UNIVERSITY OF SOUTHAMPTON

FACULTY OF NATURAL AND ENVIRONMENTAL SCIENCE

School of Ocean and Earth Science

Multiyear initial-value predictability of subsurface temperature in the North Atlantic Ocean and the impact of ensemble size

by

Guantong Lyu

Thesis for the degree of Master of Philosophy

May 2022

UNIVERSITY OF SOUTHAMPTON

Faculty of Natural and Environmental Science
School of Ocean and Earth Science

Master of Philosophy

ABSTRACT

Multiyear initial-value predictability of subsurface temperature in the North Atlantic Ocean and the impact of ensemble size

by Guantong Lyu

Apart from the well-documented relationships between North Atlantic sea surface temperature and local climate, there exists a close link connecting subsurface temperature in the North Atlantic Ocean and large-scale oceanic circulation and climate on multiple time scales. However relatively less is known about the variation and predictability of subsurface temperature due mainly to lack of long-term observations, which motivates this study. Using a hierarchy of coupled climate models, this thesis focuses on multiyear initial-value predictability of subsurface temperature in North Atlantic Ocean and impact of ensemble size on prediction skill and predictability assessment.

Analysis of outputs from two GCMs with different resolutions reveals that, horizontal development of subsurface temperature prediction uncertainty, which is characterized by large ensemble variance, is propagating clockwise in the subtropical North Atlantic while anticlockwise in the subpolar North Atlantic. The path along the Gulf Stream and the North Atlantic Current is the principle site where large prediction uncertainty emerges first. A further examination of the latitude-depth space reveals that temperature prediction uncertainty propagates downward mainly in the subpolar North Atlantic and Southern Ocean. The downward development in the North Atlantic is thought to be associated with mode water subduction and deep water formation process.

It is also discovered that ensemble size casts a substantial impact on prediction skill and assessment of initial-value predictability of North Atlantic Oceanic climate. Increasing ensemble size leads to a significant reduction of biases of initial-value predictability assessment. Though there

is slight difference, the minimal ensemble sizes required to make a ‘steady’ assessment of predictability for sea surface temperature, subsurface temperature and the Atlantic meridional overturning circulation in the North Atlantic in a coarse resolution (nominal resolution 2.75 °-by-3 ° degree) model are all approximately 20. This minimal size reduces to about 7 in an eddy-permitting model (nominal resolution 1/4 °-by-1/4 °). This is probably caused by the difference in the potential uncertainty sources, as a consequence of the difference in spatial resotluion, as well as the fact that the IPSL-CM5A-LR is a fully coupled climate model, while the eddy-permitting model is a forced ocean model. That is to say, for the case of North Atlantic Oceanic climate forecasting, increasing ensemble size could improve the prediction skill and credibility of predictability assessment; and minimal size are variable-independent and resolution dependent. We therefore do not recommend ensemble size far less than 20/7 and suggest that continuing increasing ensemble members if the size is already over 20/7 could be unnecessary when attempting to forecast North Atlantic Oceanic climate using non-eddy-permitting/eddy-permitting models.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Guantong Lyu, declare that the thesis entitled **Multiyear initial-value predictability of subsurface temperature in the North Atlantic Ocean and the impact of ensemble size** and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;

- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;

- where I have consulted the published work of others, this is always clearly attributed;

- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;

- I have acknowledged all main sources of help;

- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;

Signed:.....

Date:.....

Acknowledgements

First and foremost, I would like to thank Florian Sévellec and Sybren Drijfhout for offering me this wonderful opportunity to pursue a research degree in the United Kingdom. Their guidance and support are also sincerely appreciated. Without their patience and support, there would be much more difficulties in the past years.

Besides, I would like to give my special thanks to Bablu Sinha. Thank you very much for your incredible patience and support for both academia and spirit welling being.

My thanks go to Agathe Germe for her professional guidance and kind help on the use of NEMO output; to Simon Muller for his support on use of a quarter degree model output; to Jennifer V Mecking for help on JASMIN; to Jeffrey Blundell for kind help on use of NOCS clusters; to Mary Smith for her continuous support to my study. I cannot imagine the past years without your valuable help.

Friendships of Chuang Xuan (thank you for your very strong reference letter as well), Yuxi Jin, Jianghui Li, Naiyu Zhang, Wenhao Wang, Tianya Yin, Xiaodong Yang, Yingxu Wu, Tian Qi, Yue Wu, Yanan Wang, Yichu Lu, Jiachun Wang and Xiaode Zhou are amazing. Parties and with you guys are incredibly beautiful memories. Thanks also go to Joel Hirschi (I like our short discussions in the workstation area and the canteen), Prof. John Marshall (for always being very, very nice), Dafydd Stephenson, Ben Chichester, Zoë Holbrook, Elwyn De la Vega and Rachael Shuttleworth.

Last but not least, my little brother and my girlfriend Yanxin Wang deserve best appreciation.

I feel deep sorry to my family, especially to my mother and grandmother, I have ruined everything and make our life even harder than it was in 2016. Please give me some time, I will manage to take back what we lost.

Chapter 1 Introduction and backgrounds

1. Introduction

Due to the chaotic nature of atmosphere, forecasting instantaneous weather several weeks ahead is documented almost impossible. Climate, conventionally defined as an average of the weather over a long time (seasonal or longer), however in some cases is predictable years or even decades in advance thanks to various contributors (J Shukla, 1998). It is not surprising that known varying external forcing of Earth's climate, such as changing external forcing caused by volcano eruptions, would provide some potential of climate predictability (OH Otterå et al. 2010). Besides, internal variations of the climate system, such as ice sheet melting in Greenland and slow ocean processes are also predictability sources. Among them, the ocean is a significant internal source of climate predictability due principally to seawater's huge thermal inertia (S Manabe et al 1990), some intrinsic oscillatory modes of oceans (F Sévellec & AV Fedorov 2013) and the change of mixed layer depth (MA Alexander & C Deser 1995) and so on. A typical example is the North Atlantic Ocean, where a substantial predictability of the climate system resides in (J Marshall et al. 2001; M Visbeck 2002; C Franzke & T Woollings 2011). On the other hand, as a key component of global climate system, North Atlantic Ocean influences regional and distant climate on various time scales. For example, evidences are found that variation of North Atlantic Sea Surface Temperature (SST) influences storm track (JP Kossin et al. 2010; T Woollings et al. 2010; DJ Brayshaw et al. 2011), North Atlantic Oscillation (A Czaja & C Frankignoul 1999; MJ Rodwell et al. 1999; AW Robertson et al. 2000; RT Sutton et al. 2000; A Czaja & C Frankignoul 2002), climate of Europe and North America (RT Sutton & DLR Hodson 2005; Y Kushnir et al. 2010; S Feng et al. 2011), Pacific climate (R Zhang & TL Delworth 2007; C Sun et al. 2017), El Niño - Southern Oscillation (DB Enfield & DA Mayer 1997; YG Ham et al. 2013) and Asian summer monsoon (R Lu et al. 2006; J Zuo et al. 2013). Likewise, changes of Ocean Heat Transport (OHT; JP Grist et al. 2010; M Årthun et al. 2012), Mixed Layer Depth (MLD; R Seager et al. 2000) and the Atlantic Meridional Overturning Circulation (AMOC; S Drijfhout et al. 2012), subtly impact the climate on seasonal to multi-decadal timescales. Therefore the predictability of variation of these processes in the North Atlantic Ocean has drawn increasing scientific attention recently (MJ Rodwell & CK Folland 2002; M Collins et al. 2006;

M Latif et al. 2006; C Franzke & T Woollings 2011; L Zanna et al. 2012; MW Buckley et al. 2019). Also, evidences are found that North Atlantic is one of several regions where the climate is potentially predictable (RT Sutton & MR Allen 1997; SM Griffies & K Bryan 1997; R Msadek et al. 2010).

Several mechanisms have been proposed to (at least partially) account for the abovementioned predictability of the North Atlantic Ocean. Seasonal prediction of tropical Atlantic are documented to remotely obtain seasonal predictability from El Niño (P Chang et al. 2003) in tropical Pacific, which is a well-known episodic phenomenon (with a period of approximately 4 years) in the tropical Pacific Ocean. It is documented that some intrinsic low-frequency modes of the ocean, such as oscillatory behavior (P Cessi, F Paparella 2001; Zanna, Laure 2012; Sévellec, Florian, and Alexey V. Fedorov 2013), are key predictability sources on relatively longer time scales. Slow oceanic processes such as ocean advection are believed to provide the potential of predictability (SM Griffies & K Bryan 1997). Likewise, the predictability of heat content in the North Alantic Ocean on decadal or multi-decal times are intimately associated with low-frequency variability of the Atlantic Meridional Overturning Circulation (AMOC; X Yan et al. 2018). Another focus is on the non-normal characteristic of the ocean, which leads to optimal growth of initial error (E Tziperman et al. 2008; E Hawkins & R Sutton 2009) and hence enables us to examine predictability limit. Which mechanism(s) is/are playing the main role depends on the specific problem we are dealing with.

To quantitatively assess the consequent predictability, a number of approaches are progressively put forward and applied. One category of methods are based on long control simulations (E Tziperman and L Zanna 2008; T DelSole and MK Tippett 2009; E Hawkins and R Sutton 2009), such as Averaged Predictable Time (APT; T DelSole and MK Tippett 2009), Linear Optimal Perturbation (LOP; e.g., F Sévellec et al. 2017), Singular Value Decomposition (SVD; TN Palmer et al. 1994), Conditional Nonlinear Optimal Perturbation (CNOP; M Mu et al. 2003,), etc.. In addition, ensemble simulation proves itself to be a more straightforward yet informative approach in climate predictability study (JM Murphy 1988; Branković Čedo and T. N. Palmer 1997; DP Rowell 1998; M Collins 2006). One distinct advantage of ensemble prediction is that we can obtain both the spread and the probability distribution of predicted trajectories, in forms of ensemble spread (i.e., ensemble variance) and Probability Density Function (PDF), respectively. Probability forecasts given by ensemble predictions overcome to some extent the drawbacks induced by the deterministic control equations of numerical models. On the other side, ensemble variance analysis is widely used in climate predictability research. For example, the

initial-value predictability is considered to be lost if the ensemble variance is comparable to the climatological variance (i.e., the temporal variance of a long control run). Averaging over initialized ensemble members of one General Circulation Model (GCM) cancels out the initial error induced internal variability of the climate system, hence making temporal evolution of the ensemble mean mostly a result of external forcing. This in turn simplifies the quantitative comparison of internal and external variability. Multi-model ensemble mean, on the other hand, mitigates potential deficiency of particular model(s) to some degree. With these distinct advantages, ensemble method is broadly used in climate predictability studies recently (X Yang et al. 2013). Meanwhile, development of numerical computation capacity makes it possible to run ensembles with eddy-permitting or even eddy-resolving models (S Masina et al. 2017). Therefore in this thesis I choose ensemble approach as the fundamental research tool. In particularly, ensemble variance, which indicates forecast uncertainty, is frequently analyzed.

2. Background A: Impact of ensemble size on weather and climate prediction skill and assessment of predictability

Varying ensemble size can unsurprisingly influence weather and climate prediction skill on various time scales (Č Branković and TN Palmer 1997; R Buizza, TN Palmer 1998; A Kumar et al. 2001). Therefore one perpetual question prior to an ensemble forecasting is to estimate how many ensemble members are required to provide a trustworthy prediction. Theoretically, if the model is perfect, exhaustive simulations of all possible realizations of future weather or climate will provide 'realistic' probabilistic distribution and uncertainty range, which however is without doubt impossible and unnecessary. Insufficient ensemble size might probably lead to biases both in predicted probabilistic forecast and associated uncertainty range, while excessive ensemble size, on the contrary, would overly consume unnecessary computational resources. On the other hand, improving model spatial resolution is documented another way to enhance weather (CF Mass et al. 2002) or climate prediction (HT Hewitt et al. 2017). This is due mainly to the fact that models with higher resolution better represent small-scale processes, which are missed or parameterized in coarse resolution models. In this way predictions are promoted (MJ Roberts et al. 2016). Considering the limited computational sources, there must be a tradeoff between ensemble size and model resolution. Therefore to determine a sufficient but not superfluous ensemble size is an

unavoidable prerequisite when designing ensemble forecasts of weather or climate.

Indeed the influence of ensemble size on weather prediction has been studied and discussed for decades. JM Murphy (1988) reveals a considerable improvement in amplitude and phase skill of the ensemble mean forecast compared with individual forecasts, and the increase of significant predictability period for daily fields could be up to 50%. R. Buizza & T. N. Palmer (1998) investigate how ensemble size influence the ensemble performance using the European Centre for Medium-Range Weather Forecasts ensemble prediction system (EPS). The authors find that larger ensemble size (e.g., 32) leads to more reliable control error bound than smaller size does (e.g., 2). They consequently anticipate that higher ensemble size would continue to improve ensemble prediction performance. Another interesting finding is that, using the L^2 norm (for a vector $\mathbf{x}=[x_1, x_2, x_3, \cdots, x_n]^T$, its L^2 norm is defined as $|\mathbf{x}|=\sqrt{x_1^2+x_2^2+x_3^2+\cdots+x_n^2}$), skill-spread relationship could not be consistently promoted with higher ensemble size. It is therefore summarized that the impact of ensemble size (particularly from 8 to 16 and from 16 to 32 members) on ensemble prediction performance relies strongly on the measure utilized to evaluate, which is consistent with a more recent finding (E Hawkins et al. 2016). Another perspective was on the trade-off between improvement of model resolution and ensemble size (R. Buizza et al., 1998) of ECMWF EPS. It is unsurprising that models with higher resolution has better prediction performance due to better representation of small scale processes in the realistic ocean. The authors pointed out despite the fact that most experiments indicate a relative advantage of an enhancement in model resolution, the best performance could be obtained by combining an increase in ensemble size (from 32 to 50 members) and enhancement in model resolution (from T63L19 to T106L31). A later study (SL Mullen & R Buizza 2002) adopts probabilistic frequency difference as the measure of prediction skill to examine how ensemble size influences prediction skill of 24-h accumulated precipitation. Interestingly this study has a slightly different conclusion, that coarser-resolution, larger-member ensembles have better ability to predict rare precipitation events than higher-resolution, smaller-member ensemble. They argue that probabilistic predictions of precipitation produced by a large ensemble sizes at lower resolution provide more value to users and decision makers than that by smaller ensemble sizes at higher resolution. Adam J. Clark et al (2011) use a measure named area under the relative operating characteristic curve (ROC area) to assess the Probabilistic quantitative precipitation forecasts (PQPFs) of a storm-scale ensemble prediction system. They find that more members are needed to reach

statistically indistinguishable PQPF skill relative to the full ensemble for shorter lead time increases and coarser spatial scales. B Zhou & J Du (2010) emphasize the improvement of probabilistic forecasts induced by increasing ensemble size when the size is relatively small. But they also suggest that this promotion would decrease with the increase in ensemble size. More recently, M Leutbecher (2019) investigates how ensemble size cast its influence on the mean predictions skill (e.g., the continuous ranked probability score, the quantile score, and the Dawid–Sebastiani score converge) and what the suboptimal ensemble size secondary to infinity should be. The data used in that study is a 200-member ensemble experiment produced by the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS) model at a resolution of 29 km and a forecast range of 15 days. It is further recommended in that study, that ensemble sizes of four to eight members are anticipated to be best applicable to scientific testing.

Impact of ensemble size on climate prediction, however, is still in its early stage, due mainly to relatively later emergence of ensemble forecasts for climate. In contrast to aforementioned investigations of impact on weather prediction skill, F Sienz et al. (2016) take a different approach to examine the influence of ensemble size on decadal predictive skill. They formulate a conceptual model to simulate the decadal prediction, which are a combination of a nonlinear trend, a periodic oscillation with a 72-year periodicity and a Gaussian white noise. Their analysis reveals that small ensemble and hindcast sample sizes causes biased test performances. They further point out that to evaluate decadal prediction skill or as basis for the prediction system development, the ensemble size should not be less than 10. And much larger ensemble size is needed for regions with low signal-to-noise ratios. Another simple mathematical model is formed in Christopher A. T. Ferro et al. (2012), where the authors suggest that higher resolution should probably take priority over larger ensembles at resources of present day. However, as pointed out in the paper, whether nor not this conclusion can be applicable to general circulation models require more investigations and discussions. More recently, S. G. Yeager et al. (2018) use a bootstrapped resampling from a 40-member ensemble to investigate how Anomaly Correlation Coefficient (ACC) for boreal summer land precipitation relative to the CRU-TS4.0 dataset depends on ensemble size. It is found that a considerable improvement of ACC could be achieved by increasing ensemble members. Due to these benefits of increasing ensemble size, there is a growing interest in taking larger ensemble in climate predictions (S. G. Yeager et al. 2018). Unlike the commonly used approach of gradually increasing ensemble size, N Herger et al. (2018) take a contrary way. They reduce

the ensemble size to test whether with fewer members the ensemble could still maintain key ensemble characteristics. They show that the final size is dependent on the metrics utilized, which is in agreement with the result of R. Buizza & T. N. Palmer (1998) and E Hawkins et al. (2016). Through increasing the atmospheric resolution from 0.8 to 0.3 °, AA Scaife et al. (2019) evaluate the effect of varying atmospheric horizontal resolution on the prediction skill and fidelity of seasonal forecast of El Niño–Southern Oscillation (ENSO), North Atlantic Oscillation (NAO) and tropical cyclones. They point out that it would be a better choice for operational forecasting systems at these resolutions to increase ensemble size, due to the fact that larger ensemble size clearly promote extratropical seasonal prediction skill

In addition, in a newly noticed behavior of ensemble prediction, signal-to-noise paradox, ensemble size comes into play. The signal-to-noise paradox in climate science is a recently found counterintuitive phenomenon in ensemble forecasts (Dunstone N. et al. 2016; AA Scaife & D Smith 2018; CH O'Reilly et al. 2019; F Sévellec & SS Drijfhout 2019; W Zhang & B Kirtman 2019). This paradox refers to the phenomenon that the correlation between ensemble members is lower than that between ensemble mean and observation. This means ensemble forecasts can predict the real world better than forecast itself. A typical example is the case of NAO, and it is further pointed out that the signal-to-noise paradox of predicted sea level pressure (SLP) is principally in the North Atlantic (AA Scaife & D Smith 2018). The improvement in prediction skill of winter NAO of the real world induced by creasing ensemble size is larger than that of the model itself (AA Scaife & D Smith 2018). From this perspective, we could consider ensemble size is one of the factors that contributes to signal-to-noise paradox in climate prediction.

Despite these significant work on the effect of varying ensemble size, relatively little attention has been paid specifically to the impact of ensemble size on prediction skill and assessment of initial-value predictability of North Atlantic Ocean climate on inter-annual to multiyear time scales. Initial-value predictability is the degree to which the initial condition can influence forecasts (G Branstator et al. 2012). A commonly used metric named potential predictability (GJ Boer 2004), is referring to the predictability of internally generate variability of the climate system, i.e., due to orchestrated effect of the nonlinearity in the hydrodynamic and thermodynamic components, and linear instabilities. For an initialized ensemble, potential predictability can be used as one diagnostic of initial-value predictability. Pragmatically, potential predictability can be assessed by analyzing the spread among members (i.e., ensemble variance) of an initialized and identically forced ensemble. It is noteworthy that there is a fundamental and broadly

adopted assumption of this assessment, i.e., we assume that the model utilized perfectly represents the real world (E Hawkins et al. 2011). This is pragmatically impossible, therefore potential predictability estimated in this way is indeed the upper limit of realistic initial-value predictability.

3. Background B: Predictability of sea surface temperature and upper ocean heat content in the North Atlantic Ocean

3.1 Investigations based on observational data

North Atlantic Ocean, as mentioned above, is one of the few places where the climate is considered predictable (R Msadek et al. 2010). Pioneering work includes RT Sutton & MR Allen (1997), where the authors perform analyses of shipboard observations, revealing a substantial decadal predictability of North Atlantic sea surface temperature. This predictability is considered attributable to the advective propagation of SST anomalies as well as a regular period of 12–14 years in the propagating signals. Most recently, using gridded ocean observations, MW Buckley et al. (2019) estimate a lower bound on predictability time scales for SST and upper ocean heat content in the North Atlantic. In the subpolar gyre decorrelation time scales for both wintertime SST and UOHC are longest. Generally the decorrelation time scales for wintertime SST is analogous to that of UOHC. Approximately, 51%–73% of the regional variations in decorrelation time scales for UOHC could be explained by spatial variations in the wintertime climatological mixed layer depth. This percentile drops to only 26%–40% for wintertime SST in the extratropical North Atlantic. It is therefore suggested that to leading order, the thermal memory of the ocean determines decorrelation time scales for UOHC.

3.2 Investigations based on idealized or simplified models

H Teng et al. (2011) quantify the decadal predictability of the AMOC and associated oceanic and atmospheric fields in the Community Climate System Model, version 3 (CCSM3) with a 700-yr control run and two 40-member “perfect model” ensemble experiments. They find that the natural variability of the AMOC is predictable up to a decade, and the upper 500-m temperature in the North Atlantic is more

predictable than the AMOC by a few years. It is pointed out that subsurface and sea surface temperature (SST) anomalies propagating anticlockwise along the subpolar gyre is associated with this predictability. After three to four forecast years, North Atlantic SST remains predictable. Based on outputs from a coupled global climate model (the third climate configuration of the Met Office Unified Model, HadCM3), E Hawkins & R S utton (2009) investigate the predictability of three-dimensional Atlantic Ocean anomalies on decadal time scale. It is revealed that the evolution of temperature and salinity in the Atlantic, and the AMOC, could be properly captured by a white noise-forced linear dynamical system. For several decades, this simplified linear model even possesses higher prediction skill than other reference forecasts. L Zanna (2012) using Linear inverse modeling (LIM) and observed Atlantic SST anomalies between latitudes 20 ℃S and 66 ℃N from 1870 to 2009 to construct a linear stochastic model. The assessed prediction skill is O(3–5 yr) and is hugely reduced after that, especially in the subpolar gyre. It is further pointed out that forecast kill of annual average SST anomalies is due to four damped eigenmodes, which is associated with the optimal growth of SST error, and provides information for interannual variability and predictability of Atlantic SSTs.

3.3 Investigations based on ocean-forced or coupled models

Taking another approach, SM Griffies & K Bryan (1997) use a global coupled ocean-atmosphere model, showing that the North Atlantic climate could be predictable longer than decades. It is hence suggested that predictions of variations of the dominant multidecadal sea surface temperature is possible in the North Atlantic once there exists an adequate and sustainable system for monitoring the Atlantic Ocean. Meanwhile, SM Griffies & K Bryan (1997) use the GFDL coupled ocean-atmosphere climate model to quantify the predictability of North Atlantic climate on multi-decadal time scales. Their results reveal three physical mechanisms, namely oceanic deep convection in the subpolar region, large-scale dynamics of the thermohaline circulation and periodic anomalous fresh water transport advecting southward from the polar regions in the East Greenland Current, on which the North Atlantic predictability depends. Predictability seems to get lost fastest in regions characterized by strong convection in the northern North Atlantic. R Msadek et al. (2010) examines decadal predictability of AMOCA and its signatures, i.e., sea surface height, subsurface temperature, and upper ocean heat content anomalies, in the GFDL CM2.1 climate model. It is revealed that the AMOC could be predictable up to 20 years, and the predictability

of these signatures is analogous to that of the AMOC. LF Borchert et al. (2018) highlight the strong dependence of the interannual-to-decadal predictability of North Atlantic SSTs on the strength of subpolar ocean heat transport at the beginning of a forecasting. Specifically, they analyze three-member ensemble of initialized simulations (1901-2010) in the MPI-ESM-LR, unravelling the influences of Atlantic northward ocean heat transport (OHT) at 50 °N on sea surface temperature variability in the North Atlantic region for several years. Warm SST anomalies in the North Atlantic and cold SST anomalies in the Gulf Stream region emerge three to ten years after strong ocean heat transport phases at 50 °N. Interannual-to-decadal SST predictability of annually initialized hindcasts is linked to this SST pattern which originates from persistent upper-ocean heat content anomalies as a consequence of southward-propagating OHT anomalies in the North Atlantic. G Branstator & H Teng (2014) quantify and compare the predictability of Atlantic meridional overturning circulation (AMOC) and that of the upper-500-m heat content in the North Atlantic using control simulations in nine different climate models. The time range at which initialization cast a positive impact on the annual-mean AMOC prediction is about a decade, though this range varies substantially among models. Results show that for the averaged fields, AMOC possesses higher predictability than heat content. It is suggested that the AMOC could probably cast a positive impact on the predictability of ocean heat content. F Sévellec & AV Fedorov (2017) examine the excitation of decadal variability and predictability of the North Atlantic Oceanic climate state in a realistic ocean general circulation model using initial linear optimal perturbations (LOPs) method. The instantaneous and annual-mean values of meridional volume and heat transports are shown to be less predictable than surface-averaged sea surface temperature and volume-averaged ocean heat content. On interannual time scales the initial ocean temperature errors could result in prediction errors of ~0.1-K of North Atlantic sea surface temperature. G Branstator & H Teng (2010) quantify the initial-value decadal predictability of upper-300-m ocean temperature in the Community Climate System Model, version 3 (CCSM3), with a few 40-member climate change scenario simulations. It is found that the impact of initial states sustains no longer than a decade, and the ensemble mean signals contribute more to initial-value predictability rather than the distribution about the mean does. Using hindcast ensemble experiments, B Tiedje et al. (2012) quantify the interannual potential predictability of the meridional heat transport (MHT) in the North Atlantic Ocean. The prognostic potential predictability (PPP) of the MHT and the AMOC varies with latitude. The gyre component of the MHT influences the PPP structure of the MHT in the subpolar region, on contrast the overturning component of the MHT

shapes the PPP structure of the MHT in the subtropical gyre. MW Buckley et al. (2014) quantify the upper-ocean heat budget (1992–2010) in the North Atlantic on monthly to interannual time scales using the Estimating the Circulation and Climate of the Ocean (ECCO) project product. The maximum climatological mixed layer depth (H) are determined differently in different regions: in the interior of subtropical gyre, H could be largely explained by local forcing on all time scales; low-frequency H anomalies along the Gulf Stream are forced by geostrophic convergences and damped by air–sea heat fluxes; it is the diffusion and bolus transports that contributes the most to H variability. G Branstator et al. (2012) examine the initial-value predictability (i.e., the problem of to what degree the initial state impact forecasts) of North Pacific and North Atlantic in six atmosphere–ocean general circulation models. In both basins, the impact of the initial conditions could last for approximately a decade. The predictability limit and high predictability locations vary substantially among models. Predictability is found to be determined principally by the mean of prediction distributions rather than the spread about the mean, which is consistent with earlier study (G Branstator & H TengWhen 2010). The horizontal propagation is also highlighted in playing a sbustantial role in the evolution of these signals, therefore make itself a key factor in tell the difference of the predictability in various models. L Trenary & T DelSole (2016) examine the links between the predicted Atlantic multidecadal oscillation (AMO) and AMOC across various climate models. It is found that across all climate models, there exist three overturning patterns substantially coupled to the AMO on interannual time scales. The AMO predictability could be prolonged by 2–9 years if these structures is included in an autoregressive model extends, relative to when there is no such structures. L Zanna et al. (2012) quantify the predictability using singular vectors of the meridional overturning circulation (MOC) and upper‐ocean temperatures in an idealized ocean general circulation model (GCM). The time after which the MOC and upper‐ocean temperature anomalies experience maximum growth are 18.5 and 13 years, respectively. In this study, it is suggested that this growth could be partly explained by the westward propagation of upper‐ocean anomalies riding on the mean flow. The ocean dynamics is revealed to be working on the substantial growth of the anomalies *via* non-orthogonal eigenmodes in the non-normal system. It is also found that there is only weak correlation between optimal MOC and upper‐ocean temperature, which challenges the use of SST observations to monitor MOC variability. This results highlight the difference between the predictability of the MOC and upper-ocean heat content.

4. Background C: Variability and initial-value predictability of subsurface temperature in the North Atlantic Ocean on varying time scales

The Atlantic meridional overturning circulation (AMOC) plays a pivotal role in the transport of the heat and water masses in the Atlantic Ocean. Furthermore, the AMOC, *via* air-sea interaction (e.g., exchanges of heat, freshwater and momentum fluxes), profoundly impacts the atmosphere circulation over it, and consequently modulating the climate both locally and remotely (S Mahajan et al. 2013; LC Jackson et al. 2015; Rong Zhang et al. 2019; W Liu et al. 2020). A number of previous studies reveal that, subsurface temperature not only is a fingerprint of AMOC, but also can be utilized to skillfully reconstruct the AMOC variation (J Mignot et al. 2007; Rong Zhang 2008; R Msadek et al. 2010; Mahajan et al 2011). J Mignot et al. (2007) investigate the response of Atlantic subsurface temperature to a collapse and recovery of the AMOC using an intermediate-complexity climate model CLIMBER-3α. Within the first decade after the freshwater perturbation, subsurface temperature would either be cooled or warmed, depending on whether intermediate ventilation remains active. Furthermore, as long as the anomalous freshwater forcing is removed, the sign of subsurface temperature anomaly could be a good implication for the recovery of the AMOC. For example, on centennial time scales, the AMOC recovery from the cold case is much quicker. R Zhang (2008) analyze both altimeter data and instrumental subsurface ocean temperature data in the North Atlantic, showing that in both data there exists opposite signs between the subpolar region and the Gulf Stream path. This dipole pattern is also detected in a 1000‐year coupled ocean‐atmosphere model run, and is deemed a fingerprint of AMOC variability. The result highlight the ongoing subsurface temperature measurements, for it could be used for monitoring future AMOC variations. Later on, motivated by these associations of AMOC and its fingerprints (SSH, SST and Tsub), S Mahajan et al. (2011) develop a skillful linear statistical forecast model of observed fingerprints of AMOC. R Msadek et al. (2010) analyze output of the GFDL CM2.1 climate model to examine the potential predictability of the AMOC. It is found that potential predictability of modeled surface and subsurface signatures of AMOC variations, such as sea surface height,

subsurface temperature, and upper ocean heat content anomalies, is analogous to that of the AMOC. This result highlights a new promising approach for assessing the predictability of AMOC variations, considering that observational records for these variables are longer than for the AMOC.

EC Brady and BL Otto-Bliesner (2011) found that subsurface temperature in the subpolar region increases when they artificially add freshwater to the high latitude North Atlantic and slightly drops when freshwater is imposed to the Gulf of Mexico. In contrast, the subsurface temperature in the Southern Hemisphere rises in all cases.

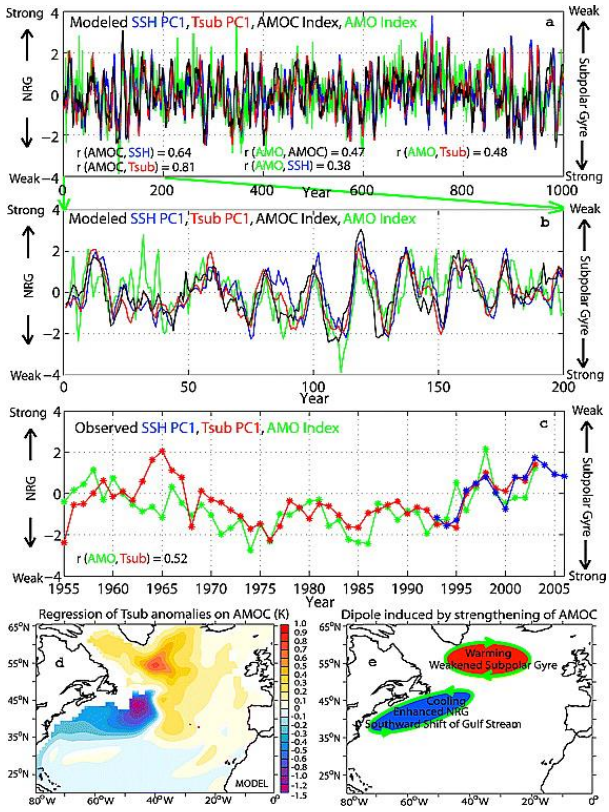

Figure 1 Modeled and observed anomalies. (a) Modeled SSH PC1 (blue), Tsub PC1 (red), AMOC index (black), and AMO Index (green) from GFDL CM2.1 1000 - year control integration (normalized). The standard deviation (SD) of modeled AMOC index is 1.8Sv. The SD of modeled AMO index is 0.16K. (b) Same as Figure 1a, but enlarged for the first 200 years. (c) Observed Tsub PC1 (red) and AMO Index (green) for 1955–2003, and observed SSH PC1 (blue) for 1993–2006 (normalized), all anomalies are relative to the climatology mean of 1993 - 2003. The long - term trend of the instrumental subsurface data is removed. The SD of observed AMO index is 0.18K. Positive anomalies in Figures 1a - 1c correspond to a weakening in the subpolar gyre and a strengthening in the NRG. (d) Regression of Tsub anomalies (K) on the AMOC Index from GFDL CM2.1 1000 - year control integration. The regression corresponds to 1 SD of the AMOC Index (1.8Sv). (e) Schematic diagram of the dipole pattern induced by the strengthening of the AMOC. From *R Zhang 2008*.

In addition, evidences are found that multiyear subsurface temperature variability in the North Atlantic Ocean (NAO) is associated with the change of the North Atlantic Oscillation (NAO), revealing the a moderate impact of surface wind on subsurface thermal field (RL Molinari et al. 1997). The North Atlantic Oscillation (NAO) represents a redistribution of atmosphere between the Arctic and the subtropical Atlantic. The NAO index is defined as the Sea-Level Pressure (SLP) difference between the Subtropical (Azores) High and the Subpolar Low (JW Hurrell et al. 2003). Likewise, a subtle subsurface-surface relationship exists on decadal time scales. It is found that decadal variations of North Atlantic subsurface temperature depends on surface forcing and upper-ocean mixing rather than gyre circulation (RW Houghton 1996).

Influences of North Atlantic subsurface temperature on basin scale variability extends to multi-decadal time scale. Rong Zhang (2008) conduct a combined analysis of instrumental records and GFDL CM2.1 1000‐year control simulation, showing that AMO Index has significant correlations with Tsub PC1 for the period 1955‐2003 in both modeling results and observations. More interesting, it is revealed that the AMO Index is in phase with Tsub variations in subpolar region and anti‐phase with Tsub variations close to the Gulf Stream path (Figure 1). Later on, an analysis of a 500-year simulation of version 2.1 of the Geophysical Fluid Dynamics Laboratory's Coupled Model (CM2.1) discovers that, the 20–30-yr variability in the North Atlantic is predominantly characterized by the westward propagation of subsurface temperature anomalies (LM Frankcombe et al. 2010). Furthermore, using an ocean general circulation model called Océan Parallélisé (OPA), Sévellec et al. (2013) reveal a damped mode of the AMOC, which is dominated by the westward propagating upper ocean temperature (upper 300m) anomaly and compensated by the salinity. Interestingly, there is also a connection between North Atlantic subsurface temperature and well-known Atlantic Mutli-decadal Oscillation (AMO, also known as Atlantic Multi-decadal Variability, AMV). AMO has huge impacts on regional climate, defined as the linear detrended mean sea surface temperature in 0-60 °N North Atlantic Ocean. Observational evidence of links between subsurface thermohaline structure and AMO are also found, though the duration and space coverage are limited (Polyakov et al. 2005, 2010).

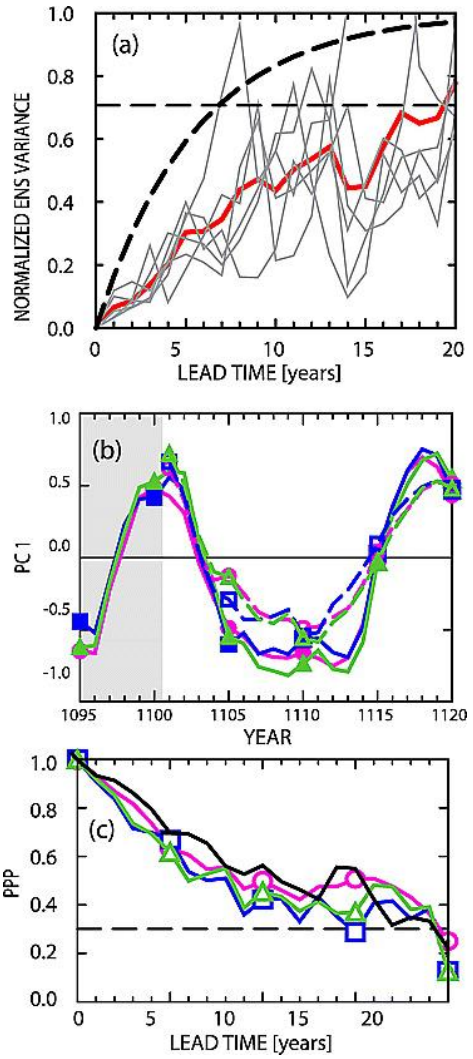

Figure 2 (a) Normalized ensemble variance of the AMOC averaged over the 60 experiments (red line) and for the 10 - member ensembles (grey lines), along with the red noise null hypothesis (dashed black line). (b) Predictability of AMOC fingerprints initialized around year 1101: the North Atlantic PC1 of the 400 m subsurface temperature Tsub (magenta circle), the upper 700 m OHC (green triangle), and the SSH (blue square). Plain lines are the control run values, dashed lines are the ensemble mean. (c) PPP of the AMOC index (black) and its fingerprints. The color scheme is as in Figure 1b. The dashed horizontal line indicates the 95% significance level. From *R Msadek et al. (2010)*

The variability of North Atlantic subsurface temperature has been investigated using both numerical models and valuable yet scarce observational data. Back in 1996 RW Houghton using instrumental data examine the subsurface quasi-decadal fluctuations in the North Atlantic. It is revealed that the quasi-decadal fluctuation seems more likely to be forced by surface heat and wind variations. Furthermore, it is pointed out that the cold phase of North Atlantic quasi-decadal fluctuations in the western sub-polar gyre is consistent with weaker mixing while in the subtropical

region it corresponds to larger vertical mixing and heat flux. Later on, RL Molinari et al. (1997) analyze subsurface temperature data obtained during 1966 and 1995 in the western North Atlantic Ocean, providing an evidence for decadal variability in the upper 400 m. Periods of positive subsurface temperature anomaly correspond to periods of positive NAO index; periods of negative subsurface temperature anomaly, on the contrast, are consistent with periods of negative NAO index. This finding extends previously found links between the NAO and western Atlantic SST on decadal timescales down to 400 m. More recently, A Ruiz-Barradas et al. (2018) reveal that the North Atlantic subsurface has a warming trend from mid-1980s to the mid-2000s while possesses a cooling trend since the mid-2000s. It is further pointed out that the spatial pattern of decadal trends of North Atlantic upper-ocean heat content (5–657 m) before and after 2006 indicates a link with variability of the Gulf Stream–Subpolar Gyre system. Specifically, the variability of the Gulf Stream leads that of upper-ocean heat content within the Subpolar gyre by approximate 13 years. Though we are aware of these rich variations as well as climatic significance of surface temperature variations, we are still short of long-term global observations of subsurface temperature. Some mapping methods are proposed to reconstruct global or basin-wide subsurface temperature (or ocean heat content) filed. In order to make a complete map of the ocean subsurface temperature (0-700m) of 1940 to 2014, L Cheng & J Zhu (2016) apply a approach named ensemble optimal interpolation with a dynamic ensemble (EnOI-DE) to the outputs from multimodel ensemble of phase 5 of the Coupled Model Intercomparison Project (CMIP5) historical and representative concentration pathway 4.5 simulations. This advanced mapping strategy provides an improved estimation of the long-term historical ocean heat content variations since 1940, which is of great significance to the evaluation of Earth's energy budget. X Wu et al. (2012) use a self-organizing map (SOM) neural network trained with Argo gridded datasets (including anomalies of SST, SSS, SSH) to estimate the North Atlantic subsurface temperature anomaly (STA). IT shown that correlation coefficient between the estimation and in situ data for the temperature of upper 700 m is over 0.8 along the Gulf Stream path.

Given the variability of North Atlantic subsurface temperature on various time scales, as well as its broad climatic impacts, it would be interesting to explore its predictability. H Teng et al. (2011) use two 40-member ensemble and a 700-yr control run to analyze the predictability of AMOC and associated 500-m temperature in the North Atlantic Ocean. They find that natural variability of the

AMOC is predictable for longer than a decade. This predictability is related to anticlockwise propagating subsurface and SST anomalies along the subpolar gyre, which dominate 10 years after the AMOC anomalies peaks. Likewise, R. Msadek et al. (2010) show that the potential predictability of subsurface temperature and upper-ocean heat content anomalies modeled by GFDL CM2.1 climate model bears a resemblance to that of the AMOC (Figure 2). The case in the North Pacific Ocean, however, is different. H Teng & G Branstator (2011) perform an analysis to three 40-member ensemble experiments and a 700-year control run to examine the initial value predictability in the North Pacific produced by Community Climate System Model version 3 (CCSM3). They reveal that an eastward propagating mode is generated by the combination of the first two empirical orthogonal functions (EOFs) of subsurface temperature variability. They point out that the predictability of EOF1 is less than 6 years. Despite these significant work, little is known about the initial-value predictability of subsurface temperature in the North Atlantic Ocean on multiyear time scale. This motivates our investigation in this thesis.

5. Structure of this thesis

Briefly speaking, the objectives of this thesis are: to assess the impact of ensemble size on the evaluation of initial value predictability of slow, internal oceanic variations (chapter 2); to illustrate the evolution process of subsurface temperature uncertainty arising from initial perturbations and to unravel sources and sinks in the development of these uncertainties (chapter 3).

There are 3 chapters following this chapter in this thesis. They are organized as follows:

Chapter 2: Impact of ensemble size on prediction skill and the assessment of initial-value predictability of the North Atlantic Oceanic climate: perspectives from a conceptual model and two GCMs. Ensemble size among many other factors is a key considered issues in present climate prediction. By gradually enlarging the ensemble size of a conceptual SST model, we have a flavor of the impact of ensemble size: larger ensemble size will ‘lead’ the ensemble mean to the theoretical resolution, and the uncertainty is reduced in this process. We further perform an analysis to the output of a 40-member ensemble from a low-resolution climate model and the output of an 11-

member ensemble from an eddy-permitting model. A diagnostic, Prognostic Potential Predictability (PPP), is used to reveal prediction uncertainty and probability distribution, respectively. Results from the three models coherently indicate that larger ensemble size will lead the mean PDF to 'an asymptotic value', and can reduce the PPP uncertainties among different sub-ensembles (i.e., uncertainty in different realizations of sub-ensembles with a same ensemble size).

Chapter 3: Understanding the growth of subsurface temperature prediction uncertainty originating from initial perturbations in the North Atlantic Ocean: the effects of mean state and meso-scale eddies. In this chapter the spatial evolution of prediction uncertainty of subsurface temperature, in the form of a decreasing Prognostic Potential Predictability (PPP), is revealed and discussed. The evolution is characterized by a clockwise propagation in the subtropical North Atlantic and an opposite propagation in the sub-polar gyre. On the other hand, downward propagation of temperature ensemble variance, which indicates prediction uncertainty, mainly takes place in the North Atlantic subpolar gyre (Labrador Sea and Irminger Sea). Furthermore, this uncertainty develops downward principally along continental shelves (e.g., Greenland shelf), which is in accordance with the strong subduction in these regions. In addition, to explore the influence of model spatial resolution on the robustness of abovementioned findings, we analyze the output of an intermediate-resolution, eddy-permitting model, NEMO in ORCA25 grid configuration. Each term plays a qualitatively same role as in the coarse resolution ensemble.

Chapter 4: Conclusions and discussions. A summary of the thesis is made in this chapter. Some caveats of the study are discussed. In addition, I outline possible future work beyond this thesis.

Chapter 2 Impact of ensemble size on prediction skill and the assessment of initial-value predictability of the North Atlantic Oceanic climate: perspectives from a conceptual model and two GCMs

Key points:

• The impact of ensemble size on prediction skill and the assessment of initial-value predictability of both conceptual and realistic models is investigated, representing predictability problems under varying external forcing and arising from internal nonlinear dynamics as a result of initial perturbation, respectively.

• Increasing ensemble size reduces uncertainty of main prediction properties, such as Residual Mean Standard Deviation (RMSD) leads to a closer distance to theoretical result from a conceptual SST model.

• Similar inverse correlathion exists between the pair of ensemble size and prognostic potential predictability, and the pair of ensemble size and PPP uncertainty. Prognostic potential predictability doesn't change significantly the contribution of individual ensemble member significantly drops as ensemble size is over 15-20 for IPSL-CM5A-LR, yet this critical size is around 7 for the eddy-permitting ocean model.

• Increase of ensemble size leads to a better resemblance of PDF of the ensemble and that of the climatology.

Abstract: In this chapter I investigate how varying ensemble sizes influence prediction skill and assessment of potential predictability of North Atlantic climate. Ensemble methods is a widely used approach to make climate predictions, whereas the impact of ensemble size on prediction skill and assessment of climate predictability remains an open question. Here we use several diagnostics to measure the effect of different ensemble sizes. Analyses of a white-noise driven, conceptual sea surface temperature (SST) model indicate that growth of ensemble size leads ensemble means to approaching to theoretical resolution and a reduction of uncertainty among ensemble means. The

results from an initially perturbed, low-resolution general circulation model (GCM) IPSL-CM5A-LR show that by increasing the ensemble size the prognostic potential predictability of Atlantic meridional overturning circulation (AMOC), sea surface temperature (SST) and subsurface temperature (Tsub) smoothly declines, however the descending rate slows down substantially as the ensemble number is larger than a particular value. On the other hand, the uncertainty of the ensemble means of different ensembles with the same ensemble size reduces as the ensemble size grows, though there is no guarantee that larger ensemble size leads to more accurate prediction. In addition, the probability density function evolves closer to the climatology with the growth of ensemble size. Analysis of output from an eddy-permitting model show similar results, differing only in 'critical size'. Both types of models reveals that the individual contribution of ensemble member drops substantially when ensemble size is sufficiently large. The findings imply that the extra increase of ensemble members should be dealt cautiously in the design of future climate prediction system for North Atlantic sector if the size is larger than 15~20 for low resolution coupled models, or larger than ~7 for eddy-permitting ocean models, efforts in alternative directions, such as improving model resolution, might be an option more worth considering.

1. Introduction

Climate prediction using numerical models inevitably includes uncertainties, due principally to errors in initial states (RA Pielke Sr 1998), boundary conditions and physical representations of the numerical model (Weaver, Andrew J., and E. S. Sarachik. 1991; M Collins 2002; Collins, Matthew, and Myles R. Allen. 2002). Among these factors the initial state errors, and the consequences of which, have attracted increasing interest. Forecasting trajectories starting from slightly different initial states diverge to a certain degree. To investigate the possibilities of the forecast results, ensemble experiments are usually performed as a powerful approach. In the phase space, these trajectories of forecasts originate from close phase points, projecting to a much broader range later on (K Fraedrich 1987; C Nicolis 1990; L Gerrit & J Schneider 1999). In the field of decadal climate prediction, how to determine ensemble size among many others remains one key challenges (Florian Sévellec and Bablu Sinha 2018). However, the choice of ensemble size does impact the properties

of the ensemble forecast, e.g., ensemble mean, ensemble members etc. (F Sienz et al. 2016; H Nadja et al. 2018; Martin Leutbecher 2018). A larger ensemble size is often competing with a higher model resolution, given the fact that both are beneficial to prediction performance while computation consuming. Ferro et al. (2012) proposed a theoretical approach, and demonstrated that within this simple framework optimal resolution smoothly decreases as the ensemble size increases. Though in this study the only measurement of forecasting performance used is root mean squared error, this study is informative for the problem of ensemble size with different initial conditions.

A key assumption of underlying our ensemble experiments is the 'perfect model framework', i.e., we assume the model perfectly represents the realistic climate system (R De Elia et al. 2002). In an initialized ensemble, the forcing for all ensemble members is identical, the only source of ensemble variance comes from initial difference; i.e., potential predictability is indeed a measure of the predictability of internal variability.

In this study, we narrow our focus on the impact of ensemble size on prediction skill and assessment of multiyear initial-value climate predictability of the North Atlantic Ocean state, which to the authors' knowledge has not been broadly studied to date. Evidences are found that the Atlantic meridional overturning circulation impacts the climate of surrounding landmasses (Carton, James A., et al. 2014; Buckley, Martha W., and John Marshall. 2016; Yan et al. 2017; Weijer, Wilbert, et al. 2019), or even of remote basins (Timmermann, Axel, et al. 2007; Liping Zhang et al. 2017; Williamson, Mark S., et al. 2018), on various time scales. Sea surface temperature is documented to have an intimate connection to AMOC (Wei Cheng et al. 2013; Frankignoul Claude et al. 2013; Muir, L. C., and A. V. Fedorov. 2015). The subsurface temperature, on the other hand, are found to be a key finger print of AMOC (J Mignot et al. 2007; Rong Zhang 2007; Rong Zhang 2008; AP Parker et al. 2015). For these reasons, we analyze impact of ensemble size on prediction skill and the assessment of initial-value predictability of AMOC, SST and Tsub in this study.

The scientific questions we are answering in this study are:

(1) Whether or not can ensemble size change prediction skill and assessment of initial-value predictability (see methods) of SST, Tsub, and AMOC, respectively? If so, how?

(2) What will the consistency and discrepancy be like among the impacts of ensemble size on prediction skill and assessment of initial-value predictability of SST, Tsub, and AMOC?

This chapter is organized as follows: In section 2, I describe a conceptual SST model, a low-

resolution coupled model, and an eddy-permitting ocean model. Analysis of the results is given in section 3. Finally I draw the conclusions and have a brief discussion of the caveats of this study and what remains to be done in the future in section 4.

2. Models and data

In this chapter we employ a simplified conceptual SST model as well as two fully coupled GCMs with different resolutions in this study. The reason of this mixed use of models is that hopefully in theoretical model we will have a flavor of what kind of ensemble size-predictability assessment relationship exists in the climate system, and an examination of this relevance in two GCMs may testify the robustness.

2.1 A theoretical model of sea surface temperature

The first model we use is a conceptual SST anomaly model in which the atmosphere temperature anomaly is taken as a zero-mean white noise and the SST anomaly is a response to SAT anomaly. This is an adapted version of Hasselmann 1976 model, where low-frequency SST variation at mid-latitude is deemed a red-noise-like response to with-noise-like atmospheric forcing (K Hasselmann 1976). Here I take the same parameters as in Florian Sévellec & Bablu Sinha (2018). The model reads:

$$\frac{\mathrm{d}\,\mathrm{SST}_t}{\mathrm{dt}} = \left(\mathrm{SAT}_t - \mathrm{SST}_t\right)/\lambda \tag{1}$$

Where SST refers to sea surface temperature anomaly at time step t, SAT the atmospheric temperature anomaly at time step t, λ the oceanic adjustment time (here λ is 10 years, same as in Florian Sévellec & Bablu Sinha, 2018). The SAT is assumed to be a zero-mean white noise with a variance per time of 0.5 K^2 yr^{-1}. This model simulates the one-way process between the SST and SAT (SAT forces SST), consequently representing a dynamical system forced by external white noise.

First of all, we solve the equation numerically by the SAT white noise SAT. For each fixed ensemble size, theoretically we can have infinite resolution of SST as long as we use various SAT. To make the computation feasible as well as statistical properties stable, we run the model N times, and then analyse the consequent N ensemble mean. Particularly, we are interested in the root mean square between the N ensemble means and the mean of the N ensemble. The root mean square difference over the 10,000-year simulation between the theoretical solution of the ensemble standard deviation evolution and the one computed N ensembles reflect the prediction skill. The ensemble size N is growing from 5 to 500, with an equal interval of 5.

It is not surprising that the spread of different ensemble means, as shown in the upper panel of Figure 1, drops substantially as ensemble size is changed from 5 (gray lines) to 100 (cyan lines). The change of spreads between the situation of N=100 and N=500 (red lines), however, is not as obvious. This can be confirmed from the lower panel, where the Root Mean Square Deviation (RMSD) drops substantially as N<=100, while reduces insignificantly as N>=100. Specifically, the RMSD drop between RMSD for N=5 and RMSD for N=100 is 0.00156, much larger than that between RMSD for N=100 and RMSD for N=500, being 0.00025. This indicates the individual contribution to the reduction of RMSD is decreasing dramatically when ensemble size is larger than 100.

The middle panel displays that mean of ensemble means oscillates around zero, meanwhile the amplitude of this oscillation is generally reduced as the ensemble size grows (the reduction is

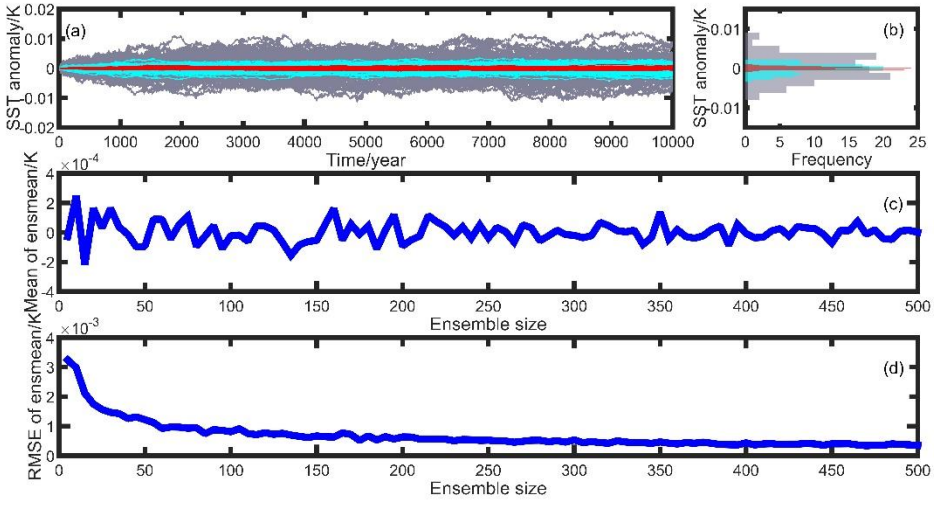

Figure 1. Sea surface temperature (SST) anomaly from adapted Hasselmann 1976 model (a). Gray lines represent ensemble means in repeated realizations with an ensemble size of 5; cyan lines are ensemble means in repeated realizations with an ensemble size of 100; red lines stand for ensemble means in repeated realizations with an ensemble size of 500. For each ensemble size, we ran the ensemble experiment for 100 times, consequently having 100 ensemble means (see details in the context). Probability density function of simulations with different ensemble size (b). Mean of the 100 ensemble means as a function of ensemble size (c). Root mean square deviations (RMSDs) of the 100 ensemble means for various ensemble size (d). Note that the results in (b), (c) and (d) are calculated using the end year of the simulation, i.e. the data of year 10000. This model represents a situation where the ocean is forced by the atmosphere, i.e., a system driven by external forcing. The results, as a consequence, display the impact of ensemble size on the prediction property of an externally-forced system. In (b)(c)(d) the data of year 10000 is used. Unit of SST anomaly: K. Unit of SST anomaly frequency: year. Unit of ensemble mean SST anomaly: K.

smaller when size is larger than 100, though). Thus it reveals another aspect of the impacts of ensemble size on one measure of prediction skill, namely the approaching of mean of ensemble means to theoretical resolution. (Obviously, in this model, the expectation of SST is zero because of the zero-mean property of the external forcing. See appendix B for more details.).

This model represents a system forced only by external field, i.e., there's no other internal dynamics in addition to the persistence of SST. To examine the robustness of findings drawn from this model, we perform an analysis to outputs of a low-resolution climate model and a forced-ocean eddy-resolving model.

2.2 A low-resolution fully coupled GCM: IPSL-CM5A-LR

The first GCM used in this study is IPSL-CM5A-LR fully coupled climate model. As the low resolution version of the IPSL-CM5A Earth system model, its atmospheric component is LMD5A and has a horizontal resolution of 1.875 °x3.75 °, and 39 vertical levels for the atmosphere component LMD5A. The ocean component, namely NEMOv3.2, has a nominal horizontal resolution of 2 °(the resolution is finer near-equator region), and 31 vertical levels for the ocean (this corresponds to the ORCA2 configuration). In addition, a sea ice model LIM2 and a biogeochemistry model PISCES are also components of the ocean model as well. The coupler between LMD5A and NEMOv3.2 is OASIS3 (Dufresne, J-L., et al. 2013, readers are also referred to visit http://icmc.ipsl.fr/). The IPSL-

CM5A-LR has been widely used for studies of climate predictability in terms of initial condition uncertainties, model performance assessment and variability of the Atlantic Ocean etc. (Mignot et al. 2013; P Ortega et al. 2015; Germe et al. 2017).

In our ensemble experiments, there are 40 members, each with a duration of 20 years spanning from 1st January 2056 to 31st December 2075 (the year only refers to date along the control simulation). In addition, a 1000-year control run is utilized as well. The 1000-year control simulation is produced in this way: the model adopts pre-industrial boundary conditions of greenhouses gases and aerosols concentrations after spin-up (Dufresne et al. 2012), and is then integrated for more than 1000 years. These 1000 years are from model year 1800 to model yeare 2799. The ensemble experiments are generated by imposing 40 perturbations to the model initial condition, which is subsequently run for 20 year (from 1st January 2056 to 31st December 2075) to get the entire 40 ensemble members (see Table 1 for details of the ensemble design, see Figure 2 for the schematic of the piControl run and the ensemble experiments). Note that 4 different methods have been used to generate 4 types of initial perturbation, therefore we have 4 10-member sub-ensembles which are differently initialized. Despite the difference in perturbation types, Germe et al. (2017) demonstrated that no significant difference of ensemble statistics, such as ensemble mean and ensemble spread of SST and AMOC, arises from the 4 different options of perturbations. Therefore we consider the 40 ensemble as an entire set in our study.

Readers should also keep in mind the fundamental assumption of our analysis with IPSL-CM5A-LR, is that the 40 member capture all possible solutions of the model. This assumption is surely questionable, however it is justifiable to adopt it for this purpose: to examine the semi-quantitative impact of varying ensemble size on the assessment of initial-value predictability.

Table 2.1 Scheme of the four sub-ensembles (Adapted from Germe et al. 2017)

Sub-ensemble name	Initial perturbation	Ensemble size	Initial state	Time span
ATM	Coupler SST	10	1st January 2056	20 yr
3D	3D oceanic temperature	10	1st January 2056	20 yr
3D10	3D oceanic temperature with tenfold magnitude	10	1st January 2056	20 yr

DEEP	Deep 3D oceanic temperature (below 2000 m)	10	1st January 2056	20 yr

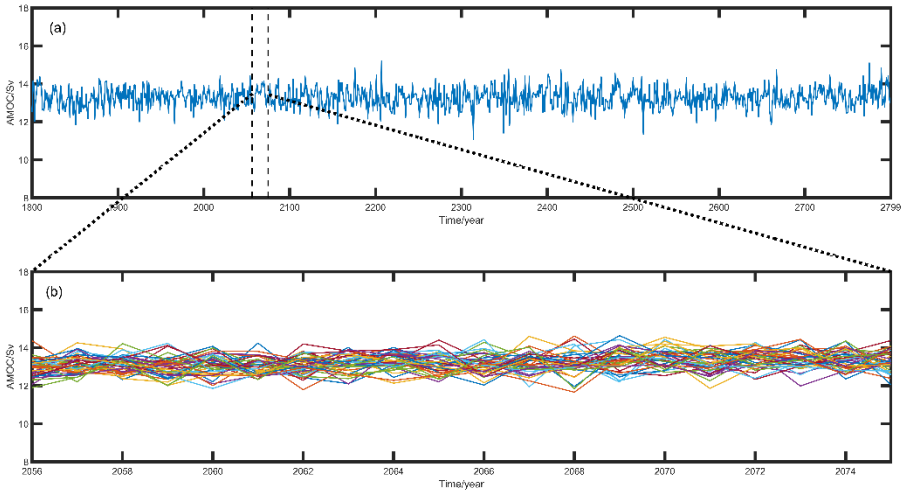

Figure 2. Schematic of the ensemble experiment design. Blue and colored curves represent 'ocean states' (the stream function indices, defined as the maximum zonally averaged stream functions between 50-700 m). Upper panel is a general view of the control run and the 40-member ensemble. Lower panel is the zoom-in of the 40-member ensemble. A statistically stable 1000-year control run is viewed as a basic state. We choose a year of the 1000 years as the initial basic state of the 40-member ensemble experiments. Forty slightly different perturbations (see Table 1 for the details of the perturbations) are imposed on the identical initial basic state, and the model is then time integrated for 20 years. As a consequence, the ocean states gradually spread from nearly identical to a widely different 20 years later. In this chapter we choose the control run as our reference, and explore the impact of ensemble size on the statistical properties of the ensemble simulations. Unit: Sv.

As mentioned above, because of their profound influence on regional and global climate, we analyze the statistical properties of Sea Surface Temperature (SST), subsurface Temperature (Tsub) in the North Atlantic Ocean and the Atlantic Meridional Overturning Circulation (AMOC). Here the subsurface depth is referring to the depth of 400 m (same to the definition in several other studies, e.g., Rong Zhang 2007, Rong Zhang 2008) and the AMOC is defined as the maximum zonally averaged overturning stream function within 50-70 ℃N and 500-3000 m, similar to that in D Swingedouw et al. (2013). The regionally averaged SST and Tsub is chosen to be consistent with the AMO (Atlantic Multi-decadal Variability) region (0-60 ℃N in the North Atlantic Ocean). The statistics we are interested in are prognostic potential predictability (PPP), which measures the spread of predicted results, and probability density function, which reveals the distribution of the

predicted trajectories and is discussed in more details above.

2.3 An eddy-permitting (a quarter degree) ocean model

To investigate whether the impact of ensemble size is dependent on model resolution, an eddy-permitting ocean model in a forced configuration is used as well. This dataset is the output of an 11-member ensemble produced by NEMO v3.4 (Nucleus for European Modelling of the Ocean, version 3.4). The details of the ocean model configuration are as follows: Implemented on the tripolar ORCA25 horizontal grid (Madec and Imbard 1996), the ocean model NEMO has a nominal horizontal resolution of 1/4 °and 75 vertical levels. The layer thickness varies from 1 m at the surface to 204 m near the ocean bottom.

There are 11 members in the ensemble, of which each is integrated from slightly different initial states. We have a relatively small ensemble size because of computational cost, yet analysis below will show that an evident asymptotic behavior has emerged even with such a small size (at least for the measures and timescales considered here). The integration period is correspond to a hindcast from 8th January 2001 to 15th June 2005. Since no long control run is available, the traditional definition of PPP (see section 3 for details) is not applicable to this data. To overcome this problem we use an arbitrary member as climatology, in this way the PPP can be calculated (using different as climatology hardly changes our results, not shown).

3. Methods, results and analyses

3.1 Methods

As mentioned above, the diagnostic we use to measure the initial-value predictability is prognostic potential predictability (PPP, P Holger et al. 2004), which is defined as:

$$PPP = 1 - \frac{\delta_{ens}^2}{\delta_{\infty}^2} \tag{2}$$

Where δ_{ens}^2 is the variance between ensemble members, δ_{∞}^2 the time variance of control simulation. *PPP* measures the spread of the ensemble, that is, a lower *PPP* represents larger spread of the ensemble, indicating broader distribution of prediction results caused by internal dynamics of the system as a result of various initial perturbations. If ensemble variance is comparable to the climatology variance, initial-value predictability is lost. This corresponds to the case when PPP is close to zero. On the contrary, if ensemble variance is much smaller than climatology variance, initial-value predictability sustains well. This corresponds to the case when PPP is close to one.

3.2 Results and analyses

3.2.1 Results and analyses from IPSL-CM5A-LR

Atlantic Multi-decadal Variability (AMV) is a pronounced climate phenomenon in the North Atlantic Ocean, it could be monitored through the detrended sea-surface temperature within 0-60 ℕ. Evidences from various climate models are found that AMV significantly impact northern hemisphere climate (Enfield et al., 2001; Knight et al., 2006; Lu et al., 2006; Zhang et al., 2007).

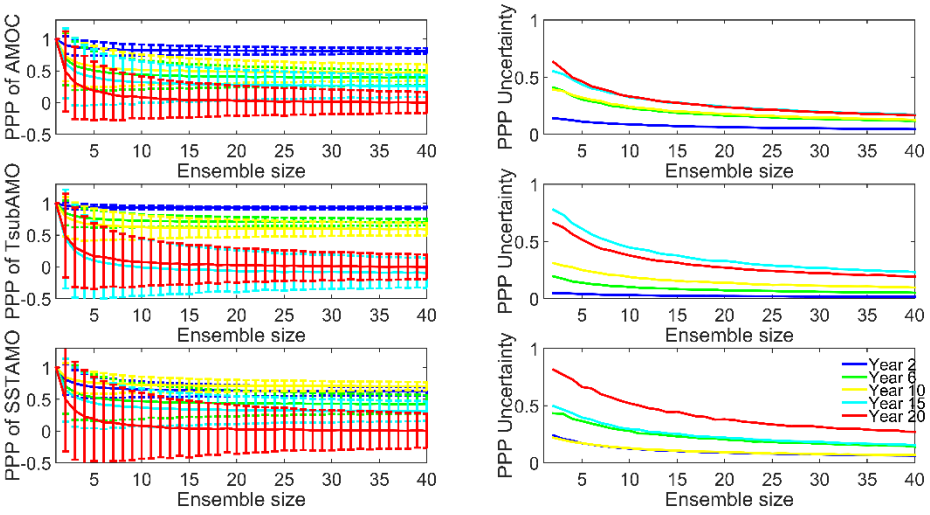

Figure 3. Prognostic potential predictability of AMOC (upper left), PPP of spatial-averaged Tsub in AMV region (middle left), PPP of spatial-averaged SST in AMV region (lower left), PPP_{AMOC} uncertainty (upper right), PPP_{Tsub} uncertainty (middle right) and PPP_{SST}

uncertainty (lower right), respectively, as a function of ensemble size. The curves in the left panel are the mean PPPs of the 5000 random realizations (see text for the details of the random selection process). The vertical bars are standard deviation of the 5000 simulations. The curves in the right panels are standard deviations as a function of ensemble size. Here the Tsub is defined as oceanic temperature at 400 m. Only year 2, 6, 10, 15 and 20 are shown here. PPP and its uncertainty are by definition dimensionless.

We hence restrict our research region in the AMV index region (0-60 ℃N). On the other hand, Sea Surface Temperature (SST), Atlantic Meridional Overturning Circulation (AMOC) and subsurface temperature (Tsub) in this region are also initimately linked with climate variability. Therefore we choose SST (within 0-60 ℃N), AMOC and Tsub (within 0-60 ℃N) as the analyzed objects of our study. Here Tsub is the oceanic temperature at 400 m within 0-60 ℃N in the North Atlantic Ocean, and AMOC is defined as the maximum meridional overturning stream function in the 0-60^{o}N and 50-700 m sector of the North Atlantic Ocean.

In figure 3 we present PPP of AMOC index, regionally averaged Tsub and regional averaged SST, and their uncertainty respectively. Readers ought to notice the horizontal axes are referring to ensemble size. We show only the situation at year 2, 6, 10, 15 and 20. The random selection of sub-ensembles and averaging process is as follows: (1). Each sub-ensemble is randomly selected from the entire 40 members; (2). Get a PPP of the sub-ensemble produced in step (1); (3) Repeat step (1) and step (2) five thousand times and get the mean PPP and the standard deviation. Note that the sampling process is with replacement, i.e., the size of sampling pool is always 40. The uncertainty is due to the repeated random selection process. For instance, when we are calculate the PPP of SST with an ensemble size of 20, we will have 5000 different PPP values. The standard deviation across the 5000 values is referred to as the PPP uncertainty. There are two points of information in this figure: (1). PPPs smoothly decrease and approach to steady values as ensemble size grows; (2). PPPs with longer lead time are lower than that with shorter lead time in most cast (one exception is that the PPP of Tsub in AMV region in year 20 is slightly larger than that in year 15; (3). The uncertainty of PPP is reduced as ensemble size grows.

It is noteworthy that, the aforementioned sub-sampling method and assessment are based on a critical assumption that the 40 member represents all possible situations. This is not exactly the real case but we can argue that, if we reveal some asymptotic properties when the size increase over a certain ‘critical value’ which is less than 40, it is hence justifiable to anticipate that, the asymptotic

properties will not experience dramatic change when the size is larger than 40. This justifies our analysis below.

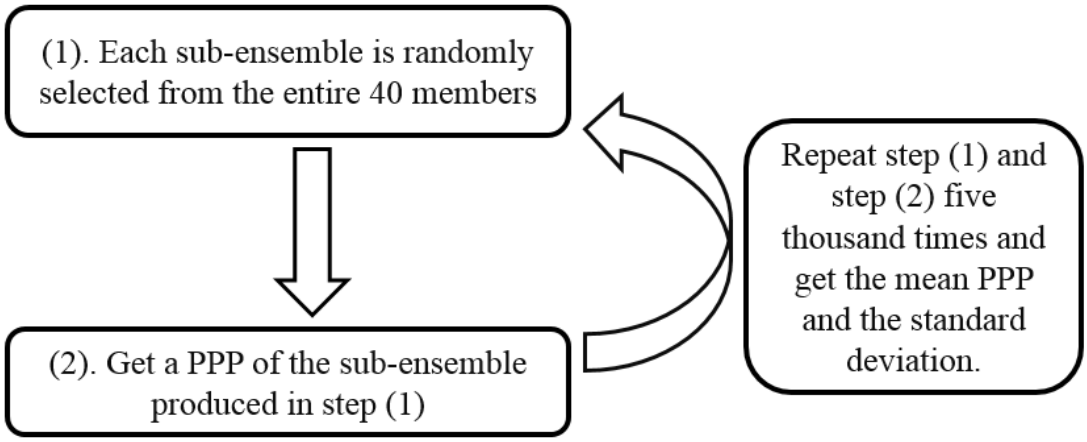

Figure 4. Flow chart of the PPP calculation method. Similar method are adopted in producing Figure 6, 7 and 8.

Unsurprisingly the PPPs decrease as the ensemble size grows, which is in agreement with our empirical speculation: more ensemble members mean more simulation results and larger variance across these simulations. What's interesting in the left panel is that mean PPPs will reach steady values when we have "enough" ensemble member. Loosely we can consider 15 as a threshold size, since mean PPPs in situations of ensemble size larger than 15 is basically not changing any more (Figure 3). Another interesting point in figure 3 is the change of PPP uncertainty. The right panel shows that the uncertainty drops as the ensemble size is enlarged. It is noteworthy that the changing patterns of PPP and PPP uncertainty as size rises are not related to the choice of variable, i.e., PPPs and PPP uncertainties of SST, Tsub and AMOC all tend to drop as ensemble size grows. As a matter of fact, the PPPs reach near-steady states when ensemble size is over 15, and the PPP uncertainties approach near-steady values when ensemble size is no less than 20. In terms of the assessment of initial-value predictability, it indicates that enlarging ensemble size no longer induces significant changes to PPPs when size if over 20. That is to say, in terms of ensemble spread, larger ensemble size leads the assessment of predictability significantly closer to the realistic predictability on when size is less than 20, and the contribution of extra ensemble members on the assessment of initial-value predictability fades to negligible when size is over 20. In addition, despite the slight difference in minimal members needed to get a near-steady state of PPP_{AMOC}, PPP_{Tsub} and PPP_{SST}, loosely the

number 20 is sufficient, i.e., the minimal ensemble size for a near-steady PPP are variable-independent.

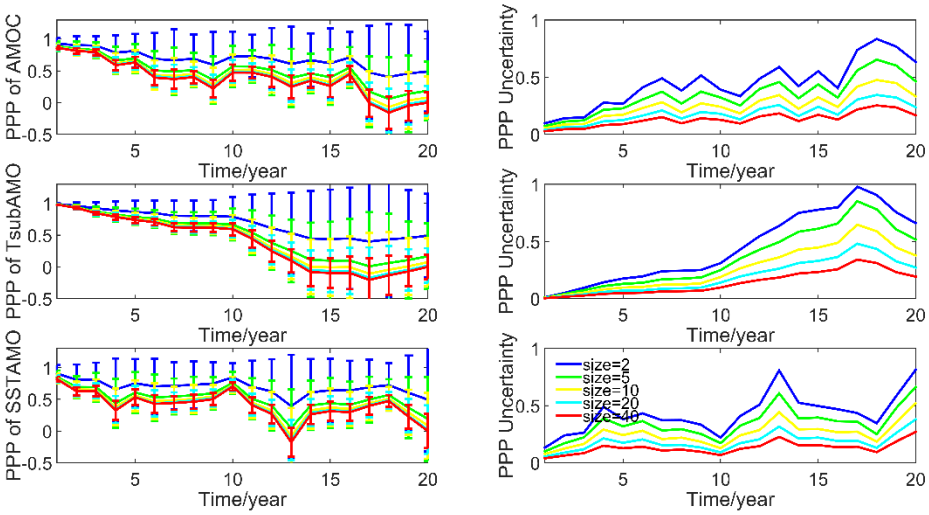

Figure 5. PPP_{AMOC} as a function of time (upper left), PPP_{Tsub} as a function of time (middle left), PPP_{SST} as a function of time (lower left). Uncertainty of PPP_{AMOC} as a function of time (upper right), uncertainty of PPP_{Tsub} as a function of time (middle right), uncertainty of PPP_{SST} as a function of time (lower right). There blue, green, yellow, cyan and red curves are size 2, size 5, size 10, size 20 and size 40, respectively. There are 4 messages in this figure: (1). PPPs decay as lead time increases. (2). The larger ensemble size, the lower PPP. (3). Uncertainty of PPP grows as lead time increases. (4). Larger ensemble size leads to smaller uncertainty of PPP. PPP and its uncertainty by definition are dimensionless.

From another perspective, we could look at the temporal evolution of PPP and PPP uncertainty (Fig. 5). The anti-correlated relationships (larger ensemble size corresponds to lower PPP; and larger ensemble size correspond to less PPP uncertainty) is in accordance with that in Figure 3. We also note that as time goes by, PPPs experience sharp changes, which do not exist in the relationship of PPP and ensemble size. A nuance deserves some attention, that the sharp drop of PPP_{AMOC} (starts in year 16) is 6 years later than that of PPP_{Tsub} and PPP_{SST} (both start in year 10). These sharp drops are likely to be associated with certain particular events (such as a sudden SST cooling event), however further study is required to clarify the mechanism. A comparison of Figure 3 and 5 reveals that the PPP decreases monotonically as ensemble size increases and exhibits an asymptotic property as the size is larger than 15 or so. While PPP doesn't change monotonically with longer lead time (see PPPs in Figure 5 occasionally jump up as the lead time increases).

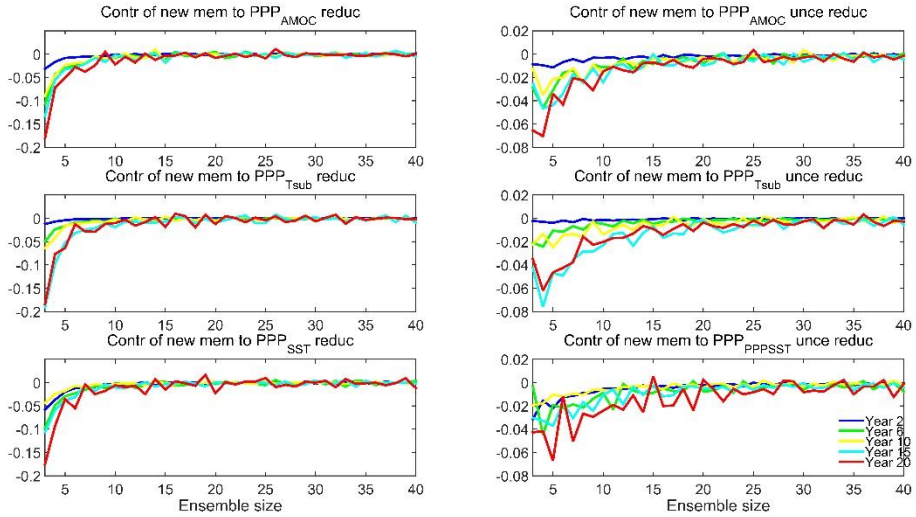

Figure 6. Left column: contribution of one new ensemble member to reduction of PPP_{AMOC} (upper), PPP_{Tsub} (middle), PPP_{SST} (lower). The contribution is defined as PPP(size=n+1)-PPP(size=n). Right column: contribution of one new ensemble member to reduction of PPP_{AMOC} uncertainty (upper), PPP_{Tsub} uncertainty (middle) and PPP_{SST} uncertainty (lower). The contribution is defined as $PPP_{uncertainty}$(size=n+1)-$PPP_{uncertainty}$(size=n). There are 2 messages in this figure: (1). Contribution of one new member decreases as ensemble size grows; (2). Contribution of one new member is larger with longer lead time. Contribution to PPPs by definition is dimensionless.

We further investigate individual ensemble member's contribution to PPP and $PPP_{uncertainty}$. The contribution is defined as

$$contribution_{PPP} = PPP_{size=n+1} - PPP_{size=n} \quad (3)$$

And,

$$contribution_{PPP_{uncertainty}} = PPP_{uncertainty\ size=n+1} - PPP_{uncertainty\ size=n} \quad (4)$$

Unsurprisingly, as the ensemble size grows, individual ensemble member contributes relatively less to the change of PPP and PPPuncertainty (Fig. 5). There are no significant difference among the effects of size-growing process on individual contribution to the metrics used (PPP_{AMOC}, PPP_{SST} and PPP_{Tsub}), i.e., the impact is variable-independent. This is in accordance with our analyses above. An integrated analysis of figure 3 and 6 reveals that, as the size is raised over 15~20, the PPP, PPP uncertainty and individual member's contribution no longer change significantly. Therefore we can loosely define the range of 15~20 as a 'critical size' of IPSL-CM5A-LR. We should be cautious about whether we should continue increasing ensemble size when it is already larger than 15~20,

considering the sharp drop of individual contribution of extra members. It should be noted, though, that 15 members seem to be a good number for the PPP, whereas it is closer to 20 for the PPP uncertainty.

Now that we have confirmed the impact of ensemble size on accurate estimation of the ensemble spread. We wonder what other impact ensemble size can have on numerical simulatons. One important aspect of ensemble simulations is the relation between the distribution of ensemble at long lead time and that of climatology. It is without doubt that the resemblance should be as high as possible. Therefore we compute the probability density functions of AMOC, SST and Tsub of the ensemble experiments for a range of sample size and test its convergence to climatology. Here the PDFs of climatology are calculated using the entire 40 members.

3.2.2 Results and analyses from NEMO in ORCA025 configuration

(a) Impact on PPPs

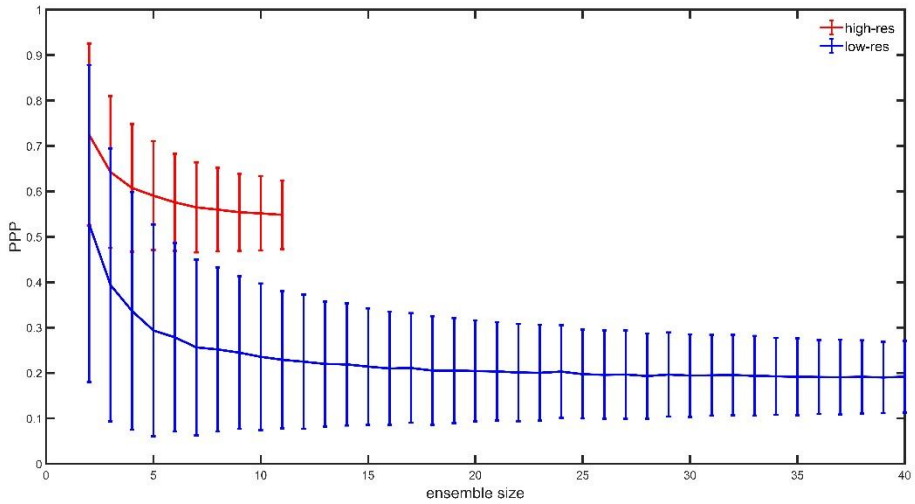

Figure 7. PPPs and corresponding uncertainties of AMOC in the ocean-forced eddy-permitting intermediate-resolution (red) (for the final complete year, i.e. year 2004) and climate model (for the final year of the ensemble experiment, i.e., year 2075) with low-resolution laminar-ocean (blue) models. Here AMOC is defined as the maximum annual mean overturning stream function between 50 °-70 °N in the Atlantic Ocean, at depth between 500-3000 m. Note here the denominator is obtained from control runs. PPP by definition is dimensionless.

Next, to clarify whether the 'critical size' is dependent on model resolution, we calculate the PPP of AMOC as a function of ensemble size in both GCMs which have different resolution. Figure 7 shows us the PPPs of AMOC in both models with varying ensemble size. We use the end points of AMOC time series (i.e. year 20) to calculate PPPs, and the detailed calculation steps are as follows: Take a sub-ensemble with a size of 6 for instance. Firstly we randomly pick a 6-member sub-ensemble from the full 11-member (or 40-member for the low-resolution model) pool, get a PPP of the 6 member. After running this process for 5000 times, we have 5000 PPPs. Take an average of the 5000 PPPs we have a mean PPP, which is shown as the red curve (or the blue curve for the low-resolution model). It is not difficult to tell that in the eddy-permitting model, as the size grows larger than 7, the mean PPP is nearly saturated. Another noteworthy point is that PPP of AMOC in the intermediate-resolution model is always higher than in low-resolution model. The reduction of PPP caused by increasing ensemble members in the intermediate-resolution model is even larger than in the low-resolution mode. These two factors together lead to a larger gap between PPPs in two models with more ensemble members. This near-saturation requires only about 6 members, which is in contrast with the 'critical size' of the low-resolution model (Figure 7). If we consider the PPP in the case of largest ensemble size in both models as the 'realistic' potential predictability of the system, then the smallest size required to reach near-saturation state of mean PPP is actually the minimal size needed to reveal the 'realistic' potential predictability. In other words, increasing ensemble size over 'critical size' does not significantly change the assessment of predictability anymore. I here owe this reduction in 'critical size' to the increase of model resolution. The eddy-permitting property makes it possible for the intermediate-resolution model to capture small scale dynamical processes, which are missed in IPSL-CM5A-LR but key to AMOC variation (Thomas, M. D., and X. Zhai 2013; Mecking, J. V., et al 2016). Another point we can tell from Figure 7 is that, the PPP for the AMOC in the eddy-permitting model is remarkably higher than in the low-resolution model. This once again can be at least partly attributed to the fact that the eddy-permitting model better represents sub-grid processes. This improvement in representation of small scale motions in turn reduces the caused uncertainty. Conventional wisdom holds that eddy is a source of uncertainty, in contrast with our result here. Further investigation is needed on this particular counter-intuitive problem. Apart from these discrepancies, both models agree qualitatively that PPPs decrease asymptotically to saturation as ensemble size grows. It is revealed that the PPP drops exponentially

with ensemble size. The evolving pattern, somewhat counter-intuitively, is independent of model resolution, lead time and variables.

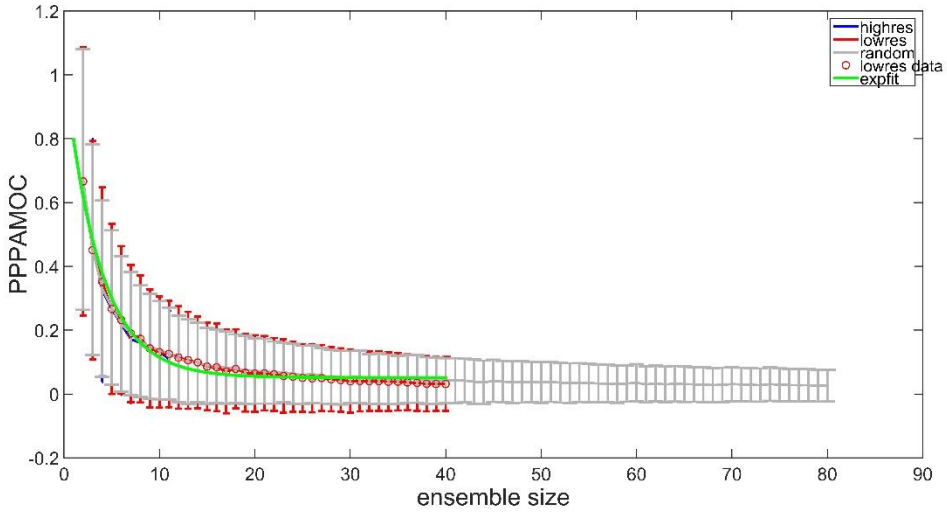

Figure 8.Mean PPPs of AMOC obtained from the high resolution model (blue, at year 3), the low-resolution climate model (red, at year 20), a 80-member Gaussian distribution (gray, with a standard of 0.5, though the PPP-ensemble size relation is indeed independent of standard deviation.), and a fitted exponential function based on the 40-member low-resolution ensemble (green). The mean PPPs is the average PPP of 5000 realizations. Note here the denominator is equation (2) is the variance of the full ensemble (in contrast to Figure 10, where the denominator is obtained from a control run, i.e., a temporal variance). PPP by definition is dimensionless.

Table 2.2 The correlation coefficients of PPPAMOCs obtained using different realizations

Correlation Coefficient	Intermediate-resolution model	Low-resolution model	Fitted exponential function	Random distribution
Intermediate-resolution	1	0.9982	0.9132	0.9992
Low-resolution	0.9982	1	0.9184	0.9985
Fitted function	0.9132	0.9184	1	0.9062
Random distribution	0.9992	0.9985	0.9062	1

It is interesting to find that the mean PPPs of the AMOC obtained from different ensembles is nearly indistinguishable. Precisely, the correlation coefficients between PPP of ocean-forced intermediate-resolution ensemble, PPP of climate-model low-resolution ensemble, PPP of random

distribution and fitted exponential function are remarkably high (Table 2). It is therefore justifiable to conclude that the PPP-ensemble size relation is nearly exponential decrease, and independent of model resolution or the existence of the coupling. We employ an exponential function to fit the PPP-ensemble size relation, the fitted equation type reads,

$$f(x) = a - be^{-cx} \tag{5}$$

Where x is the ensemble size, f(x) is the PPP. The coefficients a, b and c are 0.05277, -1.037 and 0.2933 respectively. Following the concept of e-folding time, I calculate the e-folding ensemble size. And the e-folding sizes (compared to PPP with a size of 2) for the intermediate-resolution, low-resolution and random distribution ensemble are all 6 members.

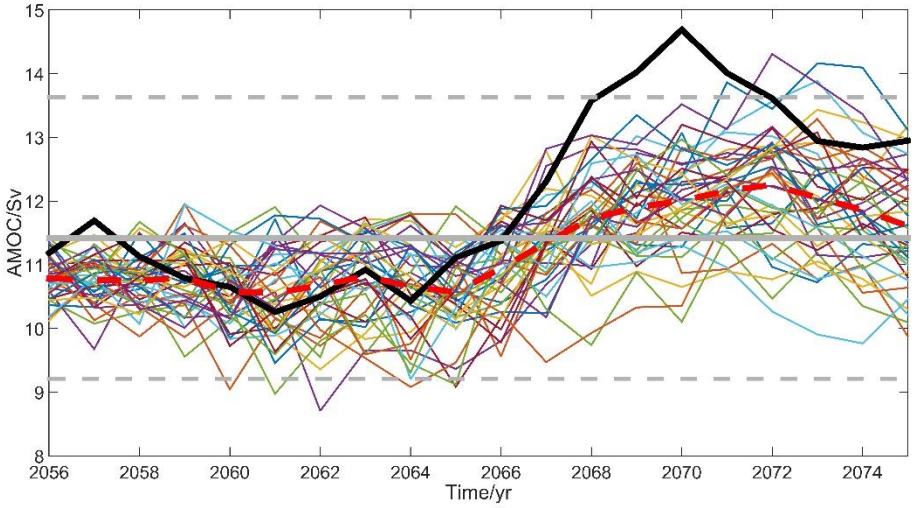

Figure 9. Time series of AMOC in IPSL-CM5A-LR. The black curve is the control run. The colorful curves are ensemble members. Climatological average is shown as the gray thick line. Two standard deviations above or below climatology average are shown as gray dashed lines.

(b) The detection of extreme AMOC events

As mentioned, there exists an AMOC peak event (year 2071) in the modelling period (Fig. 9). Therefore it would be interesting to explore how many members are required to detect the AMOC peaks. To be more precise, we define an extremely strong (weak) AMOC event as when the AMOC

is 2 standard deviation above (below) the climatological mean obtained from the 1000-year control run. In the control run there are 3 extremely strong AMOC events (at year 2070, 2071 and 2072). In contrast, there are only 4 extremely strong AMOC events in the ensemble runs. Note the emergence time of the extremely strong is lagging that in the control run. Surprisingly 6 extremely weak AMOC events occur in the ensemble runs, which never exists in the control run. The ensemble is good at capturing above-or-below average property, but poor at capturing extreme AMOC events.

4. Conclusions and discussions

Ensemble climate prediction skill can be improved through various approaches, such as increasing ensemble size (DS Richardson et al. 2001; T DelSole et al. 2014), using higher resolution models (R Buizza et al. 1998), or adopting better parameterization scheme of unresolved physical processes (J Berner et al. 2017). On the other hand, in an initialized ensemble composed of presumed perfect climate model, the assessment of initial-value predictability is unsurprisingly dependent on ensemble size (A Kumar & M Chen 2015). Note that predictability is an inherent property of the climate system (A Kumar & M Chen 2015), hence it cannot be modified by changing ensemble size. It is the assessment of predictability that varies as ensemble size changes. Among these above mentioned options, the competition between increase of ensemble size and improvement of model resolution is a long-standing topic in climate research, because of the fact that computational resources are limited while both approaches are computationally expensive. It is necessary for us to determine a relatively small but 'sufficient' ensemble size prior to any ensemble forecasting, thus a better understanding of the impact of ensemble size on ensemble prediction skill and assessment of initial-value predictability is required. To this end we perform our analyses in the study of this chapter.

Specifically, the outputs from a conceptual white-noise forced SST model, a fully coupled low-resolution climate GCM IPSL-CM5A-LR and an eddy-permitting forced-ocean GCM, i.e., NEMO in ORCA025 configuration, are used to assess the impact of ensemble size on prediction skill and the evaluation of initial-value predictability of the North Atlantic Ocean subsurface temperature variability. The conceptual model is driven only by white-noise-type high-frequency surface air

temperature, representing a dynamical system that evolves under continued external forcing. The two ensembles are initialized and then evolves under respective identical boundary condition (i.e., in one ensemble, all members share one same boundary condition). The divergence of prediction using this kind of models stands for the spread of trajectories originating from slightly different initial states, due principally to the internal nonlinearity of the system. Hence the impacts of ensemble size on both types of models (one with numerous external forcing; one with identical external forcing) are evaluated in our study.

Flaws inevitably exist in this study. Firstly, since we use only two GCMs in this study, there is no guarantee that the results we get here are universally applicable to various other GCMs (e.g., those models participating Coupled Model Intercomparison Project Phase 5 or 6, i.e., CMIP5 or CMIP6). Secondly, no ensemble size-model resolution relationship is obtained due to unavailability of datasets from models with various resolutions. Thirdly, initialization in each of the two GCMs is imposed on one single year (year 2056 for IPSL-CM5A-RL and year 2001 for the ocean eddy-permitting model). We would anticipate our conclusions to be more robust if the findings could be examined against additional ensemble members initialized in varying years (such as 2057, 2058, etc. for IPSL-CM5A-LR and 2002, 2003, etc. for the ocean eddy-permitting model).In addition, it is possible that the differences of PPP behavior in 2 ensembles are related to other factors such as the inclusion of the atmospheric turbulence (In IPSL-CM5A-LR, there exists atmospheric turbulence, which is not included in the eddy-permitting NEMO insce it is an ocean-only model). Finally, ours results would be more widely representative if we test our findings against other ocean basins, which remains to be performed in the future. However, this study provides some insights about how to determine a ‘sufficient’ ensemble size when designing climate prediction. The conclusions are summarized as follows:

(1) In a white noise forced, conceptual SST model, the mean of ensemble mean approaches toward the theoretical value as ensemble size grows. The uncertainty between ensemble means in different realizations is reduced as ensemble size grows.

(2) In terms of IPSL-CM5A-LR, ensemble size rising leads to a smooth drop of PPP, and a reduction of uncertainties of PPPs. Furthermore, these anti-correlated relations between ensemble size and PPP, ensemble size and PPP uncertainty, are independent of the examined variables.

(3) The PPP and PPP uncertainty of the AMOC no longer change remarkably as the size is over the 'critical size', which is loosely 15~20 for IPSL-CM5A-LR, while ~7 for the ocean eddy-permitting model. That is to say, the minimal size needed are resolution-dependent. Individual contribution of extra ensemble member to the change of PPP and PPP uncertainty drops dramatically as the size is over 'critical size'. It should be noted that the thresholds (i.e., 15~20 for IPSL-CM5A-LR and approximately 7 for eddy-permitting NEMO) at this point are only statistical, whether there are physical processes determing these cretiria remains to be further examined. The difference in critical sizes, however, can probably be attributed to the varying resolutions. In eddy-permitting NEMO, uncertainty related to ocean turbulence are fully captured, indicating that less ensemble members are required for the total uncertainty to saruate (i.e,. PPP uncertainty reaches a quasi-steady state).

(4) The extreme events of AMOC are not well captured in ensemble, suggesting the exceptional characteristics of this event.

(5) The mean PPP-size relation is independent of model resolution, lead time and variables. The PPPs obtained from GCMs are indistinguishable from that obtained from a Gaussian distribution. As a consequence, the e-folding size of PPP for all considered variables in two models with different resolution is 6 member.

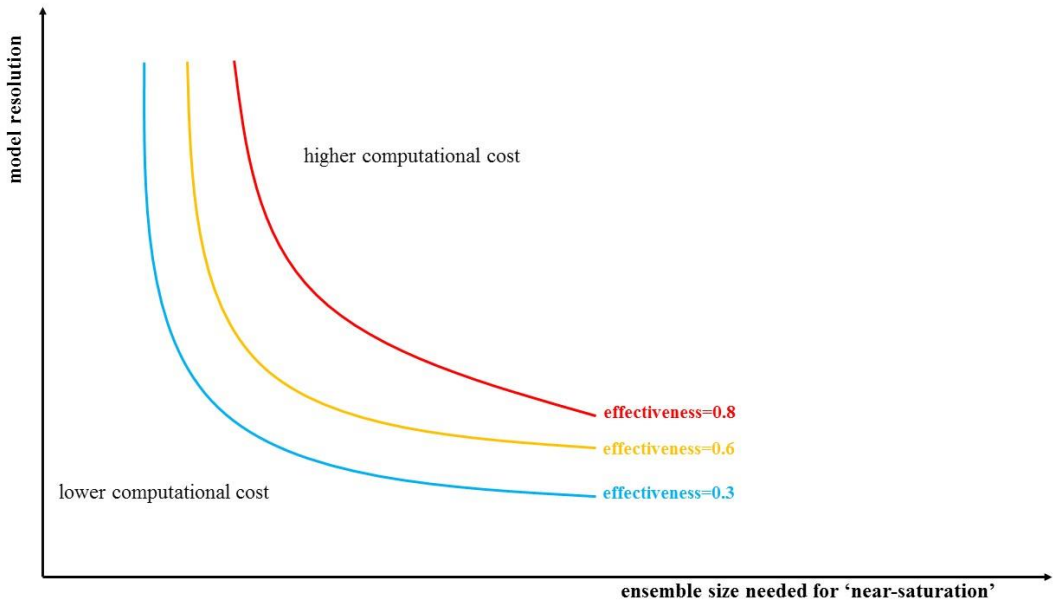

Figure 10. Schematic of the relationship between model resolution, ensemble size, ensemble

forecast effectiveness and computational cost. To achieve a particular effectiveness, we can either enlarge the ensemble size or improve the model resolution. Higher effectiveness of ensemble forecast definitely needs more ensemble members and higher model resolution, but the dependent relationships between computational cost, forecast effectiveness and ensemble size are not simply linear. Doubling ensemble size means twice computational cost and higher (whist not double for sure) effectiveness, but doubling model resolution means eightfold computational cost and higher effectiveness.

Though we suggest no more increase in members if the size is already over the range of 15~20, it is noteworthy that decision should be made depending on specific circumstances. With the intention of measuring the uncertainty caused by internal variability, we define a diagnostic 'effectiveness'. Th effectiveness equals identity if an ensemble can fully capture the internal uncertainty of a particular climate system. Evidently, the effectiveness could be promoted either by using more initialized ensemble members or improving model resotluion. Consider, for example, an ensemble with N members, and the resolution is k×m×n (k, m, and n are grid number in zonal, meridional, and vertical direction, respectively). Doubling the ensemble size will result in a double computation cost, in contrast, improving the resolution to 2k×2m×2n will lead to sixteen-fold computation cost (Figure 10). The one of the two that is the better approach relies on the specific problem we are dealing with.

Chapter 3 Understanding the growth of subsurface temperature prediction uncertainty originating from initial perturbations in the North Atlantic Ocean: the effects of mean state and meso-scale eddies

Key points:

• A clockwise propagation mode of subsurface temperature prediction uncertainty originating from initial temperature errors are found in the subtropical North Atlantic in an ensemble produced by IPSL-CM5A-LR. It can be at least partly explained by the advection of large-scale barotropic circulation in the subtropical region. In contrast, the propagation of prediction uncertainty in the sub-polar region is cyclonic and the vertical development is mainly along continents. This is thought to be associated with the subduction processes.

• Pathways of the Gulf Stream and the North Atlantic Current are the two main source regions of subsurface temperature prediction uncertainty.

• Horizontal eddy diffusive flux of subsurface temperature variance is approximately two orders of magnitude larger than its vertical counterpart, both working as the sources of subsurface temperature predictability.

• The three velocity-related terms as a whole is a source of subsurface temperature prediction uncertainty, which together with the horizontal eddy diffusive flux term are the two major components impacting the developments of subsurface temperature prediction uncertainty.

• In the subtropical gyre, subsurface temperature prediction uncertainty is amplified by the nonlinearity of Gulf Stream and North Atlantic Current, and passively propagated by the Gulf Stream and North Atlantic Current; while in the subpolar gyre, subsurface temperature prediction uncertainty actively impacts AMOC *via* inducing uncertainty in subduction along continental shelves and deep convection process. It can be explained by a salt-advection-feedback-like process in the latitude-depth space.

• Meso-scale eddies along the Gulf Stream and North Atlantic Current are crucial to ensemble spread of subsurface in addition to large-scale oceanic circulation and Rossby waves.

• On various time scales the relative importance of each term impacting the development of subsurface temperature prediction uncertainty is different.

• Vertical and horizontal eddy diffusive fluxes act to suppress the prediction uncertainty (ensemble spread) of subsurface temperature.

Abstract: Subsurface temperature (Tsub hereafter) in the North Atlantic Ocean (NA) is believed to be intrinsically related to the variation of the AMOC. Therefore the necessity of an in-depth understanding of Tsub prediction uncertainty arises in our efforts to better simulate Tsub and the AMOC. In this chapter we use the output of a 40-member ensemble together with a 1000-year control run produced by IPSL-CM5A-LR, conducting a diagnosis of the evolving process of Tsub. Here the Tsub prediction uncertainty develops in the form of ensemble spread, representing internal variability stemming from initial temperature errors. We reveal that the development of Tsub prediction uncertainty is consistent with the clockwise large-scale gyre circulation in the subtropical North Atlantic; in the subpolar region Tsub prediction uncertainty propagates anti-clockwise. In addition, a zonal dipole pattern sustains in the subtropical North Atlantic in the entire period of twenty years. To explain the development of this prediction uncertainty, a dynamical attribution of Tsub variance ($Tsub_{var}$ hereafter) is adopted, decomposing the $Tsub_{var}$ tendency into three velocity-related terms plus horizontal and vertical eddy diffusive fluxes. In this way we find that horizontal eddy diffusive flux of subsurface temperature variance is two orders of magnitude larger than its vertical counterpart. Furthermore we point out that horizontal and vertical eddy diffusive fluxes are sinks of $Tsub_{var}$, while velocity-related terms as a whole act to generate and propagate the prediction uncertainty. Both velocity-related terms and horizontal eddy diffusive flux term are of relatively higher importance to the $Tsub_{var}$ development. An analysis of zonally averaged fields support our findings from 3D data, and further points out the possibility of temperature uncertainty's active impact on meridional overturning circulation *via* inducing uncertainty in deep convection in the subpolar region. This study indicates that higher prediction accuracy of velocity at subsurface depth can contribute to less Tsub prediction uncertainty, which is beneficial for better prediction of climate.

It has been evidenced that subsurface temperature in North Atlantic Ocean plays a significant role in modulating European climate. However few attention is paid to the impact of eddy process in setting the pattern of subsurface temperature in North Atlantic Ocean. In this study, an eddy-

permitting model is used to explore the impact of eddy dynamics on subsurface temperature field. It is revealed that in addition to the large-scale oceanic circulation, meso-scale eddies also contribute to the ensemble spread of subsurface temperature, partly explaining the prediction uncertainty. The generation of meso-eddies is taken place mainly along the Gulf Stream and North Atlantic Current, making these regions important seedbeds of prediction uncertainty. This uncertainty then is propagated to almost the entire subtropical North Atlantic Basin. In this process, horizontal and vertical eddy diffusive flux of subsurface temperature variance both work to suppress the uncertainty, though to different degrees. Eddy diffusive fluxes work as a key sink of this uncertainty. This study highlights the role of mesoscale eddies in shaping the pattern of Tsub prediction uncertainty resulting from initial value uncertainty in the North Atlantic Ocean, suggesting that *via* improving model resolution to eddy-resolving level is a possible way to reduce the biases in estimating the initial value predictability of Tsub in the Labrador Sea.

1. Introduction

Acting as a crucial northward heat conveyor, the Atlantic meridional overturning circulation (AMOC) plays a pivotal role in the redistribution of the oceanic thermal energy in the Atlantic Ocean. The variability of the Atlantic, *via* air-sea interaction, profoundly impacts the atmosphere circulation over it, and consequently modifying the climate both locally and remotely (Rong Zhang et al. 2019). It is fair to say that AMOC plays a non-negligibly important role of global climate system. Evidences are found, for example, that AMOC can impact European and global climate (Jackson et al. 2015; Sutton et al. 2012); Arctic sea ice state (Mahajan, Salil, et al. 2011) and ENSO variability (Timmermann, Axel, et al. 2007; Williamson, Mark S., et al.). To make better climate prediction, it is of great significance to better understand physical processes that have effect on dynamics of AMOC.

Under circumstances of global warming, there exists the possibility of an AMOC (Manabe and Stouffer 1993; Rahmstorf and Ganopolski 1999; Houghton et al. 2001; Wei Liu et al. 2017). Enormous efforts have been made to get a clearer picture of the dynamics of AMOC, including theoretical research, field observation, and numerical simulation (MW Buckley and J Marshall 2016;

GD McCarthy et al. 2019; W Weijer et al. 2019). Generalization of pioneering work by Stommel (1961) to the AMOC pointed out its possible bi-stability. This was followed by numerous studies on the stability problem. Gnanadesikan use an idealized 2-layer model to study the global circulation, revealing the significant effect of Southern Ocean in the large scale oceanic structure. Timmermann et al 2000 investigate the impact role of climatic noise for the thermohaline component of the AMOC using the Stommel model, implying a relevance between the two. Jian Zhao uses a 2-layer linearized model and a state-of-the-art Earth simulator to reveal that large scale wind-forced oceanic internal Rossby waves leads to coherent interannual variability patterns of the AMOC across the tropics and subtropics. On longer timescales, intrinsic oceanic dynamics play a leading role (J Bjerknes 1964). A wave-pattern westward propagation of temperature anomaly, i.e., a thermal Rossby waves within 30-60°N, are documented to be responsible for the leading interdecadal oscillatory-eigenmode of the AMOC (Sévellec and Fedorov, 2013). The mode can be efficiently stimulated by external noise due to the nonnormality of the system (Sévellec et al. 2009). Other factors, such as subpolar gyre in the North Atlantic (L Hermanson et al. 2014), difference of the freshwater flux in the South Atlantic and the North Atlantic, magnitude and location of extra surface freshwater flux (RS Smith et al. 2009; Q Yang et al. 2016), perturbations in the deep water and Southern Ocean dynamic processes, have impacts on variation of the AMOC (TL Delworth and F Zeng 2008; D Swingedouw et al. 2009; R Farneti and TL Delworth 2010; F Sévellec et al. 2011; T Martin et al. 2013; C Buizert et al. 2015) . Direct observation has become available since 2004, e.g. the RAPID-MOCHA array (Hirschi, J., et al 2003; B Sinha et al 2018; McCarthy, G. D., et al 2019) and Overturning in the Subpolar North Atlantic Program (OSNAP, MS Lozier et al. 2017). Focusing on the 26 ºN, this project provides valuable information. The duration, however, is not long enough for the research on variability at timescales longer than decadal. Nowadays we know the AMOC much better than decades ago but many details remain unclear, among them is the prediction uncertainty of subsurface temperature on decadal time scale.

Unlike the case of ocean surface, subsurface ocean is less influenced by the synoptic scale disturbance of the atmosphere above, endowing subsurface thermal field a relatively higher signal to noise ratio. This characteristic implies that the study of subsurface may shed some light on the intrinsic dynamics of the ocean. Several previous studies reveal that, subsurface temperature not only play the role of the AMOC fingerprint, but also can be utilized to skillfully reconstruct the

AMOC variations (J Mignot et al. 2007; Rong Zhang 2008 ;R Msadek et al. 2010;Mahajan et al 2011).

In the second part of this chapter, an eddy-permitting model is used to explore the impact of eddy dynamics on subsurface temperature field. It is revealed that in addition to the large-scale oceanic circulation investigated using the IPSL-CM5A-LR, meso-scale eddies also contribute to the ensemble spread of subsurface temperature, partly explaining the prediction uncertainty. The generation of meso-eddies is taken place mainly along the Gulf Stream and North Atlantic Current, making these regions important seedbeds of prediction uncertainty. This uncertainty then is propagated to almost the entire subtropical North Atlantic Basin. In this process, horizontal and vertical eddy diffusive flux of subsurface temperature variance both work to suppress the uncertainty, though.

Additionally, a quasi-annual burst of intensified subsurface temperature variance in the Labrador Sea is reported for the first time. With a quasi-6-week life cycle, it emerges in the spring, spreading in the form of an expanding cycle. This is thought to be related with atmospheric activities in the spring, but further studies are required.

Here we continue our research with a study of Tsub prediction uncertainty propagation pattern, followed by an analysis of the subsurface temperature ensemble variance ($Tsub_{var}$ hereafter) budget. The purpose of the first part of this chapter is to reveal (using ouput of IPSL-CM5A-LR):

(a) What is the spatial pattern of the evolution of in IPSL-CM5A-LR?

(b) What is dominating its evolution? What are the sources of $Tsub_{var}$ and what are the sinks?

(c) What are the relative significance of contribution of those factors impacting $Tsub_{var}$'s development?

The scientific questions to be dealt with in this last part of the chapter are (using output of eddy-permitting NEMO):

(1) Is there any difference in the North Atlantic $Tsub_{var}$ evolution in a laminar and an eddy-permitting model?

(2) How can the meso-scale eddies contribute to the development of North Atlantic $Tsub_{var}$?

(3) Is the contribution of eddies dependent on time scales?

This chapter is organized as follows: data, model and methods used are described in section 2;

results and analyses are presented in section 3; we close this chapter with conclusions and discussion in section 4.

2. Data, model and methods

2.1 A coupled low-resolution ocean model IPSL-CM5A-LR

The ensemble experiment is performed using IPSL-CM5A-LR, with an ocean component being NEMO (Nucleus for European Models of the Ocean) 3.4 with a configuration of ORCA2 (Madec Gurvan and Maurice Imbard 1996). The data used so far are output from a 40-member ensemble experiments, and one 1000-year piControl run. The 40 ensemble experiments are initialized from a same year in the 1000 years, with a same background initial state and boundary conditions but only perturbed differently at the beginning. For further details of this model and the ensemble design, readers are recommended to refer to Germe et al (2017). The 1000-year piControl control is produced in this way: the model adopts pre-industrial boundary conditions of greenhouses gases and aerosols concentrations after spin-up (Dufresne et al. 2012), and then is integrated for more than 1000 years. Subsequently a year in the middle part is arbitrarily appointed to be 1800, and the following 1000 years is utilized in our analysis. The ensemble experiments are generated as follows: four different types of initial temperature perturbations (SST/3D oceanic temperature/3D oceanic temperature with tenfold magnitude/deep 3D oceanic temperature; and for each type of them, we produce 10 very similar but slightly different perturbations) are imposed on an identical background initial state, then we have 40 'complete initial states'. We subsequently run the model for 20 year (from 1st January 2056 to 31st December 2075) to get 40 ensemble members (see Table 1 in Chapter 3 for details of the ensemble design). The reason for choosing the span of 2056 to 2075 is the existence of a peak AMOC in this period, hopefully the possible factors impacting the peak can be investigated, such as to find a precursor of anomaly AMOC (F S évellec and AV Fedorov 2013, 2015; F S évellec et al. 2017). Besides, Germe et al. (2017) demonstrated that no significant difference of ensemble statistics, such as ensemble mean and ensemble spread of SST and AMOC, arises from different kinds of perturbations. Therefore we consider the 40 ensemble members equally in our study (i.e., we ignore the difference in perturbation types).

Table 3.1 Model details

Ens size	Oceanic horizontal resolution	Oceanic vertical levels	Data start time	Data end time
40	2 °×2 °	31	1st Jan 2056	31st Dec 2075

A combined analysis of numerical simulation results, satellite altimetry data and instrumental ocean temperature reveals a consistent signal of subsurface temperature and the AMOC, suggesting Tsub a good proxy to reconstruct past change of the AMOC (Zhang 2008). A follow-up research conduct predictions basing on this relationship. Using subsurface temperature and sea surface height produced by GFDL CM 2.1 model, Mahajan et al. (2011) develop two statistical models to predict the change of the AMOC, indicating a decline trend in the near future. However the results in both studies are obtained using outputs from only one model and its robustness has not been examined in other GCMs. Other studies exhibit linkages between the AMOC and Tsub on different time scales (Wang et al. 2013, Marcott et al. 2011, Schmidt et al. 2012, Mignot et al. 2007). These findings provide some confidence in the design of our future work, to deepen our understanding of decadal predictability of the AMOC in the perspective of subsurface signal.

2.1 An intermediate-resolution ocean model

We adopt NEMO 3.4 (the Nucleus for European Models of the Ocean 3.4) to perform the ensemble experiments. The oceanic component is implemented on in the ORCA025 configuration: the nominal horizontal resolution is 1/4 degree (1442-by-1021 grid points), and on the vertical there are 75 levels with varying thicknesses ranging from 1 m at the surface to around 200 m at the bottom. The ensemble size is eleven, differing only by their initial states. This version of NEMO is described and used in AT Blaker et al. (2012), CA Katsman et al. (2018), among many others. The time span of this ensemble experiment is from 8th January 2001 to 15th June 2005.

Table 3.2 Model details

Ensemble size	Model resolution	Vertical levels	Data start time	Data end time
11	1/4 °×1/4 °	75	8th January 2001	15th June 2005

The diagnostic is performed with data of 3D temperature, 3D velocities and the zonally averaged meridional overturning stream function in the Atlantic Ocean. Among them of former two are available from 8th January 2001 to 2005 15th June, the latter one from 8th January 2001 to 17th December 2004.

2.2 Prognostic potential predictability

In order to reveal ensemble spared as well as the relative magnitude of ensemble variance and climatological variance, we use Prognostic Potential Predictability (PPP, Pohlmann et al. 2004) as our first diagnostic. It is defined as

$$PPP = 1 - \frac{\sigma_{ens}^2}{\sigma_{\infty}^2} \tag{1}$$

Where σ_{ens}^2 is the variance of the ensemble, σ_{∞}^2 the variance of the 1000-year control run or the variance of the full ensemble pool. PPP is a dimensionless quantity usually between [0, 1], representing the relative difference between the climatological spread and ensemble spread. Note that PPP a bit lower that zero is possible, given a huge ensemble variance larger than the control temporal variance. If PPP equals unity, it represents a 'perfect' forecast, indicating that ensemble spread is zero, i.e., internal variation originating from initial errors vanishes (This is definitely impossible in the real world, therefore zero ensemble spread indicates overconfidence of the model). On the contrary when PPP approximately vanishes, it means the large internal variation, i.e., the forecast is not significantly better than a prediction made based on climatology (Pohlmann et al. 2004). However PPP possesses no straightforward physical nature, to this end we further analyze the evolution Tsub variance, from which the σ_{ens}^2 actually arise. Apparently, the variance of an initialized ensemble at the initial time is close to one. On the contrary, if the ensemble variance is

comparable to the total variance, PPP approximately vanishes, which suggests that the initial value predictability is 'lost' entirely.

2.3 Temperature variance budget analysis in 3D and 2D (latitude-depth) spaces

PPP_{Tub} is, by nature, a statistic metric of ensemble spread of Tsub, which in turn is essentially consequences of the climate internal variation. In order to reveal the factors responsible for the development of ensemble spread of Tsub, i.e., the dynamical elements involved in the internal variation of Tsub, I analyze the development of Tsub variance ($Tsub_{var}$ hereafter). Here the $Tsub_{var}$ refers to the variance across the forty ensemble members. Obviously, according to the definition, $Tsub_{var}$ is anti-correlated with PPP_{Tsub}. Note that unlike PPP_{Tsub}, $Tsub_{var}$ has the advantage that the budget equation can be directly derived from passive tracer equation of temperature. Consider the temperature conservation equation (Sévellec et al. 2006) for the subsurface ocean with the assumption that the forcing term and the convection term are constants for simplicity:

$$\frac{\partial T}{\partial t} = -\mathbf{U}\cdot\nabla_{3D}T + K_H\frac{\partial^2 T}{\partial y^2} + K_V\frac{\partial^2 T}{\partial z^2} + F_T + C_T \tag{2}$$

Where **U** is the 3-dimensional velocity. The temperature and its components can be decomposed into the ensemble mean $\bar{T}$ and the deviation T_i', so can the velocity, hence we have:

$$\frac{\partial(\bar{T}+T_i')}{\partial t} = -\left(\bar{\mathbf{U}}+\mathbf{U}_i'\right)\cdot\nabla_{3D}\left(\bar{T}+T_i'\right) + K_H\nabla_H^2\left(\bar{T}+T_i'\right) + K_V\frac{\partial^2(\bar{T}+T_i')}{\partial z^2} + F_T + C_T \tag{3}$$

Sum the 40 members up, we have:

$$\begin{aligned} 40\frac{\partial\bar{T}}{\partial t} + \sum_{i=1}^{40}\frac{\partial T_i'}{\partial t} = &-40\bar{\mathbf{U}}\cdot\nabla_{3D}\bar{T} - \sum_{i=1}^{40}\bar{\mathbf{U}}\cdot\nabla_{3D}T_i' - \sum_{i=1}^{40}\mathbf{U}_i'\cdot\nabla_{3D}\bar{T} - \sum_{i=1}^{40}\mathbf{U}_i'\cdot\nabla_{3D}T_i' \\ &+40\left[K_H\nabla_H^2\bar{T} + K_H\nabla_H^2T_i' + K_V\frac{\partial^2\bar{T}}{\partial z^2} + K_V\frac{\partial^2 T_i'}{\partial z^2} + \bar{F}_T + \bar{C}_T + F_T' + C_T'\right] \end{aligned} \tag{4}$$

The second term on the left hand side, the second, third, sixth and eighth terms on the right hand side all vanish. Hence we have:

$$\frac{\partial\bar{T}}{\partial t} = -\bar{\mathbf{U}}\cdot\nabla_{3D}\bar{T} - \frac{1}{40}\sum_{i=1}^{40}\left(\mathbf{U}_i'\cdot\nabla_{3D}T'\right) + K_H\nabla_H^2\bar{T} + K_V\frac{\partial^2\bar{T}}{\partial z^2} + \bar{F}_T + \bar{C}_T \tag{5}$$

Consider now the tendency of temperature variance:

$$\frac{1}{2}\frac{\partial T'^2}{\partial t} = \frac{1}{2}\left(2T'\frac{\partial T'}{\partial t}\right) = T'\frac{\partial T'}{\partial t} \tag{6}$$

Thus we have:

$$\frac{1}{2}\frac{\partial T'^2}{\partial t}=\frac{1}{40}\sum_{i=1}^{40}\left(T_i'\frac{\partial T_i'}{\partial t}\right)=\frac{1}{40}\sum_{i=1}^{40}\left\{T_i'\left[\frac{\partial\left(\bar{T}+T_i'\right)}{\partial t}-\frac{\partial\bar{T}}{\partial t}\right]\right\} \tag{7}$$

Rewrite it in a compact form,

$$\overline{\frac{1}{2}\frac{\partial T_i'^2}{\partial t}}=\underbrace{-\bar{\mathbf{U}}\cdot\overline{T_i'\nabla_{3D}T_i'}}_{\text{advection term}}\underbrace{-\overline{\mathbf{U}_i'T_i'}\cdot\nabla_{3D}\bar{T}}_{\text{generation term}}\underbrace{-\overline{\mathbf{U}_i'T_i'\cdot\nabla_{3D}T_i'}}_{\text{self advection term}}\underbrace{+K_H\overline{T_i'\nabla_H^2T_i'}}_{\text{horizontal eddy diffusive flux}}\underbrace{+K_V\overline{T_i'\frac{\partial^2T_i'}{\partial z^2}}}_{\text{vertical eddy diffusive flux}}+residual \tag{8}$$

where variables with prime symbols refer to the deviation of an individual ensemble member from the ensemble mean; variables with an overbar is the ensemble mean. The term on the left-hand side is the ensemble mean of the $Tsub_{var}$ tendency. The first term on the right-hand side is the advection term; the second term on the right-hand side is the generation term; the third term on the right-hand side is the self-advection term (Tréguier Anne-Marie et al. 1997, where equation (35) is QG eddy potential enstrophy balance, but the second the third term on the right-hand side is in similar form to second and Equation (8) in this manuscript, this leads us to make this definitions); the fourth and fifth terms are horizontal eddy diffusive flux and vertical eddy diffusive flux, respectively. A residual term is expected to exist to close the budget equation. The residual being the component that derives from the convection and forcing terms (which are not expected to be significant in depth and studied region), and high-frequency process not resolved by our original monthly mean data. This variance budget has been successfully used in a range of previous studies (Sévellec et al. 2006; Arzel et al. 2006; Buckley et al. 2012; Hochet et al, 2020; Sévellec et al. 2021). The relative importance of the right-hand-side terms in equation (2) to (8) depends on both geographical position and time scale. For example, on inter-annual timescale, in ocean interior (far from boundary regions), the advection term is predominantly important due to larege bacis flow magnitude. However in front regions, there exist strong temperature gradient exis, the generation term might be comparably important.

3. Results and analyses

3.1 Tsub as an indicator of the AMOC variation in the piControl run of IPSL-CM5A-LR

Here the subsurface temperate refers to the oceanic potential temperature at 400-m depth, same as Rong Zhang (2007, 2008, 2010). The North Atlantic mean SST and Tsub are averaged over the North Atlantic Region, for the reason that temperature field in this region has substantial impact on climate over surrounding landmasses (JR Knight et al. 2006) and remote regions (R Zhang and TL Delworth 2007; Bin Wang et al. 2013; C Sun et al. 2017). The AMOC index (hereafter AMOC) are defined as the maximum overturning stream function within 50-70 ℕ. These definitions can also be supported by Figure 1, when lag-correlation coefficients between AMOC index and temperature at different depths are displayed. Both time series are normalized by one standard deviation after being linearly detrended. Largest positive correlation coefficients (up to 0.454) occur when AMOC leads temperature by approximately 1-8 years at depth between 80-429 m; meanwhile largest negative correlation coefficients (up to -0.637) occur when AMOC lags temperature by around -2-10 years at depth between 100-1200 m. This indicates that temperature at sub-to-middle depth is can be a precursor of the change of AMOC up to 10 years later.

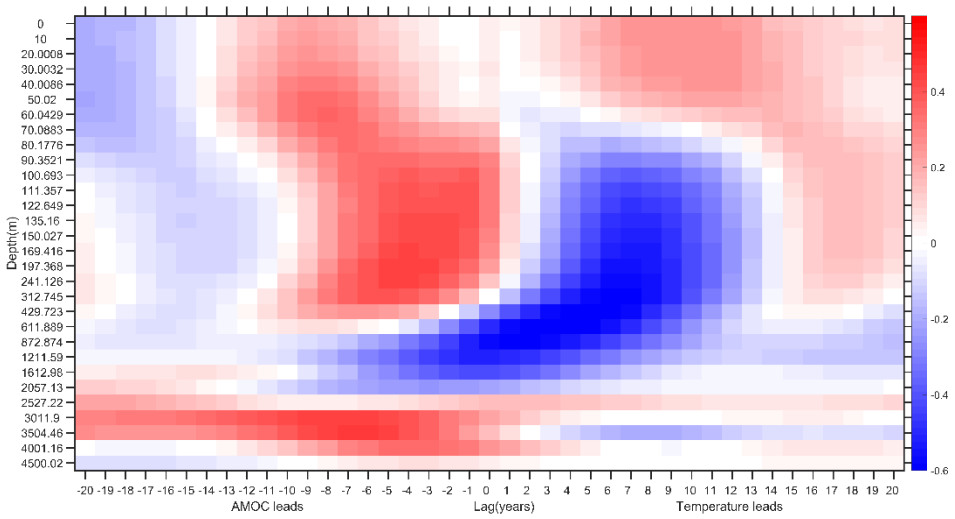

Figure 1. Lag-correlation coefficients between the AMOC and temperature at different depths in the 1000-year piControl run. Note the vertical axis is not scaled according to real depth. The temperature is locally averaged over the entire North Atlantic Ocean, and

weighted by grid cell surface area. Both AMOC and temperature time series are normalized and detrended prior to the computation. Largest positive correlation coefficients (up to 0.454) occur when AMOC leads temperature by around 1-8 years at depth between 80-429m. Largest negative correlation coefficients (up to -0.637) occur when AMOC lags temperature by around -2-10 years at depth between 100-1200m. This indicates that temperature at sub-to-middle depth could work as a precursor of the change of AMOC up to 10 years later. Lag-correlation coefficients by definition are dimensionless.

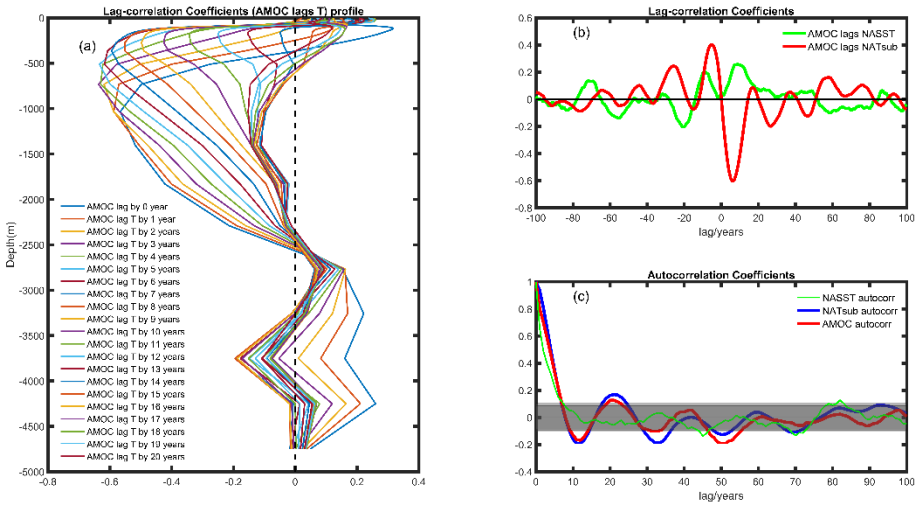

Figure 2. Lag-correlation coefficients between AMOC and temperature at different depths and with various lag time (left panel) in the 1000-year piControl run. In upper right panel are lag-correlation coefficients between AMOC-North Atlantic SST (green line), and AMOC-North Atlantic subsurface temperature (red line). In lower right panel are autocorrelation coefficients of North Atlantic SST (green line), North Atlantic subsurface temperature (blue line) and AMOC (red line). All data are normalized and detrended prior to calculation of correlation coefficients. Noticeably largest correlations coefficients occur at subsurface depth (from -200 to -1000 m), being up to around -0.6. On the other hand, the good agreement between NATsub autocorrelation coefficient and AMOC autocorrelation coefficient indicate a same periodicity between the two. Lag-correlation coefficients are dimensionless by definition.

To further support the links between AMOC and temperature shown in Figure 1, correlation coefficient profiles are calculated and displayed in Figure 2 as well. Notably largest correlations coefficients occur at subsurface to mediate depth (from -200 to -1000m), being up to around -0.6. Autocorrelation coefficients of North Atlantic SST (NASST), North Atlantic subsurface temperature (NATsub) and AMOC are also shown in right lower panel. It clearly points out that NATsub and AMOC share several same periods. This implies that the

association between subsurface-to-deep ocean temperature and AMOC on interannual time scale is stronger than that between SST and AMOC.

So we can make a small summary that Tsub could be, at least in IPSL-CM5A-LR, a good indicator of the AMOC variation, partially in accordance with previous studies (J Mignot et al. 2007; Rong Zhang 2008). Therefore the necessity of a better capture of Tsub variations in numerical simulations arises, which calls for a more in-depth understanding of Tsub prediction uncertainty. This motivates the following analysis.

3.2 Understanding the prediction uncertainty of Tsub

Shown in Figure 3 is the temporal evolution of PPP_{Tsub} in the twenty years. Prior to the calculation of PPP, the original monthly data were averaged to generate the annual mean to remove the seasonal cycle. A small chain of low-value patches emerges originally in the western boundary current region and North Atlantic Current region (for convenience chain A hereafter). Then this chain expands both meridionally and zonally within years 2-6. Afterwards, a small low-value patch (patch B hereafter) appears to be in the middle of North Atlantic Ocean. This new patch then propagates southwestward as well as gets larger within years 8-12. The two patches merge and extend to the most part of the North Atlantic Ocean within years 14-20. This temporal evolution process is characterized by 5 remarkable features: firstly, the low value patch originates in the western boundary region and North Atlantic Current (chain A); secondly, patch B propagates southwestward, in the opposite direction as chain A (A expands mainly eastwards); thirdly, there is a zonal dipole pattern of PPP_{Tsub} in the subtropical North Atlantic Ocean, i.e., high PPP is largely in the western subtropical North Alantic while low PPP is mainly in the east part (near Europe and northwestern Africa).

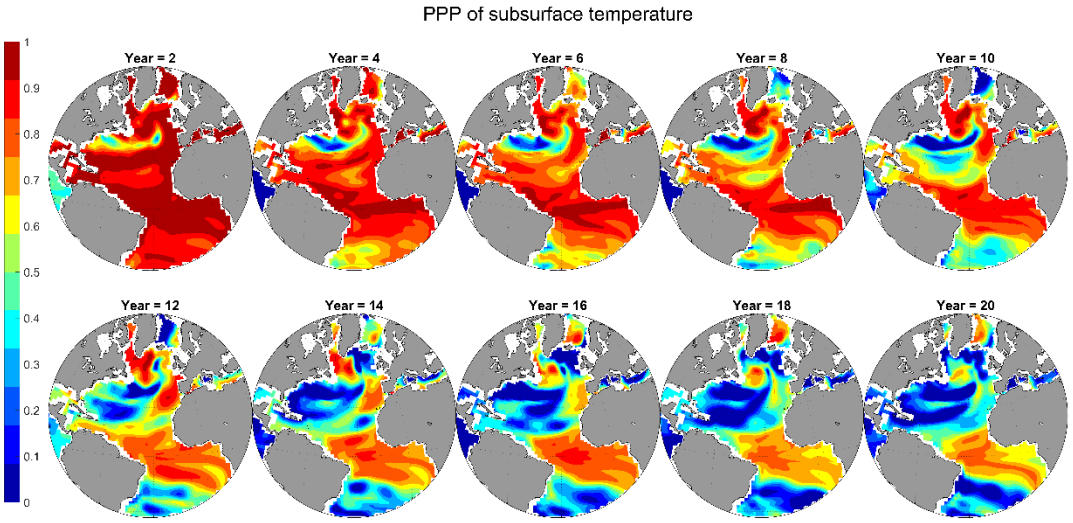

Figure 3. Prognostic potential predictability (PPP) of subsurface temperature in the 20-year ensemble run. All 40 members are taken into account when calculating ensemble variance. The control run is used to compute the historical variance. In red is above 0.5, in blue below 0.5, in white equivalent to 0.5. Small blue patches originate in the western boundary region and North Atlantic Current region, expanding clockwise to fill almost the entire subtropical NA 20 years on, except the eastern part near southwestern Europe and northwestern Africa. This data used here is the annual mean of monthly PPP. PPP by definition is dimensionless.

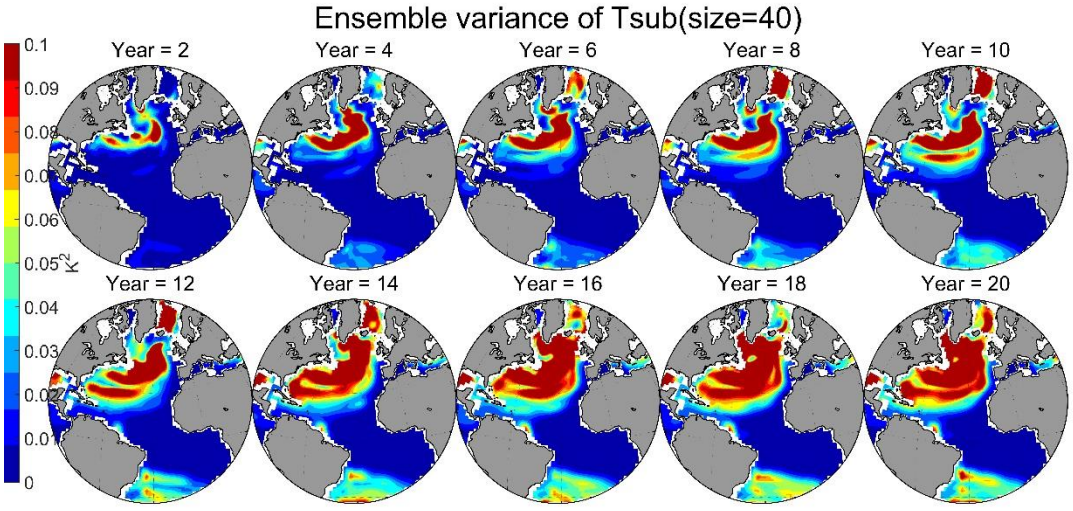

Figure 4. Evolution of subsurface temperature ensemble variance in NA in the 20-year ensemble run (unit: ℃). Here the variance is the normalized squared deviation from ensemble mean. Unit: K^2.

Figure 4 exhibits the temporal evolution of $Tsub_{var}$. We can readily notice that the development of high-value patches of $Tsub_{var}$ shear several features with the development of low-value patches of

PPP_{Tsub}, in correspondence with the fact that PPP is actually the local-normalization of ensemble variance. There are four distinct features in the development of high-value $Tsub_{var}$: 1. A high-value patch (patch A') originates in the western boundary current region and North Atlantic Current; 2. Another high-value patch (patch B') emerges and develops within years 6-8; 3. The expanding or propagation direction of patch A' and B' are opposing, i.e., A' moves eastward B' westward; 4. In the late period of the twenty year, a significant dipole pattern of $Tsub_{var}$ exists in the subtropical North Atlantic Ocean, i.e., strong variance is mainly located in the western part while weak variance is in the east subtropical North Atlantic. This is in good agreement with the dipole pattern of PPP shown in Figure 3. Now we may examine the sources and sinks of $Tsub_{var}$. This requires a thorough investigation of terms in equation (2), see in the appendix A for the details of how to generate equation (2).

Since the velocity data are unavailable, the first three terms on the right hand side cannot be calculated straightforwardly. So here I only display the left-hand-side term (Figure 5) and the last two terms on the right-hand side (Figure 6 and Figure 7). Note that the eddy diffusivities K_h and K_z are indeed associated with the velocity, however we can estimate the magnitudes of term 4 and term 5 by using typical values of K_h and K_z. Even though an exact quantification of each term cannot be realized, we can still have a flavor of the comparison of contributions by the first three terms on the right-hand side as a whole and the contributions of the fourth term and the fifth term.

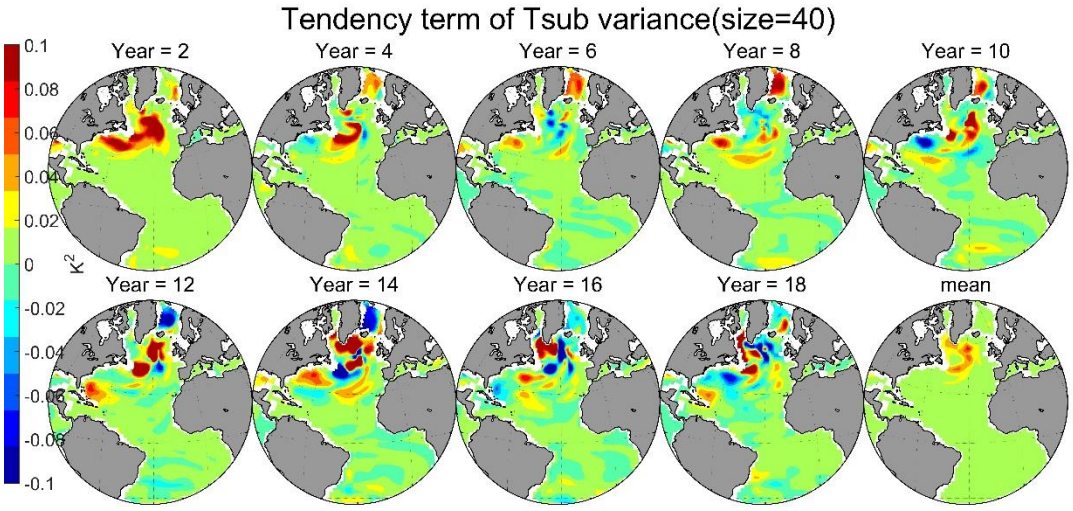

Figure 5. Development of temperature variance tendency term (i.e., the left-hand-side in Equation (8)) in the 20-year ensemble run. Note that there is no particular spatial pattern in

the growing process. Even though in individual year the left-hand-side might not be uniformly positive, the 20-year mean is positive. It is in accordance with the fact that the 20-year integration reveals a generally positive growth in the subtropical North Atlantic. Unit: K^2 yr^{-1}

Figure 5 is the temporal evolution of the left-hand-side term of Equation (8). There is no particular long-standing spatial pattern during this developing process. However the 20-year mean is predominantly positive in almost the entire North Atlantic, with high-value along the North Atlantic Current and in the SPG. Note that the magnitude is one order larger than vertical eddy diffusive flux in Figure 7 and one order smaller than horizontal eddy diffusive flux in Figure 6, indicating the insignificance of vertical eddy diffusive flux of $Tsub_{var}$.

Since the background initial states (before imposing perturbation) and boundary conditions of the 40 ensemble experiments are identical, the only source of the ensemble spread stems from the initial temperature perturbation and internal error growth process of the ocean, i.e., the terms in Equation (8). However, since we have no access to the 3D velocities, it is impossible to compute the first three terms, i.e., the velocity-related terms. Hence here I calculate the vertical and the horizontal eddy diffusive fluxes of $Tsub_{var}$. Typical values of K_h and K_v, 10^4 m^2 s^{-1} and 10^{-4} m^2 s^{-1}, respectively, are taken in the computation. Figure 6 shows development of horizontal eddy diffusive flux of $Tsub_{var}$. Noticeably the horizontal eddy diffusive flux is negative in majority of the North Atlantic Ocean (NA), except in the center of subpolar gyre (SPG) and several patches in the North Atlantic Current. The strongest values occur in the SPG, the Gulf Stream and the North Atlantic Current. In the later part of the twenty year, there emerges a zonal dipole pattern of horizontal eddy diffusive flux of $Tsub_{var}$ in the SPG. In the subtropical region, the propagation model is by and large in accordance with the clockwise large-scale circulation.

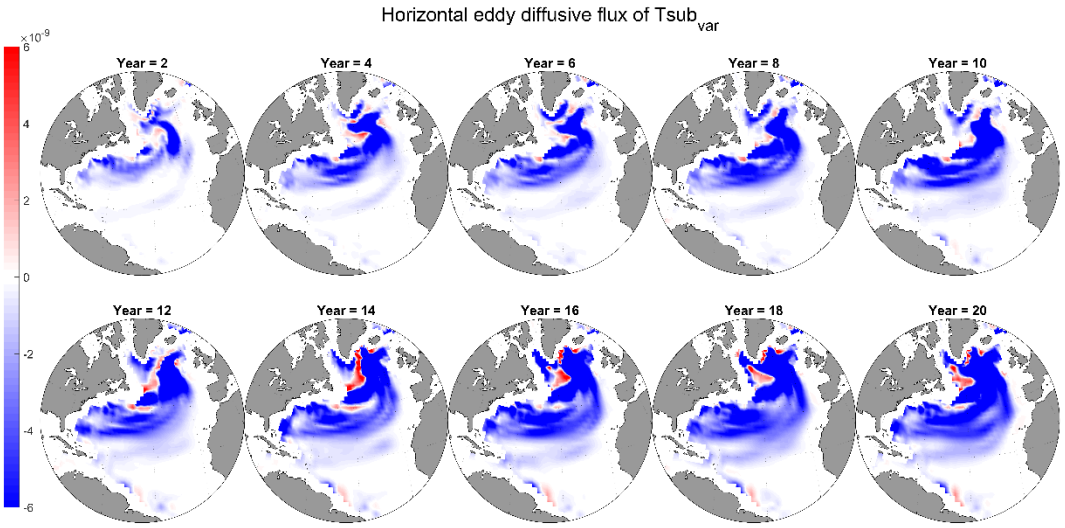

Figure 6. Temporal evolution of horizontal eddy diffusive flux of $Tsub_{var}$ in the 40-member, 20-year ensemble run. Throughout the entire twenty years the effect of horizontal eddy diffusive flux of $Tsub_{var}$ is largely opposing the growth of $Tsub_{var}$. In most regions of the North Atlantic, horizontal eddy diffusive flux of $Tsub_{var}$ is to depress the development of $Tsub_{var}$, the only exceptions are patchy areas in the center of subpolar gyre and along the North Atlantic Current. Unit:K^2 s^{-1}.

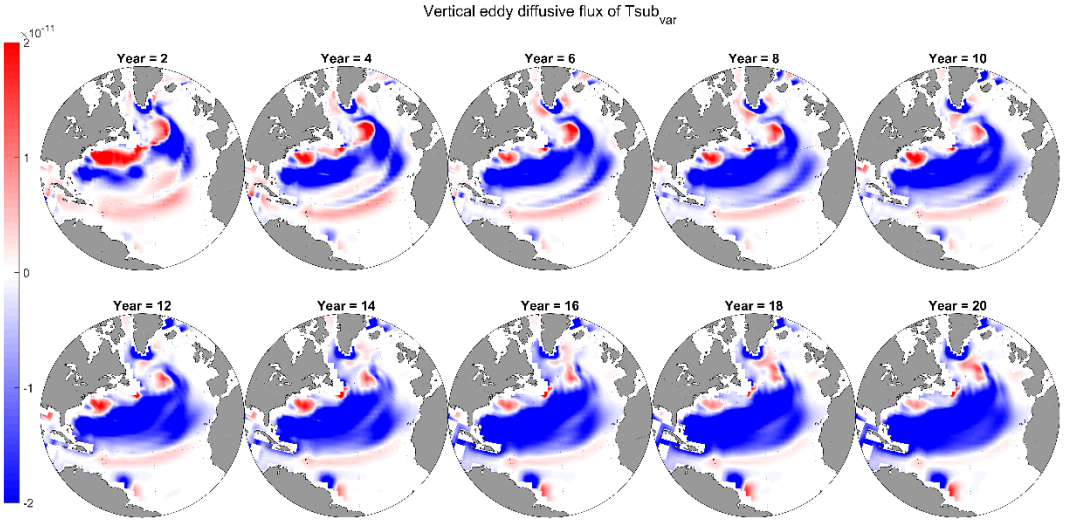

Figure 7. Temporal evolution of vertical eddy diffusive flux of $Tsub_{var}$ in the 40-member, 20-year ensemble run. Similar to its horizontal counterpart, during the twenty years, vertical eddy diffusive flux of $Tsub_{var}$ is mostly negative in the North Atlantic Ocean (Only patchy regions along the Gulf Stream and North Atlantic Current, and in the middle of sub-polar gyre are positive.). It indicates that vertical eddy diffusive flux of $Tsub_{var}$ is working to depress the $Tsub_{var}$ in most part of the subtropical North Atlantic Ocean. Note the difference of the color bar scales in Figure 6 and Figure 7, which indicates that the horizontal eddy diffusive flux of $Tsub_{var}$ is loosely two orders of magnitude larger than the vertical one. Unit: K^2 s^{-1}.

Likewise we calculate the vertical eddy diffusive flux of $Tsub_{var}$ as well to examine its effect (Figure 7). Basically it depresses the development in most regions of the North Atlantic, except that the fluxes on the north edges of the Gulf Stream and the North Atlantic Current are positive. A simple comparison of Figure 6 and 7 reveals that the horizontal eddy diffusive flux is approximately two orders of magnitude larger than the vertical diffusive flux, and roughly one order of magnitude larger than the left-hand-side term. This clearly indicates that the vertical eddy diffusive flux does not play the main role in Tsub prediction uncertainty development.

Even though we cannot directly calculate the three velocity-related terms, we can still infer their effects from a comparison of the left-hand-side term and the horizontal and vertical eddy diffusive fluxes (Figure 6 and 7). In most regions, eddy diffusive fluxes work to oppose the growth of $Tsub_{var}$, particularly along the Gulf Stream and the North Atlantic Current. It is in distinct contrast with the developing pattern of $Tsub_{var}$ (Figure 6 and 4). This suggests that the high-value region of three velocity-related terms as a whole should be along the Gulf Stream and the North Atlantic Current, to oppose the effect of eddy diffusive fluxes.

3.3 Understanding zonally averaged temperature prediction uncertainty and its association with AMOC

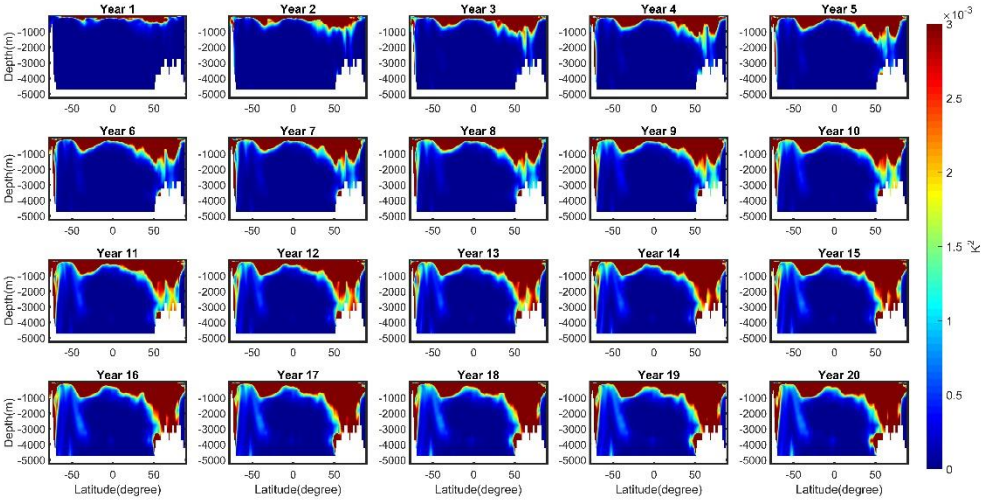

Figure 8. Ensemble variance (30 of the full 40 members excerpt the SST-perturbed sub-ensemble, because the zonally averaged temperature of the SST-perturbed sub-ensemble data is unavailable.) of zonally averaged temperature in the Atlantic Ocean in the 20-year ensemble run. Unit: K^2.

Now that we are already aware of the horizontal evolution pattern of temperature prediction uncertainty, we might be interested in its vertical propagation as well. Therefore, to investigate the development of temperature uncertainty in latitude-depth space, I calculate the ensemble variance of zonally averaged temperature in the Atlantic Ocean (Figure 8). Strong variance originates from ocean surface develops throughout the entire basin, then it penetrates downwards firstly through the wind-driven cell and then mainly in two sites: subpolar region in the North Atlantic (50 °-70 °N) and near Antarctic continental shelf. It is probably caused by deep convection process in these two sites. Surface anomalies induce deep and bottom water formation processes uncertainty and convey surface uncertainty downwards. This means that oceanic temperature within this depth is less sensitive to initial temperature perturbation, even though the perturbation is indeed three-dimensional.

To examine this hypothesis, I plot the Hovmöller diagram of zonally averaged temperature variance at depths of 5 m, 200 m, 400 m, 1000 m, 2000 m and 3000 m, in the northern North Atlantic Ocean (north of 40°N, Figure 9). Remarkable and coherent propagation modes exist in the sub-to-deep oceans. There are 2 distinct high-value region, one centred in around 73 °N, the other one around 60 °N. At 400m, the 73 °N-centered one feds away in years 12-13, meanwhile the 60 °N-centered one starts growing dramatically. The two patches are thought to be caused by different reasons. The 60 °N- centered one corresponds to a similar surface (at 5m depth) structure, indicating that this patch is probably a consequence of downward propagating surface temperature uncertainty. There is no such correspondance for the 73 °N-centered patch. This patch is likely to be a result of interior ocean process, such as local enhanced mixing at subsurface depth. It is not clear if the two high value patches are somewhat linked or not. In further investigation we will dig into this problem.

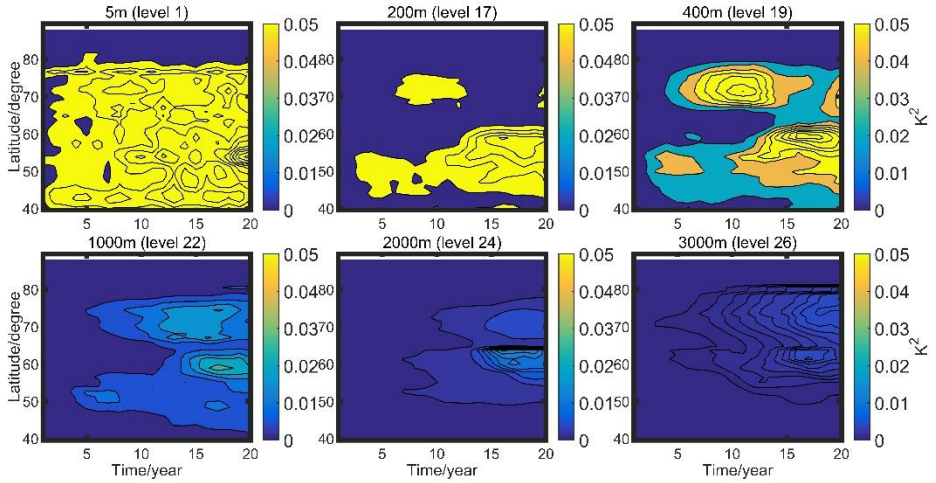

Figure 9. Hovmöller diagram of zonally averaged temperature variance at depths of 5m, 200m, 400m, 1000m, 2000m and 3000m, in the northern North Atlantic Ocean. Below 400m, there are 2 distinct high-value region, one centered in around 73 ºN, the other one around 60 ºN. At 400m, the 73 ºN-centered one feds away in years 12-13, meanwhile the 60 ºN-centered one starts growing dramatically. Unit: K^2.

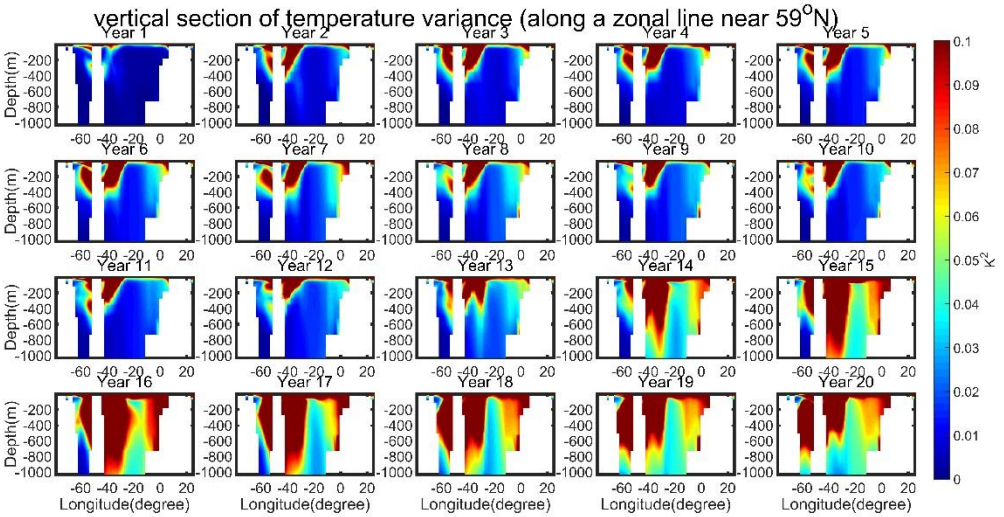

Figure 10. Vertical section of temperature variance along a zonal line near 59 ºN. The uncertainty develops initially in the surface of Irminger Sea, later on propagating into the subsurface depth (200-600m). In the late part of the 20 years, the largest uncertainty develops mainly along the continental shelves. Interestingly the downward propagation takes place mainly along continental shelves (i.e., Greenland Island shelf, Western European shelf and North American shelf). Unit: K^2.

From another perspective, the development of temperature uncertainty along a zonal grid line near 59 ºN is shown in Figure 10. Interestingly, the uncertainty originates at the surface and propagates

into subsurface depth gradually. In the late part of the 20 years, the largest uncertainty develops mainly along the continental shelves. This is consistent with the fact that subpolar mode water formation processes are mainly happening along continental shelves (Elena Brambilla et al. 2008); i.e., temperature uncertainty is transported downward by subpolar mode water formation along shelves. In the open ocean deep convection might be involved in the downward transportation of temperature uncertainty, but the contribution is significantly less than near-shore subpolar mode water formation processes.

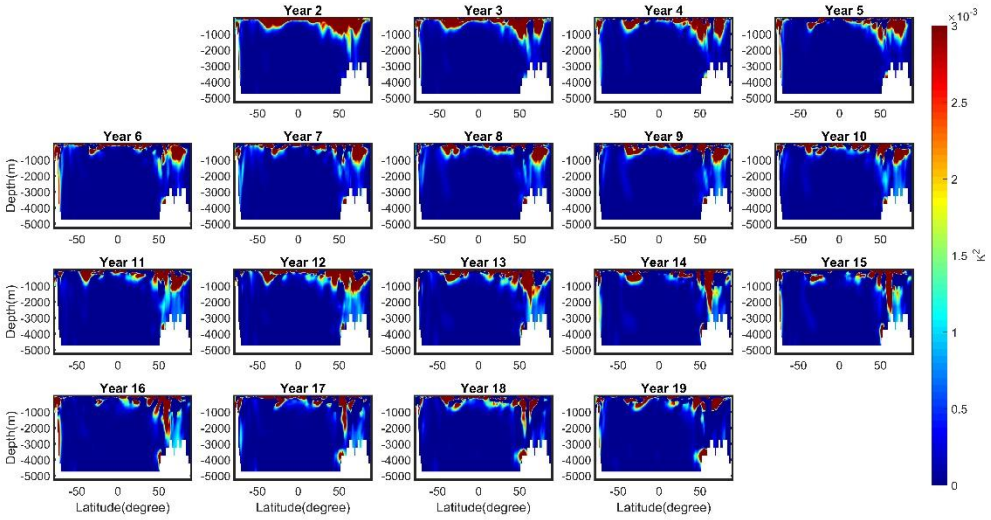

Figure 11. Zonally averaged temperature variance tendency in the 20-year ensemble run. Here the term is calculated using leap-frog scheme, with a time step length of 2 years. Unit: $K^2\ yr^{-1}$.

The tendency term of ensemble variance of zonally averaged temperature is displayed in Figure 11. Only patches of positive values can be identified mainly above -1000 m in the ex-polar regions. However we can still tell that strong surface-to-deep-ocean propagation of uncertainty occurs in the high latitude North Atlantic.

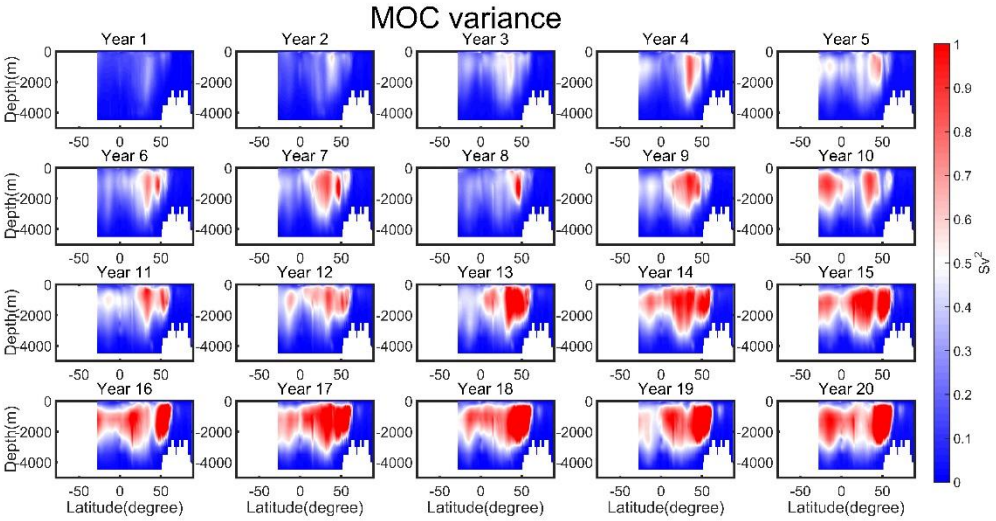

Figure 12. Evolution of ensemble variance of meridional overturning circulation in the 20 years. Large uncertainty develops mainly at sub-to-middle depth (from 100 to 2000m). Unit: Sv^2.

Now a question arises naturally that, what principally sets the pattern of temperature ensemble variance. Meanwhile we may wonder in latitude-depth space, what is the pattern of AMOC uncertainty growth, and whether or not there is resemblance and dynamical association between AMOC uncertainty pattern and temperature uncertainty pattern. To this end, I calculate the ensemble variance of meridional overturning circulation (MOC, Figure 12). As a matter of fact, the patterns of MOC bear little resemblance to that of temperature variance. A reminiscence of $Tsub_{var}$ shown in Figure 4 reminds us that, even the initial temperature perturbation has no particular spatial pattern at subsurface depth, the development of $Tsub_{var}$ is spatially non-uniform. The remarkable contrast between $Tsub_{var}$ along the Gulf Stream and North Atlantic Current and that in the interior in the early years corresponds to the difference of nonlinearity in the two regions. Velocity field along the Gulf Stream and North Atlantic Current has larger nonlinearity. One significant cause of this strong nonlinearity is the active eddy activities as a consequence of barotropic or baroclinic instabilities in these regions. In the interior subtropics (exclude boundary regions), however, the ocean is, on the first order, a consequence of linear geostrophic balance. This explains the non-uniformity of $Tsub_{var}$ developments. On the other hand, after amplified temperature uncertainty arrives at subpolar regions, it starts to expand into the deep ocean. In this way the temperature uncertainty probably intimately influences the overturning circulation, mostly likely *via* inducing uncertainty in the North Atlantic

Deep Water formation process.

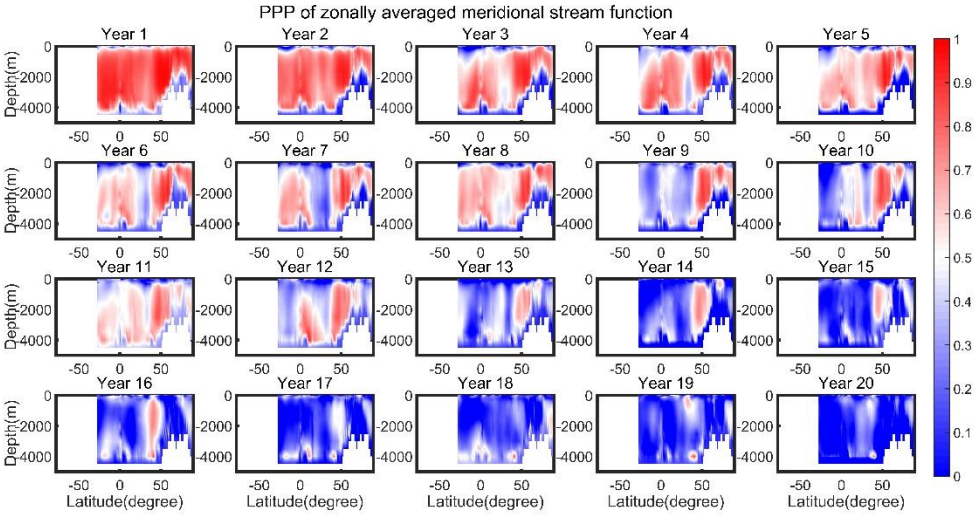

Figure 13. PPP of zonally averaged meridional stream function in the Atlantic Ocean in the 20-year ensemble run. Notably in the first year, low-value PPP are on ocean surface and at bottom. It implies that on shorter time scale (1-5 years) PPP of zonally averaged meridional stream function at subsurface-to-deep depth (except ocean bottom) is less sensitive to oceanic temperature perturbation than when it's at surface or bottom. Then the two low-value PPP patches develop downwards and upwards respectively, eroding strong PPP in the middle ocean. In the late period (year 16-20) the PPP within 50-70 ℕ drops close to zero and remains quite low. It indicates that the internal variation originating from initial temperature perturbation substantially contaminates the potential predictability of deep convection. PPP is dimensionless by definition.

Next, to investigate the potential association between PPP of zonally averaged meridional stream function and temperature spread in latitude-depth space, I compute PPP of zonally averaged meridional stream function in the Atlantic Ocean and display it in Figure 13. Notably in the first year, weak PPP are at ocean surface and at bottom. It indicates that on shorter time scale (1-5 years) zonally averaged meridional stream function at subsurface to deep depth (except ocean bottom) is less sensitive to oceanic temperature perturbation than when it's at surface or bottom. Later on the two low-value PPP patches develop downwards and upwards respectively, eroding strong PPP in the middle ocean. In the late period (years 16-20) the PPP within 50-70 ℕ drops close to zero and remains at a low level. It indicates that the internal variation originating from initial temperature perturbation substantially reduces the potential predictability of deep convection in the subpolar region. The large ensemble variance of

temperature in the subpolar region (years 10-20) in North Atlantic in Figure 8 corresponds well with the weakest PPP of zonally averaged stream function (years 16-20) shown in Figure 13.

It is not difficult to find that the ensemble variance of zonally averaged meridional overturning circulation in figure 15 is in contrast with PPP shown in Figure 13 (usually high ensemble variance is in consistency with low PPP). This paradox indicates that the climatological variance is large in regions with low PPP (and high ensemble variance) at sub-to-middle depth. The climatological variance is a consequence of both external and internal forcing, while ensemble variance is a result of internal variance only (ensemble variance is calculated at each time step, in this way external forcing is excluded).

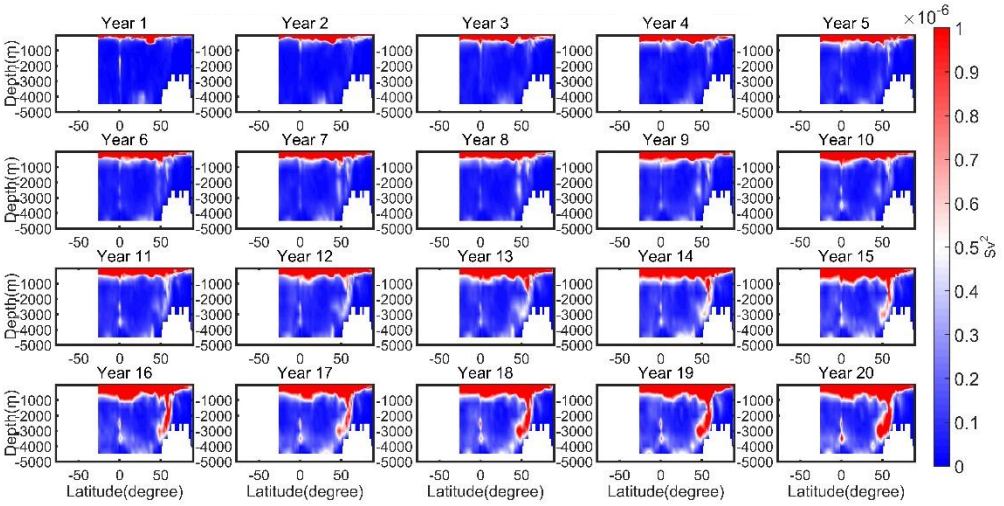

Figure 14. Evolution of ensemble variance of meridional overturning streamfunction in the 20 years. Large uncertainty develops mainly at ocean surface in the first 10 years. After that a vertical sheet of high value develops downward into deep ocean at around 60 ℃N. Unit: Sv^2.

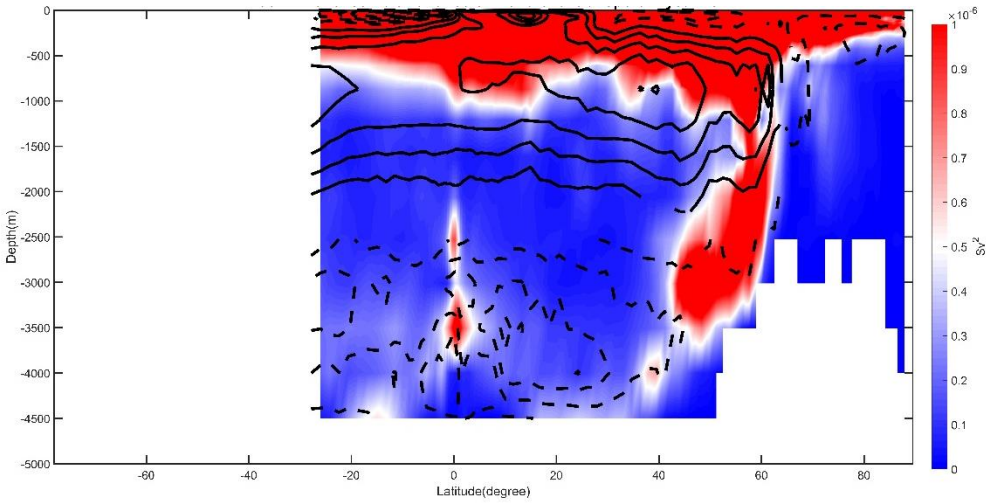

Figure 15. Zoom in of ensemble variance of overturing streamfunction in year 20, plus ensemble mean of zonally averaged overturning streamfunction. Dashed contour lines represent negative stream functions, and solid contour lines are positive stream functions. Shading stands for ensemble variance of streamfunction. Unit: Sv^2.

Figure 15 exhibits the ensemble variance of meridional overturning streamfunction in year 20, as well as mean overturning stream function. It can be found that there seems no obvious links between mean stream function and ensemble variance of meridional overturning streamfunction. This suggest that the development of meridional mass transport is not dominated by the ensemble mean signal.

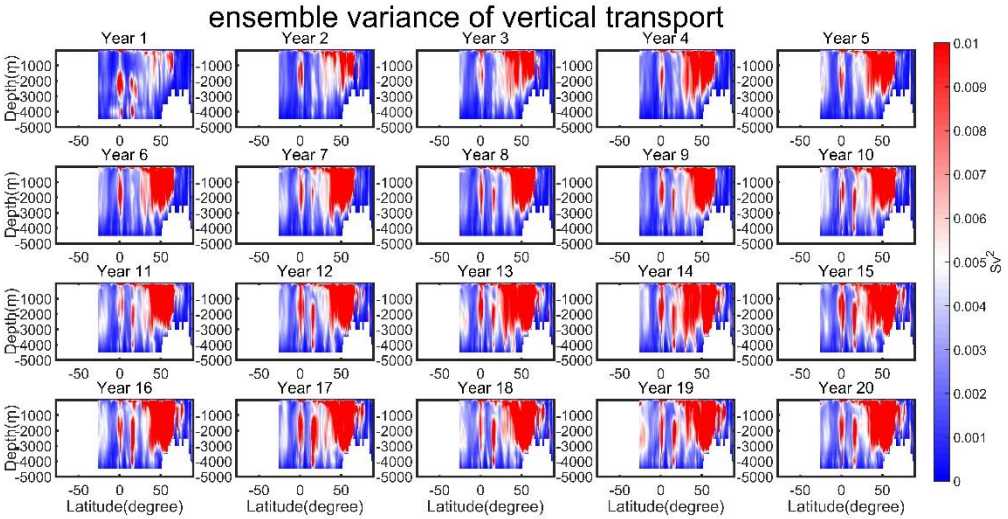

Figure 16. Same to Figure 14 but for the vertical velocity. Notably the ensemble variance of vertical velocity is 4 orders of magnitude larger than its' meridional counterpart. It rules out the possibility that the ensemble variance of meridional velocity shapes the pattern of

temperature uncertainty evolution. Unit: Sv^2.

Next, to identify whether the uncertainty of meridional or vertical transport is more important than the other one, I calculate the ensemble variance of both and display them in Figure 14 and 16, respectively. Remarkably the ensemble variance of vertical transport is 4 orders of magnitude larger than its meridional counterpart, which excludes the possibility that uncertainty in meridional transport shapes the evolution pattern of zonally averaged temperature. In addition, what's intereting is that, the patterns of ensemble variance of meridional overturning streamfunction (Figure 14) bears a resemblance to that of temperature variance (Figure 11).

In order to give an idea of how $Tsub_{var}$ in an ocean forced eddy-permitting model differs from that in a coupled laminar ocean model, we begin by calculating annual average of $Tsub_{var}$ from the output of both models (i.e., IPSL-CM5A-LR in used in Chapter 2/3 and NEMO3.4 used in this chapter) (Figure 17).

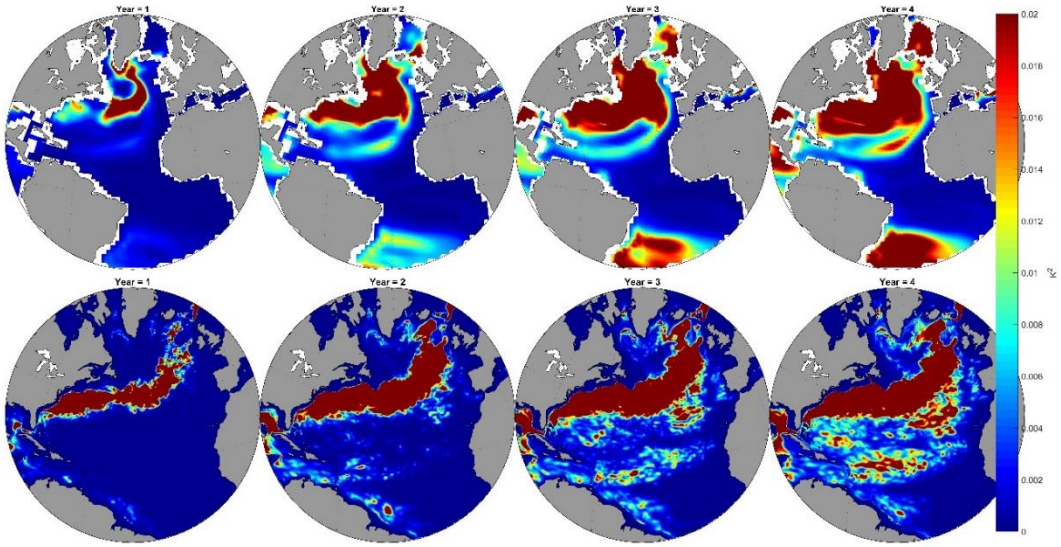

Figure 17. Comparison of subsurface temperature ensemble variance from the coupled laminar model (upper) and that from the forced eddy-permitting model (lower). Only four years are presented because the intermediate-resolution model data is merely available from 8th January to 15th December 2005. The main concensus between the two model output:1.high $Tsub_{var}$ along the Gulf Stream and the North Atlantic Current and low $Tsub_{var}$ in the southern part of the subtropical gyre. The main discrepency is that, after 2 years of integreation, SPG in the coupled laminar model is fullfilled with high-value $Tsub_{var}$, while in the forced eddy-permitting model, a diople pattern sustains throughout the four years.

In order to reveal evolution of $Tsub_{var}$ tendency, we first calculate left-hand-side term from 5-day average (Figure 18).

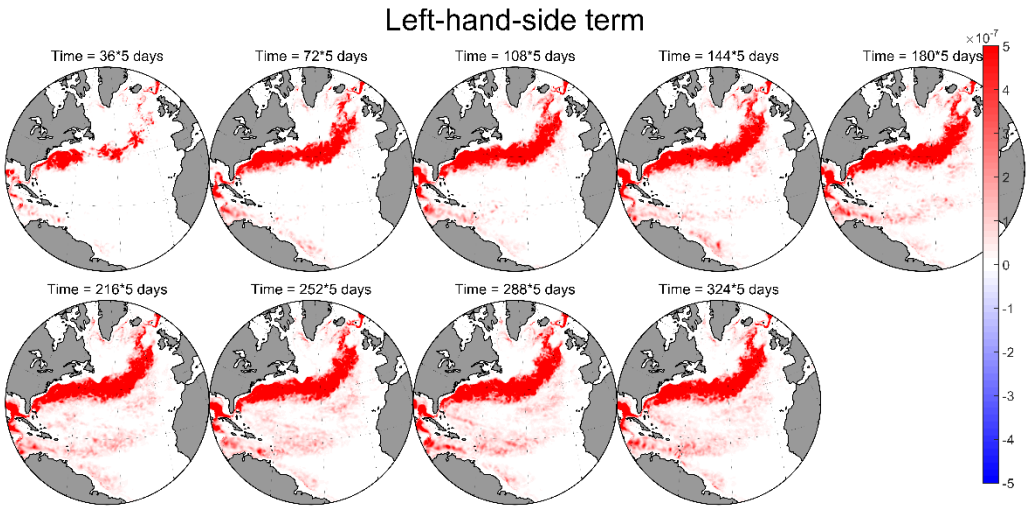

Figure 18. The left hand side of equation (8), namely tendency of $Tsub_{var}$. This term here is obtained from 5-day mean. High tendency is located mainly along the Gulf Stream and the North Atlantic Current. Unit: $K^2\ s^{-1}$.

High $Tsub_{var}$ tendency is mainly along the Gulf Stream and the North Atlantic Current, and spreading far into the Norwegian Sea. On both sides of the North Atlantic Current, the $Tsub_{var}$ tendency is much weaker. This is in contrast to the case in the coupled laminar model, where the tendency of $Tsub_{var}$ in the Labrador Sea is comparable to that in the Irminger Sea.

Next in order to explore the effect of advection term in equation (8), we then calculate it from the 5-day average (Figure 19).

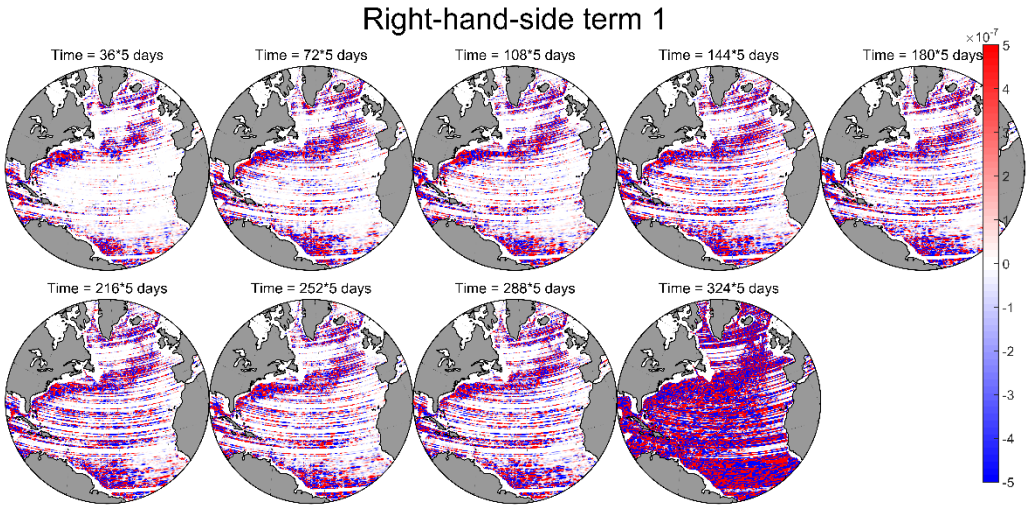

Figure 19. The first term on the right hand side of equation (8), namely the horizontal advection term. This term here is computed from 5-day mean. This term work both to damp and feed positively the growth of $Tsub_{var}$. Unit: $K^2\ s^{-1}$.

Interestingly the effect is not straightforward. The advection term feed both positively and negatively to the evolution of $Tsub_{var}$. There is no particular location for each of the both effects. That is to say, $Tsub_{var}$ both damp and favor the growth of $Tsub_{var}$ everywhere in the North Atlantic Ocean.

In order to quantify the effect of the generation term, the term 2 on the right-hand side is obtained from 5-day average (Figure 20).

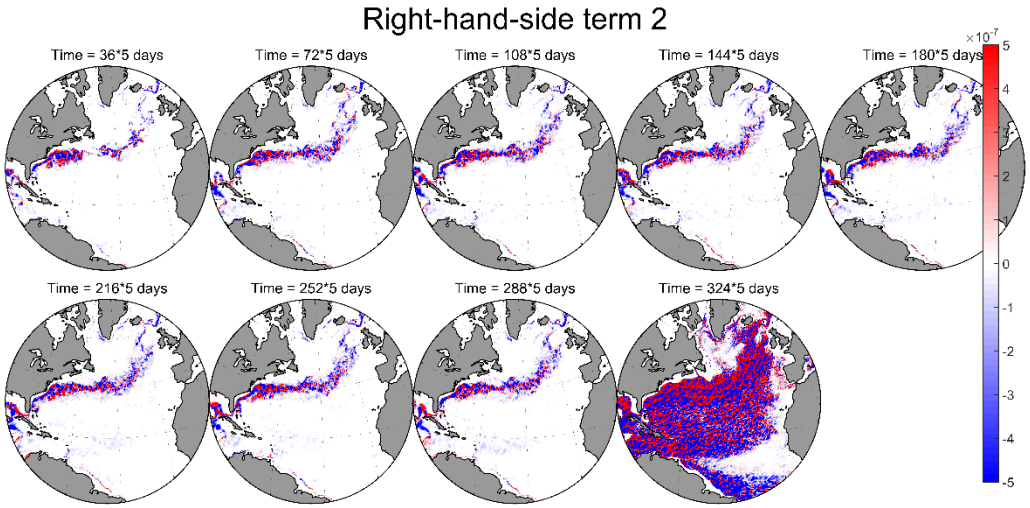

Figure 20. The second term on the right hand side of equation (8), namely the generation term. This term here is calculated from 5-day mean. This term contributes both positively and negatively to the $Tsub_{var}$ tendency. Unit: $K^2\ s^{-1}$.

Like the case of the advection term, the generation term oppose as well as flavor the Tsu_{var} growth. What is different is that, both effects are mainly along the Gulf Stream and the North Atlantic Current (the only exception is the case of 324×5 days).

To achieve a quantification of the self-advection, the third term on the right-hand side of equation (8) is then obtained from 5-day mean (Figure 21).

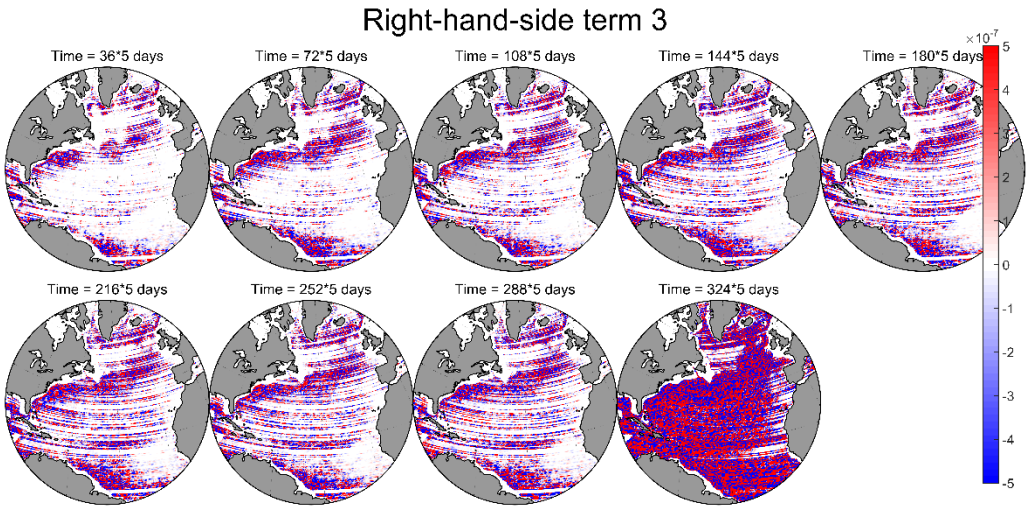

Figure 21. The third term on the right hand side of equation (8), namely the self-advection term. This term here is obtained from 5-day mean. This term work both to oppress and boost the growth of $Tsub_{var}$. Unit: $K^2\ s^{-1}$.

The self-advection term bears a resemblance to the advection term. It work both to damp and promote the development of $Tsub_{var}$, and both effects are at work everywhere in the North Atlantic Ocean.

It should be noted that even though each of the velocity-related terms has both damping and enhancing effects, the net contribution from every term is positive. Therefore the net effect of the three velocity-related terms is positive as well, i.e., the three terms work together to feed positively the growth of $Tsub_{var}$. This is clearly in qualitative consistency with the results from the coupled

laminar model.

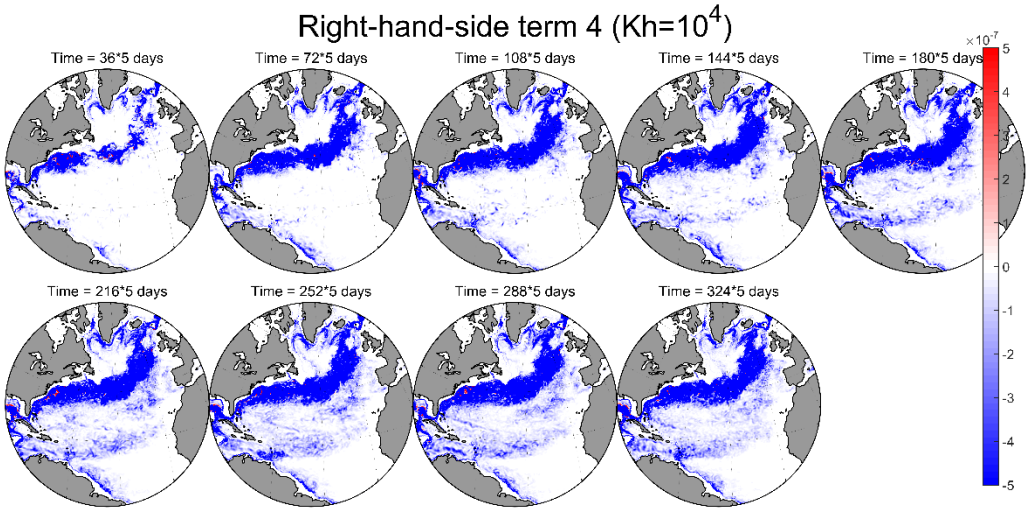

Figure 22. The fourth term on the right hand side of equation (8), namely the horizontal eddy diffusive flux. This term here is calculated from 5-day mean. Note that this term predominately contribute negatively to the growth of $Tusb_{var}$, with vanishingly few exceptions as scattered patches. Unit: $K^2 s^{-1}$.

Next, to clarify the contribution from eddy diffusive fluxes, the horizontal and vertical eddy diffusive fluxes are obtained from 5-day averages (Figure 22, 23). Unsurprisingly, the horizontal eddy diffusive flux mainly work to damp the $Tusb_{var}$, with negligibly few exceptions existing as scattered patches. High damping effect are along the Gulf Stream and the North Atlantic Current. This is qualitatively consistent with results from coupled laminar model. It should be noted that the magnitude of horizontal diffusive fluxes is 3 orders larger than that in Figure 6. This might be associated with the fact that temperature fluctuation on 5-day time scale is larger than on monthly time scale.

As for the vertical eddy diffusive flux, it predominate effect is also to suppress the growth of $Tsub_{var}$ (Figure 23). Likely, the damping are also along the Gulf Stream and the North Atlantic Current. What is distinctly different from its horizontal counterpart is that, on the north and west boundaries of the North Atlantic Current, vertical eddy diffusive flux contributes positively to the development of $Tsub_{var}$. This is qualitatively consistent with the result from a coupled laminar model in (Figure 7) even if the spatial extent is more restricted in the case of forced eddy-permitting model.

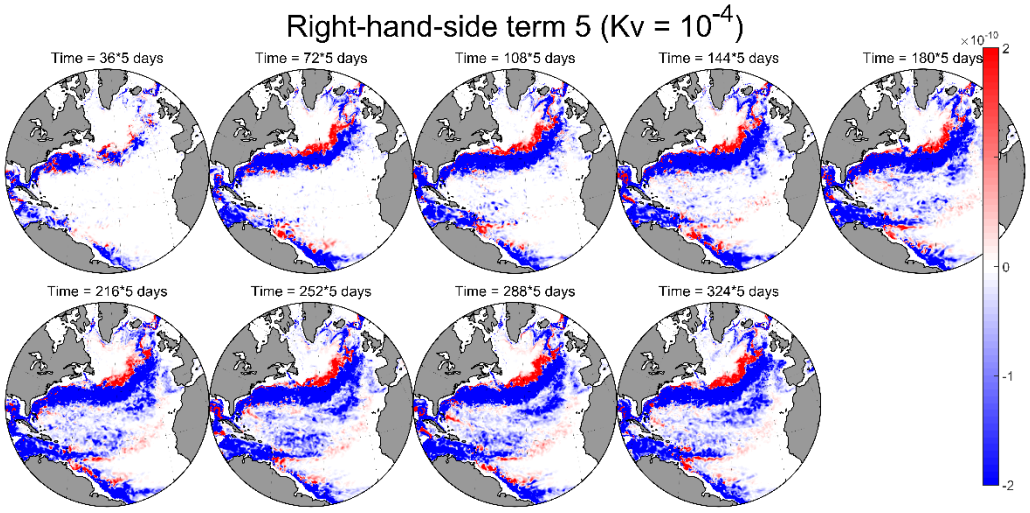

Figure 23. The fifth term on the right hand side of equation (8) in Chapter 3, namely the vertical eddy diffusive flux. This term here is calculated from 5-day mean. Note that this term mainly contribute negatively to the growth of $Tusb_{var}$, with exceptions merely on the north and west edges of North Atlantic Current and in scattered patches in other regions. Unit: $K^2 s^{-1}$.

Another interesting finding worth mentioning is a quasi-annual burst of subsurface temperature variance in the Labrador Sea (Figure B.1 in Appendix B). The whole life cycle is as follows: in the early spring, a high-value patch of temperature variance occurs in the middle of the Labrador Sea. This patch then spread into all directions in the form of an expanding cycle. Finally the cycle merge into near-shore high-value region in the northern and western edge of Labrador Sea, or is somehow damped in the eastern and southern border of the Labrador Sea. This whole evolution last for approximately 40 days ($\approx$ 6 weeks), and occurs in every spring of the 4 modelling years. As of this writing, this phenomenon has not been reported previously to the author's knowledge. However investigation into the underlying is beyond the scope of this thesis due mainly to the limitation of length.

4. Conclusions and discussions

Subsurface temperature in the North Atlantic is one of the key fingerprints of the AMOC (J Mignot et al. 2007; Rong Zhang 2008) and are connecting to variations of SST, upper ocean heat content and regional climate (G Shaffer et al. 2004; LM Frankcombe et al 2010) on multiple time scales.

Using subsurface temperature to construct the AMOC variation is documented to be skillful (Mahajan et al 2011). Besides, models with assimilated subsurface temperature field has better performance than those without (Ziwang Deng et al. 2010). It is therefore worthwhile to further investigate the factors that increase, propagate, and decrease Tsub prediction uncertainty. Here I use a forty-member ensemble and a 1000-year piControl run produced by IPSL-CM5A-LR to examine the development process of Tsub prediction uncertainty. A clockwise propagation mode of Tsub prediction uncertainty is found by diagnosing the PPP_{Tsub} evolution. Then I decompose the $Tsub_{var}$, which by nature is equivalent to opposite PPP, into six components. The first five terms on the right hand side come from tracer budget equations, the sixth term is due to unresolved sub-grid processes.

There are of course a few drawbacks in this study. For instance, despite our efforts of extracting valuable information from the reduced data, the fact that we lack velocity data makes it difficult to conduct a full investigation of all the terms in Equation (8). This attempt is then carried with the output from an eddy-permitting model. Furthermore, the data is produced by a single model, hence it would be interesting to examine the robustness of our finding against other model outputs in the future. Even though, our analysis provides a better understanding of the $Tsub_{var}$ development in the North Atlantic Ocean. This is beneficial on several aspects, for example, when designing the monitoring system of subsurface temperature, we can implement more equipment along Gulf Stream and North Atlantic Current, for the reason that largest uncertainty is arising from these regions. Another example is the reconstruction and use of upper ocean heat content: we can be clearer about the internally generated uncertainty range if we have a thorough understanding of the development of $Tsub_{var}$. In the late part of this Chapter, using an 11-member ensemble produced by a forced eddy-permitting ocean general circulation model (NEMO3.4), I have studied the subsurface temperature variance in the North Atlantic Ocean, with special interest in the effect of meso-scale eddies. We realize that some shortcomings are in this study. For instance, only one model is used in this study, therefore it would be desirable to examine the robustness of our findings using output from other models. Besides, it would be better if the integrated time of the ensemble were longer. The main findings in our study can be summarized as follows:

(1) The prediction uncertainty of subsurface temperature, represented by low-value PPP, is mainly consistent with the clockwise large-scale gyre in the subtropical North Atlantic Ocean. It originally emerges in the Gulf Stream region and North Atlantic Current region, then expands

and propagates southwestward to low-latitude subtropical North Atlantic. In the end of the twenty model years, Tsub prediction uncertainty has grown to be overwhelming in the entire subtropical North Atlantic Ocean.

(2) A zonal dipole pattern of Tsub prediction uncertainty exists in the subtropical North Atlantic. Uncertainty in the western subtropical North Atlantic, such western boundary region and interior of the subtropical gyre, is distinctly larger than in the east, especially the coastal regions along south Europe and northwestern Africa.

(3) The growth of $Tsub_{var}$, the physical nature of the prediction uncertainty, is decomposed into five different components. Among them horizontal and vertical eddy diffusive temperature variance fluxes work to depress $Tsub_{var}$. In contrast, first three terms on the right-hand side as a whole act to produce and propagate $Tsub_{var}$. These three terms are associated with velocity, implying that a more accurate prediction of velocity may contribute to reduce Tsub prediction uncertainty. Horizontal eddy diffusive flux is two orders of magnitude larger than vertical eddy flux, though both work to limit the growth of Tsub prediction uncertainty.

(4) The left-hand side is of the same magnitude than vertical eddy diffusive flux, while around two orders of magnitude smaller than the horizontal eddy diffusive flux and velocity-related terms as a whole. This indicates that the left-hand side of Equation (8), i.e. the $Tsub_{var}$ tendency, is by and large a consequence of the competition among the two larger terms (velocity-related terms as a whole and horizontal eddy diffusive flux).

(5) Downward propagation of temperature ensemble variance in the subpolar region of North Atlantic takes place mainly along continental shelves, i.e., Greenland Island shelf, Western European shelf and North American shelf. This is in accordance with the strong subduction in these regions.

(6) The triple terms work together to generate and propagate the temperature error.

(7) The meso-scale eddies damp the temperature variance. Three velocity-related terms work together to feed positively to the growth of $Tsub_{var}$. More specifically, the net effects of the advection term, the generation term and the self-advection term are all positive.

(8) It is found that the horizontal and vertical eddy diffusive fluxes mainly work to oppose the $Tusb_{var}$, and the high damping effect are along the Gulf Stream and the North Atlantic Current.

(9) A quasi-annual break-out of $Tsub_{var}$ in the Labrador Sea is reported for the first time in this

Chapter. This break-out appears in the spring and propagates as an enlarging cycle for around 6 weeks, finally disappearing on edge of the Labrador Sea. This is reckoned to be linked with atmospheric activities and sea-ice melting in the spring, however this argument needs further evidenced confirmation.

In addition to these findings, some shortcomings exist in this study. First of all, considering the limitation induced by our two data sources, it would be desirable to examine the robustness of our findings here, by analyzing outputs from various models in the future. In this way we could probably rule out the possibility that the findings are special cases in one particular model. Different details in models, for example different eddy parameterization schemes or different coordinate frameworks, may probably modify our findings to some degree. On the other hand, whether or not model resolution will impact the development of $Tsub_{var}$ remains an open question. These questions has been dealt with, at least partly. Consensuses and discrepancies of results of different models are anticipated to be explored and explained in future work as well.

As noted above, it would also be desirable to examine the robustness of these findings in the future, using outputs from other forced ocean or coupled climate models, such as those participating in CMIP5/6 (Coupled Model Inter-comparison Project phase 5/6).

Chapter 4 Conclusions and discussions

1. Overview of research content and findings

Subsurface temperature variation in North Atlantic Ocean is intimately associated with the changes of AMOC and climate over North Atlantic sector on various time scales (J Mignot et al. 2007; R Zhang 2008; R Msadek et al. 2010; S Mahajan et al. 2011; H Teng et al. 2011; CD Roberts et al. 2013). However, despite several valuable endeavors (MS Lozier et al. 2017; DA Smeed et al. 2018), we still lack an enduring, dedicated basin-wide observation of subsurface temperature in the Atlantic Ocean, besides Argo networks (not particularly design for this purpose of AMOC monitoring). Therefore numerical ensemble experiment remains a worthwhile approach to examine variations and predictability in ocean and climate sciences, though in some cases there might be biases in the predictability assessment using climate models (DM Smith et al. 2020). Besides, a key prerequisite for ensemble forecasting is to estimate the required ensemble size. So in this thesis, using 3 suites of ensemble experiments run at varying model resolutions, I investigate multiyear predictability of North Atlantic subsurface temperature, and the impact of ensemble size on prediction skill and assessment of initial-value predictability of North Atlantic Oceanic climate. To be more exact, the following questions are examined in this study:

In chapter 2:

(1) Whether or not can ensemble size change prediction skill and assessment of initial-value predictability (see methods) of SST, Tsub, and AMOC, respectively? If so, how?

(2) What will the consistency and discrepancy be like among the impacts of ensemble size on prediction skill and assessment of initial-value predictability of SST, Tsub, and AMOC? We have these findings:

(1) In an externally forced, conceptual SST model, which is an adapted version of Hasselmann (1976) model, the mean of ensemble mean approaches toward the theoretical value as ensemble size grows. The uncertainty between different realizations of ensemble means is reduced in the size-growing process.

(2) In terms of IPSL-CM5A-LR, ensemble size rising leads to a smooth drop of PPP, and a reduction

of uncertainties of PPPs. This anti-correlated relations between ensemble size and PPP, ensemble size and PPP uncertainty, don't dependent on the chosen variables (within the tested ones).(3) The PPP, PPP uncertainty and RMSD of AMOC no longer change remarkably as the size is no less than the 'critical size', which is loosely 15 to 20 for IPSL-CM5A-LR, while ~7 for the forced eddy-permitting model. Individual contribution of extra ensemble member to the change of PPP and PPP uncertainty drops dramatically as the size is over 'critical size'.

To summarize, we have answered all the scientific questions listed in Chapte 2. In addition, we also revealed through our analysis that the there is a different 'critical sizes' for the coarse ensemble and the eddy-permitting ensemble. This indicates the effect of oceanic eddies in the growth of internal uncertainty, and imply that one advantage of eddy-permitting ensemble forecasting is that it need less members to quantify the upper bound of internally generated uncertainty.

In the first part of chapter 3 the scientific questions are:

(1)What is the spatial pattern of the evolution of $Tsub_{var}$?

(2)What is dominating its evolution? What are the sources of $Tsub_{var}$ and what are the sinks?

(3)What are the relative significance of contribution of those factors impacting $Tsub_{var}$'s development?

In Chapter 3, we have also examined subsurface temperature ensemble variance by analyzing output from a forced eddy-permitting ocean model, to gain a better understanding of eddy's role in initial value predictability of subsurface temperature. The scientific questions in the second part of chapter 3 are:

(1) Is there any difference in the North Atlantic Tsubvar evolution in a coupled laminar ocean model and a forced eddy-permitting ocean model?

(2) How can the meso-scale eddies contribute the development of North Atlantic Tsubvar?

(3) Is the contribution of eddies dependent on time scales?

The main findings in Chapter 3 are:

(1) The prediction uncertainty of subsurface temperature, represented by low-value PPP, is mainly consistent with the clockwise large-scale gyre in the subtropical North Atlantic Ocean. It originally emerges in the Gulf Stream region and North Atlantic Current region, then expands and propagates

southwestward to low-latitude subtropical North Atlantic. In the end of the twenty model years, Tsub prediction uncertainty has grown to be overwhelming in the entire subtropical North Atlantic Ocean.

(2) Uncertainty in the western subtropical North Atlantic, such western boundary region and interior of the subtropical gyre, is distinctly larger than in the east, especially the coastal regions along south Europe and northwestern Africa.

(3) The growth of $Tsub_{var}$, the physical nature of the prediction uncertainty, is decomposed into five different components. Among them horizontal and vertical eddy diffusive temperature variance fluxes work to depress $Tsub_{var}$. In contrast, first three terms on the right-hand side as a whole act to produce and propagate $Tsub_{var}$. These three terms are associated with velocity, implying that a more accurate prediction of velocity may contribute to reduce Tsub prediction uncertainty. Horizontal eddy diffusive flux is two orders of magnitude larger than vertical eddy flux, though both work to oppose the growth of Tsub prediction uncertainty.

(4) The left-hand side is of the same magnitude of vertical eddy diffusive flux, while around two orders of magnitude smaller than horizontal eddy diffusive flux and velocity-related terms as a whole. This indicates that the left-hand side of Equation (8), i.e. the $Tsub_{var}$ tendency, is by and large a consequence of the competition among the two larger terms (velocity-related terms as a whole and horizontal eddy diffusive flux).

(5) Downward propagation of temperature ensemble variance in the subpolar region takes place mainly in along continental shelves, i.e., Greenland Island shelf, Western European shelf and North American shelf. This is in accordance with the strong convection in these regions.

In general, the work presented in Chapter 3 has answered the corresponding scientific questions except questions 3 of the second part of Chapter 3, due mainly to the sudden termination of the research. Some elementary work (not shown in this thesis) indicates time-depent of eddy's effects, however a quantitative evalutation is still needed.

Implications of findings in chapter 3:

In all, this thesis provides a quantification of the subsurface temperature variance development in the North Atlantic Ocean on inter-annual time scale. This is beneficial for a better understanding

of the predictability of North Atlantic Ocean state and of the AMOC.

2. Discussions of caveats and outlook of future research

2.1 What have been learnt?

In spite of our best effect to improve this study, there are several drawbacks and caveats in it. Firstly, since we use only 2 models in this study (one coupled climate model with a laminar ocean and one forced eddy-permitting ocean model), there is no guarantee that the results we get here are universally applicable to various GCMs (e.g., those models participating Coupled Model Intercomparison Project Phase 5/6, i.e., CMIP5/6), for instance, or even to the actual climate. Secondly, no ensemble size-model resolution relationship is obtained due to unavailability of datasets from models with various resolutions. Thirdly, initialization in each of the two GCMs is imposed on one single year (model-year 2056 for IPSL-CM5A-RL and actual-year 2001 for the forced eddy-permitting ocean model). We would anticipate our conclusions to be more robust if we have additional ensemble members initialized in varying years (such as 2057, 2058, etc. for IPSL-CM5A-LR and 2002, 2003, etc. for the forced eddy-permitting ocean model).

2.2 Some suggesitons

For those who are are building an AMOC preictions system, we would, based on our findings in this thesis, make some suggestions.

The first one is on the necessary ensmemem size, which is basically resolution dependent. As memtioned in Chapter 2, usually 15-20 members are enough for a non-eddy-permitting, laminar decadal prediction system. Larger ensembles seems unnecessary if computational resources are limited due mainly to substantially reduced gain by new added mmbers when size is over the 15-20 threshhold. If the forced eddy-permitting system is used to make a forecasting, we would suggest that to have approximately 7 members, and beyond that, no more members are suggested.

The second one is on the initialized approach for estimation of upper limit of subsurface temperature uncertainty on internal-annual to multi-annual time scales. Our analyses suggest that

western boundary and North Atlantic Currents regiosn are major sites where uncertainty originates, properly perturbe temperature in these regions could probably lead to better estimation of internal uncertainty of Tsub. From another perspective, targeted observations can be implemented in these 'sensitive regions' to reduce intial error, in order to diminish the consequent internal uncertainty.

2.3 Open questions

Also, despite our best efforts of extracting valuable information from the reduced data, the fact that we lack velocity data makes it difficult to conduct a full investigation of all the terms in Equation (8). This attempt is carried in the second part of Chapter 3 with the output from a forced eddy-permitting ocean model. However, the data is produced by a single model, hence it would be interesting to examine the robustness of our finding against other ocean or climate model output in the future.

In the future, it would be desirable to quantify the subsurface temperature prediction uncertainty in the sub-polar region using other initialized ensembles, such as those participating in CMIP5/6.

Appendix A Integral of adapted Hasselmann 1976 stochastic climate model

In chapter 2 a stochastic SST is used, which is an adapted version of famous Hasselmann 1976 model (Florian Sévellec & Bablu Sinha 2018; K Hasselmann 1976). Essentially, in that model, the red-noise power spectral characteristic of mixed layer temperature at mid-latitude is well reproduced by a white-noise forced slab ocean model (K Hasselmann 1976). The model reads,

$$\frac{\mathrm{d}\, SST_t}{\mathrm{d}t} = \left(SAT_t - SST_t\right)/\lambda \qquad \text{（A.1）}$$

where SST_t is instantaneous sea surface temperature anomaly, SAT_t instantaneous surface air temperature anomaly, λ the adjustment time scale of ocean (set to be 10 years, same to that in Florian Sévellec & Bablu Sinha 2018). The SAT_t here is set to be a zero-mean white noise with a variance of 0.5 K^2 (While in Florian Sévellec & Bablu Sinha 2018 it is set to be 1.6K^2).

Take an infinite integral on both sides of equation (1),

$$\int \frac{\mathrm{d}\, SST_t}{\mathrm{d}t}\mathrm{d}t = \int \left(SAT_t - SST_t\right)/\lambda \,\mathrm{d}t \qquad \text{(A.2)}$$

which is equivalent to

$$SST_t = (\int SAT_t \mathrm{d}t - \int SST_t \mathrm{d}t) / \lambda \qquad \text{(A.3)}$$

Since SAT_t is prescribed to be a zero-mean white noise, its infinite integral vanish. Therefore we have

$$SST_t = (0 - \int SST_t \mathrm{d}t) / \lambda \qquad \text{(A.4)}$$

This is equivalent to

$$SST_t + 1/(2\lambda) \cdot SST_t^2 = 0 \qquad \text{(A.5)}$$

There are two solutions of SST_t, 0 or -2λ. The latter is unreasonable in the real world, so the solution of SST_t (i.e., expectation of SST_t) must be zero.

In pragmatic calculation of Chapter 2, the integral length is 10000 years, which is long enough to be considered ‘infinite’. Therefore the expectation of instantaneous SST at year 10000 is zero.

Appendix B Some supporting figures

In this appendix, a few supporting figures for chapter 2 and 3 are presented.

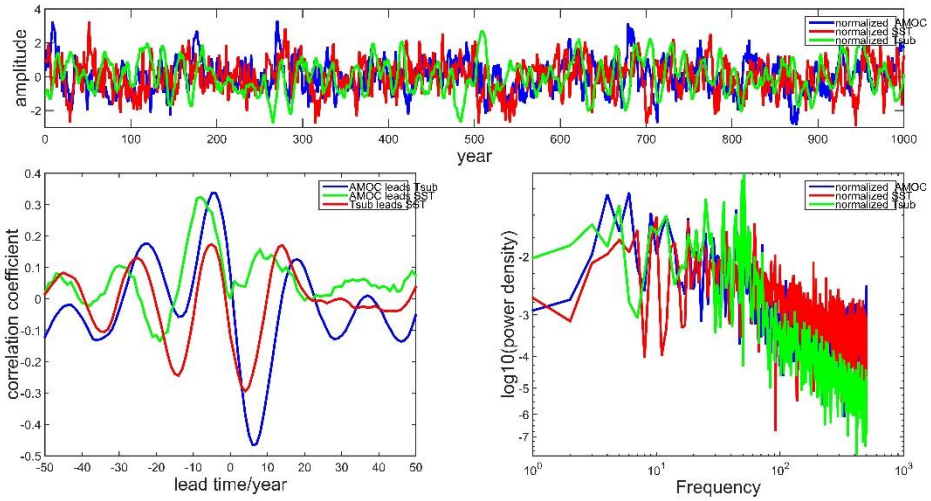

Figure B1. (a) Time series of normalized anomalies of annual mean AMOC, SST and Tsub in the North Atlantic Ocean. The anomalies are calculated by subtracting long-term mean from its corresponding original series. The normalization is carried out by dividing the anomaly by standard deviation. Blue line represents normalized AMOC anomaly; red line denotes normalized SST anomaly; normalized Tsub anomaly is shown in green line. These series by construction are dimensionless. (b) Lead-lag correlations among anomalies of AMOC, Tsub and SST. Blue line represents normalized AMOC anomaly; red line denotes normalized SST anomaly; normalized Tsub anomaly is shown in green line. These coefficients by construction is dimensionless. (c) Power spectrum of normalized AMOC, SST and Tsub. Blue line represents power spectra of normalized AMOC anomaly; red line denotes power spectra of normalized SST anomaly; power spectra of normalized Tsub anomaly is shown in green line. These power densities by construction are dimensionless.

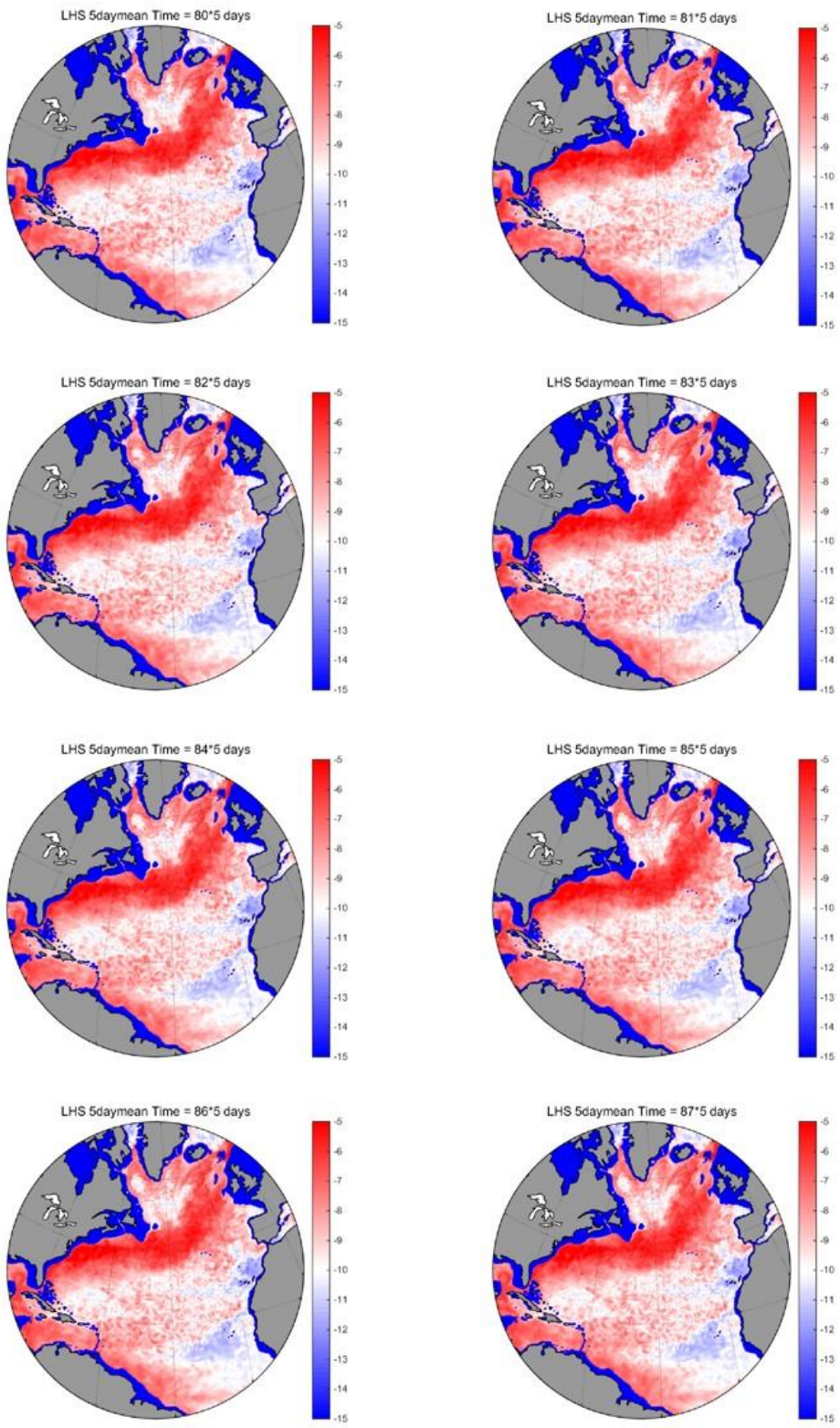

Figure B2. A life cycle of quasi-annual burst of subsurface temperature variance in the Labrador Sea. In the 80th 5-day mean, a high value patch occurs in the middle of the Labrador Sea. Afterwards it spreads in the form of a growing cycle. In the end of the life cycle it emerges on the western and northern edge of the Labrador Sea, or disappears (is damped) in the southern and eastern boundary of the Labrador Sea. Unit of temperature variance: $^{\circ}C^2$.

Bibliography

Alvarez-Solas, Jorge, and Gilles Ramstein. "On the triggering mechanism of Heinrich events." Proceedings of the National Academy of Sciences 108.50 (2011): E1359-E1360.

Amaya, Dillon J., et al. "WES feedback and the Atlantic Meridional Mode: observations and CMIP5 comparisons." Climate Dynamics 49.5-6 (2017): 1665-1679.

Andr é, et al. "Millennial-scale versus long-term dynamics in the surface and subsurface of the western North Atlantic Subtropical Gyre during Marine Isotope Stage 5." Global and planetary change 111 (2013): 77-87.

Arbic B K, M üller M, Richman J G, et al. Geostrophic turbulence in the frequency–wavenumber domain: Eddy-driven low-frequency variability[J]. Journal of physical oceanography, 2014, 44(8): 2050-2069.

Arbic B K, Scott R B, Flierl G R, et al. Nonlinear cascades of surface oceanic geostrophic kinetic energy in the frequency domain[J]. Journal of Physical Oceanography, 2012, 42(9): 1577-1600.

Arzel O, Huck T. Contributions of Atmospheric Stochastic Forcing and Intrinsic Ocean Modes to North Atlantic Ocean Interdecadal Variability[J]. Journal of Climate, 2020, 33(6): 2351-2370.

Arzel, O., T. Huck, and A. Colin de Verdi ère, 2006: The different nature of the interdecadal variability of the thermohaline circulation under mixed and flux boundary conditions. J. Phys. Oceanogr., 36, 1703-1718.Bahr, Andr é, et al. "Millennial-scale versus long-term dynamics in the surface and subsurface of the western North Atlantic Subtropical Gyre during Marine Isotope Stage 5." Global and planetary change 111 (2013): 77-87.

Bakker, Pepijn, et al. "Fate of the Atlantic Meridional Overturning Circulation: Strong decline under continued warming and Greenland melting." Geophysical Research Letters 43.23 (2016): 12-252.

Bengtsson L, Hodges K I. Can an ensemble climate simulation be used to separate climate change signals from internal unforced variability?[J]. Climate Dynamics, 2019, 52(5-6): 3553-3573.

Berner J, Achatz U, Batte L, et al. Stochastic parameterization: Toward a new view of weather and climate models[J]. Bulletin of the American Meteorological Society, 2017, 98(3): 565-588.

Bjerknes, Jf. "Atlantic air-sea interaction." Advances in geophysics. Vol. 10. Elsevier, 1964. 1-82.

Blaker, Adam T., et al. "Large near-inertial oscillations of the Atlantic meridional overturning circulation." Ocean Modelling 42 (2012): 50-56.

Boer G J. Long time-scale potential predictability in an ensemble of coupled climate models[J]. Climate dynamics, 2004, 23(1): 29-44.

Borchert L F, Müller W A, Baehr J. Atlantic ocean heat transport influences interannual-to-decadal surface temperature predictability in the north atlantic region[J]. Journal of Climate, 2018, 31(17): 6763-6782.

Boer, G. J. "A study of atmosphere-ocean predictability on long time scales." Climate Dynamics 16.6 (2000): 469-477.

Bradley A A, Schwartz S S, Hashino T. Sampling uncertainty and confidence intervals for the Brier score and Brier skill score[J]. Weather and Forecasting, 2008, 23(5): 992-1006.

Brady, Esther C., and Bette L. Otto-Bliesner. "The role of meltwater-induced subsurface ocean warming in regulating the Atlantic meridional overturning in glacial climate simulations." Climate dynamics 37.7-8 (2011): 1517-1532.Branstator G, Teng H. Two limits of initial-value decadal predictability in a CGCM[J]. Journal of climate, 2010, 23(23): 6292-6311.Branstator, Grant, et al. "Systematic estimates of initial-value decadal predictability for six AOGCMs." Journal of Climate 25.6 (2012): 1827-1846.

Branstator G, Teng H. Is AMOC more predictable than North Atlantic heat content?[J]. Journal of climate, 2014, 27(10): 3537-3550.

Branković, Čedo, and T. N. Palmer. "Atmospheric seasonal predictability and estimates of ensemble size." Monthly weather review 125.5 (1997): 859-874.

Buckley M W, Ponte R M, Forget G, et al. Low-frequency SST and upper-ocean heat content variability in the North Atlantic[J]. Journal of Climate, 2014, 27(13): 4996-5018.Buckley M W, DelSole T, Lozier M S, et al. Predictability of North Atlantic sea surface temperature and upper-ocean heat content[J]. Journal of Climate, 2019, 32(10): 3005-3023.

Buckley, M. W., D. Ferreira, J.-M. Campin, J. Marshall, and R. Tulloch, 2012: On the relationship between decadal buoyancy anomalies and variability of the Atlantic meridional overturning circulation. J. Climate, 25, 8009-8030.

Buckley, Martha W., and John Marshall. "Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review." Reviews of Geophysics 54.1 (2016): 5-63.

Buizert, Christo, and Andreas Schmittner. "Southern Ocean control of glacial AMOC stability and Dansgaard‐Oeschger interstadial duration." Paleoceanography 30.12 (2015): 1595-1612.

Buizza R, Petroliagis T, Palmer T, et al. Impact of model resolution and ensemble size on the performance of an ensemble prediction system[J]. Quarterly journal of the royal meteorological society, 1998, 124(550): 1935-1960.

Buizza R, Palmer T N. Impact of ensemble size on ensemble prediction[J]. Monthly Weather Review, 1998, 126(9): 2503-2518.

Caesar, Levke, et al. "Observed fingerprint of a weakening Atlantic Ocean overturning circulation." Nature 556.7700 (2018): 191.

Carton, James A., et al. "The Atlantic overturning circulation: More evidence of variability and links to climate." Bulletin of the American Meteorological Society 95.8 (2014): ES163-ES166.

Cessi, Paola, and Francesco Paparella. "Excitation of basin modes by ocean‐atmosphere coupling." Geophysical research letters 28.12 (2001): 2437-2440.

Chanut J, Barnier B, Large W, et al. Mesoscale eddies in the Labrador Sea and their contribution to convection and restratification[J]. Journal of Physical Oceanography, 2008, 38(8): 1617-1643.Cheng L, Zhu J. Benefits of CMIP5 multimodel ensemble in reconstructing historical ocean subsurface temperature variations[J]. Journal of Climate, 2016, 29(15): 5393-5416.

Cheung, Anson H., et al. "Comparison of low-frequency internal climate variability in CMIP5 models and observations." Journal of Climate 30.12 (2017): 4763-4776.

Cheng, Wei, John CH Chiang, and Dongxiao Zhang. "Atlantic meridional overturning circulation (AMOC) in CMIP5 models: RCP and historical simulations." Journal of Climate 26.18 (2013): 7187-7197.

Chiang, John CH, C-Y. Chang, and M. F. Wehner. "Long-term behavior of the Atlantic interhemispheric SST gradient in the CMIP5 historical simulations." Journal of Climate 26.21 (2013): 8628-8640.

Chylek, Petr, et al. "The Atlantic Multidecadal Oscillation as a dominant factor of oceanic influence on climate." Geophysical research letters 41.5 (2014): 1689-1697.

Clark A J, Kain J S, Stensrud D J, et al. Probabilistic precipitation forecast skill as a function of ensemble size and spatial scale in a convection-allowing ensemble[J]. Monthly weather review, 2011, 139(5): 1410-1418.

Clement, Amy, et al. "The Atlantic Multidecadal Oscillation without a role for ocean circulation." Science 350.6258 (2015): 320-324.

Collins, Matthew. "Climate predictability on interannual to decadal time scales: The initial value problem." Climate Dynamics 19.8 (2002): 671-692.

Collins, Mat, et al. "North Atlantic decadal predictability." Clivar Exchanges 28 (2003): 6-7.

Collins, M., et al. "Interannual to decadal climate predictability in the North Atlantic: a multimodel-ensemble study." Journal of climate 19.7 (2006): 1195-1203.

Collins, Matthew, and Bablu Sinha. "Predictability of decadal variations in the thermohaline circulation and climate." Geophysical Research Letters 30.6 (2003).

Collins, Matthew, and Myles R. Allen. "Assessing the relative roles of initial and boundary conditions in interannual to decadal climate predictability." Journal of Climate 15.21 (2002): 3104-3109.

Crivellari, Stefano, et al. "Thermal response of the western tropical Atlantic to slowdown of the Atlantic Meridional Overturning Circulation." Earth and Planetary Science Letters 519 (2019): 120-129.

Czaja, Arnaud, and Claude Frankignoul. "Observed impact of Atlantic SST anomalies on the North Atlantic Oscillation." Journal of Climate 15.6 (2002): 606-623.

Davini, Paolo, and Chiara Cagnazzo. "On the misinterpretation of the North Atlantic Oscillation in CMIP5 models." Climate dynamics 43.5-6 (2014): 1497-1511.

Davini, Paolo, Jost von Hardenberg, and Susanna Corti. "Tropical origin for the impacts of the Atlantic Multidecadal Variability on the Euro-Atlantic climate." Environmental Research Letters 10.9 (2015): 094010.

De Elia R, Laprise R, Denis B. Forecasting skill limits of nested, limited-area models: A perfect-model approach[J]. Monthly Weather Review, 2002, 130(8): 2006-2023.

DelSole, Timothy, and Michael K. Tippett. "Average predictability time. Part I: theory." Journal of the atmospheric sciences 66.5 (2009): 1172-1187.

Delworth, Thomas L., et al. "The central role of ocean dynamics in connecting the North Atlantic Oscillation to the extratropical component of the Atlantic Multidecadal Oscillation." Journal of Climate 30.10 (2017): 3789-3805.

Deppenmeier, Anna-Lena, Reindert J. Haarsma, and Wilco Hazeleger. "The Bjerknes feedback in the tropical Atlantic in CMIP5 models." Climate Dynamics 47.7-8 (2016): 2691-2707.

Desbruyères, D. G., et al. "Full‐depth temperature trends in the northeastern Atlantic through the

early 21st century." Geophysical Research Letters 41.22 (2014): 7971-7979.

Deser, Clara, et al. "Uncertainty in climate change projections: the role of internal variability." Climate dynamics 38.3-4 (2012): 527-546.

Deshayes, Julie, Ruth Curry, and Rym Msadek. "CMIP5 model intercomparison of freshwater budget and circulation in the North Atlantic." Journal of Climate 27.9 (2014): 3298-3317.

Drijfhout, Sybren S., David P. Marshall, and Henk A. Dijkstra. "Conceptual models of the wind-driven and thermohaline circulation." International Geophysics. Vol. 103. Academic Press, 2013. 257-282.

Drijfhout, Sybren, Geert Jan Van Oldenborgh, and Andrea Cimatoribus. "Is a decline of AMOC causing the warming hole above the North Atlantic in observed and modeled warming patterns?." Journal of Climate 25.24 (2012): 8373-8379.

Drijfhout, Sybren. "The relation between natural variations in ocean heat uptake and global mean surface temperature anomalies in CMIP5." Scientific reports 8.1 (2018): 7402.

Duan W, Mu M, Wang B. Conditional nonlinear optimal perturbations as the optimal precursors for El Nino–Southern Oscillation events[J]. Journal of Geophysical Research: Atmospheres, 2004, 109(D23).

Duchez, A., et al. "A new index for the Atlantic Meridional Overturning Circulation at 26° N." Journal of Climate 27.17 (2014): 6439-6455.

Dufresne, J-L., et al. "Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5." Climate Dynamics 40.9-10 (2013): 2123-2165.

Dunstone N, Smith D, Scaife A, et al. Skilful predictions of the winter North Atlantic Oscillation one year ahead[J]. Nature Geoscience, 2016, 9(11): 809-814.

Enfield D B, Mestas‐Nuñez A M, Trimble P J. The Atlantic multidecadal oscillation and its relation to rainfall and river flows in the continental US[J]. Geophysical Research Letters, 2001, 28(10): 2077-2080.

Escudier, Romain, Juliette Mignot, and Didier Swingedouw. "A 20-year coupled ocean-sea ice-atmosphere variability mode in the North Atlantic in an AOGCM." Climate dynamics 40.3-4 (2013): 619-636.

Evans J P, Ji F, Abramowitz G, et al. Optimally choosing small ensemble members to produce robust climate simulations[J]. Environmental Research Letters, 2013, 8(4): 044050.

Fernandes, Katia, et al. "Decadal covariability of Atlantic SSTs and western Amazon dry‐season hydroclimate in observations and CMIP5 simulations." Geophysical Research Letters 42.16 (2015): 6793-6801.Farneti, Riccardo, and Thomas L. Delworth. "The role of mesoscale eddies in the remote oceanic response to altered Southern Hemisphere winds." Journal of Physical Oceanography 40.10 (2010): 2348-2354.

Ferro C A T, Jupp T E, Lambert F H, et al. Model complexity versus ensemble size: allocating resources for climate prediction[J]. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2012, 370(1962): 1087-1099.

Florian Sévellec, Bablu Sinha. Predictability of Decadal Atlantic Meridional Overturning Circulation Variations. Oxford Encyclopedia of Climate Science, Oxford University Press, Oxford University Press, 2018, ff10.1093/acrefore/9780190228620.013.81ff. ffhal-02136510ff

Frajka-Williams, Eleanor, et al. "Atlantic Meridional Overturning Circulation: Observed transports and variability." Frontiers in Marine Science 6 (2019): 260.

Fraedrich, Klaus. "Estimating weather and climate predictability on attractors." Journal of the Atmospheric Sciences 44.4 (1987): 722-728.

Frankcombe, L. M., H. A. Dijkstra, and A. Von der Heydt. "Sub‐surface signatures of the Atlantic Multidecadal Oscillation." Geophysical Research Letters 35.19 (2008).Frankcombe, Leela M., Anna Von Der Heydt, and Henk A. Dijkstra. "North Atlantic multidecadal climate variability: An investigation of dominant time scales and processes." Journal of climate 23.13 (2010): 3626-3638.

Frankignoul, Claude, Guillaume Gastineau, and Young-Oh Kwon. "The influence of the AMOC variability on the atmosphere in CCSM3." Journal of Climate 26.24 (2013): 9774-9790.

Franzke, Christian, and Tim Woollings. "On the persistence and predictability properties of North Atlantic climate variability." Journal of Climate 24.2 (2011): 466-472.

Gastineau G, Mignot J, Arzel O, et al. North Atlantic Ocean Internal Decadal Variability: Role of the Mean State and Ocean‐Atmosphere Coupling[J]. Journal of Geophysical Research: Oceans, 2018, 123(8): 5949-5970.

Gent, Peter R., and James C. Mcwilliams. "Isopycnal mixing in ocean circulation models." Journal of Physical Oceanography 20.1 (1990): 150-155.

Gent, Peter R. "A commentary on the Atlantic meridional overturning circulation stability in climate models." Ocean Modelling 122 (2018): 57-66.

Germe, Agathe, et al. "On the robustness of near term climate predictability regarding initial state uncertainties." Climate dynamics 48.1-2 (2017): 353-366.

Germe, Agathe, et al. "The impacts of oceanic deep temperature perturbations in the North Atlantic on decadal climate variability and predictability." Climate dynamics 51.5-6 (2018): 2341-2357.

Goubanova, Katerina, et al. "Respective roles of remote and local wind stress forcings in the development of warm SST errors in the South-Eastern Tropical Atlantic in a coupled high-resolution model." Climate Dynamics 52.3-4 (2019): 1359-1382.

Griffies, Stephen M., and Kirk Bryan. "Predictability of North Atlantic multidecadal climate variability." Science 275.5297 (1997): 181-184.

Griffies, S. M., and K. Bryan. "A predictability study of simulated North Atlantic multidecadal variability." Climate dynamics 13.7-8 (1997): 459-487.

Grist, Jeremy P., et al. "Re-emergence of North Atlantic subsurface ocean temperature anomalies in a seasonal forecast system." Climate Dynamics (2019): 1-22.

Grötzner, A., et al. "Interannual to decadal predictability in a coupled ocean–atmosphere general circulation model." Journal of Climate 12.8 (1999): 2607-2624.

Gulev, Sergey K., et al. "North Atlantic Ocean control on surface heat flux on multidecadal timescales." Nature 499.7459 (2013): 464.

Hallberg R, Gnanadesikan A. An exploration of the role of transient eddies in determining the transport of a zonally reentrant current[J]. Journal of Physical Oceanography, 2001, 31(11): 3312-3330.

Han, Zhe, et al. "Simulation by CMIP5 models of the Atlantic Multidecadal Oscillation and its climate impacts." Advances in Atmospheric Sciences 33.12 (2016): 1329-1342.

Hasselmann, Klaus. "Stochastic climate models part I. Theory." tellus 28.6 (1976): 473-485.

Hawkins E, Robson J, Sutton R, et al. Evaluating the potential for statistical decadal predictions of sea surface temperatures with a perfect model approach[J]. Climate dynamics, 2011, 37(11-12): 2495-2509.

Haskins, Rosalind K., et al. "Temperature domination of AMOC weakening due to freshwater hosing in two GCMs." Climate Dynamics (2019): 1-14.

Hawkins, Ed, and Rowan Sutton. "Decadal predictability of the Atlantic Ocean in a coupled GCM: Forecast skill and optimal perturbations using linear inverse modeling." Journal of Climate 22.14

(2009): 3960-3978.

Hawkins E, Tietsche S, Day J J, et al. Aspects of designing and evaluating seasonal‐to‐interannual Arctic sea‐ice prediction systems[J]. Quarterly Journal of the Royal Meteorological Society, 2016, 142(695): 672-683.

Hawkins, Ed, et al. "Irreducible uncertainty in near-term climate projections." Climate Dynamics 46.11-12 (2016): 3807-3819.

Hazeleger, W., et al. "Predicting multiyear north atlantic ocean variability." Journal of Geophysical Research: Oceans 118.3 (2013): 1087-1098.

Hermanson, Leon, et al. "Forecast cooling of the Atlantic subpolar gyre and associated impacts." Geophysical research letters 41.14 (2014): 5167-5174.

Hernandez Almeida, Ivan, et al. "Subsurface North Atlantic warming as a trigger of rapid cooling events: evidence from the early Pleistocene (MIS 31–19)." Climate of the Past 11.4 (2015): 687-696.

Heuzé, Céline. "North Atlantic deep water formation and AMOC in CMIP5 models." Ocean Science 13.4 (2017): 609-622.

He C, Liu Z, Zhu J, et al. North Atlantic subsurface temperature response controlled by effective freshwater input in “Heinrich” events[J]. Earth and Planetary Science Letters, 2020, 539: 116247.

Herger N, Abramowitz G, Knutti R, et al. Selecting a climate model subset to optimise key ensemble properties[J]. Earth System Dynamics, 2018, 9(1): 135-151.

Hernandez Almeida, Ivan, et al. "Subsurface North Atlantic warming as a trigger of rapid cooling events: evidence from the early Pleistocene (MIS 31–19)." Climate of the Past 11.4 (2015): 687-696.

Heuzé, Céline, et al. "Changes in global ocean bottom properties and volume transports in CMIP5 models under climate change scenarios." Journal of Climate 28.8 (2015): 2917-2944.

Hirschi, J., et al. "A monitoring design for the Atlantic meridional overturning circulation." Geophysical Research Letters 30.7 (2003).

Hochet A, Huck T, Arzel O, et al. Direct temporal cascade of temperature variance in eddy-permitting simulations of multidecadal variability[J]. Journal of Climate, 2020, 33(21): 9409-9425.

Hochet, A., T. Huck, O. Arzel, F. Sévellec, M. Mazloff, B. Cornuelle, and A. Colin de Verdière, 2020: Direct temporal cascade of temperature variance in eddy-permitting simulations of

multidecadal variability, J. Climate, 33, 9409-9425.

Hofmann, Matthias, and Stefan Rahmstorf. "On the stability of the Atlantic meridional overturning circulation." Proceedings of the National Academy of Sciences (2009): pnas-0909146106.

Holland, David M., et al. "Acceleration of Jakobshavn Isbrae triggered by warm subsurface ocean waters." Nature geoscience 1.10 (2008): 659.

Houghton, R. W. "Subsurface quasi-decadal fluctuations in the North Atlantic." Journal of Climate 9.6 (1996): 1363-1373.

Huang, Boyin, et al. "AMOC variations in 1979–2008 simulated by NCEP operational ocean data assimilation system." Climate dynamics 38.3-4 (2012): 513-525.

Huang, Bohua, et al. "Influences of tropical–extratropical interaction on the multidecadal AMOC variability in the NCEP climate forecast system." Climate dynamics 39.3-4 (2012): 531-555.

Huang, Boyin, Jiang Zhu, and Haijun Yang. "Mechanisms of Atlantic Meridional Overturning Circulation (AMOC) variability in a coupled ocean-atmosphere GCM." Advances in Atmospheric Sciences 31.2 (2014): 241-251.

Huddart, Benjamin, et al. "Seasonal and decadal forecasts of Atlantic Sea surface temperatures using a linear inverse model." Climate Dynamics 49.5-6 (2017): 1833-1845.

Hurrell, James W., et al. "Atlantic climate variability and predictability: A CLIVAR perspective." Journal of Climate 19.20 (2006): 5100-5121.

Hurrell J W, Kushnir Y, Ottersen G, et al. An overview of the North Atlantic oscillation[J]. Geophysical Monograph-American Geophysical Union, 2003, 134: 1-36.

Jackson, L. C., et al. "Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM." Climate dynamics 45.11-12 (2015): 3299-3316.

Jackson, L. C., Robin S. Smith, and R. A. Wood. "Ocean and atmosphere feedbacks affecting AMOC hysteresis in a GCM." Climate Dynamics 49.1-2 (2017): 173-191.

Jackson, Laura C., and Richard A. Wood. "Timescales of AMOC decline in response to fresh water forcing." Climate dynamics 51.4 (2018): 1333-1350.

Jacobs, Zoe L., et al. "A sub‐annual sub‐surface pathway from the Gulf Stream to the Subpolar Gyre and its role in warming and salinification in the 1990s." Geophysical Research Letters (2019).

Johannessen O M, Korablev A, Miles V, et al. Interaction between the warm subsurface Atlantic water in the Sermilik Fjord and Helheim glacier in southeast Greenland[J]. Surveys in geophysics,

2011, 32(4-5): 387.

Katsman, Caroline A., et al. "Sinking of dense North Atlantic waters in a global ocean model: Location and controls." Journal of Geophysical Research: Oceans 123.5 (2018): 3563-3576.

Keenlyside, Noel S., et al. "North Atlantic multi-decadal variability—mechanisms and predictability." Climate change: Multidecadal and beyond. 2016. 141-157.

Kim, Jung-Hyun, et al. "Pronounced subsurface cooling of North Atlantic waters off Northwest Africa during Dansgaard–Oeschger interstadials." Earth and Planetary Science Letters 339 (2012): 95-102.

Klavans, Jeremy M., Amy C. Clement, and Mark A. Cane. "Variable External Forcing Obscures the Weak Relationship between the NAO and North Atlantic Multidecadal SST Variability." Journal of Climate 32.13 (2019): 3847-3864.

Knight J R, Folland C K, Scaife A A. Climate impacts of the Atlantic multidecadal oscillation[J]. Geophysical Research Letters, 2006, 33(17).

Kostov, Yavor, Kyle C. Armour, and John Marshall. "Impact of the Atlantic meridional overturning circulation on ocean heat storage and transient climate change." Geophysical Research Letters 41.6 (2014): 2108-2116.

Kozar, Michael E., and Vasubandhu Misra. "Evaluation of twentieth-century Atlantic Warm Pool simulations in historical CMIP5 runs." Climate dynamics 41.9-10 (2013): 2375-2391.

Latif, Mojib, et al. "Reconstructing, monitoring, and predicting multidecadal-scale changes in the North Atlantic thermohaline circulation with sea surface temperature." Journal of Climate 17.7 (2004): 1605-1614.

Latif, Mojib, et al. "A review of predictability studies of Atlantic sector climate on decadal time scales." Journal of Climate 19.23 (2006): 5971-5987.

Latif, Mojib, Klaus Arpe, and Erich Roeckner. "Oceanic control of decadal North Atlantic sea level pressure variability in winter." Geophysical Research Letters 27.5 (2000): 727-730.

Latif, Mojib, and Noel S. Keenlyside. "A perspective on decadal climate variability and predictability." Deep Sea Research Part II: Topical Studies in Oceanography 58.17-18 (2011): 1880-1894.

Latif, Mojib, Taewook Park, and Wonsun Park. "Decadal Atlantic Meridional Overturning Circulation slowing events in a climate model." Climate Dynamics 53.1-2 (2019): 1111-1124.

Leutbecher M. Ensemble size: How suboptimal is less than infinity?[J]. Quarterly Journal of the Royal Meteorological Society, 2019, 145: 107-128.

Levang, Samuel J., and Raymond W. Schmitt. "What causes the AMOC to weaken in CMIP5?." Journal of Climate 2019 (2019).

Lima, L. N., et al. "An Investigation of Ocean Model Uncertainties Through Ensemble Forecast Experiments in the Southwest Atlantic Ocean." Journal of Geophysical Research: Oceans 124.1 (2019): 432-452.

Liu, Hailong, et al. "Atlantic warm pool variability in the CMIP5 simulations." Journal of Climate 26.15 (2013): 5315-5336.

Liu, Wei, et al. "Overlooked possibility of a collapsed Atlantic Meridional Overturning Circulation in warming climate." Science Advances 3.1 (2017): e1601666.

Liu, Wei, Alexey Fedorov, and Florian Sévellec. "The mechanisms of the Atlantic meridional overturning circulation slowdown induced by Arctic sea ice decline." Journal of Climate 32.4 (2019): 977-996.

Liu, Yang, and Peter J. Minnett. "Evidence linking satellite-derived sea-surface temperature signals to changes in the Atlantic meridional overturning circulation." Remote Sensing of Environment 169 (2015): 150-162.

Liu Z, Di Lorenzo E. Mechanisms and predictability of Pacific decadal variability[J]. Current Climate Change Reports, 2018, 4(2): 128-144.

Lohmann, Gerrit, and Joachim Schneider. "Dynamics and predictability of Stommel's box model. A phase‐space perspective with implications for decadal climate variability." Tellus A 51.2 (1999): 326-336.

Long, Shang‐Min, and Shang‐Ping Xie. "Intermodel variations in projected precipitation change over the North Atlantic: Sea surface temperature effect." Geophysical Research Letters 42.10 (2015): 4158-4165.

Lozier M S, Bacon S, Bower A S, et al. Overturning in the Subpolar North Atlantic Program: A new international ocean observing system[J]. Bulletin of the American Meteorological Society, 2017, 98(4): 737-752.

Lu R, Dong B, Ding H. Impact of the Atlantic Multidecadal Oscillation on the Asian summer monsoon[J]. Geophysical Research Letters, 2006, 33(24).

Lucas-Picher, Philippe, et al. "Investigation of regional climate models' internal variability with a ten-member ensemble of 10-year simulations over a large domain." Climate dynamics 31.7-8 (2008): 927-940.

Luo, Feifei, et al. "The connection between the Atlantic multidecadal oscillation and the Indian summer monsoon in CMIP5 models." Climate dynamics 51.7-8 (2018): 3023-3039.

Lyu, Kewei, and Jin‐Yi Yu. "Climate impacts of the Atlantic Multidecadal Oscillation simulated in the CMIP5 models: A re‐evaluation based on a revised index." Geophysical Research Letters 44.8 (2017): 3867-3876.

Madec, Gurvan, and Maurice Imbard. "A global ocean mesh to overcome the North Pole singularity." Climate Dynamics 12.6 (1996): 381-388.

Mahajan S, Zhang R, Delworth T L, et al. Predicting Atlantic meridional overturning circulation (AMOC) variations using subsurface and surface fingerprints[J]. Deep Sea Research Part II: Topical Studies in Oceanography, 2011, 58(17-18): 1895-1903.

Mahajan, Salil, Rong Zhang, and Thomas L. Delworth. "Impact of the Atlantic meridional overturning circulation (AMOC) on Arctic surface air temperature and sea ice variability." Journal of Climate 24.24 (2011): 6573-6581.

Manabe S, Bryan K, Spelman M J. Transient response of a global ocean-atmosphere model to a doubling of atmospheric carbon dioxide[J]. Journal of Physical Oceanography, 1990, 20(5): 722-749.

Marcott, Shaun A., et al. "Ice-shelf collapse from subsurface warming as a trigger for Heinrich events." Proceedings of the National Academy of Sciences 108.33 (2011): 13415-13419.

Marotzke, Jochem. "Abrupt climate change and thermohaline circulation: Mechanisms and predictability." Proceedings of the National Academy of Sciences 97.4 (2000): 1347-1350.

Marotzke, Jochem. "Quantifying the irreducible uncertainty in near‐term climate projections." Wiley Interdisciplinary Reviews: Climate Change 10.1 (2019): e563.

Marsh, R., et al. "Short-term impacts of enhanced Greenland freshwater fluxes in an eddy-permitting ocean model." Ocean Science 6.3 (2010): 749-760.

Martin, Elinor R., Chris Thorncroft, and Ben BB Booth. "The multidecadal Atlantic SST—Sahel rainfall teleconnection in CMIP5 simulations." Journal of Climate 27.2 (2014): 784-806.

Mart ń-Rey, Marta, and Alban Lazar. "Is the boreal spring tropical Atlantic variability a precursor

of the Equatorial Mode?." Climate Dynamics 53.3-4 (2019): 2339-2353.

Martin, Thomas, Annika Reintges, and Mojib Latif. "Coupled North Atlantic Subdecadal Variability in CMIP5 Models." Journal of Geophysical Research: Oceans 124.4 (2019): 2404-2417.

Masina, Simona, et al. "An ensemble of eddy-permitting global ocean reanalyses from the MyOcean project." Climate Dynamics 49.3 (2017): 813-841.

Mathieu, P. P., et al. "Predictability of winter climate over the North Atlantic European region during ENSO events." Journal of Climate 17.10 (2004): 1953-1974.

McCarthy, G. D., et al. "Measuring the Atlantic meridional overturning circulation at 26 N." Progress in Oceanography 130 (2015): 91-111.

McCarthy, Gerard D., et al. "Sustainable observations of the AMOC: Methodology and Technology." Reviews of Geophysics (2019).

Mecking, J. V., et al. "Stable AMOC off state in an eddy-permitting coupled climate model." Climate Dynamics 47.7-8 (2016): 2455-2470.

Menary, Matthew B., et al. "Exploring the impact of CMIP5 model biases on the simulation of North Atlantic decadal variability." Geophysical Research Letters 42.14 (2015): 5926-5934.

Menary, Matthew B., and Richard A. Wood. "An anatomy of the projected North Atlantic warming hole in CMIP5 models." Climate dynamics 50.7-8 (2018): 3063-3080.

Mendlik, Thomas, and Andreas Gobiet. "Selecting climate simulations for impact studies based on multivariate patterns of climate change." Climatic change 135.3-4 (2016): 381-393.

Mignot J, Ganopolski A, Levermann A. Atlantic subsurface temperatures: Response to a shutdown of the overturning circulation and consequences for its recovery[J]. Journal of Climate, 2007, 20(19): 4884-4898.

Mignot, Juliette, et al. "On the evolution of the oceanic component of the IPSL climate models from CMIP3 to CMIP5: A mean state comparison." Ocean Modelling 72 (2013): 167-184.

Molinari, Robert L., et al. "Multiyear variability in the near‐surface temperature structure of the midlatitude western North Atlantic Ocean." Journal of Geophysical Research: Oceans 102.C2 (1997): 3267-3278.

Morioka Y, Doi T, Storto A, et al. Role of subsurface ocean in decadal climate predictability over the South Atlantic[J]. Scientific reports, 2018, 8(1): 8523.

Msadek R, Dixon K W, Delworth T L, et al. Assessing the predictability of the Atlantic meridional

overturning circulation and associated fingerprints [J]. Geophysical Research Letters, 2010, 37(19).

Mu M, Duan W, Wang B. Season‐dependent dynamics of nonlinear optimal error growth and El Niño‐Southern Oscillation predictability in a theoretical model [J]. Journal of Geophysical Research: Atmospheres, 2007, 112(D10).

Muir, L. C., and A. V. Fedorov. "How the AMOC affects ocean temperatures on decadal to centennial timescales: the North Atlantic versus an interhemispheric seesaw." Climate Dynamics 45.1-2 (2015): 151-160.

Muir, Les C., and Alexey V. Fedorov. "Evidence of the AMOC interdecadal mode related to westward propagation of temperature anomalies in CMIP5 models." Climate dynamics 48.5-6 (2017): 1517-1535.

Murphy, James M. "The impact of ensemble forecasts on predictability." Quarterly Journal of the Royal Meteorological Society 114.480 (1988): 463-493.

Nicolis, C. "Chaotic dynamics, Markov processes and climate predictability." Tellus A 42.4 (1990): 401-412.

O'Reilly, Christopher H., et al. "The signature of low‐frequency oceanic forcing in the Atlantic Multidecadal Oscillation." Geophysical Research Letters 43.6 (2016): 2810-2818.

O'Reilly C H, Weisheimer A, Woollings T, et al. The importance of stratospheric initial conditions for winter North Atlantic Oscillation predictability and implications for the signal‐to‐noise paradox[J]. Quarterly Journal of the Royal Meteorological Society, 2019, 145(718): 131-146.

Ortega, Pablo, et al. "Reconciling two alternative mechanisms behind bi-decadal variability in the North Atlantic." Progress in oceanography 137 (2015): 237-249.

Otterå, Odd Helge, et al. "External forcing as a metronome for Atlantic multidecadal variability." Nature Geoscience 3.10 (2010): 688-694.

Pacanowski, R. C., and S. G. H. Philander. "Parameterization of vertical mixing in numerical models of tropical oceans." Journal of Physical Oceanography 11.11 (1981): 1443-1451.

Parker, Andrew O., Matthew W. Schmidt, and Ping Chang. "Tropical North Atlantic subsurface warming events as a fingerprint for AMOC variability during Marine Isotope Stage 3." Paleoceanography 30.11 (2015): 1425-1436.

Patrizio C R, Thompson D W J. Quantifying the Role of Ocean Dynamics in Ocean Mixed-Layer Temperature Variability [J]. Journal of Climate, 2021: 1-63.

Peings, Yannick, Graham Simpkins, and Gudrun Magnusdottir. "Multidecadal fluctuations of the North Atlantic Ocean and feedback on the winter climate in CMIP5 control simulations." Journal of Geophysical Research: Atmospheres 121.6 (2016): 2571-2592.

Penland, Cécile, and Prashant D. Sardeshmukh. "The optimal growth of tropical sea surface temperature anomalies." Journal of climate 8.8 (1995): 1999-2024.

Pennell, Christopher, and Thomas Reichler. "On the effective number of climate models." Journal of Climate 24.9 (2011): 2358-2367.

Pielke Sr, R. A. "Climate prediction as an initial value problem." Bulletin of the American Meteorological Society 79.12 (1998): 2743.

Pohlmann, Holger, et al. "Estimating the decadal predictability of a coupled AOGCM." Journal of Climate 17.22 (2004): 4463-4472.Polyakov, Igor V., et al. "Multidecadal variability of North Atlantic temperature and salinity during the twentieth century." Journal of Climate 18.21 (2005): 4562-4581.

Polyakov, Igor V., et al. "North Atlantic warming: patterns of long-term trend and multidecadal variability." Climate Dynamics 34.2-3 (2010): 439-457.

Ray, Sulagna, et al. "Effect of surface restoring on subsurface variability in a climate model during 1949–2005." Climate dynamics 44.9-10 (2015): 2333-2349.

Richardson D S. Measures of skill and value of ensemble prediction systems, their interrelationship and the effect of ensemble size [J]. Quarterly Journal of the Royal Meteorological Society, 2001, 127(577): 2473-2489.

Reintges, Annika, et al. "Uncertainty in twenty-first century projections of the Atlantic Meridional Overturning Circulation in CMIP3 and CMIP5 models." Climate Dynamics 49.5-6 (2017): 1495-1511.

Roberts C D, Garry F K, Jackson L C. A multimodel study of sea surface temperature and subsurface density fingerprints of the Atlantic meridional overturning circulation [J]. Journal of Climate, 2013, 26(22): 9155-9174.

Rowell, David P. "Assessing potential seasonal predictability with an ensemble of multidecadal GCM simulations." Journal of Climate 11.2 (1998): 109-120.

Ruiz-Barradas, Alfredo, Sumant Nigam, and Argyro Kavvada. "The Atlantic Multidecadal Oscillation in twentieth century climate simulations: uneven progress from CMIP3 to CMIP5."

Climate dynamics 41.11-12 (2013): 3301-3315.

Ruiz-Barradas A, Chafik L É, Nigam S, et al. Recent subsurface North Atlantic cooling trend in context of Atlantic decadal-to-multidecadal variability [J]. Tellus A: Dynamic Meteorology and Oceanography, 2018, 70(1): 1-19.Sanderson, Benjamin M., and Reto Knutti. "On the interpretation of constrained climate model ensembles." Geophysical Research Letters 39.16 (2012).

Saunders, Mark A., and Budong Qian. "Seasonal predictability of the winter NAO from North Atlantic sea surface temperatures." Geophysical Research Letters 29.22 (2002): 6-1.

Saravanan, R., and James C. McWilliams. "Advective ocean–atmosphere interaction: An analytical stochastic model with implications for decadal variability." Journal of Climate 11.2 (1998): 165-188.

Scaife A A, Smith D. A signal-to-noise paradox in climate science [J]. npj Climate and Atmospheric Science, 2018, 1(1): 1-8.Scaife A A, Camp J, Comer R, et al. Does increased atmospheric resolution improve seasonal climate predictions? [J]. Atmospheric Science Letters, 2019, 20(8): e922.

Schleussner, Carl-Friedrich, et al. "The role of the North Atlantic overturning and deep ocean for multi-decadal global-mean-temperature variability." Earth System Dynamics 5.1 (2014): 103-115.

Schmidt, Matthew W., et al. "Impact of abrupt deglacial climate change on tropical Atlantic subsurface temperatures." Proceedings of the National Academy of Sciences (2012).

Screen, James A., et al. "Atmospheric impacts of Arctic sea-ice loss, 1979–2009: Separating forced change from atmospheric internal variability." Climate dynamics 43.1-2 (2014): 333-344.

Seidov, Dan, et al. "Eddy‐Resolving In Situ Ocean Climatologies of Temperature and Salinity in the Northwest Atlantic Ocean." Journal of Geophysical Research: Oceans 124.1 (2019): 41-58.

Sérazin G, Penduff T, Barnier B, et al. Inverse cascades of kinetic energy as a source of intrinsic variability: A global OGCM study [J]. Journal of Physical Oceanography, 2018, 48(6): 1385-1408.

Sévellec, Florian, and Alexey V. Fedorov. "Stability of the Atlantic meridional overturning circulation and stratification in a zonally averaged ocean model: Effects of freshwater flux, Southern Ocean winds, and diapycnal diffusion." Deep Sea Research Part II: Topical Studies in Oceanography 58.17-18 (2011): 1927-1943.

Sévellec, Florian, and Alexey V. Fedorov. "The leading, interdecadal eigenmode of the Atlantic meridional overturning circulation in a realistic ocean model." Journal of Climate 26.7 (2013): 2160-2183.

Sévellec, Florian, and Alexey V. Fedorov. "Optimal excitation of AMOC decadal variability: links to the subpolar ocean." Progress in Oceanography 132 (2015): 287-304.

Sévellec, Florian, Alexey V. Fedorov, and Wei Liu. "Arctic sea-ice decline weakens the Atlantic meridional overturning circulation." Nature Climate Change 7.8 (2017): 604.

Sévellec, Florian, and Alexey V. Fedorov. "Predictability and decadal variability of the North Atlantic ocean state evaluated from a realistic ocean model." Journal of Climate 30.2 (2017): 477-498.

Sévellec, Florian, and Bablu Sinha. "Predictability of decadal atlantic meridional overturning circulation variations." (2018).

Sévellec F, Drijfhout S S. The Signal‐to‐Noise Paradox for Interannual Surface Atmospheric Temperature Predictions[J]. Geophysical Research Letters, 2019, 46(15): 9031-9041.

Sévellec F, Garabato A C N, Huck T. Damping of climate-scale oceanic variability by mesoscale eddy turbulence [J]. Journal of Physical Oceanography, 2021, 51(2): 491-503.

Sévellec, Florian, Thierry Huck, and Mahdi Ben Jelloul. "On the mechanism of centennial thermohaline oscillations." Journal of Marine Research 64.3 (2006): 355-392.

Sgubin, Giovanni, et al. "Abrupt cooling over the North Atlantic in modern climate models." Nature communications 8 (2017): 14375.

Shaffer, Gary, Steffen Malskær Olsen, and Christian J. Bjerrum. "Ocean subsurface warming as a mechanism for coupling Dansgaard‐Oeschger climate cycles and ice‐rafting events." Geophysical Research Letters 31.24 (2004).Shukla, Jagadish. "Predictability in the midst of chaos: A scientific basis for climate forecasting." science 282.5389 (1998): 728-731.

Sienz, Frank, Wolfgang A. Müller, and Holger Pohlmann. "Ensemble size impact on the decadal predictive skill assessment." Meteorologische Zeitschrift 25 (2016): 645-655.

Sinha, B., et al. "The accuracy of estimates of the overturning circulation from basin-wide mooring arrays." Progress in Oceanography 160 (2018): 101-123.

Smagorinsky, Joseph. "General circulation experiments with the primitive equations: I. The basic experiment." Monthly weather review 91.3 (1963): 99-164.

Smeed D A, Josey S A, Beaulieu C, et al. The North Atlantic Ocean is in a state of reduced overturning [J]. Geophysical Research Letters, 2018, 45(3): 1527-1533.

Smith D M, Scaife A A, Eade R, et al. North Atlantic climate far more predictable than models

imply [J]. Nature, 2020, 583(7818): 796-800.

Steinman, Byron A., Michael E. Mann, and Sonya K. Miller. "Atlantic and Pacific multidecadal oscillations and Northern Hemisphere temperatures." Science 347.6225 (2015): 988-991.

Stommel, Henry. "Thermohaline convection with two stable regimes of flow." Tellus 13.2 (1961): 224-230.

Stouffer, Ronald J., et al. "CMIP5 scientific gaps and recommendations for CMIP6." Bulletin of the American Meteorological Society 98.1 (2017): 95-105.

Sun, Cheng, et al. "Contrasting spatial structures of Atlantic Multidecadal Oscillation between observations and slab ocean model simulations." Climate Dynamics 52.3-4 (2019): 1395-1411.

Sutton, Rowan T., and Buwen Dong. "Atlantic Ocean influence on a shift in European climate in the 1990s." Nature Geoscience 5.11 (2012): 788.

Sutton, R. T., and Michael R. Allen. "Decadal predictability of North Atlantic sea surface temperature and climate." Nature 388.6642 (1997): 563.

Swingedouw, Didier, et al. "Impact of transient freshwater releases in the Southern Ocean on the AMOC and climate." Climate dynamics 33.2-3 (2009): 365-381.

Swingedouw D, Mignot J, Labetoulle S, et al. Initialisation and predictability of the AMOC over the last 50 years in a climate model [J]. Climate dynamics, 2013, 40(9-10): 2381-2399.

Tandon, Neil F., and Paul J. Kushner. "Does external forcing interfere with the AMOC's influence on North Atlantic sea surface temperature?." Journal of Climate 28.16 (2015): 6309-6323.

Teng H, Branstator G. Initial-value predictability of prominent modes of North Pacific subsurface temperature in a CGCM [J]. Climate dynamics, 2011, 36(9-10): 1813-1834.

Teng H, Branstator G, Meehl G A. Predictability of the Atlantic overturning circulation and associated surface patterns in two CCSM3 climate change ensemble experiments[J]. Journal of climate, 2011, 24(23): 6054-6076.

Thomas, M. D., and X. Zhai. "Eddy‐induced variability of the meridional overturning circulation in a model of the North Atlantic." Geophysical Research Letters 40.11 (2013): 2742-2747.

Tiedje B, Köhl A, Baehr J. Potential predictability of the North Atlantic heat transport based on an oceanic state estimate [J]. Journal of climate, 2012, 25(24): 8475-8486.

Timmermann, Axel, et al. "The influence of a weakening of the Atlantic meridional overturning circulation on ENSO." Journal of climate 20.19 (2007): 4899-4919.

Ting, Mingfang, et al. "Robust features of Atlantic multi‐decadal variability and its climate impacts." Geophysical Research Letters 38.17 (2011).

Teng H, Branstator G. Initial-value predictability of prominent modes of North Pacific subsurface temperature in a CGCM [J]. Climate dynamics, 2011, 36(9-10): 1813-1834.

Terray, Laurent. "Evidence for multiple drivers of North Atlantic multi‐decadal climate variability." Geophysical Research Letters 39.19 (2012).

Toniazzo, Thomas, and Steve Woolnough. "Development of warm SST errors in the southern tropical Atlantic in CMIP5 decadal hindcasts." Climate dynamics 43.11 (2014): 2889-2913.

Tréguier, Anne-Marie, Isaac M. Held, and V. D. Larichev. "Parameterization of quasigeostrophic eddies in primitive equation ocean models." Journal of Physical Oceanography 27.4 (1997): 567-580.

Trenary, Laurie, and Timothy DelSole. "Does the Atlantic Multidecadal Oscillation get its predictability from the Atlantic Meridional Overturning circulation?." Journal of Climate 29.14 (2016): 5267-5280.

Tziperman, Eli, Laure Zanna, and Cecile Penland. "Nonnormal thermohaline circulation dynamics in a coupled ocean–atmosphere GCM." Journal of Physical Oceanography 38.3 (2008): 588-604.Von Storch H, Zwiers F W. Statistical analysis in climate research [M]. Cambridge university press, 2001.

Visbeck, Martin. "The ocean's role in Atlantic climate variability." Science 297.5590 (2002): 2223-2224.

Wang, Bin, et al. "Northern Hemisphere summer monsoon intensified by mega-El Niño/southern oscillation and Atlantic multidecadal oscillation." Proceedings of the National Academy of Sciences 110.14 (2013): 5347-5352.

Wang, Chunzai, and Liping Zhang. "Multidecadal ocean temperature and salinity variability in the tropical North Atlantic: Linking with the AMO, AMOC, and subtropical cell." Journal of Climate 26.16 (2013): 6137-6162.

Weaver, Andrew J., et al. "Stability of the Atlantic meridional overturning circulation: A model intercomparison." Geophysical Research Letters 39.20 (2012).

Weaver, Andrew J., and E. S. Sarachik. "The role of mixed boundary conditions in numerical models of the ocean's climate." Journal of physical oceanography 21.9 (1991): 1470-1493.

Weijer, Wilbert, et al. "Stability of the Atlantic Meridional Overturning Circulation: A review and synthesis." Journal of Geophysical Research: Oceans 124.8 (2019): 5336-5375.

Wen, Na, Claude Frankignoul, and Guillaume Gastineau. "Active AMOC–NAO coupling in the IPSL-CM5A-MR climate model." Climate Dynamics 47.7-8 (2016): 2105-2119.

Williamson, Mark S., et al. "Effect of AMOC collapse on ENSO in a high resolution general circulation model." Climate dynamics 50.7-8 (2018): 2537-2552.

Wills R C J, Armour K C, Battisti D S, et al. Ocean–atmosphere dynamical coupling fundamental to the Atlantic multidecadal oscillation [J]. Journal of Climate, 2019, 32(1): 251-272.

Wouters, B., et al. "Multiyear predictability of the North Atlantic subpolar gyre." Geophysical Research Letters 40.12 (2013): 3080-3084.

Wu X, Yan X H, Jo Y H, et al. Estimation of subsurface temperature anomaly in the North Atlantic using a self-organizing map neural network [J]. Journal of Atmospheric and Oceanic Technology, 2012, 29(11): 1675-1688.

Xu, Xiaobiao, Eric P. Chassignet, and Fuchang Wang. "On the variability of the Atlantic meridional overturning circulation transports in coupled CMIP5 simulations." Climate dynamics 52.11 (2019): 6511-6531.

Yan, Xiaoqin, Rong Zhang, and Thomas R. Knutson. "The role of Atlantic overturning circulation in the recent decline of Atlantic major hurricane frequency." Nature communications 8.1 (2017): 1695.

Yan, Xiaoqin, Rong Zhang, and Thomas R. Knutson. "Underestimated AMOC variability and implications for AMV and predictability in CMIP models." Geophysical Research Letters 45.9 (2018): 4319-4328.

Yan, Xiaoqin, Rong Zhang, and Thomas R. Knutson. "A multivariate AMV index and associated discrepancies between observed and CMIP5 externally forced AMV." Geophysical Research Letters 46.8 (2019): 4421-4431.

Yang, Qian, et al. "Recent increases in Arctic freshwater flux affects Labrador Sea convection and Atlantic overturning circulation." Nature communications 7 (2016): 10525.

Yeager S G, Danabasoglu G, Rosenbloom N A, et al. Predicting near-term changes in the Earth System: A large ensemble of initialized decadal prediction simulations using the Community Earth System Model[J]. Bulletin of the American Meteorological Society, 2018, 99(9): 1867-1886.

Zanna, Laure. "Forecast skill and predictability of observed Atlantic sea surface temperatures." Journal of Climate 25.14 (2012): 5047-5056.

Zanna L, Heimbach P, Moore A M, et al. Upper‐ocean singular vectors of the North Atlantic climate with implications for linear predictability and variability[J]. Quarterly Journal of the Royal Meteorological Society, 2012, 138(663): 500-513.

Zappa, Giuseppe, et al. "A multimodel assessment of future projections of North Atlantic and European extratropical cyclones in the CMIP5 climate models." Journal of Climate 26.16 (2013): 5846-5862.

Zhang, Liping, and Chuanhu Zhao. "Processes and mechanisms for the model SST biases in the N orth A tlantic and N orth P acific: A link with the A tlantic meridional overturning circulation." Journal of Advances in Modeling Earth Systems 7.2 (2015): 739-758.

Zhang, Liping, and Chunzai Wang. "Multidecadal North Atlantic sea surface temperature and Atlantic meridional overturning circulation variability in CMIP5 historical simulations." Journal of Geophysical Research: Oceans 118.10 (2013): 5772-5791.

Zhang, Liping, Thomas L. Delworth, and Fanrong Zeng. "The impact of multidecadal Atlantic meridional overturning circulation variations on the Southern Ocean." Climate Dynamics 48.5-6 (2017): 2065-2085.

Zhang R. Anticorrelated multidecadal variations between surface and subsurface tropical North Atlantic [J]. Geophysical Research Letters, 2007, 34(12).

Zhang, Rong, and Thomas L. Delworth. "Impact of the Atlantic multidecadal oscillation on North Pacific climate variability." Geophysical Research Letters 34.23 (2007).

Zhang, Rong. "Latitudinal dependence of Atlantic meridional overturning circulation (AMOC) variations." Geophysical Research Letters 37.16 (2010).

Zhang, Rong, et al. "A Review of the Role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and Associated Climate Impacts." Reviews of Geophysics (2019)

Zhang R. Coherent surface‐subsurface fingerprint of the Atlantic meridional overturning circulation[J]. Geophysical Research Letters, 2008, 35(20).

Zhang, Rong, et al. "Have aerosols caused the observed Atlantic multidecadal variability?." Journal of the Atmospheric Sciences 70.4 (2013): 1135-1144.

Zhang, Rong, et al. "Comment on “The Atlantic Multidecadal Oscillation without a role for ocean circulation”." Science 352.6293 (2016): 1527-1527.

Zhang, Shaoqing, et al. "Impact of enthalpy-based ensemble filtering sea ice data assimilation on decadal predictions: simulation with a conceptual pycnocline prediction model." Journal of Climate 26.7 (2013): 2368-2378.

Zhang W, Kirtman B. Estimates of decadal climate predictability from an interactive ensemble model[J]. Geophysical Research Letters, 2019, 46(6): 3387-3397.

Zhang W, Kirtman B. Understanding the Signal‐to‐Noise Paradox with a Simple Markov Model[J]. Geophysical Research Letters, 2019, 46(22): 13308-13317.

Zhao J. Basinwide response of the Atlantic Meridional Overturning Circulation to interannual wind forcing[J]. Climate dynamics, 2017, 49(11-12): 4263-4280.

Zhou B, Du J. Fog prediction from a multimodel mesoscale ensemble prediction system[J]. Weather and Forecasting, 2010, 25(1): 303-322.

Zhu, Jiang, et al. "AMOC response to global warming: dependence on the background climate and response timescale." Climate Dynamics 44.11-12 (2015): 3449-3468.

Zuidema P, Chang P, Medeiros B, et al. Challenges and prospects for reducing coupled climate model SST biases in the eastern tropical Atlantic and Pacific Oceans: The US CLIVAR Eastern Tropical Oceans Synthesis Working Group [J]. Bulletin of the American Meteorological Society, 2016, 97(12): 2305-2328.