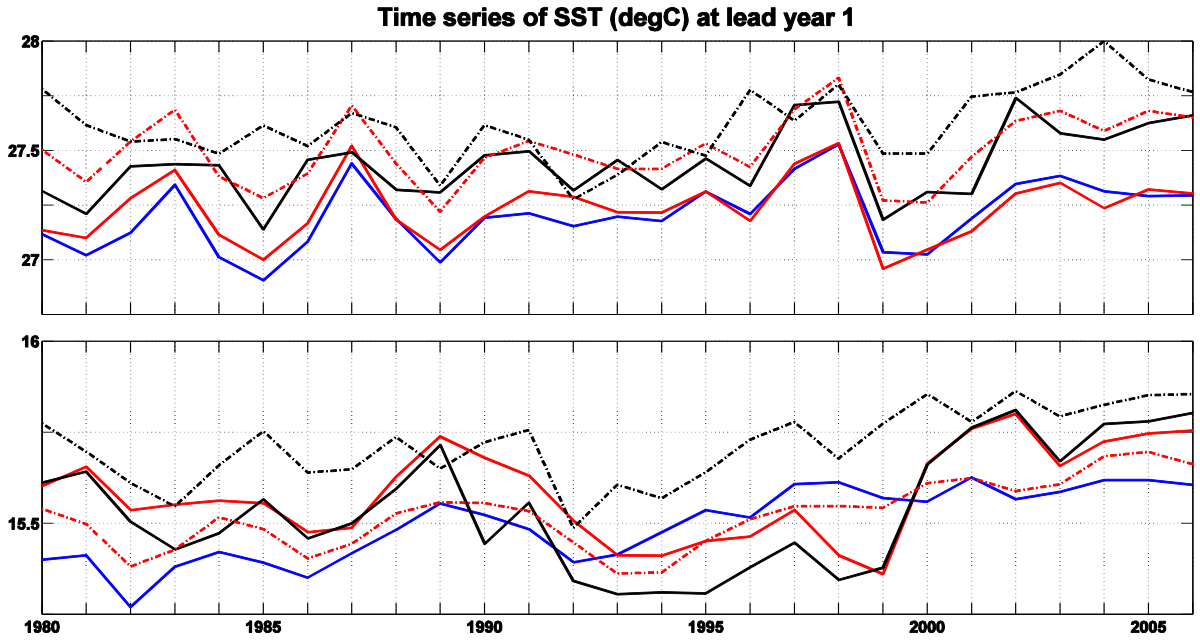

Fig 3.1.6 Time series of annual mean SST at the first lead year averaged over 20°S-20°N (upper) and 60°S-20°S & 20°N-60°N (bottom) in observations (blue, HadISST data set), hindcasts (solid black: CIH; dashed black: GIH), CDA (solid red) and GECCO2 (dashed red).

For the 4-yr average, SST predictive skill is also found over part of the Indian Ocean and the western tropical Pacific. Both GIH and CIH fail to capture the variation of SST over the North Atlantic, either interannunal or long-term variability. However, other studies indicate robust skill over the North Atlantic (NA), for forecasts initialized with GECCO Synthesis through different modeling systems, (e.g. Matei et al., 2012b; Polkova et al., 2014). Other studies also find predictive skill of NA SST (e.g. Pohlmann et al., 2009; Keenlyside et al., 2008; Matei et al., 2012c; Doblas-Reyes et al., 2013; Yang et al., 2013; Ho et al., 2012; Ham et al., 2014; Hazeleger et al., 2013a). In fact, Pohlmann et al. (2009) point out that SST predictive skill at decadal time scales is mainly due to the low North Atlantic SST in the early 1970s. As is indicated above, the evaluation period in our solution is restricted within 27-years period (i.e. 1980-2006), due to the available CDA synthesis, which is only half size of the 50-years period that is commonly used by previous studies. To investigate to what extent the reduced predictive skill in the North Atlantic in our solution may result from relatively short hindcast runs, predictive skill of CIH/GIH will be assessed against previous results later.

In the North Atlantic, SST is reported to be influenced little by radiative forcing at least in the first few years of decadal predictions (Branstator and Teng, 2010). Other studies (e.g. Enfield et al., 2001; Sutton and Hodson, 2005; Knight et al., 2006) indicate that, by the end of the twentieth century, externally forced components and internally generated variability have equal impact on North Atlantic SST, although contradictory results of a dominant impact from external forcing are found by a different methodology (Ting et al., 2009). Improved predictive skill of initialized forecasts over the tropical North Atlantic was found to be largely due to external forcing (e.g. Dunstone et al., 2013; Vecchi et al., 2013). Actually, a number of models suggest that anthropogenic aerosols may have a prominent influence on North Atlantic predictability in recent decades (Booth et al., 2012; Dunstone et al., 2013; Villarini and Vecchi, 2013), and the role of previous volcanic eruptions is also non-negligible (Otterå

et al., 2010). But the combination of skill from internally generated variability and externally forced response is usually region dependent (Meehl et al., 2014). Further analysis on the spatial distribution of GECCO2 initialized SST without trend removal indicates significant prediction skill (Fig 3.1.7) over the North Atlantic on decadal scale, and the skill is still significant until lead years 6-9. Compared with the relatively low predictive skill achieved for non-detrended SST from the initialized hindcasts, the trend gives rise to the predictability of SST over North Atlantic in GIH.

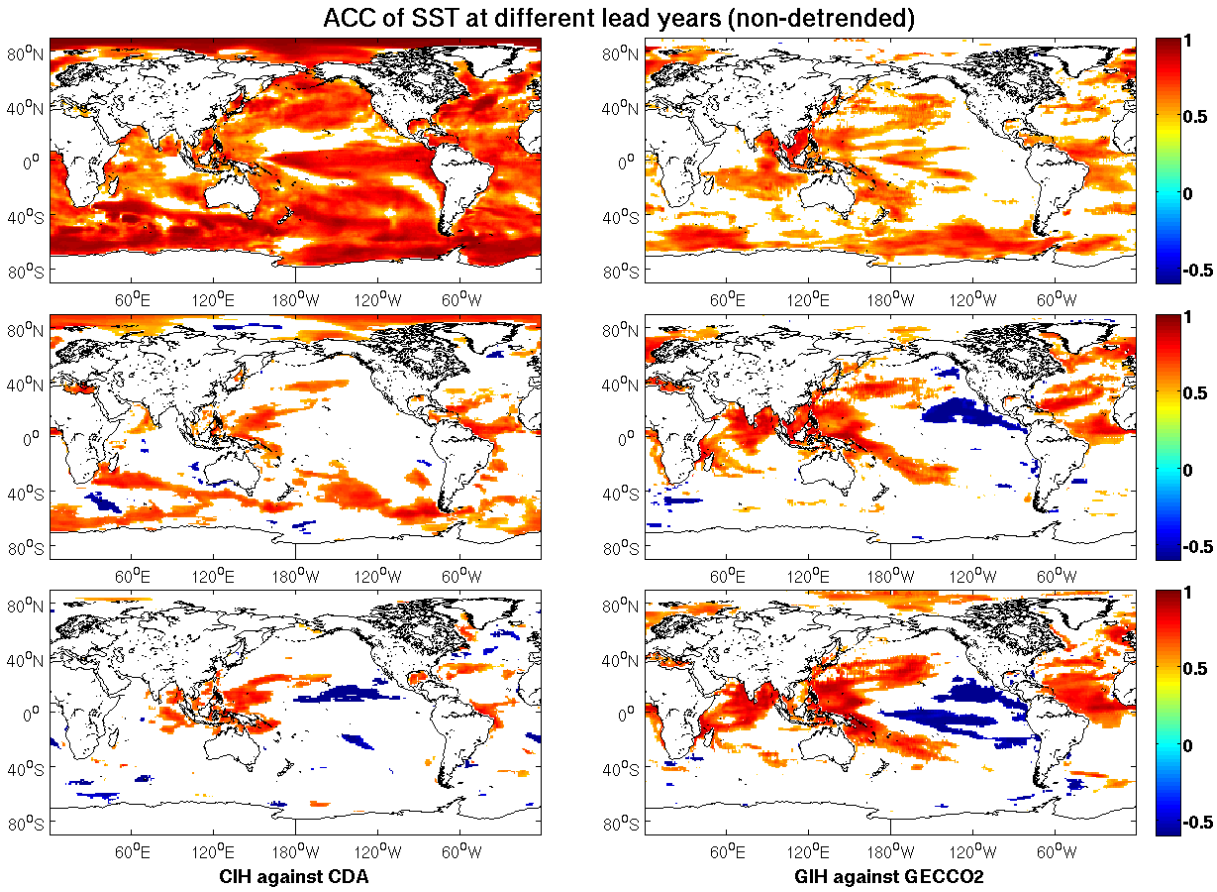

Fig 3.1.7 Spatial distribution of the non-detrended SST anomaly correlation coefficient between CIH and CDA (left), GIH and GECCO2 (right) separately, at lead year 1 (top panels), averages of yr 2-5(middle panels) and yr 6-9 (bottom panels). Only the significant coefficients (at 95% level) are shown here. Compared with corresponding spatial distribution of detrended SST ACC of left panels in Fig 3.1.4 and Fig 3.1.5, non-detrended SST from both CIH and GIH shows significant skill in decadal prediction over larger areas. The most significant improvement in predictive skill is found for GIH over the North Atlantic, western Pacific and the Indian Ocean.

On the other hand, some studies based on multi-model initialized forecasts expect that skill in NA SST prediction is related with predictability of AMOC due to initialization (e.g. Swingedouw et al., 2012). Cross correlation of the NA SST and maximum Atlantic MOC indicates that the AMOC predictability leads NA SST predictability by several years (e.g. Pohlmann et al., 2009). This lag is also supported by previous study of Latif et al. (2004). The poor predictive skill of the 4-yr average for the presented hindcasts could possibly be related

to poor predictive skill of AMOC in CIH/GIH through oceanic heat transport. To explore this, we focus further exploration on the AMOC predictability in the two sets of hindcasts.

Over the Indian Ocean and western tropical Pacific, significant predictive skill is also observed up to lead years 6-9 for GIH. In state-of-the-art decadal climate predictions, the regions with highest SST predictive skill are agreed by researchers in the Indian Ocean, which is possibly more due to externally forced trends but less due to internal climate variability (Ho et al., 2013; Guémas et al. 2013a), which is consistent with the larger areas with skill in the Indian Ocean for the non-detrended SST for the 4-yr averages of lead years 2-5 and 6-9, especially for GIH.

3.2 Evaluation of predictive skill through AMOC

3.2.1 Atlantic Meridional Overturning Circulation

As a part of the large-scale ocean circulation driven by wind and global density gradients resulting from surface heat and freshwater fluxes, the meridional overturning circulation (MOC) plays an important role for the ocean in carrying a substantial part of the heat transport towards the polar region, which has a significant impact on the Earth's radiation budget (e.g. Rahmstorf, 2003). The variability of Atlantic meridional overturning circulation (AMOC) is an important factor of the climate variability in the Atlantic Ocean and over northwest Europe on decadal timescales (Hurrell et al., 2006). Its variability provides potential prediction skill on interannual to decadal time scales (Collins et al., 2006). To date, no agreement has been reached in the AMOC evolution. Modelling results by Gregory et al. (2006) indicate that increased external forcing lead to decreased AMOC. However, due to the relatively short observation record, this discovery could not be evaluated, and the study of model response to historical forcing by Menary et al. (2013) reported a weakening downward trend of AMOC, or even strengthened AMOC, with a more realistic estimation of anthropogenic aerosols forcing. Lately, several studies showed prospect for AMOC prediction on decadal time scale from initialized hindcast experiments (e.g. Collins et al., 2006; Msadek et al., 2010). Pohlman et al. (2009) observed improved forecast skill of climate predictions up to decadal time scales for North Atlantic SST and AMOC through initialization from the GECCO synthesis, and predictability up to 6 years due to initialization was found in the later study of Pohlmann et al. (2012). The study by Matei et al. (2012a, b) reported multiyear AMOC predictability through initialization, and highlighted the importance of observed atmospheric state as a successful strategy for initialization in a skillful climate prediction. Polkova et al. (2014) explored the impacts of different initialization approaches on the evolution of AMOC, and found that the full state initialization provides higher predictive skill for AMOC than the anomaly initialization. So far, nobody has studied how the self-consistent initial conditions impact the AMOC evolution and predictability. To explore this, the AMOC that is initialized with self-consistent (CIH) and non-self-consistent (GIH) initial conditions are analyzed. Due to the lack of AMOC observations (e.g. Cunningham et al., 2007; Kanzow et al., 2007), the predictive skill of AMOC will be evaluated with respect of AMOC from CDA/GECCO2 separately. The AMOC predictability will also be verified in order to find out the underlying reason of poor NA SST predictive skill.

The AMOC is defined by integrating the velocity zonally in the Atlantic Basin, with the meridional mass transport streamfunction depending on latitude and depth defined as

$$\psi(y,z) = \int_{-H}^{z} \int_{east}^{west} v(x,y,z)dxdz$$

where v is the monthly/annual mean meridional velocity,$v = v(x,y,z,t)$, and H represents the depth along the zonal section.

3.2.2 Assessment of Atlantic MOC predictability

As indicated above, both sets of the initialized hindcasts in this paper fail to capture the long-term variability of SST anomalies over the North Atlantic. The predictability of the NA SST is model-dependent (Kim et al., 2012; Ham et al., 2014), and as is found by many studies, SST is more predictable in the subpolar gyre of NA than in other areas (e.g. García-Serrano et al., 2012; Terray, 2012; Yang et al., 2013). Nevertheless, the key role of AMOC in driving SST of the North Atlantic through heat transport is widely accepted (e.g. Latif et al., 2006; Marshall et al., 2001; Msadek et al., 2013; Pholmann et al., 2004; Srokosz et al., 2012). A potential of predictability up to decadal time scales resides in the variability of AMOC (e.g. Collins et al., 2006; Keenlyside et al., 2008; Kröger et al., 2012; Msadek et al., 2010). Predictability of AMOC is model-dependent and it varies with initial model states to some extent (IPCC AR5, 2013). An accurate initialization of AMOC is highly likely to improve predictive skill achieved in NA SST predictions (e.g. Latif and Keenlyside, 2011; Srokosz et al., 2012). The influence of a significant and realistic AMOC "oscillatory" variability translates into longer predictability (Meehl et al., 2014). Considering the different initial conditions used for the two sets of hindcasts, different variability and accuracy of AMOC from CDA/GECCO2 could possibly contribute to the skill patterns for NA SST prediction of CIH (Fig 3.1.4) and GIH (Fig 3.1.5) through initialization. To evaluate the AMOC predictability in CIH and GIH, the anomaly correlation coefficient of AMOC will be illustrated between hindcasts and the synthesis used as initial conditions (i.e. CDA and GECCO2 respectively), due to the lack of observational AMOC (Cunningham et al., 2007). Impact of consistent or non-consistent initial conditions to AMOC predictive skill will be provided. The AMOC predictive skill of initialized hindcasts will also be assessed against that of the uninitialized forecast.

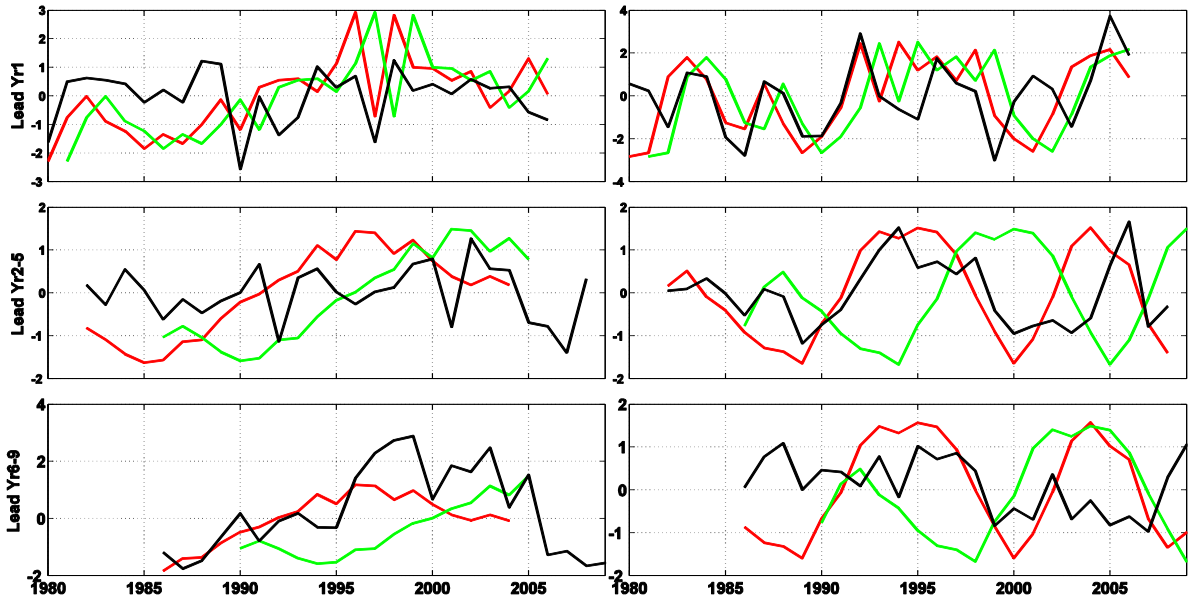

Fig 3.2.1 Time series of annual mean AMOC (Sv) anomaly (at 1000m) at 26.5°N from hindcast experiments (black), CDA (red in left panels) and GECCO2 (red in right panels), and persistence (green). The left panels are CIH and CDA used for initialization, while the right panels are GIH and GECCO2 used for initialization respectively. The top panels are AMOC at lead year 1. The 4-yr mean of lead years 2-5 is plotted with the time series plotted centered in year 3(middle). Averages over lead years 6-9 are shown in the bottom panels, with time series plotted centered in year 7.

As is mentioned above, AMOC from GECCO2/CDA will be used to evaluate the AMOC predictability of GIH/CIH. The AMOC anomalies of the upper 1000m at 26.5°N from initialized hindcasts CIH and GIH reproduce its variability in the lead year 1 (upmost panels of Fig 3.2.1), despite the different AMOC evolutions in CDA and GECCO2, from which the model is initialized. However, for the average over lead years 2-5, an abrupt negative anomaly is observed around 1992, followed by a positive phase afterward for CIH, whereas there are two abrupt negative anomalies in the early 1990's at the first lead year. In GIH, AMOC of lead years 2-5 resembles that of GECCO2 quite well. As is revealed by studies of Ishii and Kimoto (2009) and Levitus et al. (2009), a large abrupt warming of the high-latitude upper ocean in the North Atlantic happened in the mid-1990s, preceded by anomalously cool conditions in the subpolar gyre region of 50°-10°W, 50°-60°N and warm condition in 70°-30°W, 32°-42°N about three pentads before (Yeager et al, 2012). The AMOC fluctuation is highly related with the subpolar gyre (Eden and Willebrand, 2001; Pohlmann et al., 2009). The abrupt shifts in the North Atlantic circulation observed in Fig 3.2.1 lead to lagged changes in the heat content of the mid-1990s. For lead years 6-9, AMOC in CIH captures the increasing trend before the mid-1990s and decreasing trend afterwards of that in CDA, despite the poor simulation of low-frequency variability. Overall, AMOC evolution in CDA is quite different from that in GECCO2, which can lead to incompatibility of the initial conditions to the model in GIH. A striking difference between time series of AMOC is that GECCO2 contains more multi-year variability than CDA, as is shown in the right panels of 4-yr average of Fig 3.1.7. Köhl and Stammer (2008) explored the sources of the decadal and long-term AMOC variability from GECCO, and showed that the intensifying AMOC is related to different dynamical processes. The more "oscillatory" AMOC variability from GECCO2

gives rise to predictability on decadal prediction of AMOC in GIH through persistence. Acting together with the external forcing embedded within the model, hindcasts initialized with GECCO2 reproduce the long-term variability of AMOC well. Therefore higher AMOC predictability from GIH is expected at lead yr2-5 than that of CIH.

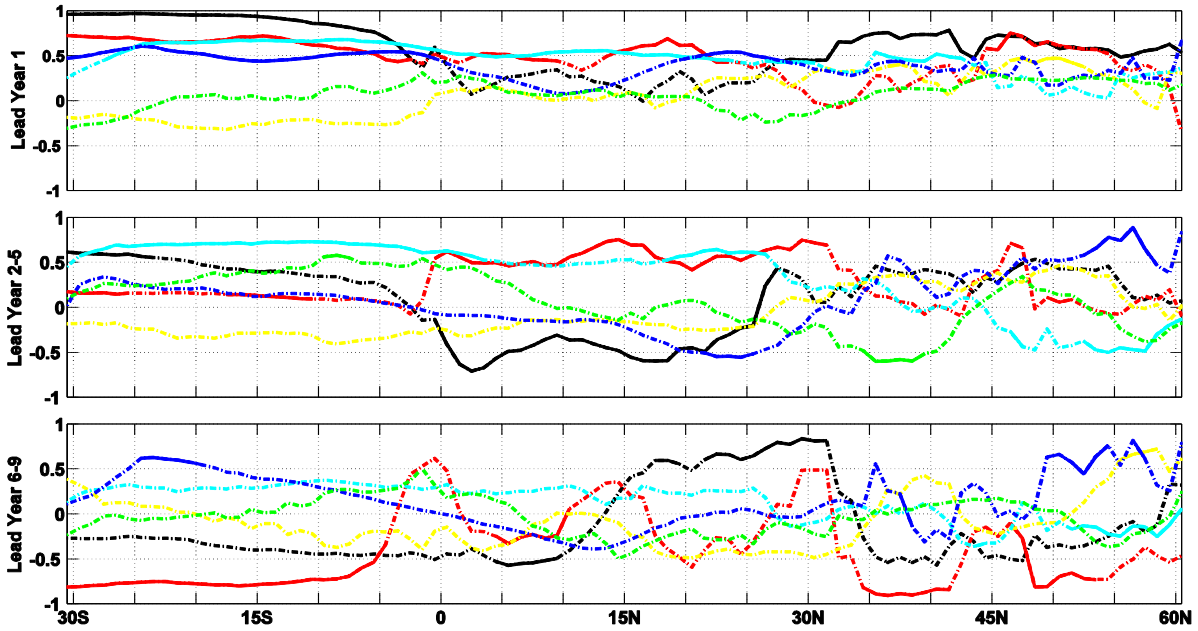

Fig 3.2.2 Anomaly correlation coefficient of the non-detrended Atlantic MOC (Sv, $Sv = 10^6 m^3 s^{-1}$) at 1000m depth between CIH and CDA (black), CDA and persistence (red), CIH and 20C (yellow), GIH and CECCO2 (cyan), GECCO2 and persistence (blue) and GECCO2 and 20C (green) along different latitudes of 30°S-60°N. ACCs are calculated for the first lead year (upper panel), 4-yr average of lead years 2-5 (middle) and lead years 6-9 (bottom panel) respectively. The dashed lines denote the hindcasts that don't obtain significant skill at 95% significant level. All the AMOC data in this figure are not detrended.

We evaluate both the "total" predictive skill (before detrending, Fig 3.2.2) and the "residual" skill (with the long-term linear trend removed, Fig 3.2.3) (van Oldenborgh et al., 2012) of the AMOC along different latitudes at different lead years through ACC skill. The degrees of freedom (dof) of AMOC is latitude-dependent and therefore are calculated through application of the effective sample size factor along different latitudes. The correlations that are below the 95% threshold of statistical significance (*t*-test) are shown as dashed lines.

Before removing the linear trends, anomaly correlation coefficients of AMOC along latitudes (Fig 3.2.2) reveal better AMOC predictive skill for both initialized hindcast and persistence at the first lead year, despite the possible incompatibility of AMOC through initialization in GIH. On the contrary, the externally-forced uninitialized simulation hardly shows any skill in AMOC predictability, which also holds for the detrended AMOC. Therefore, external forcing provides almost no skill in AMOC predictability in our expriments, as in other studies (e.g. Pohlmann et al., 2013). For the non-detrended AMOC at lead year 1, initialized hindcasts outperform the persistence forecast over the tropical South Atlantic for both CIH and GIH, as well as for a large area between 28°N-60°N for CIH. However, predictability at this time scale is related to persistence, indicated by the significant

correlations of persistence. For lead years 2-5, the patterns are different: GIH outperforms CIH over the southern hemisphere and partially in the extra tropical NA of 17°-26°N. Initialized hindcasts still show larger regions with high correlation than persistence forecast for GIH, with the regions pretty much similar as the first lead year. However, an evident decrease of initialized hindcasts for CIH is observed, while persistence forecast has significant predictive skill in the tropical North Atlantic. The time series of AMOC at 26.5°N shows that CIH resembles the AMOC variability only at lead year 1, while GIH captures the AMOC variability of lead years 2-5 (Fig 3.2.1). Hence, GIH shows significant skill in AMOC predictability of lead years 2-5. The high skill for persistence and more multi-decadal AMOC variability shown in GECCO2 indicate that the variability of AMOC is important for AMOC predictability, through the initialization procedure, despite the possible incompatibility for the model introduced through initial conditions.

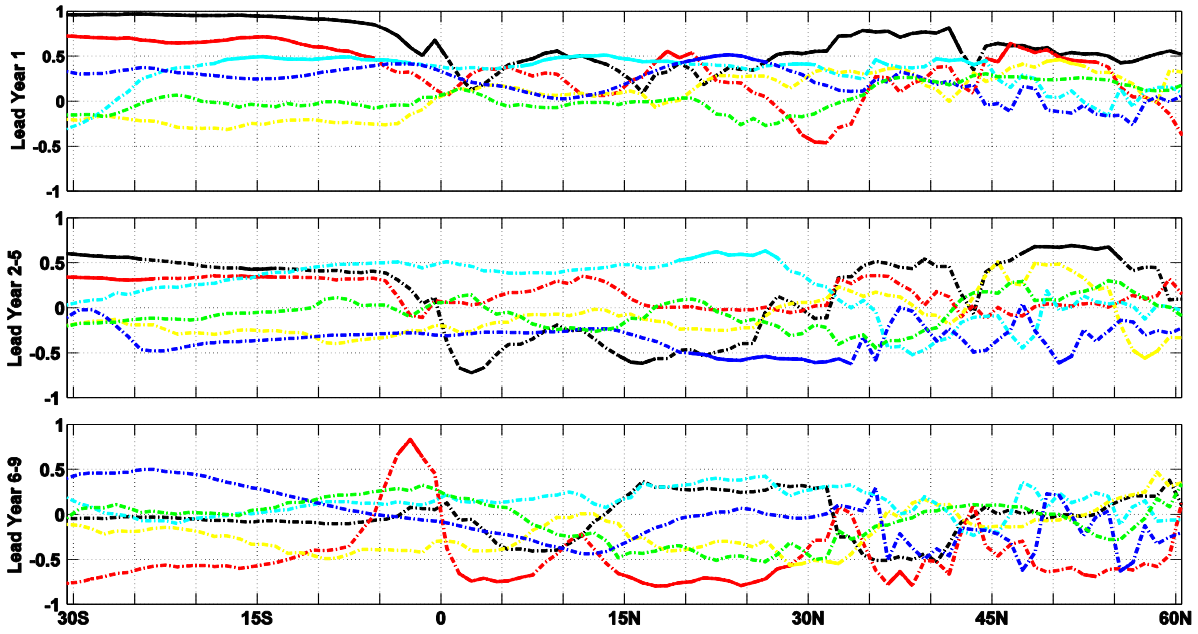

Fig 3.2.3 Anomaly correlation coefficient of the detrended Atlantic MOC(Sv, $Sv = 10^6 m^3 s^{-1}$) at upper 1000m between CIH and CDA (black), CDA and persistence (red), CIH and 20C (yellow), GIH and CECCO2 (cyan), GECCO2 and persistence (blue) and GECCO2 and 20C (green) along different latitudes of 30°S-60°N. ACCs are calculated for the first lead year (upper panel), 4-yr average of lead years 2-5 (middle) and lead years 6-9 (bottom panel) respectively. The dashed lines denotes the hindcasts fail to obtain significant skill at 95% significant level.

When the long-term trend is removed from AMOC, different patterns of skill are revealed. A reduction of predictive skill is detected for CIH, while a much more significant decrease is found for the persistence at lead years 2-5. The reduction of predictive skill is also observed for GIH, and the most pronounced decrease of skill in GIH is at lead years 2-5, as well as the persistence forecast at lead year 1 (only a small area between 20°N and 26°N with prediction skill), when the long-term trend is removed. Overall, the most significant AMOC predictive skill observed is between CIH and CDA at the first lead year, e.g. 30°-0°S, 25°N-60°N and around 10°N, with correlations up to 0.5 for detrended AMOC (upper panel of Fig 3.2.3). The persistence also shows significant although relatively small correlations compared with initialized hindcasts at the first lead year. The 4-yr average of lead years 2-5 shows similar

pattern as lead year 1 but with smaller correlations. Consistent with previous results (Pohlmann et al., 2009), significant AMOC predictive skill is gained in CIH at high latitudes until lead yr 2-5. The most evident decrease is observed for GIH when the data is detrended. This indicates that the trend in GECCO2 AMOC (Fig 3.2.1) has a strong influence on the climate prediction of North Atlantic, while that in CDA is probably small and has a lower impact on AMOC predictability. The trend of the AMOC imposed in the hindcasts is crucial to increase the predictive skill, which is most likely associated with the initialization precudure. However, due to the lack of observations for AMOC, it's difficult to evaluate whether the trend of AMOC is true. For the average of lead years 6-9, no predictive skill of AMOC is achieved. The low skill of the 20C forecast indicates that the contribution to predictive skill from the forced response is low.

Illustration of the time series of AMOC indicates more significant correlations of AMOC between GIH/GECCO2 than CIH/CDA at least near 26.5^{o}N, since more oscillatory decadal variations observed in GECCO2 possibly lead to predictability even in persistence forecasts. Consistent with that, significant predictive skill is observed at 26.5^{o}N until lead years 2-5, for both the non-detrended and detrended GIH. However, the ACC of detrended AMOC between GIH and GECCO2 reveals much smaller area with significant correlations until lead years 2-5 than before the removal of the linear long-term trend. High prediction skill between GIH and GECCO2 is shown only along the latitudes of 20^{o}-27.5^{o}N, when the trend is removed. Moreover, the persistence forecast (blue lines) shows significant predictability of AMOC only over part of the North Atlantic (around 15^{o}-27^{o}N) at the first lead year, while the non-detrended persistence shows larger area with skill in the South Atlantic (30^{o}N-0^{o}). The initialization of GIH does improve predictive skill of AMOC, compared to low predictive skill of 20C, as is proved by other researchers (e.g. Doblas-Reyes et al., 2013). But the long-term trend of AMOC from GECCO2 has stronger influence on decadal predictability, while the influence of AMOC trend from CDA is negligible.

Analysis on the previous version GECCO indicates that the strengthening trend of AMOC is likely related to the North Atlantic Oscillation (NAO; Köhl and Stammer, 2008). The mechanism has also been explored by Eden and Willebrand (2001), and they pointed out that an increased AMOC was largely due to the lagged response (about 2-3 years) to enhanced atmospheric activity associated with high NAO. A similar conclusion by Gastineau and Frankignoul (2012) reported an increase of AMOC to be followed by the negative phase of NAO. Hence, the trend of AMOC as a lagging response of NAO leads to predictability of NA climate change, possibly by inducing basinwide SST anomalies (Branstator and Teng, 2014). The strong trend of AMOC in GECCO2 imprints on GIH through initializing the model and leads to high predictive skill as is shown in Fig 3.2.2. In addtion, more multi-year variability residing in GECCO2 AMOC leads to AMOC predictability of GIH through persistence. Therefore, high predictability of non-detrended AMOC in GIH is still observed in the average of lead years 2-5. While for CIH, the consistent initial conditions lead to the high predictive skill at the first year for both the detrended and non-detrended AMOC. The relative small trend of AMOC in CDA and less multi-year variability result in poor predictive skill at longer lead times. Considering the high predictive skill in GIH of non-detrended AMOC and non-detrended NA SST, the trend contributes to North Atlantic climate predictability. Further study about impact of different periods on decadal climate predictions reveals that the trend has a strong influence to North Atlantic predictive skill in a short period (1960 to present), which is actually multi-decadal variability if longer period is considered (Müller et al., 2014).

Previous studies indicate a connection between maximum MOC and NA SST predictability through heat transport, as the former leads the latter several years, in initialized hindcasts (e.g. Pohlmann et al., 2009) and simulations forced with constant external forcing (e.g. Latif et al., 2004). The variability of AMOC, if the variability is more oscillatory and realistic, may translate into predictability of NA SST through its impact on the overlying atmosphere (e.g. Gastineau et al., 2012). The structure of Atlantic MOC in CDA, GECCO2, and the initialized/uninitialized hindcasts are different. The maximum MOC is located at 35.5^{o}N in GECCO2 Synthesis and CDA, whereas it occurs in the South Atlantic in the uninitialized hindcast (22.5^{o}S, not shown). For the initialized hindcasts, the maximum MOC is located at 37^{o}N/28^{o}N for GIH/CIH. As is revealed by the spatial distribution of SST predictive skill, significant predictive skill for NA SST is only observed in CIH at the first lead year. For AMOC predictability, significant predictive skill at 35.5^{o}N is observed only at the first lead year in CIH, while no significant skill is observed for AMOC at 35.5^{o}N in GIH. Over longer lead periods of yr 2-5, large correlation coefficient is gained for CIH there, but not statistically significant. The reduction of AMOC predictive skill in the subtropical NA over longer lead periods is possibly due to the dominating role of Rossby waves which are not predictable (Köhl, 2005; Hirschi et al., 2007; Köhl and Stammer, 2008). The poor predictive skill of SST in the North Atlantic is consistent with low AMOC predictability in our solutions. There is no evident lagged response of AMOC predictability with NA SST predictive skill as is found in some other studies (e.g. Latif et al., 2004; Pohlmann et al., 2009), suggesting again that the relation between NA SST and AMOC predictability is model-dependent (e.g. Pohlmann et al., 2009). Studies (e.g. Branstator and Teng, 2014; Tiedje et al., 2012, 2014) on predictability between AMOC and meridional heat transport indicate that meridional heat transport is less predictable than AMOC, despite their close relationship. The poor NA SST predictive skill is possibly due to low performance of the system in capturing the variability of meridional heat transport.

To summarize, significant predictive skill of AMOC is gained at the first lead year for both CIH and GIH, with the former outperforming the latter, indicating again that initializing a model with self-consistent initial conditions can improve predictive skill, although that is restricted to the first lead year. At longer lead times, GIH lose AMOC predictive skill in the subpolar NA, where higher predictability is assumed. Different from that, significant AMOC predictive skill is obtained in the subpolar area of North Atlantic until lead yr 2-5 in CIH. Atlantic MOC structure is different in CDA and GECCO2. High predictive skill for basinwide maximum AMOC is only obtained in CIH until lead year 2-5 (significant skill at the first lead year, while lead yr 2-5 not significant). Therefore, no evident relation is denoted between NA SST and AMOC predictive skill. The low performance of the system in capturing AMOC multi-year variability is consistent with low predictive skill of NA SST. AMOC in GECCO2 has a strong increasing trend and constains more multiyear variability, which is possibly a lagged response to the negative NAO phase (Köhl and Stammer, 2008). Multi-year variability of AMOC in GECCO2 gives rise to predictability of GIH though persistence, as is observed for detrended MOC at 26.5^{o}N in Fig 3.2.3 at lead years 2-5. The strong increasing trend for the AMOC in GECCO2 improves predicitve skill of GIH through initialization, and the influence remains significant until lead years 2-5. However, due to the lack of observation for MOC, the trend in GECCO2 is difficult to prove to be true. On the other hand, the trend impacts AMOC predictive skill less in CIH than in GIH, possibly due to a relative small trend from CDA. External forcing does not contribute to AMOC predictability in either CIH or GIH.

3.3 Assessment on the influence of verification period and ensemble members

So far we have investigated the influence of different initial conditions on SST/AMOC predictive skill. It's also a necessary requirement to evaluate the skill of the system against previous results. Given the CDA Synthesis is available only for the period 1980-2006, the evaluation of skill is restricted to only half the size of the 50-yr period that typically has been used in previous studies for evaluating the skill of hindcasts. To compare with previous skill, we investigate the impact of the time period of the model runs that provide initial conditions, and that can be used for evaluation of the hindcasts on decadal prediction (i.e. verification period). Besides, there is only one ensemble numbers for hindcasts at each initialized dates instead of 3 (e.g. Keenlyside et al., 2008; Pohlmann et al., 2009; Müller et al., 2014) or more as is used in previous results (e.g. Smith et al., 2008). Generation of ensemble forecasts is also required in decadal prediction, which samples the spread due to uncertainty (Meehl et al., 2009). To explore this, full-state initialized hindcasts dffGcE performed as part of the Mittelfristige Klimaprognosen (MiKlip hereafter) project based on the Max-Planck-Institute Earth System Model (MPI-ESM) and initialized from the GECCO2 Synthesis are compared to the hindcasts presented here. In addition, during the period 1975-2006, AMO, which is associated with the connection between the AMOC and surface temperature, is mainly in a negative phase with a slowly increasing trend. The hindcast period of 1980-2006 in CIH/GIH may lose some internal variability in the North Atlantic, due to the trend removal.

In decadal climate predictions, larger numbers of initial dates and a larger number of realizations in an ensemble are important in improving decadal prediction (Meehl et al., 2014): the former helps detect predictability due to lower significance levels, and the latter increases predictability by using ensemble means instead of individual realization. MiKlip hindcasts dffGcE cover the period 1961 to 2009, with 3 realizations at each start date. Hence, a comparison over the same period as CIH/GIH and the investigation of the influence of the number of initialization dates are feasible. Better predictable skill is expected for dffGcE. Comparison between performances of ensemble means of dffGcE over 1980-2006 with that of individual ensemble member offers us insights on how ensemble members impact predictability. An extra comparison between performances of dffGcE of 1961-2009 and that of 1980-2006 will be carried out to provide information on the sensitivity of predictive skill to verification periods.

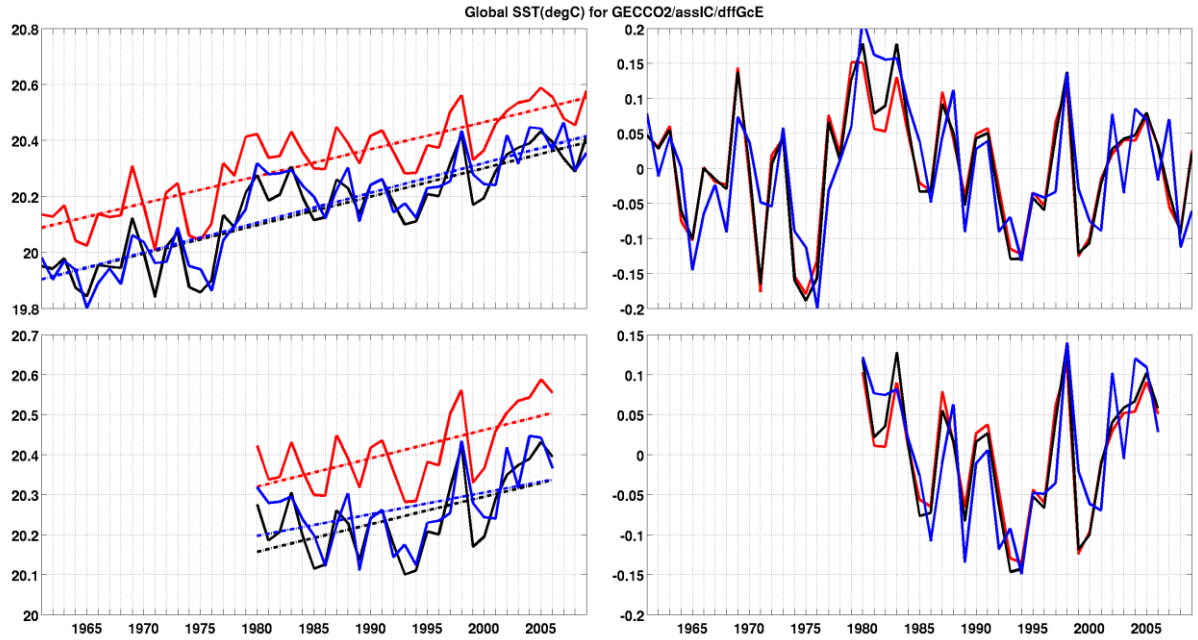

Fig 3.3.1 Time series of annual mean global SST (averaged over 60^{o}S-60^{o}N) from GECCO2 (red), assimilated GECCO2 (balck) that is used as initial conditions for dffGcE, and ensemble mean of hindcast dffGcE at the first lead year (blue). The trends are denoted in dashed lines of left panels. The lines in the right panels denote the time series after the long-term linear trend removed. As is indicated by the black lines, the warm biased GECCO2 is assimilated towards the observations (HadISST, not shown).

As is illustrated in Fig 3.3.1, after nugding the model MPI-ESM towards GECCO2, the initial conditions are less warmer than GECCO2 and much closer to observations (not shown). Initialized from the close-to-observation oceanic state, the ensemble mean of dffGcE (blue in Fig 3.3.1) at lead year 1 resembles the assimilated GECCO2 (black in Fig 3.3.1) quite well overall. A much warmer phase during 1980-1985 is observed for dffGcE. Hence, a less stronger trend is derived in dffGcE of 1980-2006 compared with that of GECCO2. When the trend removal is applied in the skill of correlation coefficient between dffGcE and hindcasts, large values are derived for dffGcE of the shorter period (0.774) and of the longer period (0.736). Therefore, high predictive skill is expected at the first lead year for dffGcE.

We start the analysis from the impact of verification period to predictive skill in terms of SST. As is indicated by comparison between the left and right panels in Fig 3.3.2, spatial distributions of ACC for SST in the dffGcE hindcasts with longer verification period (1961-2009) show similar patterns with that of the shorter verification period at lead year 1 and 6-9. Large areas of skill are observed all over the ocean at lead year 1, and larger areas of significant skill overall are shown for the period 1961-2009 despite the relatively smaller value of the correlation coefficients at some regions compared with that of the period 1980-2006. More initial dates lead to lower thresholds for the statistical significance of skill due to more samples. Therefore, despite the smilar pattern of SST predictive skill in dffGcE of 1961-2009/1980-2006, the relative small coefficients in some areas observed for the former (left panels of Fig 3.3.2) are still significant. For longer lead time, larger areas with predictive skill for SST in dffGcE of 1961-2009 are achieved than that of 1980-2006, and the significant skill remains until lead years 6-9 in the North Atlantic. But in the other regions, smaller

coefficients are revealed for SST from dffGcE of 1961-2008, compared with that of shorter period (1980-2006). This indicates that a larger number of initialization dates in decadal climate prediction help to reduce the threshold of significance. Hence, larger areas with skill are achieved until lead years 6-9, but with low correlation coefficients. The predictive skill with smaller coefficients in terms of SST from dffGcE of 1961-2009 against GECCO2 is possibly due to the application of trend removal, given that the trend after the mid 1970s is much stronger than before.

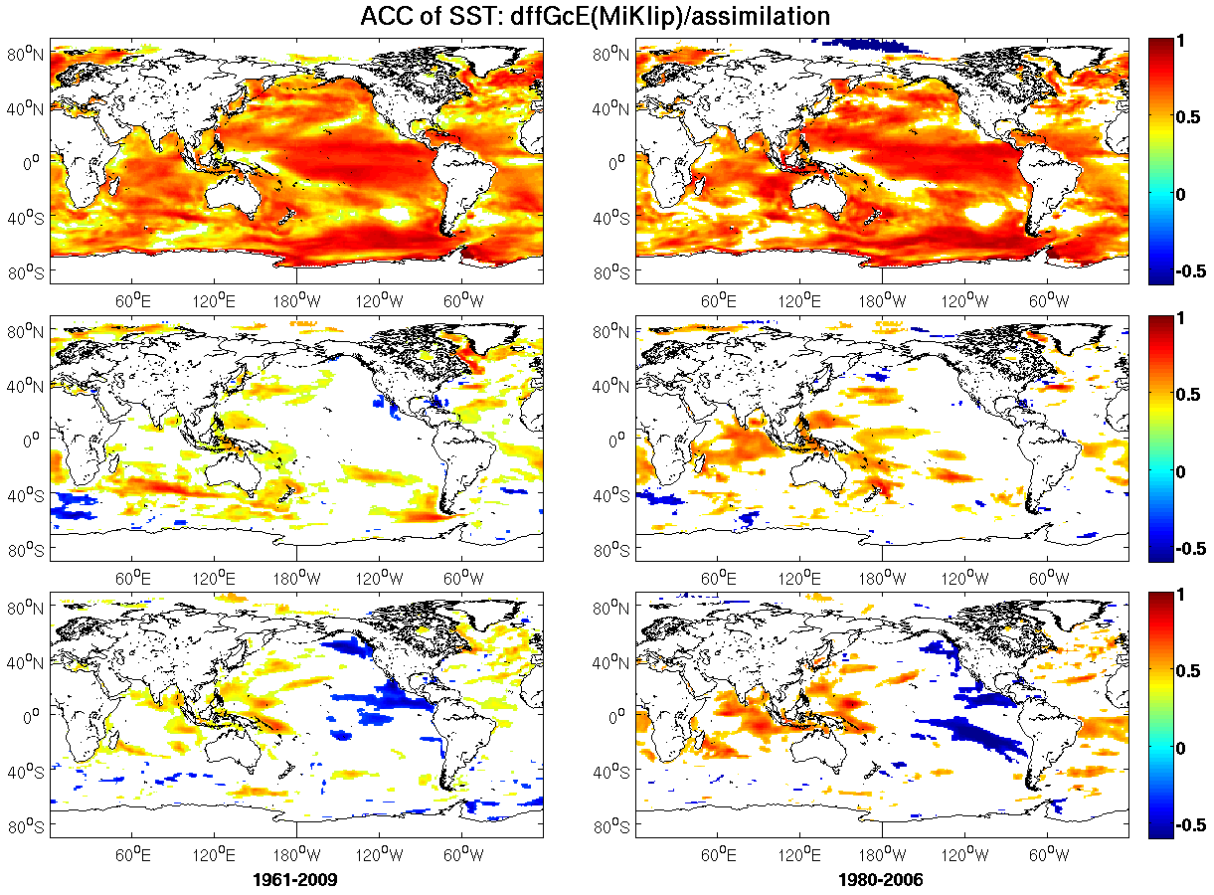

Fig 3.3.2 Spatial distribution of the hindcasted SST anomaly correlation coefficients between dffGcE (MiKlip hindcasts) and assimilated GECCO2 at lead year 1 (top panels), averages of lead years 2-5 (middle panels) and 6-9 (bottom panels). The right column corresponds to the period in our hindcasts from 1980 to 2006, and the left column is from 1961-2009. All the SST used is detrended. Only the significant coefficients (at 95% level) are shown here.

For NA SST, a larger number of initial dates doesn't only improve credibility of predictive skill, but also gives rise in prediction skill. In the North Atlantic, which is identified as a key region for decadal climate predictions by previous results (e.g. van Oldenborgh et al., 2012; Matai et al., 2012), SST predictive skill is found to be high with verification period of 1960 to present day (e.g. Doblas-Reyes et al., 2013). Consistent with that, robust skill over the North Atlantic is also found at lead year 1 and the skill remains until the 4-yr average in the subpolar area of North Atlantic. Maps of ACC for SST predictive skill at lead years 2-5 reveal significant, large correlation coefficients for dffGcE of 1961-2009, whereas predictive skill is lost over the middle of North Atlantic for dffGcE of 1980-2006. This is possibly due to more

intrinsic multi-year variability of the North Atlantic SST in dffGcE of 1961-2009 than that of 1980-2006 (not shown), which leads to predictive skill through persistence. A study by Müller et al. (2014) shows improvement of NA SST predictive skill while extending the verification period from 1901 to 2010. Therefore, for the North Atlantic predictability, longer verification period is necessary.

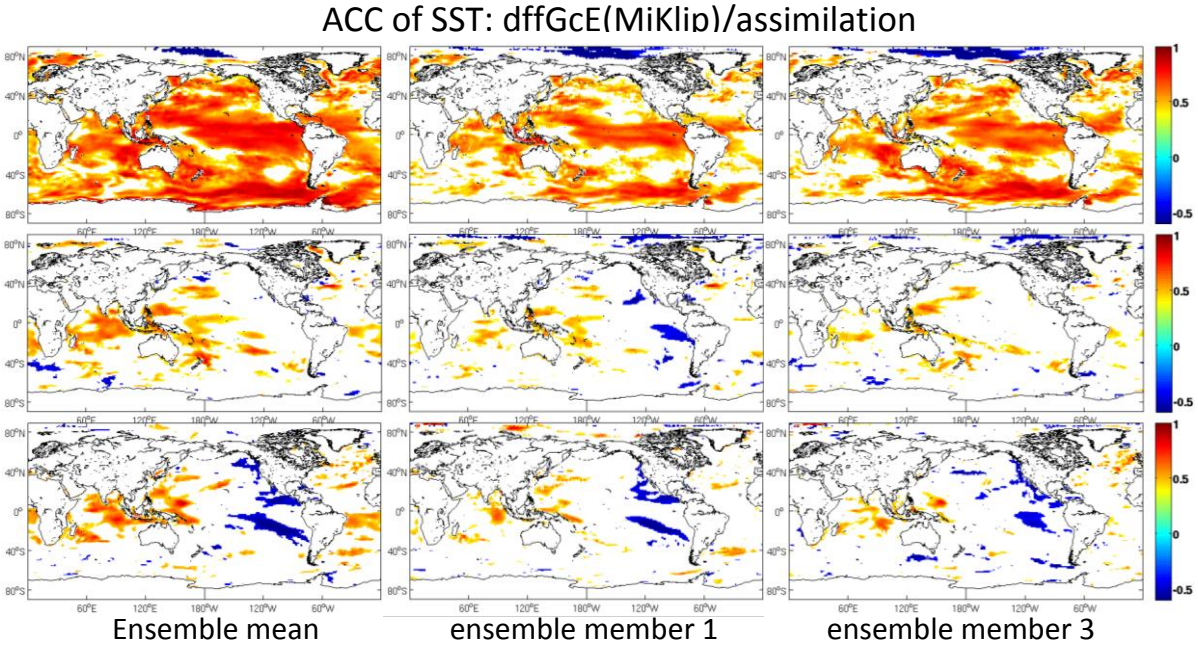

Fig 3.3.3 Spatial distribution of the hindcasted SST anomaly correlation coefficients between dffGcE (MiKlip hindcasts) and assimilated GECCO2 at lead year 1 (top panels), averages of lead years 2-5 (middle panels) and 6-9 (bottom panels). The left column is from ensemble mean of dffGcE, and the middle and right columns are from ensemble member 1 and 3 of dffGcE respectively, all with the same verification period of 1980-2006. All the SST used is detrended. Only the significant coefficients (at 95% level) are shown here.

For evaluation of SST predictive skill, more realizations at each initial date (i.e. more ensemble members) give rise to preditive skill, due to usage of ensemble means instead of individual realizations. As is displayed in Fig 3.3.3, higher predictive skill is obtained in ensemble means of dffGcE (left columns) at all time scales than that of individual ensemble member (middle and right columns). At the first lead year, high predictive skill is gained over vast areas of the ocean in ensemble mean of dffGcE. Over longer lead periods, significant predictive skill is gained in the Indian Ocean by using emsemble means of hindcasts. Therefore, given the impact from verification period and ensemble members, it's necessary to make sure that the hindcasts are with the same verification period and same ensemble size while comparing the performance between different hindcasts.

Further information on how the consistent initial conditions impact the decadal predictive skill can be derived by comparing the middle or right columns in Fig 3.3.3 with those in Fig 3.1.4/Fig 3.1.5. The ACC maps for hindcast SST from dffGcE for the same period (i.e. 1980-2006) as CIH/GIH reveal compatible spatial distribution of significant correlations with that of CIH/CDA, but larger correlations than that of GIH/GECCO2, at lead year 1. An improvement in SST predictive skill over the eastern part of tropical Pacific as is observed in CIH is obvious for dffGcE. Considering that only GIH shows low performance in the eastern

Pacific, further analysis is needed to evalute decadal predictability of regional SST in the Atlantic and the tropical Pacific.

For the 4-yr averages up to lead year 9, a similar significant drop in predictive skill is found for the three hindcasts, of which dffGcE is comparable to GIH. Hence, given the same verifiaction period and same ensemble numbers of hindcasts, the performance of our system is as good as that of dffGcE through MPI-ESM system. The reduced predictive skill in SST over longer lead periods is related to small ensemble memebers and short verification period. In the North Atlantic, SST predictive skill is strongly influenced by verification period. Former study (Bjerknes, 1964) also indicates that the long-term changes of NA SST may be forced by variations in ocean dynamics. AMOC predictability may lead to an improvement of NA SST predictability through heat transport (e.g. Pohlmann et al., 2009), with a lag of several years. To explore this, a regional assessment on decadal predictability of the Atlantic MOC is needed. Time series of AMOC evaluation will be illustrated. The performance of hindcast AMOC will be evaluated through anomaly correlation coefficients against GECCO2, as is the case for GIH.

The time series of annual mean AMOC at 1000m depth at 45^{o}N is shown in Fig 3.3.4. There is a drop up to 6 Sv after nudging MPI-ESM-LR towards GECCO2 for the AMOC (not shown), although the assimilation run reproduces the decadal variations well (Fig 3.3.4). A slow positive trend is observed, with higher values after around 1987. However, recent multi-model retrospective predictions by Pohlmann et al. (2013) revealed that there was an intensifying trend of AMOC strength at 45^{o}N from the 1960s to mid-1990s, and a decreasing trend afterwards. The regime shift of AMOC observed for dffGcE and GECCO2 is around 1987 as is shown in Fig 3.3.4. The AMOC strength starts to decrease around 2000, with a lag of around 5 years compared with the multi-model ocean synthesis results. For GECCO2, there is another AMOC strength decrease in 1993 or so, while by nudging the MPI-ESM towards GECCO2, the decrease of AMOC strength in 1992 becomes smaller and is followed by a larger decrease after the mid-1990s.

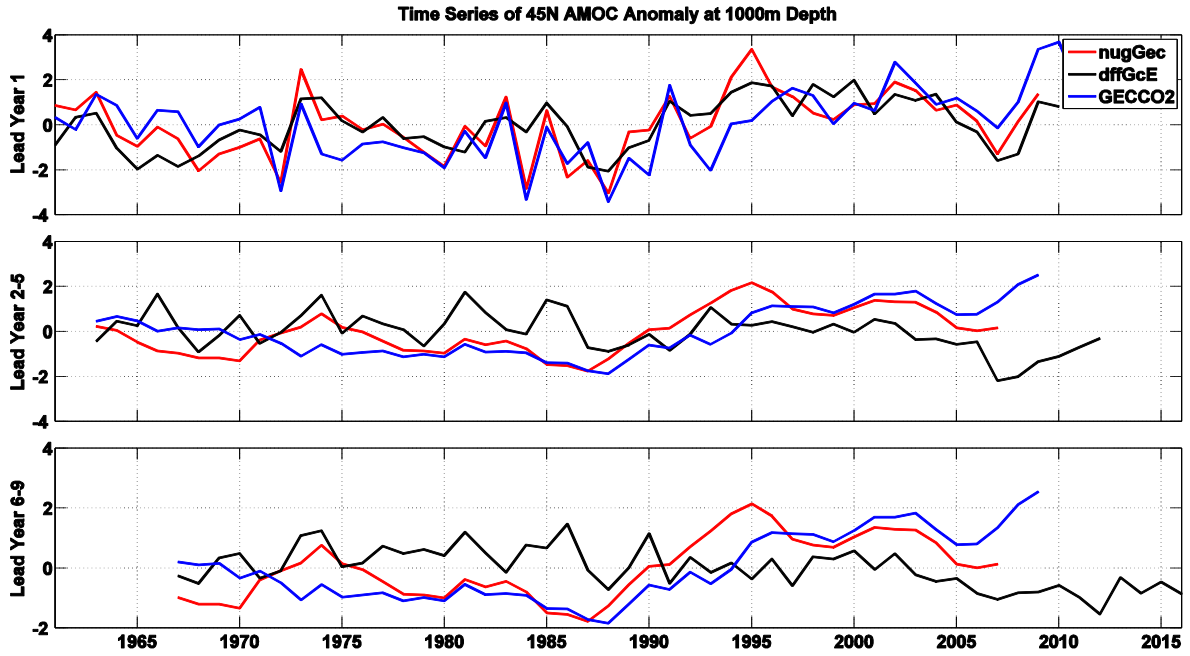

Fig 3.3.4 Time series of annual mean NA MOC (Sv) anomaly of upper 1000m at 45^{o}N from dffGcE hindcast (black), the assimilation run used for initialization (red, i.e. after nudging the MPI-ESM to GECCO2), and GECCO2 Synthesis (blue) at different lead years. The top panel is AMOC at lead year 1. The 4-yr mean of lead year 2-5 is plotted with the time series centered in year 3 (middle). Averages over lead year 6-9 are shown in the bottom panel, with time series plotted centered in year 7.

Both GECCO2 and assGec (data derived by nudging MPI-ESM-LR towards GECCO2 and used for initialization of dffGcE) show substantial multi-year variability for the 4-yr averages, which may lead to the possibility of decadal prediction skill of the AMOC through persistence. The initialized dffGcE shows much smaller trend but more low-frequency variations compared to either the GECCO2 Synthesis or the assGec, especially for the 4-yr average of lead years 6-9. For the average of lead years 2-5, hindcast AMOC resembles the variability of assGec well, but fails to capture the variability for average of lead years 6-9. Consistent with this, hindcast AMOC with more initial dates (1961-2009) shows better significant correlation coefficients with GECCO2 up to lead years 6-9 than that of period 1980-2006 (Fig 3.3.5). The correlation in the South Atlantic (30^{o}S-0^{o}) and over the tropical North Atlantic (0^{o}-35^{o}N) remains significant until lead years 2-5, with values larger than 0.5, for the 1961-2008 dffGcE hindcasts (black lines, Fig 3.3.5). This may contribute to the high SST predictive skill observed in the North Atlantic at lead years 2-5 observed for dffGcE of 1961-2009. For the AMOC predictability of dffGcE during 1980-2006, significant predictive skill (red lines, Fig 3.3.5) is achieved over the tropical Atlantic at lead year 1 and the average of lead years 2-5, similar to GIH.

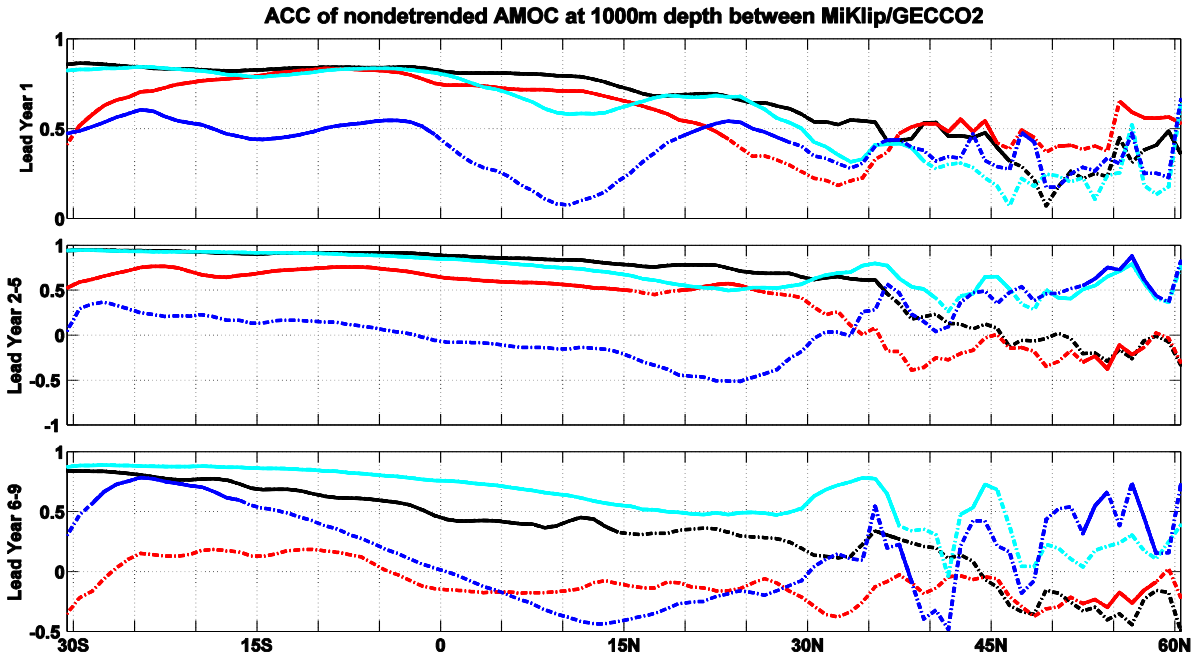

Fig 3.3.5 ACC of non-detrended Atlantic MOC(Sv) at 1000m depth between dffGcE hindcasts and GECCO2 (black: 1961-2009, red: 1980-2006), GECCO2 and persistence (cyan: 1961-2009, blue: 1980-2006) at different lead times along latitudes. The top panels are at lead year 1, the middle and bottom panel are 4-yr average of lead year 2-5 and 6-9 seprately. The dashed part of the lines denotes the correlation fails to pass the 95% significance level.

However, at lead years 6-9, no prediction skill of AMOC is observed (consistent with GIH) before the long-term trend is removed. Note, for the 1961-2008 AMOC predictability, persistence already shows large significant correlations with GECCO2 until lead years 6-9. For lead years 6-9, the persistence outperforms the initialized hindcasts, indicating that for AMOC predictability, much of the predictability probably comes from the persistence, at least for 1961-2008. When the prediction evolves to lead years 6-9, the persistence tends to provide more prediction skill. However, a large decrease of skill from persistence of 1980-2006 is observed, with correlation around 0.5 in the southern tropical Atlantic. For the 4-yr average, no significant skill is observed for persistence of 1980-2006.

Predictive skill in terms of detrended AMOC is also evaluated through anomaly correlation coefficient skill, as is shown in Fig 3.3.6. The most significant and robust skill on decadal prediction of AMOC is still achieved at lead year 1 over a large area, indicated by the large coefficient up to 95% significant level. Despite the larger area (about 22^{o}S-30^{o}N) with significant predictive skill achieved for dffGcE during 1961-2009, significant correlations with larger values are achieved for dffGcE during 1980-2008 in the tropical Atlantic (25^{o}S-20^{o}N) at lead year 1. Therefore, the ratio of AMOC predictive skill due to the trend is larger in dffGcE of 1961-2009 than that of 1980-2006. Indicated by the red and black lines of middle panel in Fig 3.3.6, a larger number of initial conditions don't improve predictive skill of AMOC if the linear long-term trend is removed, but increases credibility of predictive skill due to more samples. Compared with prediction skill of the non-detrended AMOC, a reduction in the skill is observed in both initialized hindcasts and persistence. The most significant decrease is the persistence forecast of 1961-2009. This is consistent with the

AMOC predictive skill in GIH and that of persistence forecast. The strong increasing trend of AMOC in GECCO2 contributes to predictive skill.

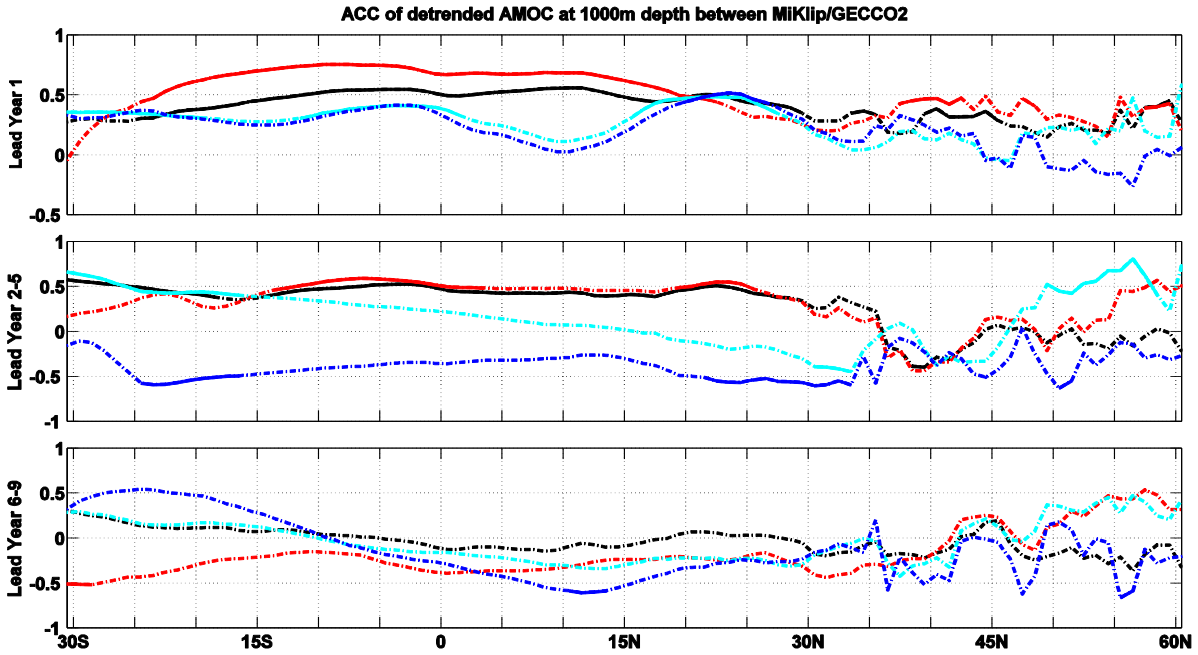

Fig 3.3.6 ACC of detrended Atlantic MOC(Sv) at 1000m between dffGcE hindcasts and GECCO2 (black:1961-2009, red: 1980-2006), GECCO2 and persistence (cyan: 1961-2009, blue: 1980-2006) at different lead times along latitudes. The top panels are at lead year 1, the middle and bottom panel are 4-yr average of lead year 2-5 and 6-9 seperately. The dashed part of the lines denotes the correlation fails to pass the 95% significance level.

Larger areas and higher correlations are derived for detrended AMOC from dffGcE during 1980 to 2006 than for GIH, at lead year 1 and lead years 2-5. In addition to the difference in the model used for hindcasts, the main difference between dffGcE and GIH is the atmosphere and initial conditions. For the initialization of dffGcE, the model is first nudged towards the GECCO2 Synthesis. The "assimilated" data is thereafter used to initialize the model. Such strategy improves the consistency of initial conditions to the model, particularly because the velocities are not nudged and are free to adjust. Hence a better balanced state is achieved at the initialization time. More realizations at each initialization date for dffGcE help to reduce the threshold of the significance of prediction skill. However, for the GIH, by directly initializing the model with the interpolated GECCO2 Synthesis, there are likely initial shocks at the beginning of the prediction. Meanwhile, by initializing with the atmosphere from CDA, a less balanced state between the atmosphere and the ocean is highly likely at the initialization time. The air-sea interaction through the coupled mechanism of the model may result in poor reproduction of AMOC variability as the prediction evolves. Therefore, better predictability in terms of AMOC is achieved for dffGcE than GIH. Different from any of them, the initial shock and difference in "observed"/model-biased climatology are all reduced through the initial conditions that are more-dynamically consistent for CIH. Therefore, better AMOC predictability is observed significantly at lead year 1 for CIH. But for lead years 2-5, dffGcE (1980-2006) outperforms CIH. Comparison among the AMOC predictive skill at the first lead

year from CIH, GIH and dffGcE (1980-2006) highlights the importance of the consistency of the initial conditions to the model in the initialized hindcasts.

For the 4-yr average of predictions, the 1961-2009 dffGcE outperforms dffGcE of the period 1980-2006. Skillful AMOC predictions for the 1961-2009 dffGcE appear between 30^{o}S and around 27^{o}N up to lead years 2-5, with significant correlations up to 0.5, due to the lower significance threshold. Hence, in order to gain larger area with predictive skill, hindcasts with more initial dates are necessary, which lead to longer verification periods due to more samples. Long term variability is crucial for AMOC predictions on decadal time scales if successfully captured, which improve predictive skill through persistence. However, due to the lack of observations, it's difficult to prove the low-frequency oscillatory variability of AMOC in GECCO2 to be true. When comparing performances between different predictions, one should always make sure that the hindcasts are used over the same period or at least are of the same length. The trend of AMOC from GECCO2 increases the predictive skill in AMOC variability for all initialized hindcasts.

To summarize, evaluation of predictive skill in CIH/GIH against dffGcE of 1980-2006 indicates that the performance of CIH is compatible with that of individual member of dffGcE, with the most significant predictive skill over most of the ocean observed at the first lead year, in terms of SST. Meanwhile, dffGcE outperforms GIH, especially at the first lead year. More realizations at each initial date improve predictive skill in dffGcE, due to the usage of ensemble mean instead of individual realization. Comparison between dffGcE of 1980-2006 and that of 1961-2009 in terms of SST predictive skill reveals similar patterns, but smaller correlation coefficients in some areas for the latter. A larger number of initial dates help reduce the significance level, but in some areas leads to smaller correlation coefficients, which is possibly related with the application of trend removal. An exception in the North Atlantic like Labrador Sea with higher predictive skill is observed in dffGcE of 1961-2009 compared to that of 1980-2006 at lead years 2-5, possibly due to more multi-year variability in SST variations and partially due to higher predictive skill of AMOC. Illustration on predictive skill of Atlantic MOC indicates that the trend of AMOC give rise to predictive skill of dffGcE, with that of 1961-2009 outperforms that of 1980-2006. Robust predictive skill is observed until lead years 2-5 over large area. However, when the linear long-term trend is removed, a reduction is observed in predictive skill. Despite the slightly smaller values of correlation coefficient for AMOC, dffGcE of 1961-2009 still shows larger areas with significant predictive skill than of 1980-2006, due to longer period of verification. The trend and low-frequency variability contained in the AMOC lead to predictive skill in dffGcE and GIH through initialization with GECCO2, although it is difficult to evaluate.

Chapter 4

Mechanism leading to low prediction skill

So far, we have evaluated the regional SST predictability on decadal time scales. The most striking impression is that CIH shows significantly high correlations over the tropical Pacific and North Pacific at the first lead year, while predictive skill over the same region is much lower for GIH. Less predictive skill over the North Pacific is also observed by Goddard et al (2012a) and Doblas-Reyes et al. (2013). On the other hand, initializtion of the Model for Interdisciplinary Research On Climate (MIROC) provids predictive skill over the extropical North Pacific (Mochizuki et al., 2010, 2012). Studies also show that prediction skill of the climate in the North Pacific is model-dependent (Mochizuki et al., 2010; Chikamoto et al., 2012a): the model's ability to reproduce subsurface temperature changes gives rise to predictive skill. Predictability of the Pacific is more sensitive to the initial state (Branstator et al., 2012; Branstator and Teng, 2012) and intrinsic decadal climate variability (Meehl et al., 2014). In the Pacific, interannual climate variability is highly characterized by ENSO (El Niño-Southern Oscillation). The relationship between ENSO and the Pacific dacadal climate variability has been explored by many studies (eg. Power et al., 1999a; Meehl and Hu, 2006; Matei et al., 2008; Meehl et al., 2010), but no agreement has been reached. In this thesis, CIH is initialized with initial conditions that are dynamical self-consistent with the CFES model. This "self-consistent" initial conditions may lead to predictabilities of ENSO. Hence, we focus the following analysis on ENSO predictability and the possible reasons that give rise to predictive skill.

4.1 Predictive skill on El Niño events

4.1.1 El Niño Southern Oscillation (ENSO)

Abbreviated from the phrase "El Niño Southern Oscillation", ENSO is usually referred to a quasi-periodic fluctuation (i.e., every 2–7 years) in the sea surface temperature of the tropical eastern Pacific Ocean and in air surface pressure in the tropical western Pacific. The term "El Niño" was originally applied to a warm oceanic phase that naturally occurred in the equatorial region off the Pacific coast of South America. It is usually accompanied with high air surface pressure in the western Pacific. The oscillation in surface air pressure between the tropical eastern and the western Pacific Ocean was hence termed "Southern Oscillation" as the atmospheric component of ENSO (Trenberth, 1997). Since the quasi-periodical events were observed around Christmas, the Spanish term "El Niño" was used to name it (El Niño means Christ child, Jesus). The term "El Niño" now is used to refer to the related Pacific Ocean sea-surface temperature changes and regional weather events, rather than just the ocean part of these events (Glantz, 1996). Ever since the 1950's, scientists have been studying the mechanisms that cause the surface air pressure see-saw shift. In the early 1990s, researchers started to make initialized ENSO predictions based on models and the physical mechanisms of ENSO (Stockdale et al., 1998). Recent efforts made by ENSO community and decadal

climate prediction community attempt to provide reliable ENSO predictions based on coupled global climate models (e.g. McPhaden et al., 2010; Smith et al., 2012b). Researchers now have better understanding of the global weather effects caused by the formation of an El Niño, and they have made a lot of efforts on ENSO forecasting in the future and recent ENSO changes, either based on observations (e.g. An and Wang, 2000; Cobb et al., 2003; Yeh and Kirtman, 2005; Power and Smith, 2007; Braganza et al., 2009; Li et al., 2011c; Yan et al., 2011) or model results (Lau et al., 2008; Wittenberg, 2009). IPCC AR5 (2013) reported that multi-modeled statistics of El Niño-Southern Oscillation (ENSO) have improved since AR4, and the further development of ENSO event is predictable up to 1 year in advance (e.g. Ludescher et al., 2014). For example, Newman (2007) and Alexander et al. (2008) indicated predictability of ENSO at a year through a multivariate empirical Linear Inverse Models (LIMs) from observational SSTs. Ludescher (2014) developed an alternative approach upon network analysis and reported 1 year ahead for El Niño projection. The ENSO is the most important factors to the interannual variability of the tropical Pacific, and it will remain the most significant mode of the natural climate variability in the 21st century (Collins et al., 2010; Guilyardi et al., 2012; Kim and Yu, 2012; Stevenson, 2012). However, the understanding of the processes responsible for the development of El Niño still remains incomplete, and the relationship between ENSO and decadal predictability in the Pacific is not fully understood (e.g. Meehl and Hu, 2006; Matei et al., 2008; Meehl et al., 2010).

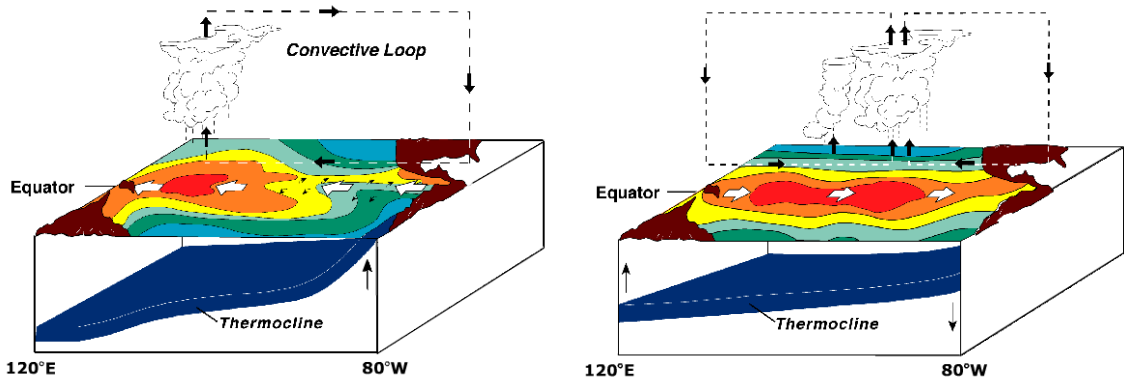

Fig 4.1.1 The Pacific Pattern under non-El Niño conditions (left) and El Niño conditions (right) from NOAA. In normal situation, equatorial winds gather warm water pool toward the west, and cold water upwells along the South American coast. During an event of El Niño, warm water pool approaches the South American coast. The absence of cold upwelling increases warming in the eastern Pacific. The absence of cold upwelling increases warming in the eastern Pacific.

Widely accepted, the formation of an El Niño is linked with the Walker Circulation. Under normal conditions (non-El Niño conditions), the strong easterly trade winds pile up warm surface water in the western Pacific and carry warm air heated up by the sun toward the west along the equator. As a result, the sea surface in the western Pacific is higher than in the eastern Pacific. The western equatorial Pacific shows a warm, wet, low-pressure weather pattern, while in the east side of Pacific Ocean, upwelling develops off the coast near Peru and Ecuador, and cold ocean current flows northward along the coast of Chile and Peru (i.e. the Humboldt Current). The first signs of an El Niño are a weakened Walker circulation and strengthened Hadley circulation (IPCC AR4, 2007). In an El Niño, the strong easterly trade winds are reduced along the equator (Fig 4.1.1, right) or head east in the Southern Pacific, due to the change in pressure pattern between eastern and western Pacific. Warm ocean water

approaches the eastern Pacific along the coast of Chile and Peru, resulting in a deeper thermocline and higher surface there. Hence a warmer pattern is observed along the equatorial coast of South America instead of the cold pattern under normal conditions.

Considering the significant influence of ENSO on interannual climate variability, the high/low predictive skill of tropical Pacific SST from CIH/GIH at the first lead year is possibly related to ENSO events. Therefore, we now explore the possible mechanism regarding the predictive skill of hindcasted SST through analyzing key characteristics of the El Niño. For this purpose, the El Niño index based on the sea surface temperature in region 3.4 (Fig 4.1.2; 170°W-120°W, 5°S-5°N) will be used.

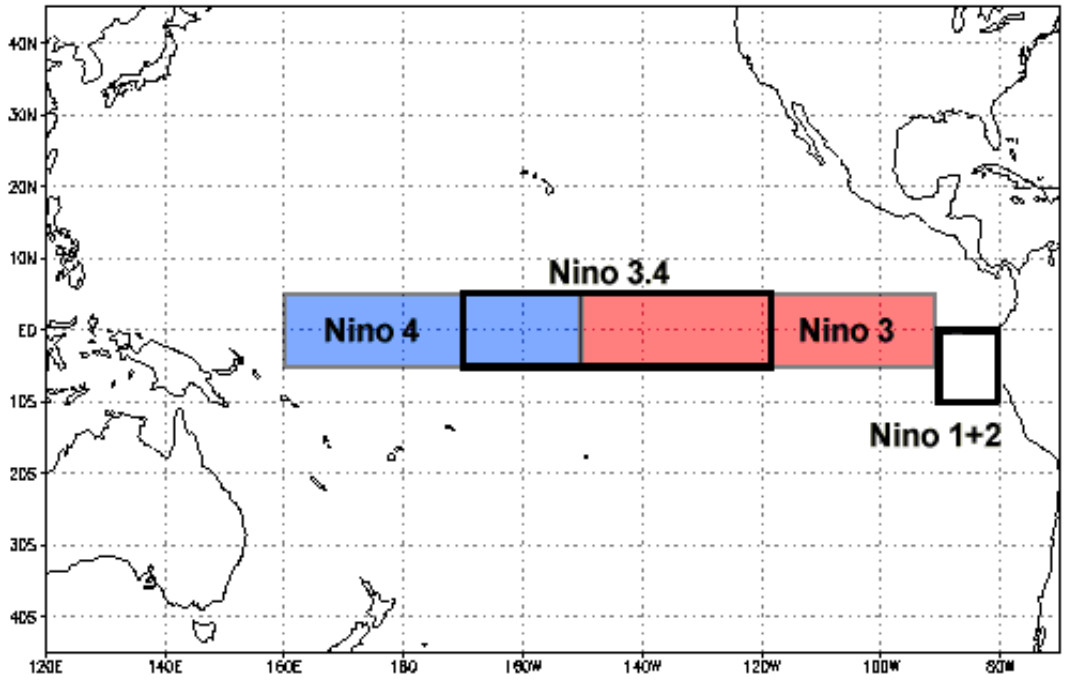

Fig 4.1.2 Location of Niño sea surface temperature regions (Niño 1, 2, 3, 3.4 and 4 regions) from NOAA. In this thesis, Niño 3.4 Index based on sea surface temperature of Niño 3.4 region is used to analyze the characteristic of ENSO. Figure is available from official website of NOAA (www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php).

4.1.2 Assessment on predictive skill of El Niño events

To identify and characterize the nature of El Niño-Southern Oscillation (ENSO), different indices have been introduced based on sea surface temperature (SST) anomalies in different regions of the Pacific (Trenberth and Stepaniak, 2000; Fig 4.1.2). The Niño 3.4 Index derived from SST anomalies in the Niño 3.4 region (referred to as N3.4) can be thought of as the mean SST throughout the equatorial Pacific east of the dateline. In this paper, Niño 3.4 Index is chosen to characterize the evolution of the ENSO event. To get the index, the area averaged total SST from Niño 3.4 region ($NSST_{im}$) is computed first. Later the monthly climatology

(1980-2006) of total SST averaged over the same area ($\overline{NSST_m}$) is subtracted to get time series of SST anomalies ($NSST_{im}^a$). The equation is listed below:

$$NSST_{im}^a = NSST_{im} - \overline{NSST_m}$$

Here in this formula, m denotes the month, i denotes the year, a represents anomaly, and $NSST$ denotes the value of SST averaged over Niño 3.4 region. Therefore, $\overline{NSST_m}$ is the climatology of month m over the period of 1980-2006, $NSST_{im}^a$ is the SST anomaly (over Niño 3.4 region) of month m in year i. The 5-months running method is applied to the anomalies in order to smooth out the possible intraseasonal variations from $NSST_{im}^a$ (Trenberth, 1997). Finally the smoothed SST anomalies are normalized by its standard deviation over the climatological period. The Niño 3.4 Index from NCAR CGD's Climate Analysis Section is used as a reference (www.cgd.ucar.edu/cas/catalog/climind/TNI_N34/). For the NCAR index, a base period climatology of 1950-1979 is chosen to get the time series. Restricted by the available period of 1980-2006 from CDA, we choose the base period of 1980-2006.

Table 4.1.1 Standard deviations (s) for SST in the Niño 3.4 region over the base period 1980-2006 (CDA, CIH, GECCO2, and GIH)/1950-1979 (NCAR) in ^{o}C.

	s(Niño 3.4)
CDA	0.95
CIH	0.90
GECCO2	0.98
GIH	0.95
NCAR	0.70

The different choices of background climatology (base period) may lead to difference in the index, since offset of the response in the atmosphere to tropical SST anomaly is possible (Trenberth, 1997; Tab 4.1.1). As is shown in Tab 4.1.1, the standard deviations of the background chosen in this thesis are larger than that of NCAR. This is because the mean SST in the tropical Pacific is biased warmer after 1979 (e.g. Trenberth and Hoar 1996a). Despite the difference shown, a commonly accepted threshold of indication to ENSO events is ±0.4^{o}C for Niño 3.4 (Trenberth, 1997; Trenberth and Stepaniak, 2000). The index calculated from different hindcasted SST (black) and the corresponding initialization data (red) are shown in Fig 4.1.3, in comparison with the NCAR index (green).

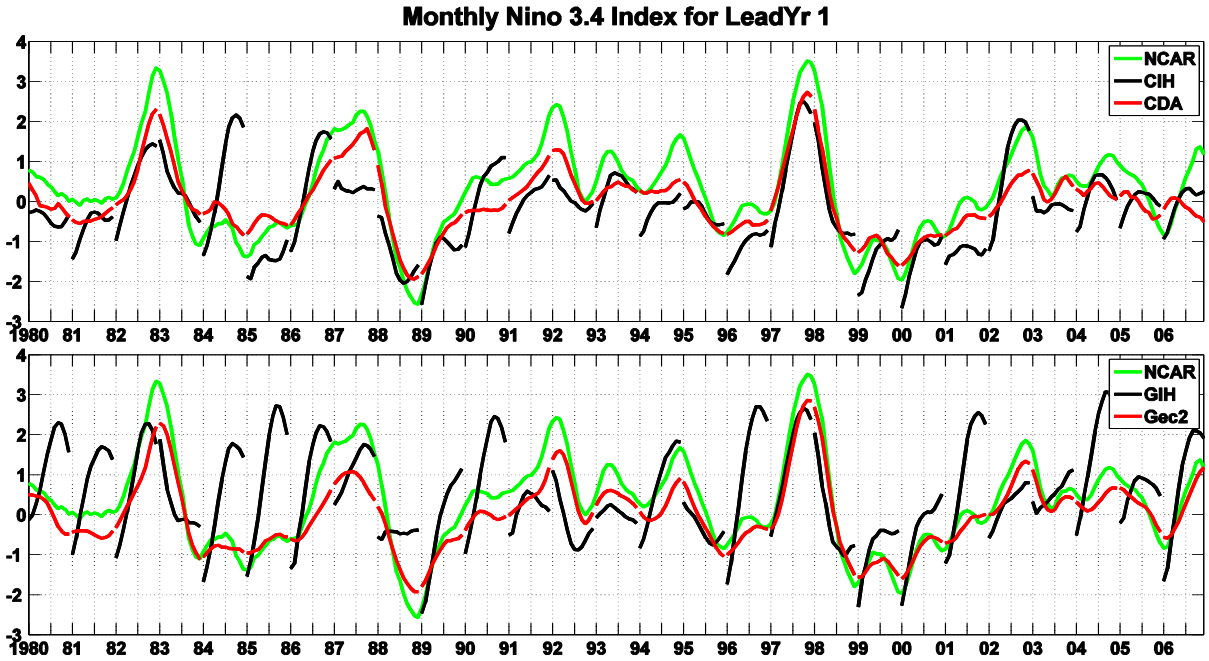

Fig 4.1.3 Niño 3.4 Index of lead year 1 from the hindcasts (black), NCAR (green) and the initialization data (red, CDA and GECCO2 respectively). The upper panel is for CIH and CDA. The bottom panel is GIH and GECCO2.

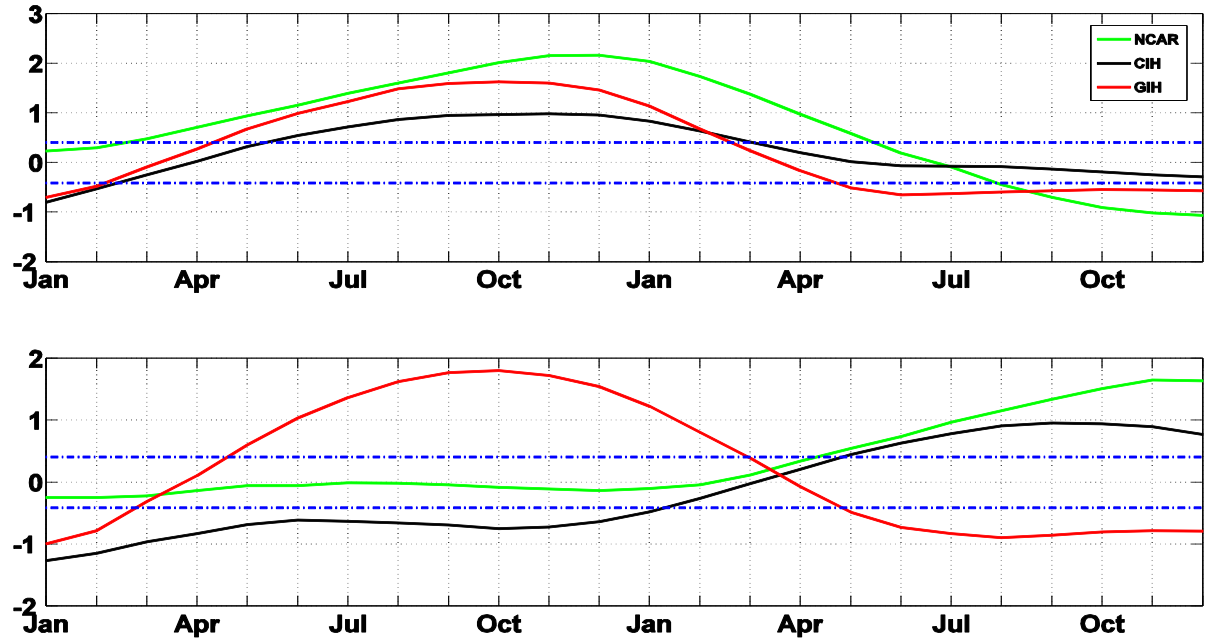

Fig 4.1.4 Climatological monthly Niño 3.4 Index over the first 24 months of historical El Niño years (i.e. 1982, 1986, 1991, 1994, 1997, 2002 and 2006; upper) and GIH produced El Niño-like years (i.e. 1980, 1981, 1985, 1989, 1996, 2001, 2003, 2005; bottom) from CIH (black), GIH (red) and NCAR (green). For the hindcasts CIH and GIH, all the monthly indices shown are calculated from the climatology of the first 24 months of forecasts started from the initialization date. The index from NCAR is used as a reference. The blue dashed lines represent the threshold for ENSO event for CIH and GIH. The value exceeding ±0.4^{o}C indicates an El Niño event.

By comparing the time series of the hindcasted monthly Niño 3.4 Index during the first lead year to the NCAR data, it becomes obvious that CIH reproduces ENSO events for most of the hindcasts in the first lead year (Fig 4.1.3). However, besides the reproduced historical ENSO events, some additional erroneous El Niño events are also detected, especially for GIH. CDA initialized hindcasts produce only one El Niño-like event in 1984 while surprisingly, GIH produces 8 more—1980, 1981, 1984, 1985, 1989, 1996, 2001, 2003 and 2005. The climatology of the Nino 3.4 Index from GIH over the first 18 months after initialization shows that El Niño-like events (Fig 4.1.4, middle panel) develop around May, reach peak intensity during September-October and tend to weaken during the following February, lasting around 10 months. The developing phase is almost the same as that of the reproduced historical El Niño years from CIH/GIH (Fig 4.1.4, top panel), with the peak during September to October, which is slightly different from the NCAR's index peaking during November-December.

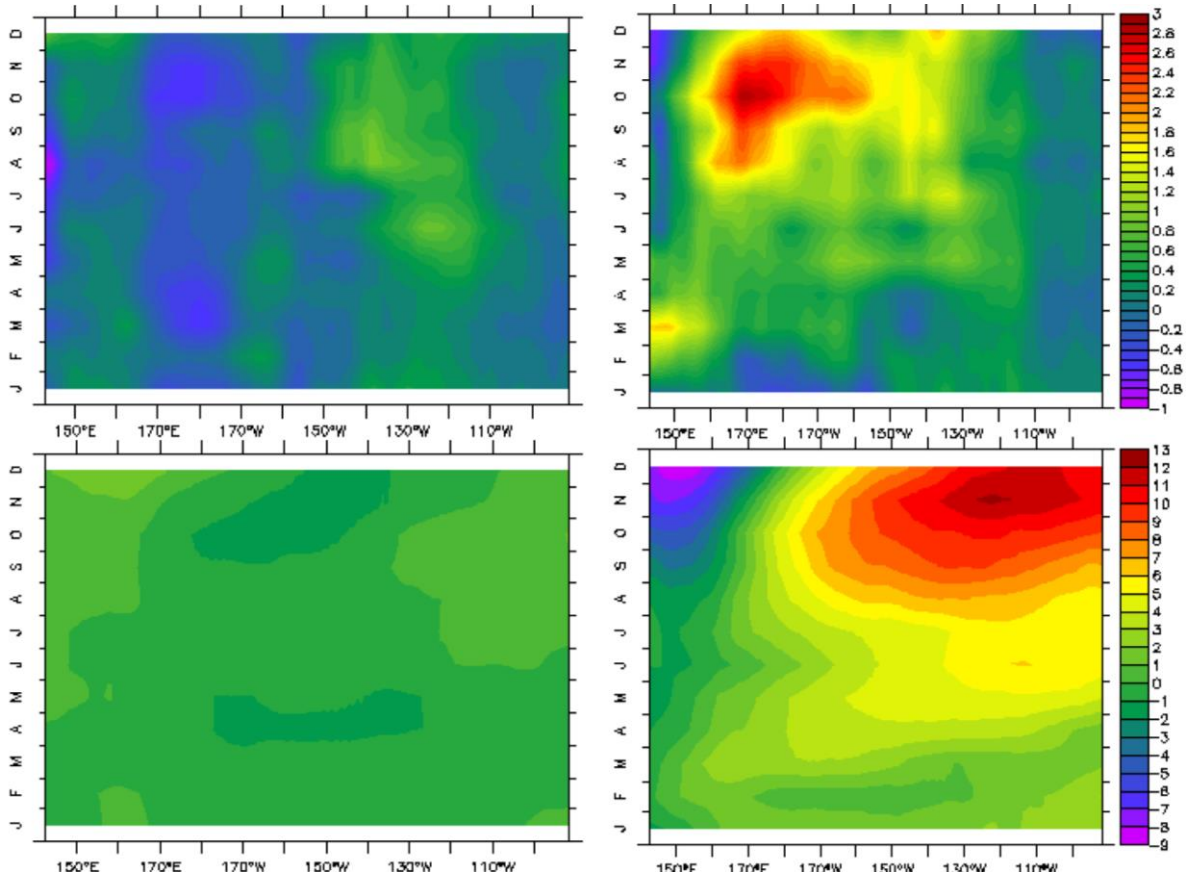

Fig 4.1.5 Hovmöller diagrams of climatological SSH anomaly (bottom panels) and zonal wind stress *taux* anomaly (upper panels) along the equator from GECCO2 (1980-2006) over historical El Niño years (right panels) and the non-El Niño years (left panels). The units for SSH anomaly and *taux* anomaly are centimeters (*cm*) and $100N/m^2$ respectively.

Observations of historical ENSO events indicate that the average interval of El Niño occurrence is 5 years, with one event lasting typically 9-24 months. The first sign of an El Niño usually develops around March-June, indicated by a weakening of the Walker circulation and strengthening of the Hadley circulation. Hovmöller diagrams (Hovmöller, 1949) of zonal wind stress *taux* (Fig 4.1.5) from GECCO2 show an occurrence of outbursts of westerly wind in March on average during El Niño years. The strongest westerly wind is

observed in October. Around November, warm water reaches its peak warmth in the eastern Pacific, and the highest sea surface height is observed there.

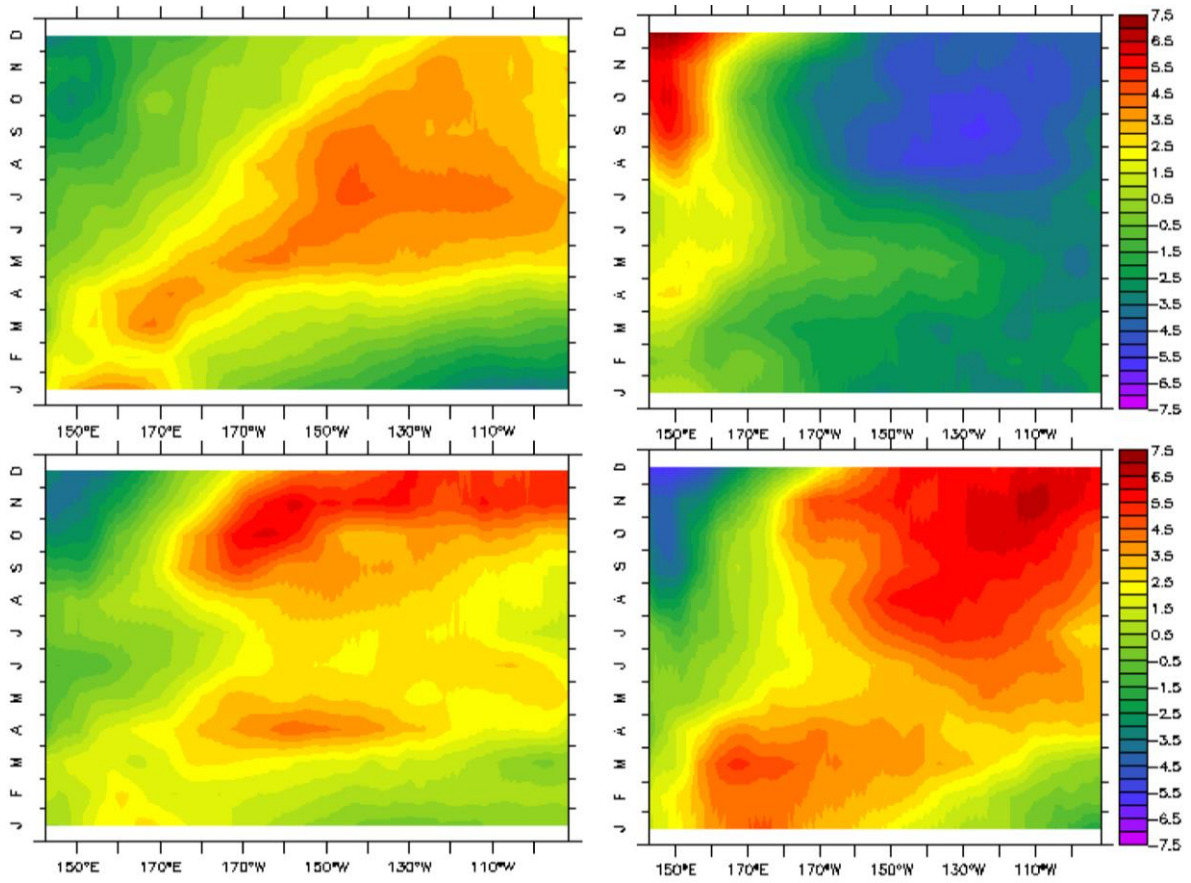

Fig 4.1.6 Hovmöller diagrams of climatological SSH anomaly along the equator from the first 12-months of CIH (right panels) and GIH (left panels) over historical El Niño years (bottom panels) and the GIH produced El Niño-like years (upper panels). The units are centimeters (*cm*).

For the CDA initialized hindcasts, the historical El Niño events are well reproduced, the warm core shifts from western Pacific under normal conditions to east coasts starting around March (bottom right of Fig 4.1.6), and the outbursts of westerly winds happen also around March (bottom right of Fig 4.1.7). The intensity peak of warm core reaches the eastern Pacific around November, in agreement with the peak of Niño 3.4 index. The surface component of Hadley circulation is observed as the strong and months-lasting westerly trade winds that peak around October (bottom right of Fig 4.1.7). The developing phase of El Niño is almost the same as that from GECCO2, but with relatively large/small values of both zonal wind stress and SSH than the latter in January/at the peak of intensity. Therefore, these ENSO events reproduced by CIH are realistic. The GECCO2 initialized hindcasts also perform well in capturing historical ENSO events, with different phases matching the observations (bottom left of Fig 4.1.6 and 4.1.7). Compared with Hovmöller diagrams of CIH, both SSH and *taux* from GIH have smaller values in January of historical El Niño years, which holds until when it peaks. Since smaller values are already observed in GECCO2 than that of CDA, the relative smaller values of SSH in GIH is probably due to oceanic initialization with GECCO2, compared with that of CIH. The response in the atmosphere results in weaker zonal wind stress from GIH as is found in Fig 4.1.7. Overall, the initialization strategy of coupled climate

system with self-consistent initial conditions in this paper does a remarkable job in reproducing the ENSO events.

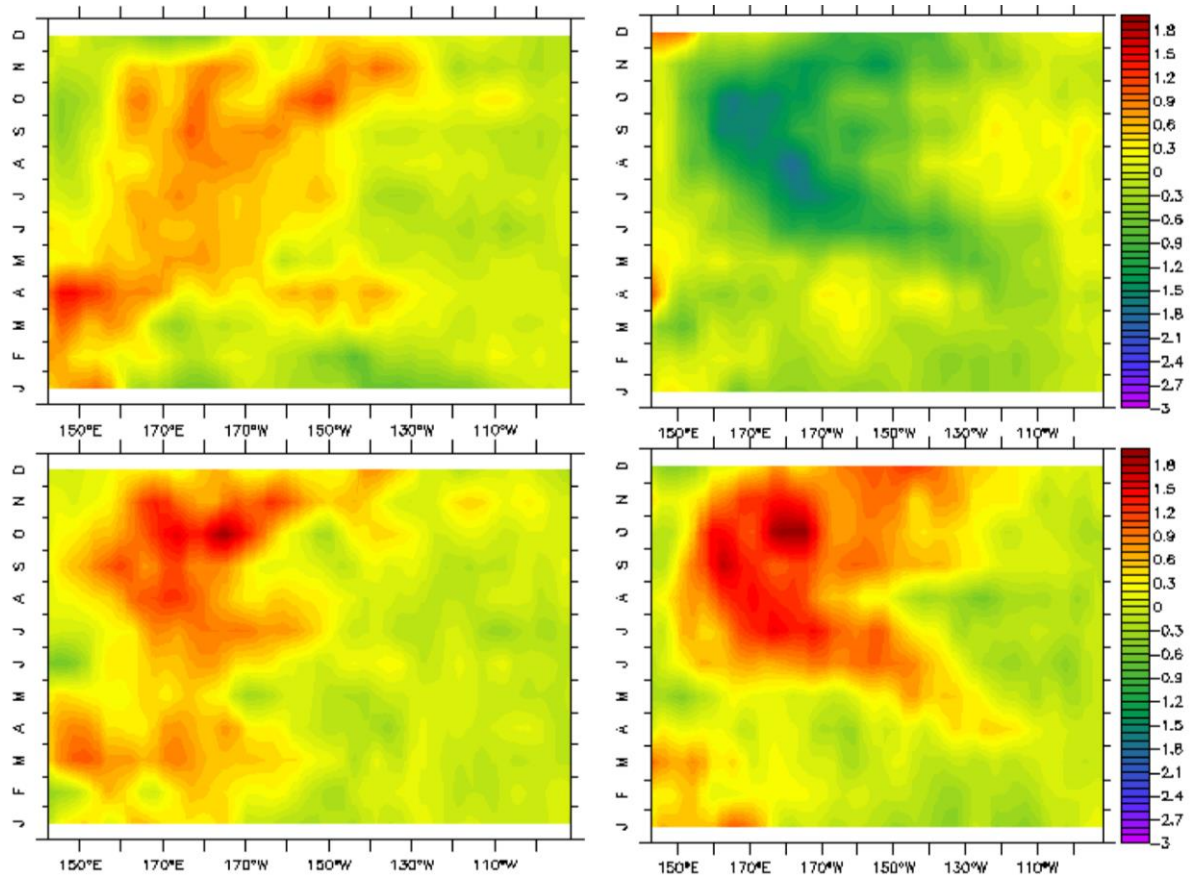

Fig 4.1.7 Hovmöller diagrams of climatological zonal wind stress (taux) anomaly along the equator from the first 12-months of CIH (right panels) and GIH (left panels) over historical El Niño years (bottom panels) and the GIH produced El Niño-like years (top panels). The units are $100N/m^2$.

Besides the successful reproduction of the historical ENSO events for the GECCO2 initialized hindcasts, analysis on the Niño 3.4 Index (middle panel of Fig 4.1.4) and Hovmöller diagrams of related variables (zonal wind stress in upper left of Fig 4.1.6; sea surface height in upper left of Fig 4.1.7) of the El Niño-like events (Fig 4.1.4, Fig 4.1.6 and 4.1.7) indicate that their characteristics match the features and development procedure of El Niño quite well. However, observations of tropical Pacific SST do not indicate occurrence of El Niño during these years. This leads to a conclusion that GECCO2 initialized hindcasts produce pseudo El Niños, and the pseudo ones account for one third of the overall ensemble numbers of the forecasts. As is discussed above, spatial distribution of SST anomaly correlation coefficients of the differently initialized hindcasts reveals that GIH does a less good job than both CIH and dffGec at lead year 1 in the tropical Pacific. The skill of the first year for decadal prediction is strongly associated with ENSO predictability. These pseudo-ENSO events are highly likely to reduce predictability of the forecasts at the first lead year in the eastern equatorial Pacific. From comparison between the four panels of Fig 4.1.6 and Fig 4.1.7, the significant similarities between the historical El Niño events (bottom panels) and GIH produced pseudo El Niños (upper left panels) are the high SSH anomaly core off-coast and westerly zonal wind stress anomaly in the western Pacific, which already exist in January and further develop as time evolves. A larger SSH zonal gradient along the equator is obvious

for GIH in January of pseudo El Niño years than that of CIH in historical El Niño years. On the contrary, the CDA initialized hindcasts show totally different developing phase during the years of GIH-pseudo El Niño events.

As is mentioned above, the CIH and GIH are sharing the same atmospheric initial conditions. The main difference between CIH and GIH resides in the oceanic initialization. CIH is initialized with the CDA ocean states, which is the optimal estimations of the ocean state through assimilating the same climate model as is used for the hindcasts. However, GIH is initialized with interpolated ocean estimations from GECCO2 Synthesis. When initializing the model with an oceanic state that has different topography and dynamics from interpolated GECCO2, the model needs to adjust to the initial conditions at the beginning of the forecasts. If the imbalance is large, this could lead to conditions resembling the early phase of an El Niño event. To explore the hypothesis, more inputs are needed.

4.2 Zonal momentum balance in the upper equatorial Pacific

The difference in the skill over the tropical Pacific near the equator and the first signs of El Niño observed in January for the GIH-pseudo El Niño years is significant. In climate prediction, assimilating data into ocean model is commonly used for a more consistent initialization to the model. However, the study by Bell et al. (2004) pointed out that special care should be taken for the balance between pressure gradient and wind stress near the equator, when assimilating thermal data into ocean model in seasonal prediction. A dynamical imbalance with unrealistic deep overturning circulation is highly likely to happen near the equator due to the assimilation. This is because for such experiments, the ocean models will drift away from the observation data assimilated into them towards the model preferred climatological unless special care is taken. For GIH, by initializing the CFES coupled model with interpolated GECCO2 ocean estimations and atmospheric conditions that result from the same CFES coupled model, different dynamics between initial conditions and the model may lead to imbalance between the zonal wind stress and pressure gradient force in the key regions with variability near the equator. In order to explore the low performance of GIH in predicting the equatorial Pacific SST, we will focus the following study on the balance along the equator between the pressure gradient force (PGF hereafter) and the zonal wind stress (*taux* hereafter) for CIH/GIH of the upper Pacific.

4.2.1 Equations of Zonal momentum Balance along the equator

The equatorial Pacific zonal momentum balance was brought up by Sverdrup in 1947. Later studies refined the balance between the vertically integrated zonal pressure gradient and zonal wind stress on seasonal to annual scale (Bryden and Brady, 1985; Mcphaden and Taft, 1988). Hebert et al. (1991) expected a distinct vertical structure of the momentum balance based on upper ocean data sets measured in spring 1987. Their examinations on different terms of the momentum balance equation at the equatorial Pacific showed that the balance only held on longer (seasonal) time scale. In this thesis, the zonal momentum balance will be studied based on the two different sets of initialized hindcasts to explore the effects from assimilation and

initialization. Following Bryden et al. (1985), the governing equation of momentum balance in the upper equatorial Pacific (Reynolds-averaged Navier-Stokes equations under the hydrostatic and Boussinesq assumption) in the x-direction under balanced situation can be written as:

$$A_v \frac{\partial u}{\partial z} = \frac{1}{\rho_0}(\tau^x - \int_0^z \frac{\partial P}{\partial x} dz)$$

$$P = \rho g \eta + \int_z^{\eta} \rho' dz$$

where A_v is the eddy viscosity, u is the zonal velocity. The left-hind term represents parameterization of the vertical mixing due to uncompensated surface wind stress. Here τ^x is the surface zonal wind stress, ŋ is the sea surface height, P is the pressure and ρ is the density of the sea water. ρ' in the equation is the difference between ρ and ρ_0 ($\rho_0 = 1026 kg/m^3$). At the depth where the vertical shear of zonal velocity becomes zero, the wind stress is compensated by integrated pressure gradient force (Bryden et al., 1985). Note, due to the atmospheric pressure forcing embedded in the coupled climate model CFES, the effects of sea surface air pressure on sea surface height shall be removed before calculation of pressure gradient force.

4.2.2 Zonal momentum Balance along the equator for GECCO2 Synthesis

We start to explore the balance of monthly-mean GECCO2 Synthesis as a reference, following the governing equation of momentum balance in the upper equatorial Pacific by Bryden et al. (1985) in the x-direction. The residual between pressure gradient force and zonal wind stress along the equatorial Pacific will be illustrated in two categories: historical El Niño years and non- El Niño years in Fig 4.2.1.

As is indicated in Fig 4.2.1, in historical El Niño years a negative imbalance is already observed in December in the western Pacific (Fig 4.2.1 (a)). However, a positive imbalance follows in January (Fig 4.2.1 (b)) and only from February onward a consistently developing negative anomaly can be seen. From March on (Fig 4.2.1 (d)), a strong negative imbalance has developed in the western equatorial Pacific. The uncompensated negative imbalance could possibly result from 1) relatively strong easterly trade winds; 2) small pressure gradient force (due to the tilted SSH along the equatorial Pacific) or the combination of 1) and 2). More analysis will be devoted in following part on these possibilities. The imbalance propagates eastwards as Kelvin wave until it reaches the eastern pacific, which is in agreement with the developing phase of El Niño. As a result, warm pattern is observed in the eastern Pacific. Around November (Fig 4.2.1, (j)), the El Niño reaches the peak intensity, corresponding to the disappearance of imbalance or a positive imbalance. If combined with the development of El Niño events, it seems that the imbalance between zonal wind stress and pressure gradient force along the equatorial Pacific provides information on the occurrence of an El Niño several months prior to the El Niño, and the imbalance is observed only from March onward. However, all the simulations of CIH and GIH are initialized in January. The imbalance between zonal wind stress and pressure gradient force is probably not a good

criteria to identify an ENSO event. Despite this, we still want to check whether the imbalance is introduced in January already for GIH, due to adjustment of the model to different oceanic dynamics between the model itself and initial conditions. This could possibly explain the occurrence of those pseudo El Niños observed for GIH.

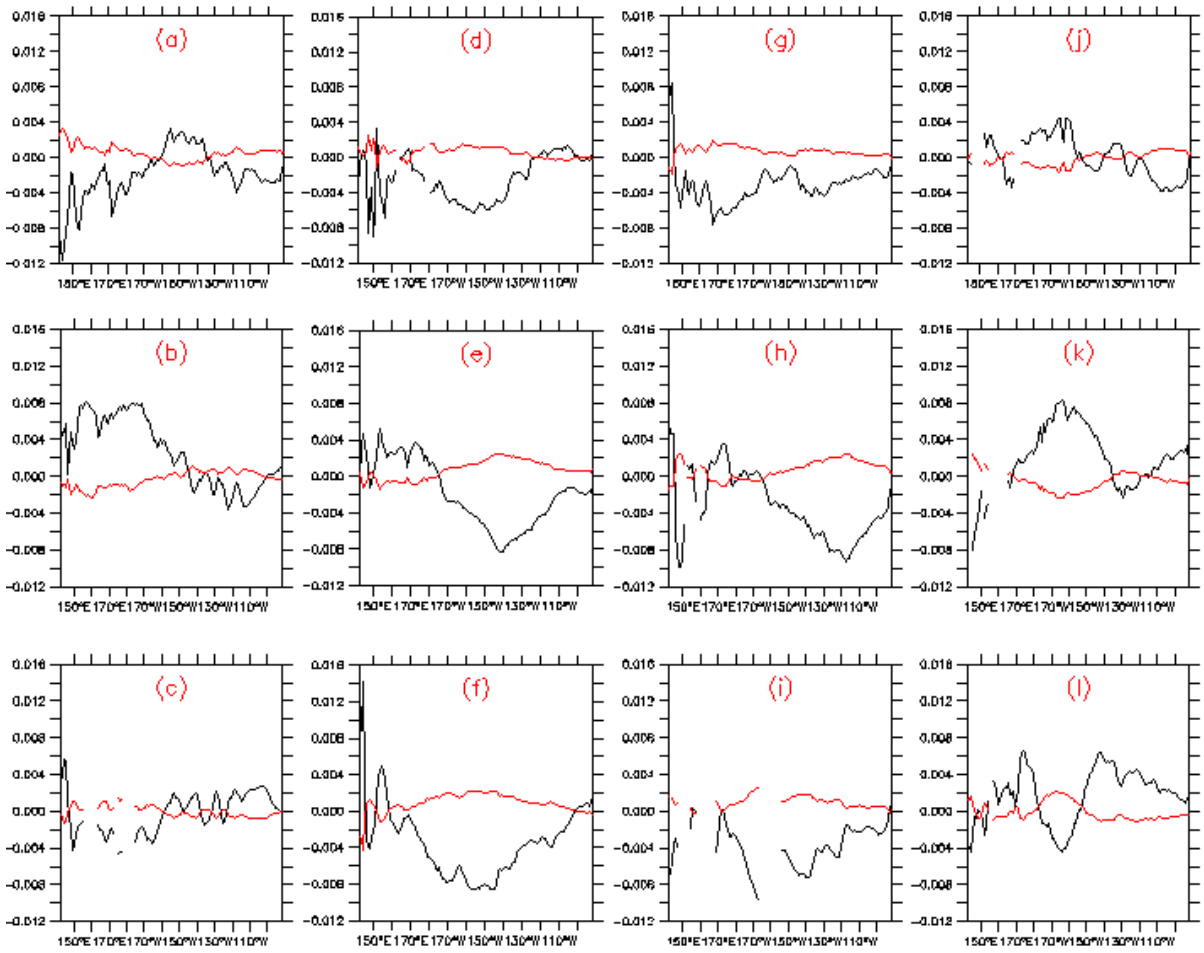

Fig 4.2.1 Zonal momentum balance of upper equatorial Pacific between pressure gradient force and zonal wind stress from GECCO2 Synthesis (1980-2006) in historical El Niño years (black) and the non-El Niño years (red) at: (a) the former December, (b) January, (c) February, (d) March, (e) May, (f) July, (g) August, (h) September, (i) October, (j) November , (k) December and (l) the following January.

4.2.3 Zonal momentum Balance along the equator for initialized hindcasts CIH/GIH

Motivated by the hypothesis above, we now explore the balance between zonal wind stress and pressure gradient force for CIH/ GIH. Considering the pseudo El Niño events produced in GIH, we study the zonal momentum balance of GIH in three categories 1) historical El Niño years 2) pseudo El Niño years and 3) the non-El Niño years. The momentum balance of CIH will be assessed in two categories as the GECCO2 Synthesis. Since CIH produces one pseudo El Niño in 1984, this year is not counted in for any of the two categories of CIH. As indicated

by the green line of Fig 4.2.2, in historical El Niño years, a negative imbalance between PGF and *taux* is observed in the central and eastern equatorial Pacific in January for CIH. The negative imbalance lingers there until May. In June ((f) of Fig 4.2.2) the negative imbalance appears in the western equatorial Pacific and it moves eastwards as is observed for GECCO2 Synthesis. However, the evolution is not as consistent as observed in the GECCO2 Synthesis. The behavior of the climatological imbalance in the non-El Niño years for CIH is almost the same as in GECCO2, with a balanced state year round (cyan line).

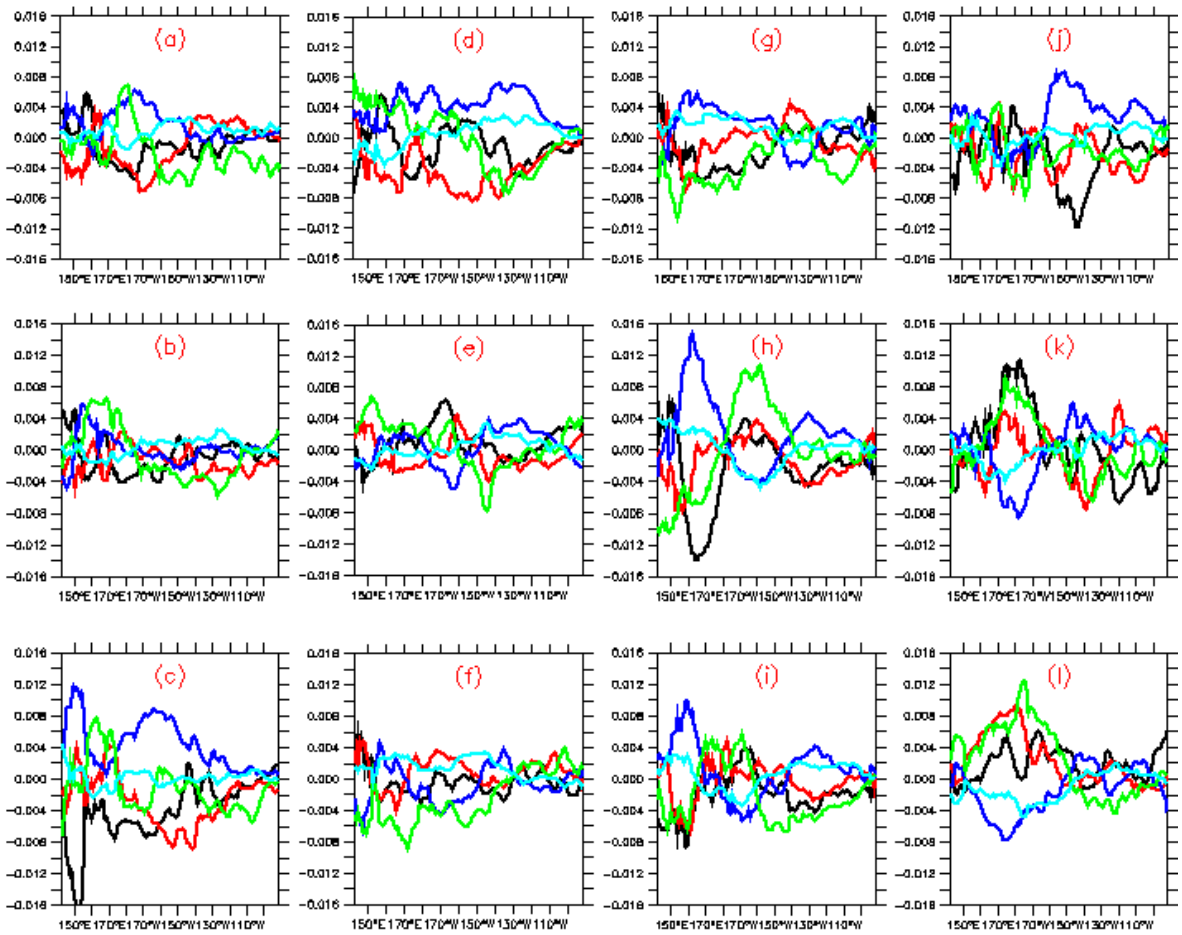

Fig 4.2.2 Zonal momentum balance of upper equatorial Pacific between pressure gradient force and zonal wind stress: 1) from CIH in historical El Niño years (green) and the rest years (cyan); 2)from GIH in historical El Niño years (black); GIH-produced pseudo El Niño years (red) and the non-El Niño years (blue). The climatological first 12 months of the corresponding years during 1980-2006 are shown from (a) to (l) respectively (i.e. January in (a), February in (b), … , December in (l)).

The time evolution of climatological zonal momentum imbalance of GIH produced pseudo El Niño years indicates that a negative imbalance is there over large area of western equatorial Pacific in January, when the forecast is initialized. However, the imbalance is quite sensitive to perturbations and develops quite inconsistent, with the negative area mainly in the middle of equatorial Pacific in March and April. Different from the GECCO2, there is hardly any eastwards transportation of the imbalance visible along the equator. Such situation also holds true for the climatological imbalance of historical El Niño years. In this thesis, the hindcasted period is 1980-2006, restricted by the available CDA initial conditions. Within such a short period, there are only 7 historical El Niño events and 8 pseudo El Niño events. The small

sample size makes it difficult to test the hypothesis that the zonal imbalance between wind stress and pressure gradient force of upper equatorial Pacific triggers the pseudo El Niño events. Nevertheless, the imbalance observed in January is most likely related to pseudo El Niño events observed in GIH, since it doesn't disappear as the prediction evolves. The relatively large zonal SSH gradient and different state of wind stress along the equator from Hovmöller diagram for GIH pseudo El Niño years observed in January may all contribute to the poor predictive skill in the tropical Pacific of GIH at the first lead year.

The characteristics of ENSO is highly related with the tropical mean climate conditions, including the warm-pool of the tropical ocean (Sun 1997), the thermocline depth and wind stress (e.g. An and Jin, 2001; Wang and An, 2001). Studies on modulation of ENSO show that equatorial zonal SST gradient is an important factor (e.g. Sun, 2003; Kim and An, 2011) and the atmospheric process has stabilizing effect on ENSO process (e.g. Zhang et al., 2008). Therefore, to better understand the mechanism leading to erroneous El Niño, it's still useful to explore the elements that contribute to the imbalance at the initialization state, despite the restriction due to sample size. For this purpose, the properties of climate variables will be illustrated in three categories: 1) pseudo El Niño years, 2) historical El Niño years, and 3) non-El Niño years.

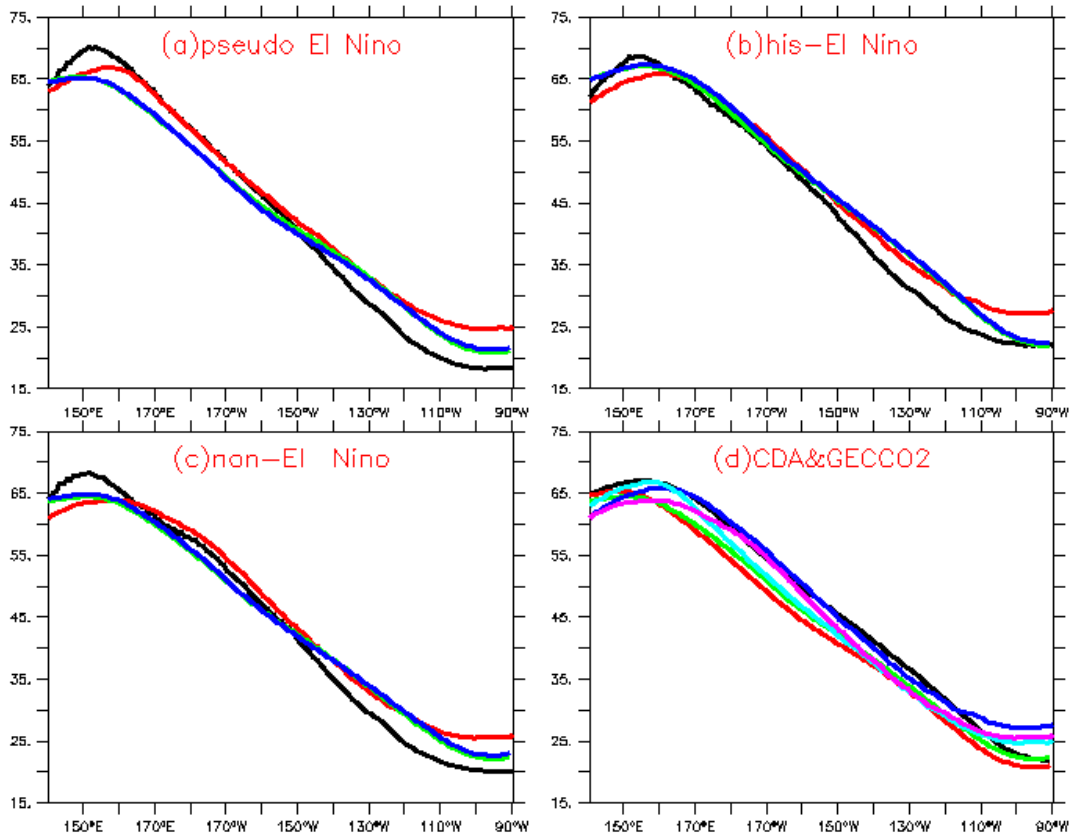

Fig 4.2.3 Climatological SSH along the equatorial Pacific in January of (a) GIH produced pseudo El Niño years, (b) historical El Niño years, and (d) non-El Niño years for GIH (black), GECCO2(red), CDA (green) and CIH (blue). Climatological SSH in January of CDA are displayed in (d) in three categories of historical El Niño years (black), pseudo El Niño years (red) and non-El Niño years (green), as well as GECCO2 of the same categories (blue, cyan, magenta respectively). The units are *cm*.

The climatological SSH (*cm*) along the equatorial Pacific in January is shown in (a), (b) and (c) of Fig 4.2.3 from four different data: GECCO2 (red), CDA (green), GIH (black), and CIH (blue), in three categories as is mentioned above. The blue lines are identical to the green lines in the three panels, suggesting that CIH well resembles CDA. The SSHs shown can be characterized by the position of highest value and zonal gradient along the equator. As is indicated in (d) of Fig 4.2.3, the highest value of SSH in GECCO2 is observed much closer to the central Pacific (i.e. further away from the coast) in the January of pseudo El Niño years than in CDA (Tab 4.2.1). Moreover, larger slop (i.e. larger zonal SSH gradient) in GECCO2 is shown for SSH along the equator in pseudo El Niño years than that of CDA. Meanwhile, both the position of highest SSH and zonal SSH gradient is pretty much similar in GECCO2 as in CDA in historical El Niño years.

Table 4.2.1 Characteristics of climatological SSH in January along the equator

	data	position of highest SSH	largest SSH difference	Highest SSH	Lowest SSH
pseudo-El Niño	GIH	152°E	52cm	70cm	18cm
	CIH	150°E	34cm	55cm	21cm
	GECCO2	161°E	32cm	57cm	25cm
	CDA	150°E	32cm	55cm	21cm
his-El Niño	GIH	155°E	46cm	68cm	22cm
	CIH	160°E	45cm	67cm	22cm
	GECCO2	161°E	37cm	65cm	28cm
	CDA	160°E	45cm	67cm	22cm
non-El Niño	GIH	150°E	47cm	67cm	20cm
	CIH	155°E	43cm	65cm	22cm
	GECCO2	160°E	40cm	65cm	25cm
	CDA	155°E	43cm	65cm	22cm

When initializing the model with two different oceanic initial conditions from CDA and GECCO2, different patterns of SSH along the equator are gained, with CIH resembling CDA quite well whereas GIH not following GECCO2 well. As is indicated in (a), (b) and (c) of Fig 4.2.3, CIH basically overlies upon CDA for the three categories. During historical El Niño years, the largest SSH appears in 160°E and an average SSH difference up to 45cm (67cm in the west and 22cm in the east) is observed in CIH and CDA, with a slightly smaller zonal SSH gradient than that of GIH (Tab 4.2.1). While for non-El Niño years, the average difference of SSH along the equator drops to 43cm (65cm in the west and 22cm in the east) for CIH, still with a relatively smaller zonal SSH gradient than that of GIH. The largest difference of zonal SSH gradient between CIH and GIH is observed in January of pseudo El Niño years.

For the GECCO2 Synthesis that is used as initial conditions for GIH, the largest SSH appears much closer to the central Pacific (150°E for CDA and 160°E for GECCO2), and the zonal SSH gradient along the equator of is larger than that of CDA, in pseudo El Niño years; whereas these difference in historical El Niño years are quite small. The difference of the ocean state (e.g. SSH) between GECCO2 and the model makes the interpolated GECCO2 a non-consistent and incompatible initial condition with the model for GIH. The relatively

larger zonal SSH gradient along the equator from GECCO2 will need stronger wind in the central Pacific to reach balance in the forecast initialized with GECCO2, in pseudo El Niño years. Therefore, adjustment between the ocean and atmosphere through coupled air-sea interaction for GIH is needed, as is revealed by the different patterns of SSH along the equator between GIH and GECCO2. The adjustment is most likely responsible for the inconsistent imbalance shown in Fig 4.2.2 for GIH as the prediction evolves. The larger zonal SSH gradient may lead to pseudo El Niño events, if not balanced by the wind.

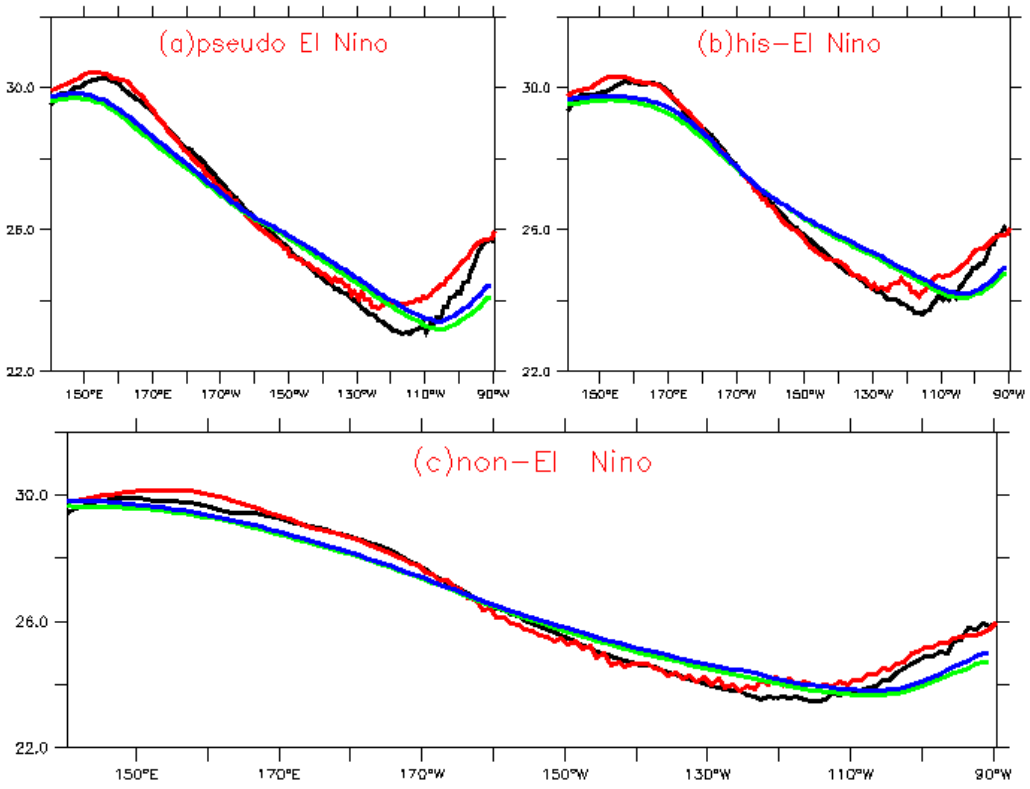

Fig 4.2.4 Climatological SST along the equatorial Pacific in January for (a) GIH produced pseudo El Niño years, (b) historical El Niño years and (c) non- El Niño years, from GIH (black), GECCO2 (red), CDA (green), and CIH (blue). The units are ^{o}C.

Differences of ocean states are also characterized by SST. As is shown in Fig 4.2.4, comparison between GECCO2 (red lines) and CDA (green lines) indicates that the average of SST in January along equator is warmer in GECCO2 than that of CDA, for either pseudo or historical Niño years. The position and range of warm/cold pool is also different between the two data sets. Overall, warmer SST in the western Pacific and a cold core closer to the central Pacific are observed in GECCO2 compared with that of CDA, indicating a larger zonal SST gradient along the equator in GECCO2 (Fig 4.2.5). SST in CDA is much closer to the observations (i.e. HadISST) than that of GECCO2.

The different ocean states between CDA and GECCO2 implies a different atmospheric response. As is shown in Fig 4.2.6 (a), stronger zonal wind stress is observed in GECCO2 in the area of 180^{o}-140^{o}W compared with CDA/20C. Lager mean zonal pressure gradient than CDA/20C in the central (175^{o}-130^{o}W) Pacific (Fig 4.2.6 (b)) is also observed in GECCO2. In the eastern Pacific, SST from GECCO2 is also warmer than that of CDA, with a smaller zonal SST gradient in the former than the latter. Consistent with that, wind stress from CDA is

stronger than GECCO2 there. After initializing the model with warm-biased SST from GECCO2, which also has a larger zonal gradient in the central but a smaller zonal gradient in eastern Pacific, colder pattern is observed over eastern Pacific in GIH (Fig 4.2.4), compared with GECCO2 during historical and pseudo El Niño years. Such difference is most likely due to the coupled air-sea feedbacks of the model (Fig 4.2.6 (a)) to the initialized smaller zonal SST gradient (through GECCO2 initialization) and the strengthened upwelling due to the stronger wind in the eastern Pacific (through atmospheric initialization). Overall, zonal SST gradient is increased due to initialization for GIH in the western and central Pacific (i.e. the west of 120^{o}W). Warmer SST with smaller zonal gradient and stronger wind stress are set in the area of 120^{o}-110^{o}W through initialization. Interaction between the ocean and atmosphere may lead to shallower thermocline and strengthened upwelling in that area (i.e. colder SST) for GIH. In the central Pacific around 175^{o}-140^{o}W (Fig 4.2.5), the model is initialized with a relatively weaker zonal wind stress but larger zonal pressure gradient, leading to a negative imbalance at the initialization state for GIH. The imbalance propagates eastwards as Kelvin waves along the equator. This could further lead to transportation of warm water to the eastern Pacific, and therefore occurrence of El Niño, jointly with perturbations due to different topography between the initial conditions and the model system.

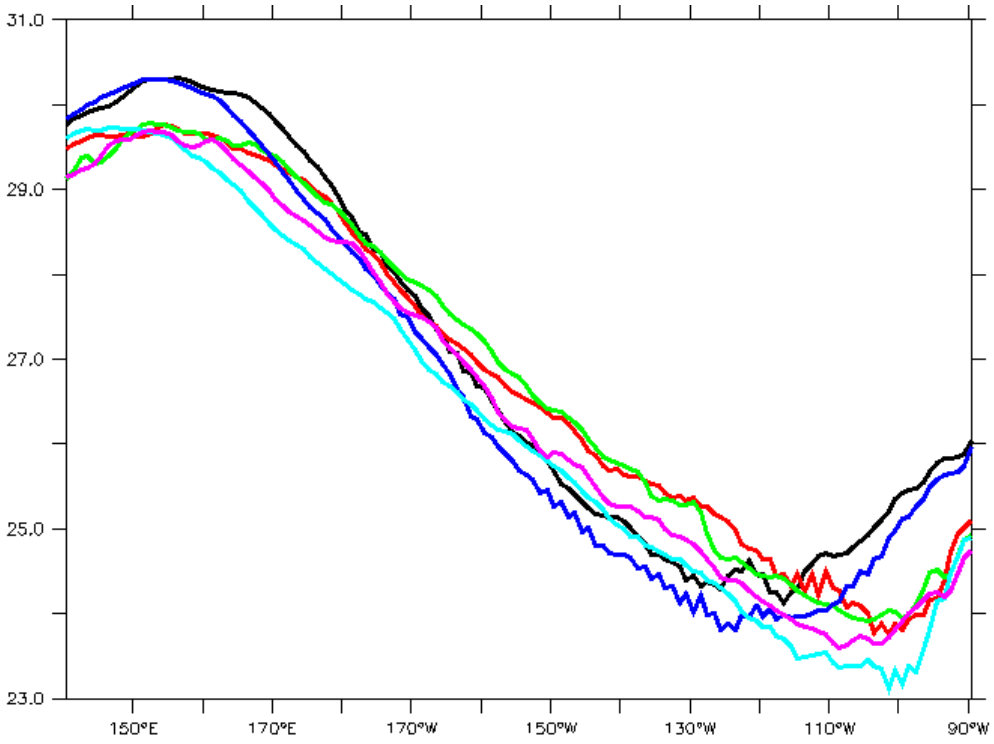

Fig 4.2.5 Climatological SST along the equatorial Pacific in January for 1) historical El Niño years from GECCO2 (black), CDA (red), and HadISST (green), 2) non-El Niño years from GECCO2 (blue), CDA (cyan), and HadISST (magenta) and CDA (dark green). The units are ^{o}C.

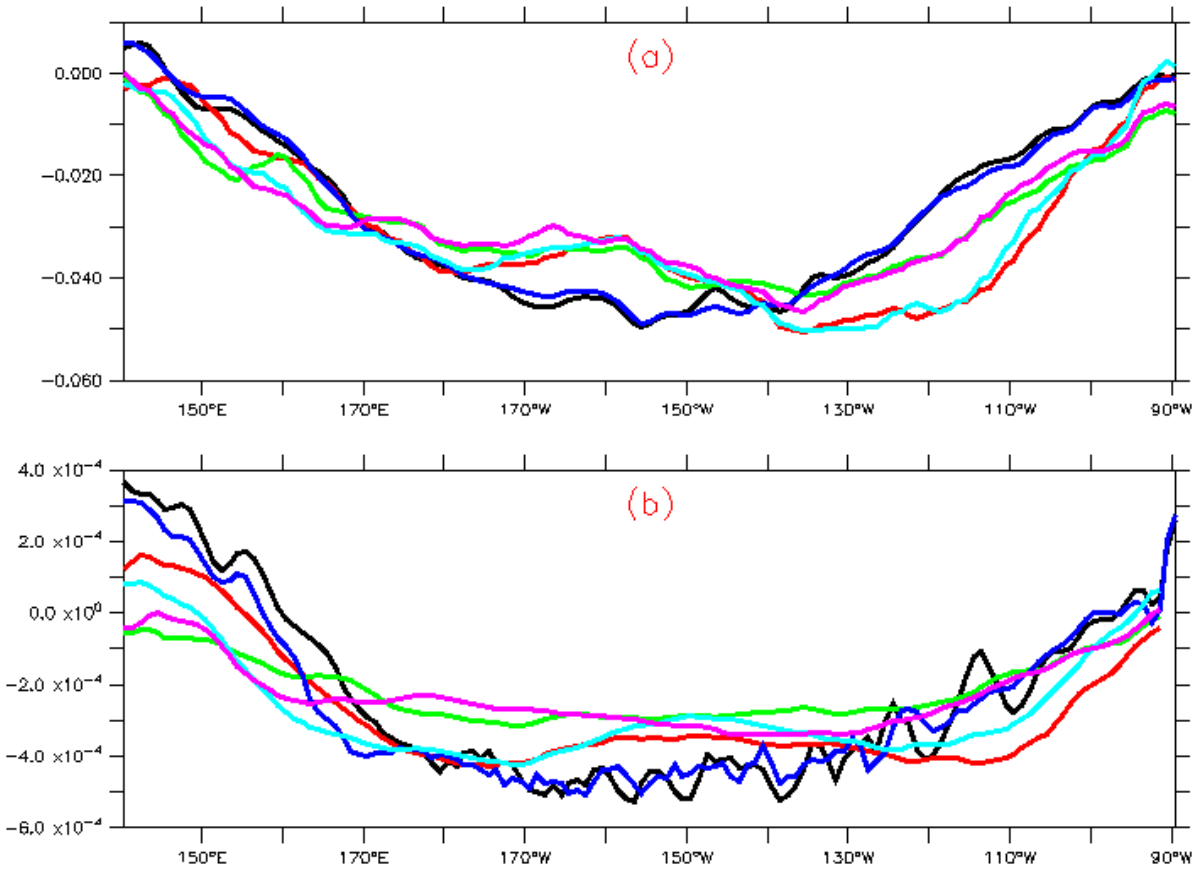

Fig 4.2.6 Climatological zonal wind stress (a) and zonal pressure gradient (b) along the equatorial Pacific in January for 1) historical El Niño years from GECCO2 (black) , CDA (red) and 20C (green); 2) non-El Niño years from GECCO2 (blue) CDA (cyan) and 20C (magenta). The units for wind stress are N/m^2, and for pressure gradients are P_a/m .

To summarize, CIH well reproduces historical El Niño events. Compared with CIH, the simulation of ENSO events is much poorer for GIH, with 8 pseudo El Niño detected at the first lead year. These pseudo El Niño events are most likely related with the poor predictive skill of SST for GIH over the tropical Pacific at the first lead year. GIH is initialized with interpolated GECCO2 ocean estimations and atmospheric conditions that result from the same coupled model as used in the hindcasts. Through initialization with GECCO2, zonal SST/SSH gradient is increased in the central Pacific and SST is warmed along the equator for GIH. However, the wind stress in the eastern/central Pacific remains relatively stronger/weaker. The different states of the ocean part (SSH/SST) leads to imbalance between the zonal wind stress and pressure gradient force through coupled air-sea interaction, and adjustment takes place as the prediction evolves. Negative imbalance is detected for GIH in the central Pacific, due to the warmer SST with larger SST/SSH zonal gradient in the central Pacific of ocean through initialization with GECCO2, which propagates eastwards as Kelvin waves. The reduced predictive skill of SST in GIH in the equatorial Pacific is therefore found to be mainly related to dynamical imbalance between zonal wind stress and pressure gradient over the central Pacific. Further imbalance may result from the differences in topography between the model system and GECCO2, and perturbations may give rise to propagating Kelvin waves. All these elements contribute to the poor predictive skill for GIH over the tropical Pacific as is observed. Hence, people shall be careful with the momentum balance between zonal wind stress and pressure gradient force along the equatorial Pacific, when initializing a coupled

model from any oceanic state. Initializing a coupled model with self-consistent initial conditions improves the predictive skill of decadal climate prediction in the tropical Pacific.

Chapter 5

Conclusions and future work

The primary goal of this thesis is to investigate the influence of model consistency between the model used to prepare initial conditions and that used for the hindcasts with the coupled model CFES developed by JAMSTEC. We estimated the relative predictive skill of SST and AMOC, in the two differently initialized hindcasts, on decadal time scales and predictability of ENSO at the first lead year. Using multiple diagnostics, we provided information on spatial distribution of predictive skill and further on the regional predictive skill, as well as the duration of predictability. These will help to gain insight into how to improve decadal prediction.

The following part of the thesis will be separated into two parts. The first part in Section 5.1 will be the main results and conclusions. In the second part, we will outline the possible future work.

5.1 Overview of main conclusion

5.1.1 Influence of model consistency on decadal climate prediction

In this thesis, the influence of model consistency in decadal climate prediction is explored by initializing an updated version (including external forcing) of the coupled model CFES developed by JAMSTEC with ocean state estimations from CDA (estimated with the CFES model without external forcing embedded) and the ocean-only forced GECCO2 Synthesis, as well as atmospheric initial conditions from CDA through control variables of air-sea fluxes.

In terms of predictability of SST, we first evaluate the SST predictability of initialized hindcasts against the uninitialized hindcasts to gain insight into whether initialization gives more predictive skill than the externally forced prediction. We find that both sets of initialized hindcasts outperform the uninitialized hindcast, indicating that full-state initialization does give rise to predictive skill on decadal time scale. Verification on relative prediction skill of SST from CIH (CDA initialized hindcasts) and GIH (GECCO2 initialized hindcasts) shows that CIH provides the most significant predictive skill at lead year 1 over vast areas, suggesting that the dynamical consistency of initial conditions can indeed improve the predictive skill of climate predictions at least in the first lead year. The most significant improvement is observed over the eastern tropical Pacific for CIH. On the contrary, by initializing the model with interpolated GECCO2, a warm bias over the tropical Pacific at the first lead year is introduced through initialization. The incompatibility of the initial conditions to the model results in poor predictability there. The initial shock, different topography and dynamics also contribute to the bias of the forecasts and reduce predictive skill. For lead year 2-5, predictive skill is gained in part of Indian Ocean for both, while for lead year 6-9 almost

no skill is observed. For prediction skill at longer lead times, the response of externally forced change starts to increase. The only difference in external forcing between the model CFES used to derive CDA and that used in model simulations (the former doesn't include external forcing and the later does) may lead to slow drift in CIH and therefore poor predictive skill at longer lead times. Comparison between spatial distributions of SST before and after the removal of linear long-term trend reveals that the trend due to increasing GHGs (green house gases) gives rise to SST predictability.

The decrease in SST predictive skill at longer lead times also holds true over the North Atlantic (NA), where previous studies indicate that the predictive skill of SST is associated with the predictability of Atlantic meridional circulation (AMOC), with the latter leading the former several years. The poor predictive skill of NA SST is possibly related to low AMOC predictive skill in CIH/GIH. Illustrations of detrended AMOC reveal significant and robust skill at lead year 1 for both CIH and GIH, with the former outperforms the latter. At longer lead times, both CIH and GIH lose predictive skill in the subpolar region of NA, suggesting that the poor predictive skill of NA SST is partially due to poor predictability of AMOC. However, due to the strong increasing trend and more multi-year variability of AMOC from initialization with GECCO2, high predictive skill of AMOC is achieved until lead years 2-5 in GIH, when the linear trend is not removed. Nevertheless, the linear long-term trend of AMOC for CIH impacts predictive skill less than that of GIH, possibly due to the small trend and less multi-year variability of AMOC from CDA. Despite the possible incompatibility of the initial conditions to the model, the increasing trend of AMOC from GECCO2 as a lagged response to negative NAO improves the predictive skill of AMOC for GIH through initialization.

To compare predictive skill between different sets of hindcasts, one shall always make sure that the predictions used are at the same period. Comparison with MiKlip dffGcE (decadal hindcasts initialized with full-fields GECCO2 for ocean and full-fields ERA Reanalysis for atmosphere) hindcasts of the same length (i.e. 1980-2006) indicates that dffGcE outperforms GIH, while its performance is compatible with CIH, in terms of SST. Predictive skill of AMOC is increased through initialization with GECCO2 until lead years 2-5, which resides more multi-year variability (i.e. low-frequency oscillatory variability). The trend of AMOC from GECCO2 has a strong influence on North Atlantic climate predictability, which can be understood as a lagged response to NAO variability. However, due to the lack of observations, it's difficult to verify whether the AMOC evolution is realistic. More realizations at each initialization dates improve the predictive skill, due to the usage of ensemble mean instead of individual realization. Comparison between the performances of dffGcE of 1980-2006 and that of 1961-2009 indicates larger areas with significant skill and longer duration in terms of SST predictability in the latter than that of the former. However, relative small correlation coefficients are observed in dffGcE of 1961-2008 in the area where significant skill is also observed in dffGcE of 1980-2008, except in the Labrador Sea at lead years 2-5. Larger number of initial dates increases the areas with predictive skill by reducing the level of significance, but doesn't increase predictive skill for most of the area.

The significant, high SST predictive skill over the tropical Pacific in CIH indicates a good reproduction of El Niño events while GECCO2 initialized hindcasts (GIH) produce additional erroneous El Niño events, thus leading to the poor predictive skill of SST over the tropical Pacific at lead year 1. During balanced states, a zonal momentum balance between the wind stress and pressure gradient force of the upper equatorial Pacific exists. The GIH is initialized

with interpolated GECCO2 ocean estimations and atmospheric conditions that result from the same coupled model as used in the hindcasts. However, the ocean state of GECCO2 is characterized by a warmer SST with lager zonal gradient than that of CDA. When initializing the model with different ocean states, the adjustment though coupled air-sea interaction results in imbalance between the zonal wind stress and pressure gradient force, as is observed in the central Pacific for GIH. Further imbalance may also result from the differences in topography between GECCO2 and the model system. The imbalance propagates eastwards as Kelvin waves. The reduced predictive skill of GIH in the Pacific equatorial region is found mainly related to the dynamical imbalance due to inappropriate oceanic initialization of the predictions. Hence, in climate predictions, it is necessary to initialize a model with self-consistent initial conditions in order to improve prediction skill in the tropical Pacific. People shall be careful with the momentum balance between zonal wind stress and pressure gradient force along the equatorial Pacific when initializing the model from any oceanic state.

5.1.2 Verification of the results

In this thesis, we estimate the predictive skill of initialized hindcasts following the multi-diagnostic approach suggested for decadal climate predictions (Goddard et al., 2013). The results shown as based on detrended hindcasts (with the trend removal that varies with lead times). We use the scheme of initializing the model every year and letting each hindcast running for 9 years. To better evaluate the performance and define reliability of decadal prediction, larger number of initial dates and/or more realization of each initial date are usually worthwhile (e.g. Meehl et al., 2014), although Ho et al. (2013) found out that limited start dates may result in relatively robust verification metric. However, limited by the initial conditions from CDA, we can only initialize the model from 1980 to 2006. Due to limitations in computer time, it's more practical to have only one realization for each initial date. We also initialize the model with interpolated GECCO2 Synthesis. For the sake of less initial shock at the beginning of predictions, an assimilation of GECCO2 may work as more appropriate initial conditions. However, due to the same computer time, we first let the model run for three days to adjust to different topography for these experiments. The outputs are then used as oceanic initial conditions.

The SST predictive skill in this thesis is verified by application of anomaly correlation skill (ACC) and the root-mean-squared-error skill score (RMSS) at time scales of lead year 1, average over lead years 2-5 and 6-9. When these skills are used, it is always necessary to carry out a significant test. For all the significant tests of the results involved, we choose the Student's *t*-test. Physical process is always auto-correlated. Hence the effective sample size (von Storch and Navarra, 1999) is used to estimate the degrees of freedom. However, since there is only 27 hindcasts for CIH and GIH, the *t*-test might not be an appropriate method, especially for the 4-yr average. The non-parametric bootstrap significant test proposed by Goddard et al. (2013) is more suitable for predictions with small ensemble size. This method does not need distribution assumptions or temporal/spatial independent assumption of hindcasts, simply by re-sampling the simulations by random selection and estimating the statistics between subsamples. Hence the bootstrap is probably a better method as is used in other studies (e.g. Polkova et al., 2014). However, more computational time is needed for generating subsamples of simulations with this method.

5.2 Outlook

This study provides information regarding the influence of model consistency on decadal prediction in terms of SST and AMOC, as well as the influence from initial dates and sample size. Along with insight into this, new questions also emerge. Further efforts can be devoted to aspects below:

1) As is mentioned in section 3.1.1, the biased warmer global SST at lead year 1 for CIH is possibly due to a slow drift in the hindcasts, which are initialized every year. However, CDA from JAMSTEC was reinitialized every 9-months, with an overlap of 1.5 months at both margins. This strategy helps to prevent slow drift in CDA. But such difference may lead to possible drift as predictions evolve. Hence, expanding the re-initialization window to longer time (e.g. 5 years) for CDA needs to be explored. This will offer more dynamical-consistent initial conditions for the prediction, and possibly higher predictive skill.

2) Predictability of the Pacific climate is sensitive to uncertainty from initial state and internal climate variability (Branstator et al., 2012; Branstator and Teng, 2012). As is discussed in section 4.2, the different ocean states (larger SST/SSH zonal gradient) of GIH through initialization with GECCO2 ocean estimation leads to imbalance between the atmosphere and therefore adjustment through the coupled mechanism. The non-self-consistent initial conditions (the biased mean state) lead to poor predictive skill of SST in the tropical Pacific. A study by Boer and Lambert (2008) pointed out that potential predictability over the extratropical and high latitude oceans is mainly controlled by contribution from internal variability. Over the mid-to-high latitude ocean regions, the surface climate is connected to the deep ocean (Boer, 2008). Since the oceanic initialization in the tropical ocean is more biased than that of the mid-and-high latitude ocean (Fig 3.1.6), removing the mean state bias along the tropical Pacific is necessary. Motivated by previous studies by Polkova et al. (2014) and Volpi et al. (2013), improved SST predictive skill in the tropical Pacific is gained with the anomaly initialization strategy. Therefore, one possible solution is to use the anomaly initialization over the tropical Pacific. By initializing the model with observed anomalies added to the model climate, the warm bias along the tropical Pacific could be reduced and could therefore lead to possible improvement of predictive skill in GIH.

3) So far the relationship between ENSO and Pacific decadal climate variability still needs more exploration (Meehl et al., 2014). Some people argue that other climate variability of Pacific, which is related with Pacific Decadal Oscillation (PDO) or Interdecadal Pacific Oscillation (IPO), is just a residual patterns of spatially asymmetric ENSO, while others suggest that many mechanisms contribute to the tropical Pacific decadal change, and the mean state change will affect ENSO (e.g. Power et al., 1999a; Meehl and Hu, 2006; Matei et al., 2008). In this thesis we only explore the influence of model consistency on decadal prediction and ENSO respectively. The possible mechanism between ENSO and the Pacific decadal variability is also interesting to explore. The sensitivity of other climate variability

such as PDO and IPO to model consistency can also be explored. Besides, due to small sample size of ENSO events within the hindcasted period, statistical significance of the results is weak. Hence the effect of ENSO on decadal predictability is not sufficiently explored. Therefore, larger number of initial dates is needed.

4) The AMOC predictability is related with NA SST predictability probably through AMO and other atmospheric variability such as NAO (e.g. Eden and Willebrand, 2001; Gastineau et al., 2012). This variability has strong influence on the climate in the North Atlantic. In this thesis, the positive trend of AMOC from GECCO2 leads to high predictive skill of AMOC. The trend is possibly a lagged response to positive phase of NAO (Köhl and Stammer, 2008). Therefore, different NAO phase at the initial state may lead to decrease/increase of AMOC in 2-3 years. Model consistency can influence AMOC predictability on decadal time scale through initialization via NAO. Other variability such as PDO is also interesting to look into.

Figure Captions

Fig 1.1 Schematic illustrating different portion of impact from initial value and boundary conditions on climate forecasts at different scale from IPCC (2013, based on Meehl et al. 2009). This figure shows us the progression with daily weather forecasts at one end (left), seasonal and decadal prediction in between, and multi-decadal to century projections as a forced boundary condition problem at the other end.

Fig 1.2 Schematic showing relative importance of uncertainty from different sources in decadal mean global surface air temperature by the fractional uncertainty ($\text{fractional uncertainty} = \frac{prediction\ uncertainty}{expected\ mean\ change}$) from Hawkins and Sutton (2009). Model uncertainty acts as the main source for uncertainty up to lead time 50yr. Internal variability is another important source of uncertainty for the first decade, while scenario uncertainty being important for multi-decadal lead time.

Fig 2.2.1 The time evolution of global CO2 concentrations (ppm) from HadGEM2 Model Development Team (2011) supported by CMIP5 for historical period (1860-2005) and for four RCPs (2006-2100). In our hindcasts (CIH, GIH and 20C), the historical (black) and RCP 4.5 (orange) CO2 concentrations from 1946 to 2045 are used. (Figure available from website: http://cordex-ea.climate.go.kr/main/modelsPage.do)

Fig 2.4.1 Predictability of upper 300-m temperature of the North Atlantic for 12 CMIP5 models resulting from initialization (dashed blue lines) and the response to RCP4.5 forcing (dashed red lines) from Branstator and Teng (2012). The solid lines are multi-model averages. Predictability is measured by relative entropy. The crossover point near year 8 for the multi-model averages indicates that after 8 years of the hindcasts, information resulting from external forcing (red line) becomes more important than information originating in the initial conditions (blue line).

Fig 2.4.2 Spatial distribution of anomaly correlation coefficient (ACC) between hindcasted SST (hindcasts initialized with CDA, short as “CIH”) and CDA (estimations derived by assimilation through the coupled climate model CFES) at different lead years: 1) lead year 1 (top panels); 2) averages of lead years 2-5 (middle panels); and 3) years 6-9 (bottom panels). The SST data shown in left panels are all detrended and these in the right panels are not. Only the significant ACC (at 95% level) are shown here. The trend gives rise to predictability on decadal time scale regionally from comparison of the two columns of figures. For the first lead year, larger areas with predictive skill are observed over the North Atlantic for the non-detrended SST (upper panels). For the 4-yr average, the trend gives rise to predictability of SST over part of the North Atlantic and western Pacific (middle and bottom panels).

Fig 2.4.3 Spatial distribution of SST RMSS (RMSE skill score) for CIH (CDA initialized hindcast) against 20C (un-initialized hindcast) at different lead years: 1) lead year 1 (upper panels); 2) averages of yr2-5 (middle panels); and 3) yr6-9 (bottom panels). SST from HadISST is used as the observation. The SST data used in right panels are all detrended and these in the left panels are not. Red area in the figure indicates improvement due to initialization. Improvement in predictive skill for initialized hindcasts is mainly due to the internal variability and externally forced climate changes (left panels). With an application of

trend removal, improved predictability for initialized hindcasts is only about the internal variability, as is shown in right panels. Comparison between the left panels and right panels indicates that the trend plays an important role in improving predictive skill on decadal prediction.

Fig 2.4.4 Time series of global-averaged annual mean SST in the HadISST dataset (green), initialized (colours, CIH) and un-initialized (blue) predictions and CDA (black). The upper panel shows the raw initialized hindcasts before detrending. The bottom panel shows the evolution of global mean SST with the linear long-term trend removed.

Fig 2.4.5 Time series of global-averaged annual mean SST in the HadISST dataset (green), initialized (colours, GIH) and GECCO2 (black). The upper panel shows the raw initialized hindcasts before bias correction. The bottom panel shows the evolution of global mean SST with mean bias removed.

Fig 2.4.6 Scaling factor applied in computation of effective sample size from original sample size of auto-correlated time series.

Fig 3.1 Evolution of global-averaged annual mean SST in the HadISST dataset (green), raw initialized hindcasts (colours) and ocean estimation (CDA/GECCO2) from which the hindcasts are initialized (black). The colored lines in top/bottom panel are raw CIH/GIH.

Fig 3.2 Spatial distribution of RMSS for annual mean SST for raw hindcast SST against the 20C at lead year 1(upper panels), averages over yr 2-5 (middle panels), and 6-9 (bottom panels). Left panels are CIH against 20C and right panels are GIH against 20C, without bias correction. Red areas indicate that the initialized hindcasts are more accurate than the 20C and blue indicates the opposite.

Figure 3.1.1 The time series of annual-mean global average sea surface temperature anomalies relative to 1961-1990 in observations (black line) from Met Office. The calculation is based on the sea surface temperature of dataset HadCRUT4 (abbreviated as HadSST3). The grey shade indicates the ranges that are 95% significant. As is shown in the figure, the SST warming trend is much larger after the 1900s. There is a warm pause from 1950-1980. After 1980, a large warm trend is found in SST again. (figure from Met Office website, available at: http://www.metoffice.gov.uk/research/monitoring/climate/surface-temperature)

Fig 3.1.2 Time series of annual mean global SST at different lead years in observations (blue, HadISST data set), persistence forecast (yellow), 20C (green), initialized hindcasts (black, left panels CIH, right panels GIH), CDA (left panels, red) and GECCO2 (right panels, red). The top panels are SST at lead year 1. The 4-yr means of lead years 2-5 are shown in middle panels, with the time series plotted centered in year 3. Averages over lead years 6-9 are shown in the bottom panels, with time series plotted centered in year 7.

Fig 3.1.3 Spatial distribution of RMSS for hindcast SST against the 20C at lead year 1(upper panels), averages over yr 2-5 (middle panels), and 6-9 (bottom panels). Left panels are CIH against 20C and right panels are GIH against 20C. Red areas indicate that the initialized hindcasts are more accurate than the 20C and blue indicates the opposite.

Fig 3.1.4 Spatial distribution of SST anomaly correlation coefficient between CIH and CDA (left), CIH and observed SST (HadISST, middle), and CDA and persistence (right), at lead year 1 (top panels), averages of lead yr 2-5(middle panels) and lead yr 6-9 (bottom panels). Only the significant coefficients (at 95% level) are shown here. A linear trend removal is applied to all the SST involved before calculation of ACC.

Fig 3.1.5 Spatial distribution of the SST anomaly correlation coefficient between GIH and GECCO2 (left), GIH and observed SST(middle), and GIH and persistence (right), at lead year 1 (top panels), averages of yr2-5 (middle panels) and yr6-9 (bottom panels). Only the significant correlation coefficients (at 95% level) are shown here.

Fig 3.1.6 Time series of annual mean SST at the first lead year averaged over 20^{o}S-20^{o}N (upper) and 60^{o}S-20^{o}S & 20^{o}N-60^{o}N (bottom) in observations (blue, HadISST data set), hindcasts (solid black: CIH; dashed black: GIH), CDA (solid red) and GECCO2 (dashed red).

Fig 3.1.7 Spatial distribution of the non-detrended SST anomaly correlation coefficient between CIH and CDA (left), GIH and GECCO2 (right) separately, at lead year 1 (top panels), averages of yr 2-5 (middle panels) and yr 6-9 (bottom panels). Only the significant coefficients (at 95% level) are shown here. Compared with corresponding spatial distribution of detrended SST ACC of left panels in Fig 3.1.4 and Fig 3.1.5, non-detrended SST from both CIH and GIH shows significant skill in decadal prediction over larger areas. The most significant improvement in predictive skill is found for GIH over the North Atlantic, western Pacific and the Indian Ocean.

Fig 3.2.1 Time series of annual mean AMOC (Sv) anomaly (at 1000m) at 26.5^{o}N from hindcast experiments (black), CDA (red in left panels) and GECCO2 (red in right panels), and persistence (green). The left panels are CIH and CDA used for initialization, while the right panels are GIH and GECCO2 used for initialization respectively. The top panels are AMOC at lead year 1. The 4-yr mean of lead years 2-5 is plotted with the time series plotted centered in year 3(middle). Averages over lead years 6-9 are shown in the bottom panels, with time series plotted centered in year 7.

Fig 3.2.2 Anomaly correlation coefficient of the non-detrended Atlantic MOC (Sv, $Sv = 10^{6}m^{3}s^{-1}$) at 1000m depth between CIH and CDA (black), CDA and persistence (red), CIH and 20C (yellow), GIH and CECCO2 (cyan), GECCO2 and persistence (blue) and GECCO2 and 20C (green) along different latitudes of 30^{o}S-60^{o}N. ACCs are calculated for the first lead year (upper panel), 4-yr average of lead years 2-5 (middle) and lead years 6-9 (bottom panel) respectively. The dashed lines denote the hindcasts that don't obtain significant skill at 95% significant level. All the AMOC data in this figure are not detrended.

Fig 3.2.3 Anomaly correlation coefficient of the detrended Atlantic MOC(Sv, $Sv = 10^{6}m^{3}s^{-1}$) at upper 1000m between CIH and CDA (black), CDA and persistence (red), CIH and 20C (yellow), GIH and CECCO2 (cyan), GECCO2 and persistence (blue) and GECCO2 and 20C (green) along different latitudes of 30^{o}S-60^{o}N. ACCs are calculated for the first lead year (upper panel), 4-yr average of lead years 2-5 (middle) and lead years 6-9 (bottom panel) respectively. The dashed lines denotes the hindcasts fail to obtain significant skill at 95% significant level.

Fig 3.3.1 Time series of annual mean global SST (averaged over 60°S-60°N) from GECCO2 (red), assimilated GECCO2 (balck) that is used as initial conditions for dffGcE, and ensemble mean of hindcast dffGcE at the first lead year (blue). The trends are denoted in dashed lines of left panels. The lines in the right panels denote the time series after the long-term linear trend removed. As is indicated by the black lines, the warm biased GECCO2 is assimilated towards the observations (HadISST, not shown).

Fig 3.3.2 Spatial distribution of the hindcasted SST anomaly correlation coefficients between dffGcE (MiKlip hindcasts) and assimilated GECCO2 at lead year 1 (top panels), averages of lead years 2-5 (middle panels) and 6-9 (bottom panels). The right column corresponds to the period in our hindcasts from 1980 to 2006, and the left column is from 1961-2009. All the SST used is detrended. Only the significant coefficients (at 95% level) are shown here.

Fig 3.3.3 Spatial distribution of the hindcasted SST anomaly correlation coefficients between dffGcE (MiKlip hindcasts) and assimilated GECCO2 at lead year 1 (top panels), averages of lead years 2-5 (middle panels) and 6-9 (bottom panels). The left column is from ensemble mean of dffGcE, and the middle and right columns are from ensemble member 1 and 3 of dffGcE respectively, all with the same verification period of 1980-2006. All the SST used is detrended. Only the significant coefficients (at 95% level) are shown here.

Fig 3.3.4 Time series of annual mean NA MOC (Sv) anomaly of upper 1000m at 45°N from dffGcE hindcast (black), the assimilation run used for initialization (red, i.e. after nudging the MPI-ESM to GECCO2), and GECCO2 Synthesis (blue) at different lead years. The top panel is AMOC at lead year 1. The 4-yr mean of lead year 2-5 is plotted with the time series centered in year 3 (middle). Averages over lead year 6-9 are shown in the bottom panel, with time series plotted centered in year 7.

Fig 3.3.5 ACC of non-detrended Atlantic MOC(Sv) at 1000m depth between dffGcE hindcasts and GECCO2 (black: 1961-2009, red: 1980-2006), GECCO2 and persistence (cyan: 1961-2009, blue: 1980-2006) at different lead times along latitudes. The top panels are at lead year 1, the middle and bottom panel are 4-yr average of lead year 2-5 and 6-9 seprately. The dashed part of the lines denotes the correlation fails to pass the 95% significance level.

Fig 3.3.6 ACC of detrended Atlantic MOC(Sv) at 1000m between dffGcE hindcasts and GECCO2 (black:1961-2009, red: 1980-2006), GECCO2 and persistence (cyan: 1961-2009, blue: 1980-2006) at different lead times along latitudes. The top panels are at lead year 1, the middle and bottom panel are 4-yr average of lead year 2-5 and 6-9 seperately. The dashed part of the lines denotes the correlation fails to pass the 95% significance level.

Fig 4.1.1 The Pacific Pattern under non-El Niño conditions (left) and El Niño conditions (right) from NOAA. In normal situation, equatorial winds gather warm water pool toward the west, and cold water upwells along the South American coast. During an event of El Niño, warm water pool approaches the South American coast. The absence of cold upwelling increases warming in the eastern Pacific. The absence of cold upwelling increases warming in the eastern Pacific.

Fig 4.1.2 Location of Niño sea surface temperature regions (Niño 1, 2, 3, 3.4 and 4 regions) from NOAA. In this thesis, Niño 3.4 Index based on sea surface temperature of Niño 3.4

region is used to analyze the characteristic of ENSO. Figure is available from official website of NOAA (www.ncdc.noaa.gov/teleconnections/enso/indicators/sst.php).

Fig 4.1.3 Niño 3.4 Index of lead year 1 from the hindcasts (black), NCAR (green) and the initialization data (red, CDA and GECCO2 respectively). The upper panel is for CIH and CDA. The bottom panel is GIH and GECCO2.

Fig 4.1.4 Climatological monthly Niño 3.4 Index over the first 24 months of historical El Niño years (i.e. 1982, 1986, 1991, 1994, 1997, 2002 and 2006; upper) and GIH produced El Niño-like years (i.e. 1980, 1981, 1985, 1989, 1996, 2001, 2003, 2005; bottom) from CIH (black), GIH (red) and NCAR (green). For the hindcasts CIH and GIH, all the monthly indices shown are calculated from the climatology of the first 24 months of forecasts started from the initialization date. The index from NCAR is used as a reference. The blue dashed lines represent the threshold for ENSO event for CIH and GIH. The value exceeding $\pm 0.4^{o}$C indicates an El Niño event.

Fig 4.1.5 Hovmöller diagrams of climatological SSH anomaly (bottom panels) and zonal wind stress *taux* anomaly (upper panels) along the equator from GECCO2 (1980-2006) over historical El Niño years (right panels) and the non-El Niño years (left panels). The units for SSH anomaly and *taux* anomaly are centimeters (cm) and $100N/m^2$ respectively.

Fig 4.1.6 Hovmöller diagrams of climatological SSH anomaly along the equator from the first 12-months of CIH (right panels) and GIH (left panels) over historical El Niño years (bottom panels) and the GIH produced El Niño-like years (upper panels). The units are centimeters (cm).

Fig 4.1.7 Hovmöller diagrams of climatological zonal wind stress (taux) anomaly along the equator from the first 12-months of CIH (right panels) and GIH (left panels) over historical El Niño years (bottom panels) and the GIH produced El Niño-like years (top panels). The units are $100N/m^2$.

Fig 4.2.1 Zonal momentum balance of upper equatorial Pacific between pressure gradient force and zonal wind stress from GECCO2 Synthesis (1980-2006) in historical El Niño years (black) and the non-El Niño years (red) at: (a) the former December, (b) January, (c) February, (d) March, (e) May, (f) July, (g) August, (h) September, (i) October, (j) November , (k) December and (l) the following January.

Fig 4.2.2 Zonal momentum balance of upper equatorial Pacific between pressure gradient force and zonal wind stress: 1) from CIH in historical El Niño years (green) and the rest years (cyan); 2)from GIH in historical El Niño years (black); GIH-produced pseudo El Niño years (red) and the non-El Niño years (blue). The climatological first 12 months of the corresponding years during 1980-2006 are shown from (a) to (l) respectively (i.e. January in (a), February in (b), … , December in (l)).

Fig 4.2.3 Climatological SSH along the equatorial Pacific in January of (a) GIH produced pseudo El Niño years, (b) historical El Niño years, and (d) non-El Niño years for GIH (black), GECCO2(red), CDA (green) and CIH (blue). Climatological SSH in January of CDA are displayed in (d) in three categories of historical El Niño years (black), pseudo El Niño years

(red) and non-El Niño years (green), as well as GECCO2 of the same categories (blue, cyan, magenta respectively). The units are *cm*.

Fig 4.2.4 Climatological SST along the equatorial Pacific in January for (a) GIH produced pseudo El Niño years, (b) historical El Niño years and (c) non- El Niño years, from GIH (black), GECCO2 (red), CDA (green), and CIH (blue). The units are ^{o}C.

Fig 4.2.5 Climatological SST along the equatorial Pacific in January for 1) historical El Niño years from GECCO2 (black), CDA (red), and HadISST (green), 2) non-El Niño years from GECCO2 (blue), CDA (cyan), and HadISST (magenta) and CDA (dark green). The units are ^{o}C.

Fig 4.2.6 Climatological zonal wind stress (a) and zonal pressure gradient (b) along the equatorial Pacific in January for 1) historical El Niño years from GECCO2 (black) , CDA (red) and 20C (green); 2) non-El Niño years from GECCO2 (blue) CDA (cyan) and 20C (magenta). The units for wind stress are N/m^2, and for pressure gradients are P_a/m .

Table Captions

Acronyms

ACC anomaly correlation coefficient

AFES Atmospheric GCM for the Earth Simulator (AFES)

AMOC Atlantic Meridional Overturning Circulation

AMO Atlantic multidecadal oscillation

CDA coupled data assimilation synthesis through CFES

CFES Coupled GCM of Earth Simulator

CIH CDA initialized hindcasts

CMIP5 Coupled Model Intercomparison Projection phase 5

dffGcE full-state initializated hindcasts from MiKlip through MPI-ESM

ECHAM6 European Centre-Hamburg model version 6

ENSO El Niño-Southern Oscillation

GCM general circulation model

GECCO2 German contribution to Estimating the Circulation and Climate Ocean project

GIH GECCO2 initialized hindcasts

GHGs green house gases

HadISST Met Office Hadley Centre's sea surface temperature

MiKlip Mittelfristige Klimaprognosen project

MOM3 GFDL Modular Ocean Model

MPI-M Max Planck Institute for Meteorology

MPI-ESM Max-Planck-Institute Earth System Model

MPI-OM Max Planck Institute ocean-sea ice model

NA North Atlantic

NAO North Atlantic Oscillation

JAMSTEC Japan Agency for Marine-Earth Science and Technology

OIFES Ocean-Sea Ice GCM for the Earth Simulator

PGF pressure gradient force

RMSE root mean square error

RMSS root mean square error skill score

taux zonal wind stress

SI seasonal to interannual

SSH sea surface height

SST sea surface temperature

Reference

Alexander, M. A., Matrosova, L., Penland, C., Scott, J. D., and Chang, P., 2008: Forecasting Pacific SSTs: Linear inverse model predictions of the PDO, Journal of Climate, 21, 385-402.

An, S. I., and Jin, F. F., 2001: Collective role of thermocline and zonal advective feedbacks in the ENSO mode, Journal of Climate, 14, 3421–3432.

An, S. I., and Wang, B., 2000: Interdecadal change of the structure of the ENSO mode and its impact on the ENSO frequency, Journal of Climate, 13, 2044-2055.

Arakawa, A., and Schubert, W. H., 1974: Interactions of cumulus cloud ensemble with the large-scale environment. Part I, Journal of the Atmospheric Society, 31, 674-701.

Baehr, J., 2011: Influence of the 26°N RAPID–MOCHA Array and Florida Current Cable Observations on the ECCO–GODAE State Estimate, Journal of Physical Oceanography, 40, 865-879.

Balmaseda, M. and Anderson, D., 2009: Impact of initialization strategies and observations on seasonal forecast skill, Geophysical Research Letters, 36, L01701, doi: 10.1029/2008GL035561.

Barnett, T. P., Pierce, D. W., Hidalgo, H. G., Bonfils, C., Santer, B. D., Das, T., Bala, G., Wood, A. W., Nozawa, T., Mirin, A. A., Cayan, D. R., Dettinger, M. D., 2008: Human-induced changes in the hydrology of the western United States, Science, 319, 1080-1083, doi: 10.1126/science.1152538

Barsugli, J., Anderson, C., Smith, J. B., and Vogel, J. M., 2009, Options for improving climate modeling to assist water utility planning for climate change, Western Utilities Climate Alliance (WUCA) White Paper, 144 pp. [available online at www.wucaonline.org/assets/pdf/actions_whitepaper_120909.pdf].

Bell, M. J., Martin, M. J., and Nichols, N. K., 2004: Assimilation of data into an ocean model with systematic errors near the equator, Quarterly Journal of the Royal Meteorological Society, 130, 873-893, doi: 10.1256/qj.02.109.

Bellucci, A., Haarsma, R., Gualdi, S., Athanasiadis, P., Caian, M., Cassou, C., Germe, A., Jungclaus, J., Kröger, J., Matei, D., Müller, W., Pohlmann, H., Melia, D. S., Sanchez, E., Smith, D., Terray, L., and Wyser, K., 2012: An assessment of a multi-model ensemble of decadal climate predictions performed within the framework of the COMBINE project, COMBINE Technical Report No. 2

Berner, J., Doblas-Reyes, F. J, Palmer, T. N., Shutts, G., and Weisheimer, A., 2008: Impact of a cellular automaton backscatter scheme on the systematic error and seasonal prediction skill of a global climate model, Philosophysical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 366, 2561-2579, doi: 10.1098/rsta.2008.0033.

Boer, G., 2000: A study of atmosphere-ocean predictability on long time scales, Climate Dynamics, 16, 469-477.

Boer, G., 2004: Long time-scale potential predictability in an ensemble of coupled climate models, Climate Dynamics, 23, 29-44.

Boer, G., and Lambert, S., 2008: Multi-model decadal potential predictability of precipitation and temperature, Geophysical Research Letters, 35, L05 706, doi: 10.1029/2008GL033234.

Boer, G. J., 2011: Decadal potential predictability of twenty-first century climate, Climate Dynamics, 36, 1119-1133, doi: 10.1007/s00382-010-0747-9

Boer, G. J., Kharin, V., and Merryfield, W. J., 2013: Decadal predictability and forecast skill, Climate Dynamics, 41, 1817-1833, doi: 10.1007/s00382-013-1705-0.

Booth, B. B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N., 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability, Nature, 484, 228-232.

Braganza, K., Gergis, J., Power, S., Risbey, J., and Fowler, A., 2009: A multiproxy index of the El Niño-Southern Oscillation, A.D. 1525-1982, Journal of Geophysical Research Atmosphere, 114, D05106.

Branstator, G., and Teng, H., 2010: Two limits of initial-value decadal predictability in a CGCM, Journal of Climate, 23, 6292-6311.

Branstator, G., and Teng, H., 2012: Potential impact of initialization on decadal predictions as assessed for CPMI5 models, Geophysical Research Letters, 39, L12703, doi: 10.1029/2012GL051794.

Branstator, G., and Teng, H., 2014: Is AMOC more predictable than North Atlantic heat content? Journal of Climate, 27, 3537–3550, doi: 10.1175/JCLI-D-13-00274.1

Branstator, G., Teng, H., Meehl, G., Kimoto, M., Knight, J., Latif, M., and Rosati, A., 2012: Systematic estimates of initial value decadal predictability for six AOGCMs, Journal of Climate, 25, 1827-1846.

Bryden, H. L., Brady, E. C., 1985: Diagnostic Model of the Three-Dimensional Circulation in the Upper Equatorial Pacific Ocean, American Meteorological society, 1255-1273.

Carton, J., and Santorelli, A., 2008: Global decadal upper-ocean heat content as viewed in nine analyses, Journal of Climate, 21, 6015-6035.

Cazenave, A., Dominh, K., Gennero, M. C., and Ferret, B., 1998: Global mean sea level change observed by TOPEX/POSEIDON and ERS-1, Physics and Chemistry of the Earth, 23, 1069-1075, doi: 10.1016/S0079-1946(98)00146-3

Chikamoto, Y., Kimoto, M., Ishii, M., Watanabe, M., Nozawa, T., Mochizuki, T., Tatebe, H., Sakamoto, T. T., Komuro, Y., Shiogama, H., Mori, M., Yasunaka, S., Imada, Y., Koyama, H., Nozu, M., and Jin, F., 2012a: Predictability of a stepwise shift in Pacific climate during the late 1990s in hindcast experiments using MIROC, Journal of Meteorological Society of Japan, 90A, 1-21, doi:10.2151/jmsj.2012-A01

Cobb, K. M., Charles, C. D., Cheng, H., and Edwards, R. L., 2003: El Nino/Southern Oscillation and tropical Pacific climate during the last millennium, Nature, 424, 271–276.

Collins, M., 2002: Climate predictability on interannual to decadal time scales: the initial value problem, Climate Dynamics, 19, 671-692.

Collins, M., Botzet, M., Carril, A., Drange, H., Jouzeau, A., Latif, M., Masina, S., Otteraa, O., Pohlmann, H., Sorteberg, A., and Coauthors, 2006: Interannual to decadal climate predictability in the North Atlantic: A multi-model ensemble study, Journal of Climate, 19, 1195-1203.

Collins, M., An, S., Cai, W., Ganachaud, A., Guilyardi, E., Jin, F., Jochum, M., Lengaigne, M., Power, S., Timmermann, A., Vecchi G., and Wittenberg, A., 2010: The impact of global warming on the tropical Pacific ocean and El Niño, Nature Geoscience, 3, 391-397.

Corti, S., Weisheimer, A., Palmer, T., Doblas-Reyes, F., and Magnusson, L., 2012: Reliability of decadal predictions, Geophysical Research Letters, 39, L21712, doi: 10.1029/2012GL053354

Cunningham, S. A., and Coauthors, 2007: Temporal variability of the Atlantic meridional overturning circulation at 26.5^{o}N, Science, 317, 935-938.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553-597, doi: 10.1002/qj.828

Delworth, T. L., Zhang, R., and Mann, M. E., 2007: Decadal to centennial variability of the Atlantic from observations and models, Ocean Circulation: Mechanisms and Impacts, Geophysical Monography Series 173, American Geophysical Union, 131-148, doi: 10.1029/173GM10.

Doblas-Reyes, F. J., Balmaseda, M., Weisheimer, A., and Palmer, T., 2011a: 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: 10.1029/2010JD015395.

Doblas-Reyes, F. J., Andreu-Burillo, I., Chikamoto, Y., García-Serrano, J., Guemas, V., Kimoto, M., Mochizuki, T., Rodrigues, L. R. L, and van Oldenborgh, G. J., 2013: Initialized near-term regional climate change prediction, Nature Communications, 4, 1715, doi:10.1038/ncomms2704.

Doblas-Reyes, F. J, Weisheimer, A., Déqué, M., Keenlyside, N., MacVean, M., Murphy, J., Rogel, P., Smith D., and Palmer, T. N., 2009: Addressing model uncertainty in seasonal and annual dynamical ensemble forecasts, Quarterly Journal of Royal Meteorological Society, 135, 1538-1559, doi: 10.1002/qj.464

Du, H., Doblas-Reyes, F. J., García-Serrano, J., Guémas, V., Soufflet, Y., and Wouters, B., 2012: Sensitivity of decadal predictions to the initial atmospheric and oceanic perturbations, Climate Dynamics, 39, 2013-2023, doi: 10.1007/s00382-011-1285-9.

Dunstone, N. J., Smith, D. M., Booth, B. B. B., Hermanson, L., and Eade, R., 2013: Anthropogenic aerosol forcing of Atlantic tropical storms, Nature Geoscience, 6, 534-539.

Eden, C., and Willebrand, J., 2001: Mechanism of interannual to decadal variability of the North Atlantic circulation, Journal of Climate, 14, 2266-2280.

Enfield, D. B., Mestas-Nuñez, A. M., and Trimble, P. J., 2001: The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental U.S, Geophysical Research Letters, 28, 2077-2080.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using monte-carlo methods to forecast error statistics, Journal of Geophysical Research: Oceans, 99(C5), 10143-10162.

Fedorov, A., and Philander, S. G., 2000: Is El Niño changing? Science, 288, 1997-2001.

García-Serrano, J., and Doblas-Reyes, F. J., 2012: On the assessment of near-surface global temperature and North Atlantic multi-decadal variability in the ENSEMBLES decadal hindcast, Climate Dynamics, 39, 2025-2040, doi: 10.1007/s00382-012-1413-1.

Giering, R., and Kaminski, T., 1998: Recipes for adjoint code construction, ACM Transactions on Mathematical Software, 24(4), 437-474.

Giering, R., and Kaminski, T., 2003: Applying TAF to generate efficient derivative code of Fortran 77–95 programs, Proceedings in Applied Mathematics and Mechanics, 2(1), 54-57.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5, Journal of Advances in Modelling Earth Systems, 5, 572-597, doi: 10.1002/jame.20038.

Glantz, M. H., 1996: Currents of Change: El Niño's Impact on Climate and Society, Cambridge University Press, 194pp.

Goddard, L., Kumar, A., Solomon, A., Smith, D., Boer, G., Gonzalez, P., Kharin, V., Merryfield, W., Deser, C., Mason, S. J., and Coauthors, 2013: A verification framework for interannual-to-decadal predictions experiments, Climate Dynamics, 40, 245-272.

Goldenberg, S., Landsea, C., Mestas-Nuñuz, A., and Gray, W., 2001: The recent increase in Atlantic hurricane activity: Causes and implications, Science, 293, 474-479.

Gregory, J., Dixon, K., Stouffer, R., Weaver, A., Driesschaerter, E., Eby, M. Fichefet, T., Griffies, S., and Bryan, K., 1997: Predictability of North Atlantic multidecadal climate variability, Science, 275, 181-184.

Guémas, V., Corti, S., García-Serrano, J., Doblas-Reyes, F. J., Balmaseda, M. A., and Magnusson, L., 2013a: The Indian Ocean: The region of highest skill worldwide in decadal climate prediction, Journal of Climate, 26, 726-739.

Guémas, V., Doblas-Reyes, F. J., Andreu-Burillo, I., and Asif, M., 2013b: Retrospective prediction on the global warming slowdown in the past decade, Nature Climate Change, 3, 649-653, doi: 10.1038/nclimate1863.

Guilyardi, E., Bellenger, H., Collins, M., Ferrett, S., Cai, W., and Wittenberg, A., 2012: A first look at ENSO in CMIP5, CLIVAR Exchanges, 58, 29-32.

HadGEM2 Model Development Team, 2011: The HadGEM2 family of Met Office Unified Model Climate configurations, Geoscientific Model Development Discussions, 4, 765-841, 2011 (http://www.geosci-model-dev-discuss.net/4/765/2011/gmdd-4-765-2011.pdf)

Hagendorn, R., Doblas-Reyes, F. J., and Palmer, T. N., 2005: The rationale behind the success of multi-model ensembles in seasonal forecasting-I. Basic concept, Tellus A, 57, 219-233.

Ham, Y.-G., Rienecker, M. M., Suarez, M. J., Vikhliaev, Y., Zhao, B., Marshak, J., Vernieres, G., and Schubert, S. D., 2014: Decadal prediction skill in the GEOS-5 forecast system, Climate Dynamics, 42, 1-20.

Hasumi, H., Hu, A., Jungclaus, J., and Coauthors, 2005: A model intercomparision of changes in the Atlantic thermohaline circulation in response to increasing atmospheric CO_2 concentration, Geophysical Research Letters, 32, L12703, dio:10.1029/2005GL023209.

Hawkins, E., and Sutton, R., 2009a: The potential to narrow uncertainty in regional climate predictions, Bulletin of the American Meteorological Society, 90, 1095-1107.

Hazeleger, W., and Coauthors, 2013a: Predicting multi-year North Atlantic Ocean Variability, Journal of Geophysical Research, 118, 1087-1098, doi: 10.1002/jgrc.20117.

Hazeleger, W., Guemas, V., Wouters, B., Corti, S., Andreu-Burillo, I., Doblas-Reyes, F., Wyser, K., and Caian, M., 2013b: Multiyear climate prediction using two initialization strategies, Geophysical Research Letters, 40, 1794-1798.

Hebert, D., Moum, J. N., Paulson, C. A., Caldwell, D. R., Chereskin, T. K., and Mcphaden, M. J., 1991: The Role of the Turbulent Stress Divergence in the Equatorial Pacific Zonal Momentum Balance, Journal of Geophysical Research, 96 (4), 7127-7136.

Hibler, W. D., 1980: Modeling a variable thickness sea ice cover, Monthly Weather Review, 108, 1943-1973.

Hirschi, J. J. M., Killworth, P. D., and Blundell, J. R., 2007: Subannual, seasonal, and interannual variability of the North Atlantic meridional overturning circulation, Journal of Physical Oceanography, 37, 1246-1265.

Ho, C. K., Hawkins, E., Shaffrey, L., and Underwood, F., 2012: Statistical decadal predictions for sea surface temperatures: A benchmark for dynamical GCM predictions, Climate Dynamics, 41, 917–935, doi: 10.1007/s00382-012-1531-9.

Ho, C. K., Hawkins, E., Shaffrey, L., Broecker, J., Hermanson, L., Murphy, J. M., and Smith, D. M., 2013: Examining reliability of seasonal to decadal sea surface temperature forecasts: The role of ensemble dispersion, Geophysical Research Letters, 40, 5770-5775.

Hovmöller, E. A., 1949: The Trough-and-Ridge Diagram, Tellus, 1(2), 62-66.

Hurrell, J. W., Visbeck, M., Busalacchi, A., Clarke, P., Delworth, T., Dickson, R., Johns, W. E., Koltermann, K., Kushnir, Y., Marshall, D., and Coauthors, 2006: Atlantic climate variability and predictability: A CLIVAR perspective, Journal of Climate, 19, 5100-5121.

Hurrell, J. W., and Coauthors, 2009: Decadal climate predictions: opportunities and challenges, Ocean Obs'09 Community White Paper, 1-21.

ICPO (International CLIVAR Project Office), 2011: Data and bias correction for decadal climate predictions, International CLIVAR Project Office, CLIVAR Publication Series No. 150, p 6. Available from http://eprints.soton.ac.uk/171975/1/ICPO150_Bias.pdf

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, 2013, [Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, USA.

IPCC: Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007, [Trenberth, K. E., Jones, P. D., Ambenje, P., Bojariu, R., Easterling, D., Klein Tank, A., Parker, D., Rahimzadeh, F., Renwick, J. A., Rusticucci, M., Soden, B., and Zhai, P.], Cambridge University Press, Cambridge, UK, pp. 235-336.

Ishii, M., and Kimoto, M., 2009: Reevaluation of historical ocean heat content variations with time-varying XBT and MBT depth bias corrections, Journal of Oceanography, 65, 287-299.

Jolliffe, I. T., and Stephenson, D. B., 2003: Forecast Verification. A Practitioner's Guide in Atmospheric Science, Willey and Sons Ltd, 240 pp.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D., and von Storch, J. S., 2013: Characteristics of the ocean simulations in MPIOM, the ocean component of the MPI-Earth system model, Journal of Advances in Modelling Earth Systems, 5, 422-446, doi: 10.1002/jame.20023.

Kanzow, T., and Coauthors, 2007: Observed flow compensation associated with the MOC at 26.5^{o}N in the Atlantic, Science, 317, 938-941.

Keenlyside, N., Latif, M., Jungclaus, J., Kornblueh, L., and Roeckner, E., 2008: Advancing decadal-scale climate prediction in the North Atlantic sector, Nature, 453, 84-88.

Kennedy, J., and Coauthors, 2010, "How do we know the world has warmed? in: 2. Global Climate, in: State of the Climate in 2009", Bulletin of the American Meteorological Society, 91 (7), 26-27

Kharin, V., Boer, G., Merryfield, W., Scinocca, J., and Lee, W.-S., 2012: Statistical adjustment of decadal predictions in a changing climate, Geophysical Research Letters, 39, L19705, doi: 10.1029/2012GL052647.

Kim, B. M., and An, S. I., 2011: Understanding ENSO regime behavior upon an Increase in the warm-pool temperature using a simple ENSO model, Journal of Climate, 24, 1438-1450.

Kim, H.-M., Webster, P. J., and Curry, J. A., 2012: Evaluation of short-term climate change prediction in multi-model CMIP5 decadal hindcasts, Geophysical Research Letters, 39, L10701, doi: 10.1029/2012GL051644.

Kim, S. T., and Yu, J.-Y., 2012: The two types of ENSO in CMIP5 models, Geophysical Research Letters, doi: 10.1029/2012GL052006.

Knight, J. R., Allan, R. J., Folland, C. K., Vellinga, M., and Mann, M. E., 2005: A signature of persistent natural thermohaline circulation cycles in observed climate, Geophysical Research Letters, 32, L20708, doi: 10.1029/2005GL024233.

Knight, J. R., Folland, C. K., and Scaife, A. A., 2006: Climate impacts of the Atlantic Multidecadal Oscillation, Geophysical Research Letters, 33, L17706, doi: 10.1029/2006GL026242.

Köhl, A., 2005: Anomalies of meridional overturning: Mechanisms in the North Atlantic, Journal of Physical Oceanography, 35, 1455-1472.

Köhl, A., 2014: Evaluation of the GECCO2 ocean synthesis: transports of volume, heat and freshwater in the Atlantic, Quarterly Journal of the Royal Meteorological, DOI:10.1002/qj.2347

Köhl, A., and Stammer, D., 2007: Decadal sea level changes in the 50-year GECCO ocean synthesis, Journal of Climate, 21, 1876-1890.

Köhl, A., and Stammer, D., 2008: Variability of the meridional overturning in the North Atlantic from the 50-year GECCO state estimation, Journal of Physical Oceanography, 38, 1913-1930.

Krishnamurti, T. N., Bedi, H. S., and Hardiker, V. M., 1998: An Introduction to Global Spectral Modeling, Oxford University Press, Oxford, United Kingdom.

Kröger, J., Müller, W. A., and von Storch, J.-S., 2012: Impact of different ocean reanalyses on decadal climate prediction, Climate Dynamics, 39, 795-810, doi: 10.1007/s00382-012-1310-7.

Latif, M., Collins, M., Pohlmann, H., and Keenlyside, N., 2006: A review of predictability studies of Atlantic sector climate on decadal time scales, Journal of Climate, 19, 5971-5987.

Latif, M., and Keenlyside, N. S., 2011: A perspective on decadal climate variability and predictability, Deep Sea Research part II, 58, 1880-1894.

Lau, K., and Coauthors, 2008: The Joint Aerosol-Monsoon Experiment—A new challenge for monsoon climate research, Bulletin of the American Meteorological Society, doi: 10.1175/BAMS-89-3-369, 369-383.

Levitus, S., Antonov, J. I., Boyer, T. P., Locarnini, R. A., Garcia, H. E., and Mishonov, A. V., 2009: Global ocean heat content 1955–2008 in light of recently revealed instrumentation problems, Geophysical Research Letters, 36, L07608, doi:10.1029/2008GL037155.

Li, J., and coauthors, 2011c: Interdecadal modulation of El Niño amplitude during the past millennium, Nature Climate Change, 1, 114–118.

Ludescher, J., Gozolchiani, A., Bogachev, M. I., Bunde, A., Havlin, S., and Schellnhuber, H. J., 2014: Very early warning of next El Niño, Proceedings of the National Academy of Sciences of the United States of America, 111, 2064-2066, doi: 10.1073/pnas.1323058111.

Magnusson, L., Alonso-Balmaseda, M., Corti, S., Moltenie, F., and Stockdale, T., 2012a: Evaluation of forecast strategies for seasonal and decadal forecasts in in presence of systematic model errors, Climate Dynamics, 41, 2393-2409, doi:10.1007/s00382-012-1599-2.

Magnusson L., Alonso-Balmaseda, M., and Moltenie, F., 2012b: On the dependence of ENSO simulation on the coupled model mean state, Climate Dynamics, 41, 1509-1525, doi:10.1007/s00382-012-1572-y.

Manganello, J. V. and Huang, B., 2009: The influence of systematic errors in the Southeast Pacific on ENSO variability and prediction in a coupled GCM, Climate Dynamics, 32, 1015-1034.

Mankiewicz, R., 2004: The Story of Mathematics (Paperback ed.), Princeton, NJ: Princeton University Press, 158 pp.

Marsland, S. J., Haak, H., Jungclaus, J. H., Latif, M., and Roeske, F., 2003: The Max-Planck-Institute global ocean/sea ice model with orthogonal curvilinear coordinates, Ocean Modelling, 5, 91-127, doi: 10.1016/S1463-5003(02)00015-X.

Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R., and Visbeck, M., 2001: North Atlantic climate variability: phenomena, impacts and mechanisms, International Journal of Climatology, 21, 1863-1898.

Masuda, S., Awaji, T., Sugiura, N., Ishikawa Y., Baba, K., Horiuchi, K., and Komori, N., 2003: Improved estimates of a dynamical state of the North Pacific Ocean from a 4 dimensional variational data assimilation, Geophysical Research Letters, 30(16), 1868, doi: 10.1029/2003GL017604

Masuda, S., Awaji, T., Sugiura, N., Toyoda, T., Ishikawa, Y., and Horiuchi, K., 2006: Interannual variability of temperature inversions in the subarctic North Pacific, Geophysical Research Letters, 33, 24610, doi: 10.1029/2006GL027865.

Matei, D., Baehr, J., Jungclaus, J. H., Haak, H., Müller, W. A., and Marotzke, J., 2012a: Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5°N, Science, 335, 76-79.

Matei, D., Keenlyside, N., Latif, M., and Jungclaus, J., 2008: Subtropical forcing of tropical Pacific climate and decadal ENSO modulation, Journal of Climate, 21, 4691-4709.

Matei, D., Baehr, J., Jungclaus, J. H., Haak, H., Müller, W. A., and Marotzke, J., 2012b: Response to Comment on "Multiyear prediction of monthly mean Atlantic Meridional Overturning Circulation at 26.5°N", Science, 338, 604, doi: 10.1126/science.1223200.

Matei, D., Pohlmann, H., Jungclaus, J., Müller, W., Haak, H., and Marotzke, J., 2012c: Two tales of initializing decadal climate prediction experiments with the ECHAM5/MPI-OM model, Journal of Climate, 25, 8502-8523.

Mcphaden, M. J., and Taft, B. A., 1988: Dynamics of seasonal and intraseasonal variability in the eastern equatorial Pacific, Journal of Physical Oceanography, 18, 1713-1732.

Means, E. M., and Coauthors, 2010: Decision support planning methods: Incorporating climate change uncertainties into water planning, Western Utilities Climate Alliance White Paper, 102 pp. [available online at www.wucaonline.org/assets/pdf/pubs_whitepaper_012110.pdf].

Meehl, G. A., and Hu, A., 2006: Megadroughts in the Indian monsoon region and southwest North America and a mechanism for associated multidecadal Pacific sea surface temperature anomalies, Journal of Climate, 19, 1605-1623.

Meehl, G. A., Goddard, L., Murphy, J., Stouffer, R. J., Boer, G., Danabasoglu, G., Dixon, K., Giorgetta, M. A., Greene, A. M., Hawkins, E., Hegerl, G., Karoly, D., Keenlyside, N., Kimoto, M., Kirtman, B., Navarra, A., Pulwarty, R., Smith, D., Stammer, D., and Stockdale, T., 2009: Decadal Prediction: Can it be skillful? Bulletin of the American Meteorological Society, 90, 1467-1485

Meehl, G. A., Hu, A., and Tebaldi, C., 2010: Decadal prediction in the Pacific region, Journal of Climate, 23, 2959-2973.

Meehl, G. A., Goddard, L., Boer, G., Burgman, R., Branstator, G., Cassou, C., Corti, S., Danabasoglu, G., Doblas-Teyers, F., Hawkins, E., Karspeck, A., Kimoto, M., Kumar, A., Matei, D., Mignot, J., Msadek, R., Navarra, A., Pohlmann, H., Rienecker, M.,

Rosati, T., Schneider, E., Smith, D., Sutton, R., Teng, H., van Oldenborgh, G. J., Vecchi, G., and Yeager, S., 2014: Decadal Climate Prediction: An Update from the Trenches, Bulletin of the American Meteorological Society, 95, 243-267. doi: http://dx.doi.org/10.1175/BAMS-D-12-00241.1.

Meehl, G. A., and Teng, H., 2012: Case studies for initialized decadal hindcasts and predictions in the Pacific regions, Geophysical Research Letters, 39, L22705, doi: 10.1029/2012GL053423.

Meehl, G. A., and Teng, H., 2014: CMIP5 multi-model initialized hindcasts for the mid-1970s shift and early 200s hiatus and predictions for 2016-2035, Geophysical Research Letters, in press.

Meehl, J., Arblaster, M., Matthes, K., Sassi, F., and van Loon, H., 2009b: Amplifying the Pacific climate system response to a small 11 year solar cycle forcing, Science, 325, 1114-1118.

Mehta, V. M., and Lau, K.-M., 1997: Influence of solar irradiance on the Indianmonsoon ENSO relationship at decadal-multidecadal time scales, Geophysical Research Letters, 24, 159-162.

Mellor, G. L., and Yamada, T., 1974: A hierarchy of turbulence closure models for planetary boundary layers, Journal of the Atmospheric Science, 31, 1791-1806.

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., 2013: Mechanisms of aerosol-forced AMOC variability in a state of the art climate model, Journal of Geophysical Research: Oceans, 118, 2087-2096.

Met office Hadley Centre, 2014: Decadal forecast covering the period from 2014–2018.

Mochizuki, T., Chikamoto, Y., Kimoto, M., Ishii, M., Tatebe, H., Komuro, Y., Sakamoto, T., Watanabe, M., and Mori, M., 2012: Decadal prediction using a recent series of MIROC global climate models, Journal of the Meteorological Society of Japan, 90A, 373-383.

Mochizuki, T., Ishii, M., Kimoto, M., Chikamoto, Y., Watanabe, M., Nozawa, T., Sakamoto, T., Shiogama, H., Awaji, T., Sugiura, N., and Coauthors, 2010: Pacific Decadal Oscillation hindcasts relevant to near-term climate prediction, Proceeding of the National Academy of Science, 107, 1833-1837.

Mochizuki, T., Miyama, T., and Awaji, T., 2007b: A simple diagnostic calculation of marine stratocumulus cloud cover for use in general circulation models, Journal of Geophysical Research, 112, D06113, doi: 10.1029/2006JD007223.

Molteni, F., Buizza, R., Palmer, T. N., and Petroliagis, T., 1996: The ECMWF ensemble prediction system: methodology and validation, Quarterly Journal of Royal Meteorological Society, 122, 73-119.

Msadek, R., Dixon, K. W., Delworth, T. L., and Hurlin, W., 2010: Assessing the predictability of the Atlantic meridional overturning circulation and associated Fingerprints, Geophysical Research Letters, 37, L19608, doi: 10.1029/2010GL044517.

Msadek, R., Johns, W. E., Yeager S. G., Danabasoglu, G., Delworth, T. L., and Rosati, A., 2013: The Atlantic meridional heat transport at 26.5^{o}N and its relationship with the MOC in the RAPID array and the GFDL and NCAR coupled models, Journal of Climate, 26, 4335-4356, doi: 10.1175/JCLI-D-12-00081.1.

Müller, W. A., Baehr, J., Haak, H., Jungclaus, J. H., Kröger, J., Matei, D., Notz, D., Pohlmann, H., von Storch, J. A., and Marotzke J., 2012: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology, Geophysical Research Letters, 39, L22707, doi: 10.1029/2012GL053326.

Müller, W. A., Pohlmann, H., Sienz, F., and Smith, D., 2014: Decadal climate predictions for the period 1901-2010 with a coupled climate model, Geophysical Research Letters, 41, 2100-2107, doi: 10.1002/2014GL059259.

Murphy, J. M., Booth, B. B. B., Collins, M., Harris, G. R., Sexton, D. M. H., and Webb, M. J., 2007: A methodology for probabilistic predictions of regional climate change from perturbed physics ensembles, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 1993-2028.

Murphy, J., Kattsov, V., Keenlyside, N., Kimoto, M., Meehl, G., Mehta, V., Pohlmann, H., Scaife, A., and Smith, D., 2010: Towards prediction of decadal climate variability and change, Procedia Environmental Sciences, 1, 287-304, doi: 10.1016/j.proenv.2010.09.018

Nakajima, T., and Tanaka, M., 1986: Matrix formulation for the transfer of solar radiation in a plane-parallel scattering atmosphere, Journal of Quantitative Spectroscopy and Radiative Transfer, 35, 13-21,doi: 10.1016/0022-4073(86)90088-9

Nakajima, T., Tsukamoto, M., Tsushima, Y., Numaguti, A., and Kimura, T., 2000: Modeling of the radiative process in an atmospheric general circulation model, Applied Optics, 39, 4869-4878.

National Research Council, 2010, Advancing the Science of Climate Change, Washington, D.C., The National Academies Press.

National Research Council, 2011, America's Climate Choices, Washington, D.C., The National Academies Press

Newman, M., 2007: Interannual to decadal predictability of tropical and North Pacific sea surface temperatures, Journal of Climate, 20, 2333-2356.

Noh, Y., 2004: Sensitivity to wave breaking and the Prandtl number in the ocean mixed layer model and its dependence on latitude, Geophysical Research Letters, 31, L23305, doi: 10.1029/2004GL021289.

Ohfuchi, W., Nakamura, H., Yoshioka, M. K., Enomoto, T., Takaya, K., Peng, X., Yamane, S., Nishimura, T., Kurihara, Y., and Ninomiya, K., 2004: 10-km mesh meso-scale resolving simulations of the global atmosphere on the Earth Simulator: Preliminary outcomes of AFES (AGCM for the Earth Simulator), Journal of the Earth Simulator, 1, 8-34.

Orszag, S. A., 1970: Transform method for calculation of vector-coupled sums: Application to the spectral form of the vorticity equation, Journal of the Atmospheric Science, 27, 890-895.

Otterå, O. H., Bentsen, M., Drange, H., and Suo, L., 2010: External forcing as a metronome for Atlantic multidecadal variability, Nature Geoscience, 3, 688-694, doi: 10.1038/NGEO955.

Pacanowski, R. C., and Griffies, S. M., 1999: The MOM3 manual (Tech. Rep. 4), Geophysical Fluid Dynamics Laboratory, Princeton, N. J.

Palmer, T. N., and Coauthors, 2004: Development of a European multi-model ensemble system for seasonal to inter-annual prediction, Bulletin of the American Meteorological Society, 85, 853-872, doi: 10.1175/BAMS-85-6-853

Phillips, N. A., 1957: A coordinate system having some special advantages for numerical forecasting, Journal of Meteorology, 14, 184-185.

Pierce, D., Barnett, T., Tokmakian, R., Semtner, A., Maltrud, M., Lysne, J., and Craig, A., 2004: Initializing a coupled climate model from observed conditions, Climate Change, 62, 13-28.

Pohlmann, H., Botzet, M., Latif, M., Roesch, A., Wild, M., and Tschuck, P., 2004: Estimating the decadal predictability of a coupled AOGCM, Journal of Climate, 17, 4463-4472.

Pohlmann, H., Jungclaus, J. H., Köhl, A., Stammer, D., and Marotzke, J., 2009: Initializing decadal climate predictions with the GECCO oceanic synthesis: Effects on the North Atlantic, Journal of Climate, 22, 3926-3938.

Pohlmann, H., Müller, W. A., Kulkarni, K., Kameswarrao, M., Matei, D., Vamborg, F. S. E., Kadow, C., Illing, S., and Marotzke, J., 2013: Improved forecast skill in the tropics in the new MiKlip decadal climate predictions, Geophysical Research Letters, 40, 5798-5802, doi: 10.1002/2013GL058051.

Pohlmann, H., Sienz, F., and Latif, M., 2006: Influence of the multidecadal Atlantic meridional overturning circulation variability on European climate, Journal of Climate, 19, 6062-6067.

Polkova, I., 2014: Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model and related mechanisms for predictability, PhD thesis, Universität Hamburg, Hamburg, 107 pp.

Polkova, I., Köhl, A., and Stammer, D., 2014: Impact of initialization procedures on the predictive skill of a coupled ocean-atmosphere model, Climate Dynamics, 42, 3151-3169, doi: 10.1007/s00382-013-1969-4.

Power, S., Casey, T., Folland, C., Colman, A., and Mehta, V., 1999a: Inter-decadal modulation of the impact of ENSO on Australia, Climate Dynamics, 15, 319-324.

Power, S. B., and Smith, I. N., 2007: Weakening of the Walker Circulation and apparent dominance of El Nino both reach record levels, but has ENSO really changed? Geophysical Research Letters, 34, L18702.

Pulwarty, R., 2003: Climate and water in the West: Science, information and decision-making, Water Resources, 124, 4-12

Rahmstorf, S., 2003: The concept of the thermohaline circulation, Nature, 421 (6924), 699, doi: 10.1038/421699a.

Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A., 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century, Journal of Geophysical Research, 108, 4407, doi: 10.1029/2002JD002670.

Robson, J., Sutton, R., Lohmann, K., Smith, D., and Palmer, M. D., 2012a: Causes of the rapid warming of the North Atlantic Ocean in the mid-1990s, Journal of Climate, 25, 4116-4134.

Robson, J., Sutton, R., and Smith, D., 2012b: Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s, Geophysical Research Letters, 39, L19713, doi: 10.1029/2012GL053370.

Roeckner, E., and Coauthors, 2003: The atmospheric general circulation model ECHAM5. Part I: Model description, Max-Planck-Institute für Meteorologie Report No. 349, 140pp.

Rosati, A., Miyakoda, K., and Gudgel, R., 1997: The impact of ocean initial conditions on ENSO forecasting with a coupled model, Monthly Weather Review, 125, 754-772.

Seager, R., Kushnir, Y., Ting, M., Cane, M., Naik, N., and Miller, J., 2008: Would advance knowledge of 1930s SSTs have allowed prediction of the dust bowl drought? Journal of Climate, 21, 3261-3281.

Schneider, E., Huang, B., Zhu, Z., DeWitt, D., Kinter III, J., Kirtman, B., and Shukla, J., 1999: Ocean data assimilation, initialization, and predictions of ENSO with a coupled GCM, Monthly Weather Review, 127, 1187-1207.

Smith, D. M., Cusack, S., Colman, A. W., Folland, C. K., Harris, G. R., et al., 2007: Improved surface temperature prediction for the coming decade from a global climate model, Science, 317, 796-799, doi: 10.1126/science.1139540.

Smith, D., Scaife, A. A., and Kirtman, B. P., 2012a: What is the current state of scientific knowledge with regard to seasonal and decadal forecasting? Environmental Research Letters, 7, 015602, doi: 10.1088/1748-9326/7/1/015602.

Smith, D. M., Eada, R., and Pohlmann, H., 2013: 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.

Solomon, A., and coauthors, 2011: Distinguishing the roles of natural and anthropogenically forced decadal climate variability: Implications for prediction, Bulletin of the American Meteorological Society, 92, 141-156.

Srokosz, M., Baringer, M., Bryden, H., Cunningham, S., Delworth, T., Lozier, S., Marotzke, J., and Sutton, R., 2012: Past, present, and future changes in the Atlantic meridional overturning circulation, Bulletin of the American Meteorological Society, 93, 1663-1676.

Stainforth, D. A., and Coauthors, 2005: Uncertainty in predictions of the climate response to rising levels of greenhouse gases, Nature, 433, 403-406.

Stammer, D., Cazenave, A., Ponte, R. M., and Tamisiea, M. E., 2013: Causes for Contemporary Regional Sea Level Changes, Annual Review of Marine Science, 5, 21-46, doi: 10.1146/annurev-marine-121211-172406.

Stammer, D., Köhl, A., Awaji, T., Balmaseda, M., Behringer, D., Carton, J., Ferry, N., et al., 2009: Ocean information provided through ensemble ocean synthesis, Proceedings of OceanObs09: Sustained Ocean Observations and Information for Society, Venice, Italy, 21-25 September 2009, 2.

Stan, C., and Kirtman, B. P., 2008: The influence of atmospheric noise and uncertainty in ocean initial conditions on the limit of predictability in a coupled GCM, Journal of Climate, 21, doi: 10.1175/2007JCLI2071.1.

Stevens, B., and Coauthors, 2013: Atmospheric component of the MPI-M Earth System Model: ECHAM6, Journal of Advances in Modelling Earth Systems, 5, 146-172, doi: 10.1002/jame.20015.

Stevenson, S. L., 2012: Significant changes to ENSO strength and impacts in the twenty-first century: Results from CMIP5, Geophysical Research Letters, doi: 10.1029/2012GL052759.

Stockdale, T., 1997: Coupled ocean-atmosphere forecasts in the presence of climate drift, Monthly Weather Review, 125, 809-818.

Sugiura, N., Awaji, T., Masuda, S., Mochizuki, T., Toyoda, T., Miyama, T., Igarashi, H., and Ishikawa, Y., 2008: 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, 113, C10017, doi: 10.1029/2008JC004741.

Sun, D. Z., 1997: El Niño: A coupled response to radiative heating? Geophysical Research Letters, 24, 2031-2034.

Sun, D. Z., 2003: A possible effect of an increase in the warm-pool SST on the magnitude of El Niño warming, Journal of Climate, 16, 185-205.

Sutton, R. T., and Hodson, D. L. R., 2005: Atlantic Ocean forcing of North American and European summer climate, Science, 309, 115-118.

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

Takata, K., Emori, S., and Watanabe, T., 2003: Development of the minimal advanced treatments of surface interaction and runoff, Global Planetary Change, 38, 209-222.
Taylor, K. E., Stouffer, R. J., and Meehl, G. A., 2008: A summary of the CMIP5 experimental design, CLIVAR, 30 pp.
Taylor, K. E., Stouffer, R. J., and Meehl, G. A., 2012: An overview of CMIP5 and the experiment design, Bulletin of the American Meteorological Society, 93, 485–498.
Tebaldi, C., and Knutti, R., 2007: 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.
Tebaldi, C., Smith, R. W., Nychka, D., and Mearns, L. O., 2005: Quantifying uncertainty in projections of regional climate change: A Bayesian approach to the analysis of multi-model ensembles, Journal of Climate, 18, 1524-1540.
Teng, H., Branstator, G., and Meehl, G., 2011: Predictability of the Atlantic Overturning Circulation and Associated Surface Patterns in Two CCSM3 Climate Change Ensemble Experiments, Journal of Climate, 24, 6054-6076.
Terray, L., 2012: Evidence for multiple drivers of North Atlantic multi-decadal climate variability, Geophysical Research Letters, 39, L19712, doi: 10.1029/2012GL053046.
Tiedje, B., and Baehr, J., 2014: Limitations of the potential predictability of meridional mass and heat transports in the North Atlantic, Geophysical Research Letters, 41, 4270-4276. doi:10.1002/2014GL060174.
Tiedje, B., Köhl, A., and Baehr, J., 2012: Potential Predictability of the North Atlantic Heat Transport Based on an Oceanic State Estimate, Journal of Climate, 25(24), 8475-8486. doi:10.1175/JCLI-D-11-00606.1.
Ting, M., Kushnir, Y., Seager, R., and Li, C., 2009: Forced and Internal Twentieth-Century SST Trends in the North Atlantic, Journal of Climate, 22 (6): 1469-1481, doi: 10.1175/2008JCLI2561.1
Toth, Z., and Kalnay, E., 1997: Ensemble forecasting at NCEP and the breeding method, Monthly Weather Review, 125, 3297-3319.
Trenberth, K. E., 1997: The Definition of El Niño, Bulletin of the American Meteorological Society, 78, 2771-2777.
Trenberth, K. E., and Hoar, T. J., 1996a: The 1990-1995 El Niño-Southern Oscillation Event: Longest on record, Geophysical Research Letters, 23, 57-60.
Trenberth, K. E., and Stepaniak, D. P., 2000: Indices of El Niño Evolution, Journal of Climate, 14, 1697-1701.
Troccoli, A., and Palmer, T., 2007: Ensemble decadal predictions from analysed initial conditions, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 365, 2179-2191.
Uppala, S. M., et al., 2005: The ERA-40 re-analysis, Quarterly Journal of the Royal Meteorological Society, 131, 2961-3012, doi: 10.1256/qj.04.176.
van der Linden, P., and Mitchell, J. F. B.(Eds.), 2009: ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project, Met Office Hadley Centre, 160 pp. [Available from Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, United Kingdom.]
van Oldenborgh, G., Doblas-Reyes, F., Wouters, B., and Hazeleger, W., 2012: Decadal prediction skill in a multi-model ensemble, Climate Dynamics, 38, 1263-1280.
Vecchi, G. A., and Coauthors, 2013: Multiyear predictions of North Atlantic hurricane frequency: Promise and limitations, Journal of Climate, 26, 5337-5357.
Vera, C., Barange, M., Dube, O. P., Goddard, L., Griggs, D., Kobysheva, N., Odada, E., Parey, S., Polovina, J., Poveda, G., Seguin, B., and Trenberth, K., 2009: Needs assessment

for climateinformation on decadal time scales and longer, White Paper, World Climate Conference 3(WCC-3), Geneva, Switzerland, 31 August-4 September, 2009.
Vikhliaev, Y., Kirtman, B., and Schopf, P., 2007: Decadal North Pacific Bred Vectors in a Coupled GCM, Journal of Climate, 20, 5744-5764.
Villarini, G., and Vecchi, G. A. 2013: Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models, Journal of Climate, 26, 3231-3240.
von Storch, H., and Navarra, A., 1999: Analysis of climate variability: applications of statistical techniques, Springer, 342 pp.
von Storch, H., and Zwiers, F. W., 1999: Statistical Analysis in Climate Research, Cambridge University Press, 484 pp.
Volpi, D., Carrassi A., Doblas-Reyes F., Guemas, V., 2013, Comparison of initialization methods in global dynamic decadal climate forecasts, in preparation
Wang, B., and An, S. I., 2001: Why the properties of El Niño changed during the late 1970s, Geophysical Research Letters, 28, 3709-3712.
Wang, B., and coauthors, 2012: Preliminary evaluations of FGOALS-g2 for decadal predictions, Advances in Atmospheric Sciences, 30, 674-683, doi: 10.1007/s00376-012-2084-x.
Wilks, D. S., 2006: Statistical Methods in the Atmospheric Science, Second Edition, International Geophysical Series, 91, Academic Press.
Wittenberg, A. T., 2009: Are historical records sufficient to constrain ENSO simulations? Geophysical Research Letters, 36, L12702.
Wu, Z., Huang, N. E., Long, S. R., and Peng, C., 2007: On the trend, detrending, and variability of nonlinear and nonstationary time series, Proceedings of the National Academy of Sciences of the United States of American, 104(38), 14889-14894, doi: 10.1073/pnas.0701020104
Yan, H., Sun, L., Wang, Y., Huang, W., Qiu, S., and Yang, C., 2011: A record of the Southern Oscillation Index for the past 2,000 years from precipitation proxies, Nature Geoscience, 4, 611–614.
Yang, X., and Coauthors, 2013: A predictable AMO-like pattern in the GFDL fully coupled ensemble initialization and decadal forecasting system, Journal of Climate, 26, 650-661.
Yeager, S., Karspeck, A., Danabasoglu, G., Tribbia, J., and Teng, H., 2012: A decadal prediction case study: Late 20th century North Atlantic Ocean heat content, Journal of Climate, 25, 5173-5189.
Yeh, S. W., and Kirtman, B. P., 2005: Pacific decadal variability and decadal ENSO amplitude modulation, Geophysical Research Letters, 32, L05703.
Zhang, Q., Guan, Y., and Yang, H., 2008: ENSO amplitude change in observation and coupled models, Advances in Atmospheric Sciences, 25, 361-366.
Zhang, R., Delworth, T. L., and Held, I. M., 2007b: Can the Atlantic Ocean drive the observed multidecadal variability in Northern Hemisphere mean temperature? Geophysical Research Letters, 34, L02709.

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 Xueyuan Liu