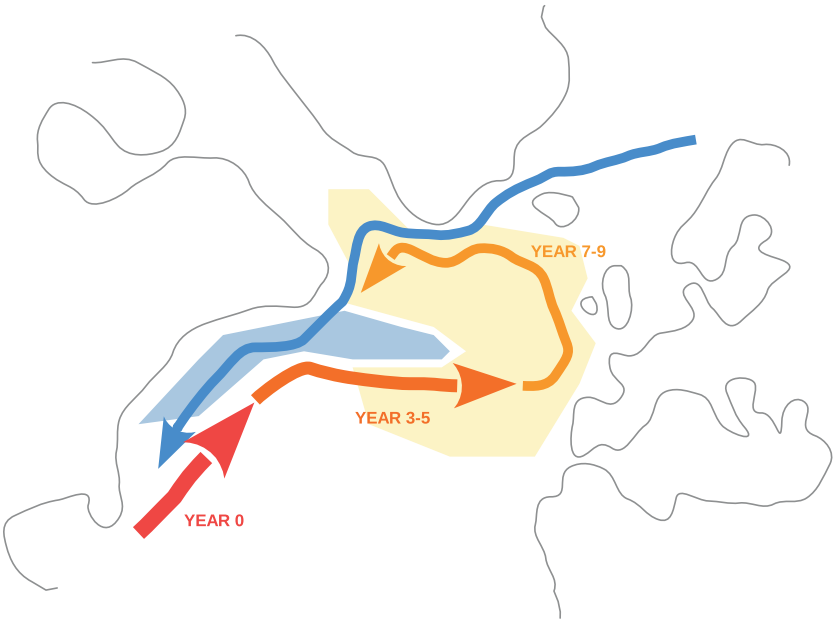
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Decadal Climate Predictions in the North Atlantic Region:
The Role of Ocean Heat Transport



Leonard Friedrich Borchert

Hamburg 2018

Berichte zur Erdsystemforschung
Reports on Earth System Science

208
2018

Hinweis

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

Sie enthalten wissenschaftliche und technische Beiträge, inklusive Dissertationen.

Die Beiträge geben nicht notwendigerweise die Auffassung des Instituts wieder.

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

Notice

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

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

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

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

Anschrift / Address

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

Tel./Phone: +49 (0)40 4 11 73 - 0
Fax: +49 (0)40 4 11 73 - 298

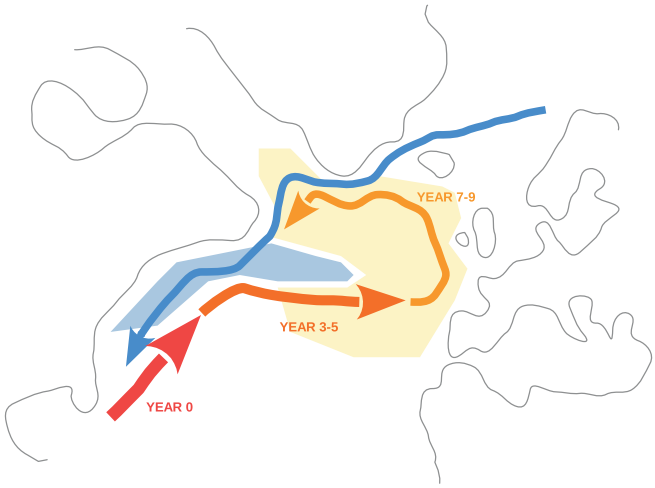
name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Decadal Climate Predictions in the North Atlantic Region:
The Role of Ocean Heat Transport



Dissertation with the aim of achieving a doctoral degree at the
Faculty of Mathematics, Informatics and Natural Sciences
Department of Earth Sciences of Universität Hamburg

submitted by

Leonard Friedrich Borchert

Hamburg 2018

Leonard Friedrich Borchert

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 06.07.2018

Folgende Gutachter empfehlen die Annahme der Dissertation:

Prof. Dr. Johanna Baehr
Dr. Wolfgang A. Müller

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 208
Reports on Earth System Science / Max Planck Institute for Meteorology 2018

ISSN 1614-1199

‘Everything changes, permanently. How boring if it wouldn’t.’
- *Klaus Schulze*

Abstract

The possibility of credible climate forecasts for several years ahead - on the *decadal* time scale - has received considerable public and economic attention. Scientific studies quantify the credibility of such forecasts by evaluating the average predictive quality (*skill*) over the last 50-60 years (in so-called *hindcasts*). Decadal hindcasts of surface temperatures were shown to be on average particularly skillful in the North Atlantic region. However, the reason for the high skill of these hindcasts is still unclear. Meanwhile, North Atlantic sea surface temperatures (SSTs) are on the decadal time scale strongly influenced by subpolar ocean heat transport (OHT) variability. I here connect OHT variability and SST predictability and test whether the knowledge of the strength of subpolar OHT at the beginning of a single SST forecast can improve its credibility.
By using initialized global climate simulations of the twentieth century, I confirm previous studies in that OHT variability influences SST variability for 3-10 years. A characteristic SST pattern of warm anomalies in the northeast Atlantic and cold anomalies in the Gulf Stream region emerges after strong OHT phases and vice versa. This pattern originates from persistently growing upper ocean heat content anomalies that arise from Southward propagating OHT anomalies in the North Atlantic. Extending previous work, I analyze strong and weak OHT phases at 50°N separately. This reveals an asymmetry between strong and weak phases of ocean heat transport: When subpolar OHT is strong, North Atlantic SSTs show stronger and more persistent decadal anomalies than when subpolar OHT is weak.
For the first time I show that the hindcast skill of northeast Atlantic SSTs 3-10 years ahead is linked to the characteristic SST pattern, and therefore OHT variability in the subpolar North Atlantic. When subpolar ocean heat transport is strong at the initialization of a hindcast, the skill of SST hindcasts in the northeast Atlantic 2 to 9 years into the future is significantly higher than when the ocean heat transport is weak at initialization. The asymmetric effect of strong and weak phases of subpolar OHT on SST variability that preconditions asymmetric hindcast skill is robust in non-initialized versions of the same climate model. The skill of decadal SST predictions therefore robustly depends on the climate state at the start of a prediction.
I show in this dissertation that hindcast skill changes over time and thus cannot be immediately translated into the credibility of a forecast. Instead, the credibility of a decadal climate forecast depends on the climate state at the start of the forecast. For North Atlantic SST forecasts, the strength of subpolar North Atlantic OHT at the start of the forecast can be used to estimate its credibility. Findings presented in this dissertation suggest that physical mechanisms might be used to improve conventional estimates of the credibility of a climate forecast on the economically and politically relevant decadal time scale.

Zusammenfassung

Die Möglichkeit, das Klima für einige Jahre glaubhaft vorherzusagen, erweckte zuletzt umfangreiches öffentliches und ökonomisches Interesse. Wissenschaftliche Studien quantifizieren die Glaubwürdigkeit solcher Vorhersagen, indem sie die durchschnittliche Vorhersagequalität der letzten ca. 50 Jahre diagnostizieren. Solche *dekadischen* Vorhersagen von Erdoberflächentemperaturen in der Nordatlantikregion zeigten besonders hohe Qualität. Der Grund für diese hohe Vorhersagequalität in der Nordatlantikregion ist bisher jedoch unbekannt. Indes beeinflussen Schwankungen im Transport von Wärme aus dem tropischen in den subpolaren Nordatlantik (*ocean heat transport, OHT*) nordatlantische Wasseroberflächentemperaturen (*sea surface temperatures, SSTs*) für etwa zehn Jahre. In dieser Dissertation zeige ich Verbindungen von niederfrequenten Schwankungen des OHT zu der Qualität dekadischer SST-Vorhersagen auf. Weiterhin diskutiere ich, wie die Kenntnis der Stärke des OHT im subpolaren Nordatlantik zu Beginn einer einzelnen SST Vorhersage genutzt werden kann, um die erwartbare Qualität dieser Vorhersage abzuschätzen.

Mit Hilfe initialisierter numerischer Modellsimulationen des gesamten zwanzigsten Jahrhunderts bestätige ich frühere Studien, indem ich zeige, dass OHT-Schwankungen die Variabilität von SSTs für bis zu 3-10 Jahre beeinflussen können. Ein charakteristisches SST-Muster mit warmen Temperaturen im nordost-Atlantik und kalten Temperaturen in der Golfstromregion erscheint nach starken OHT-Phasen und anders herum. Dieses Muster entsteht aus stetig wachsenden Wärmeanomalien im oberen Ozean, welche aus OHT-Anomalien resultieren, die sich im Nordatlantik südwärts fortpflanzen. Basierend auf diesen Analysen erweitere ich bisherige Studien und analysiere starke und schwache OHT-Phasen separat. Dies offenbart einen asymmetrischen Effekt starker und schwacher OHT-Phasen: starke OHT Phasen bei 50°N beeinflussen SSTs stärker und nachhaltiger als schwache.

Ich zeige hier erstmals, dass die Qualität von SST-Vorhersagen für 3-10 Jahre in die Zukunft mit diesem charakteristischen SST-Muster, und somit mit ozeanischem Wärmetransport, zusammenhängt. Wenn OHT zu Beginn einer Vorhersage in einer starken Phase ist, ist die Vorhersagequalität von SSTs für 2-9 Jahre in die Zukunft signifikant besser, als wenn der Ozean zu Beginn der Vorhersage wenig Wärme transportiert. Diese Asymmetrie ist robust in unterschiedlichen Realisationen des selben Klimamodells. Die Qualität dekadischer SST-Vorhersagen hängt daher vom klimatischen Zustand zu Beginn der Vorhersage ab.

Ich zeige in dieser Dissertation, dass die Qualität dekadischer Temperaturvorhersagen zeitabhängig ist, und daher Qualitätsabschätzungen für die Vergangenheit für Vorhersagen der Zukunft nicht anwendbar sind. Tatsächlich bedingt der klimatische Zustand zu Beginn einer Vorhersage deren Qualität. Bei der dekadischen Vorhersage nordat-

lantischer SSTs kann der ozeanische Wärmetransport im Nordatlantik als Kriterium zur Abschätzung der erwarteten Qualität einer Vorhersage genutzt werden. Ergebnisse, die ich in dieser Dissertation präsentiere, deuten darauf hin, dass physikalische Mechanismen genutzt werden können, um konventionelle Abschätzungen der Qualität von Klimavorhersagen für den ökonomisch und politisch interessanten dekadischen Zeitraum zu verbessern.

Teilveröffentlichungen dieser Dissertation

Pre-Published Work Related to this Dissertation

Borchert, Leonard, Wolfgang A. Müller and Johanna Baehr (2018). 'The Influence of the Strength of Atlantic Ocean Heat Transport on Interannual-to-Decadal Surface Temperature Predictability in the North Atlantic'. *Journal of Climate*, 31, 6763–6782, https://doi.org/10.1175/JCLI-D-17-0734.1.

Borchert, Leonard, Sebastian Brune, André Düsterhus, Wolfgang A. Müller and Johanna Baehr (2018). 'Hindcast Skill versus Prediction Skill: The Necessity for a Physical Mechanism'. *In preparation.*

Acknowledgements

This dissertation is in many ways the achievement of many people. There is no way I will be able to thank everyone who contributed, but I want to try nonetheless.
Without the guidance by the supervisor of this thesis, Johanna Baehr, I would certainly not have developed into the person I had to become to write this dissertation. Her wisdom, constant support, and our many discussions have formed a substantial part of the scientist I am today. I could not be more grateful and consider myself lucky that I could work with her in the past three years. I also owe significant gratitude to Wolfgang Müller, who with his calm way of providing brilliant insights constantly inspired me to keep going. His support means a great deal to me. I would further like to thank Hermann Held, whose instinct to challenge the right findings at the right time improved this dissertation a lot.
I am very thankful towards the entire Climate Modelling research group at Universität Hamburg, researchers and students, for offering bright insights and providing support when needed. Particular thanks go to my office mate Nele for discussion, laughter, lunching, and putting up with my regular video calls. Another 'thank you' goes to Susan Lozier and Pablo Ortega for challenging discussions that inspired me to drill deeper. I would also like to thank Julianna Carvalho Oliveira, whom I had the pleasure to work with during her internship, for delightful and productive cooperation. For valuable comments on several drafts of this thesis I am thankful towards Alex Winkler, Francois Lux, Elina Plesca, Larissa Schultze, Sebastian Brune, and Dirk Olonscheck.
During my time in the IMPRS-ESM I met many amazing people. I would like to thank everyone who is active in the Max Planck PhDnet for their commitment to the cause. I particularly thank Jana, Teresa, Lisa, Gabe, and Rafa for making the best PhDnet Steering Group ever! Not only their commitment, but also that of members of the Leibniz and Helmholtz Associations within the network N^2, constantly motivates me. Moreover, I would like to thank the MPI football team for the excercise I needed; Alex and Lukas for the music; Jairo for his constant enthusiasm; Larissa and Elina for ice cream; Dirk for sharing the burden of finishing the thesis; and Antje, Connie, Wiebke, and Michi for the best administrative and emotional support imaginable.
Some thanks remain unmentioned so far. I would like to thank Paula for the extra motivation I needed to quickly finish this *book*. I am also grateful towards Jim and John for, albeit being mildly unscientific, making my research memorable.
Finally, I could not be more thankful for the friends and family that stuck with me throughout this process. Laure's friendship means the world to me. Without the support from my family - Marlene, Andreas, Ursel, and all the others - I would not be where I am today. And then there is Eve, who has no idea how much she contributed to this; I am unbelievably lucky to have you in my life!

Contents

Acronyms

20CR: Twentieth Century Reanalysis
ACC: Anomaly Correlation Coefficient
AMOC: Atlantic Meridional Overturning Circulation
AMV: Atlantic Multidecadal Variability
ASSIM: Assimilation Model Experiment
EOF: Empirical Orthogonal Functions
HadISST: Hadley Center Ice and Sea Surface Temperature Data Set
HC: Hindcast Model Experiment
HIST: Historical Model Experiment
MPI-ESM-LR: Max-Planck-Institute Earth System Model in the Low Resolution Setup
NAO: North Atlantic Oscillation
OHT: (meridional) Ocean Heat Transport
piControl: Pre-Industrial Control Model Experiment
RCP: Representative Concentration Pathway
RCP4.5: Representative Concentration Pathway 4.5 Model Experiment
RCP8.5: Representative Concentration Pathway 8.5 Model Experiment
SAT: Surface Air Temperature
SHF: (ocean-atmosphere) Surface Heat Flux
SST: Sea Surface Temperature
UOHC: Upper Ocean Heat Content of the top 700m

1 | Decadal Predictions of the Climate System

1.1 Introduction

We do not know how warm or cold it is going to be in Europe in 8 years. Temperature predictions for the next few months or years - on the *seasonal-to-decadal* timescale - are still largely uncertain and subject to intense scientific debate. As a result, predictions on this time scale have received increasing scientific attention in the past decade, not just because of the scientific challenge, but also because predicting the climate on the (sub-)decadal time scale is particularly interesting for both policy makers and economic decision makers. It remains largely unknown, however, to which degree a single forecast of the temperature development over the next few years can be expected to be credible. In this dissertation, I present an approach to estimate the credibility of a forecast of North Atlantic surface temperatures several years into the future, on the *decadal* time scale, using a physical process in the ocean.

Without an estimate of the credibility of a climate forecast, any forecast is essentially useless. This credibility is in climate prediction studies usually referred to as *skill*. In the absence of knowledge of the future, series of predictions are commonly performed for the past (in so-called *hindcasts*) and evaluated against known past climate (e.g. Boer et al., 2016). The *hindcast skill* that is found for the past is then used as an estimate of how reliable decadal predictions are. There is still substantial scientific discourse on the reliability of decadal predictions after more than 10 years of research (e.g. Marotzke et al., 2016). In the face of the high confidence scientists put into both weather forecasts (on the time scale of days) and climate projections (on the scale of multiple decades to centuries), this dispute about the reliability of decadal predictions might appear puzzling. However, the climate variability on a decadal time scale results from a combination of factors that makes setting up successful hindcast climate simulations particularly challenging.

Weather forecast is considered an *initial value problem* (e.g. Pielke Sr. et al., 1999). Thus, the quality of a weather forecast mainly depends on using the correct current climate state to start the weather model from. If the initial conditions are correct and the

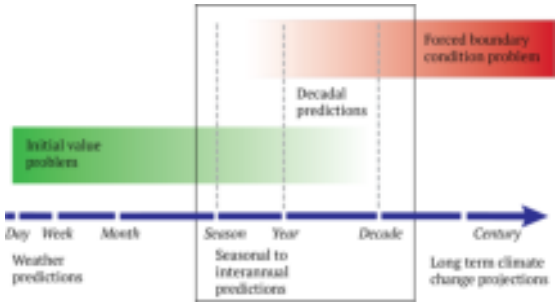


Figure 1.1: Seasonal-to-decadal climate predictions are both an initial value problem and a forced boundary problem (from Boer et al., 2016).

computational model reproduces the physics of the atmosphere somewhat accurately, weather forecasts are likely to be accurate for several days into the future. The notion of weather forecast as an initial value problem is also mirrored in the expectation towards it: a weather forecast is commonly expected to accurately capture weather patterns on very small spatial scales of only few kilometres.

For a successful climate projection, starting the climate model from the correct climate state is less important than the capability of the model to accurately reproduce climate variability patterns on the decadal to multidecadal time scale; climate projections are considered a *boundary value problem* (e.g. Boer et al., 2016). Unlike weather forecasts, climate projections are not expected to resolve small spatial scales: the expectation towards a climate projection is to accurately capture trends or changes in climate variability on spatial scales of several hundred kilometers, or even globally. It is therefore paramount for a climate projection model to accurately reproduce trends and variability in the underlying physics on (multi-)decadal time scales to produce a skillful climate projection.

Decadal climate predictions combine the expectations towards weather forecasts and climate projections: they are expected to capture long-term trends and variability changes in the climate system while producing predictions that can be used by decision makers. These decision makers mostly need information on relatively small spatial scales to use. Studies showed that both the initial conditions and low-frequency variability in the climate system should be accurately represented in a climate model to produce skillful decadal climate predictions (e.g. Palmer et al., 2004; Matei et al., 2012; Doblas-Reyes et al., 2013). Climate predictions on the decadal timescale can therefore be considered both an initial value problem and a boundary value problem (fig. 1.1).

So-called *assimilation experiments* combine the good representation of initial conditions from weather forecasting with the good representation of boundary forcing in climate models (e.g. Keenlyside et al., 2008). Assimilation experiments use a global climate projection model, or *general circulation model (GCM)*, and use (*assimilate*) observations to constrain the GCM to observed past climate states and variability. Usually, because

of the availability of observations, assimilation experiments cover the time between 1960 and today. For this time horizon, studies showed skillful decadal climate predictions in many regions around the globe (e.g. Smith et al., 2007; Doblas-Reyes et al., 2013). The North Atlantic region has been subject to particularly high attention (e.g. Smith et al., 2007; Yeager et al., 2012; Müller et al., 2012; Monerie et al., 2017).
The highest skill in decadal climate hindcasts was found for temperatures over the oceans, whose inertia forms a *memory* in the climate system that preconditions decadal climate predictability (e.g. Branstator and Teng, 2010; Matei et al., 2012; Collow et al., 2015). For other climate variables, e.g. precipitation, the skill of decadal hindcasts was found to be modulated by the ocean as well (e.g. Gaetani and Mohino, 2013). Specifically, decadal temperature skill can be expected to be high in areas that are strongly influenced by low-frequency ocean variability, like ocean overturning dynamics (e.g. Yeager et al., 2012; Robson et al., 2013). This indicates that mechanisms of low-frequency variability in the ocean precondition skill of decadal climate hindcasts. Understanding why decadal climate predictions are successful therefore requires understanding decadal climate variability.

1.2 Ocean Overturning and North Atlantic Temperatures

Sea surface temperature (SST) variability reflects a balance between the low-frequency variability of upper ocean heat content and the higher-frequency variability of ocean-atmosphere surface heat fluxes. This was first described for the North Atlantic by Bjerknes (1964), who interpreted SSTs as a coupled ocean-atmosphere mode with a driving role for the ocean and a dampening role for the atmosphere on decadal time scales. Later, simulations using computational models supported Bjerknes' hypothesis by showing that on decadal time scales and longer, SST fluctuations in the North Atlantic are driven by ocean overturning variability (e.g. Eden and Willebrand, 2001; Gulev et al., 2013). Specifically, studies underscored that, in the North Atlantic, the variability of the ocean mass transport, the Atlantic Meridional Overturning Circulation ($AMOC$), and associated oceanic heat transport (OHT) strongly shapes decadal surface temperature variability (e.g. Timmermann et al., 1998).
A robust climatic feature that was found consistently is the lagged relationship between North Atlantic deep water formation, AMOC/OHT, and SSTs in that order (Yeager and Robson, 2017). The formation of deep water in the North Atlantic, particularly in the Labrador Sea region, was shown to be connected to persistent forcing from the Atmosphere, specifically the North Atlantic Oscillation (NAO, e.g. Marshall et al., 2001). This deep water formation is one of the drivers of AMOC variability in the North Atlantic. The AMOC and OHT are therefore important features of the climate system

that connect atmospheric variability to the variability of SSTs in the North Atlantic on the decadal time scale (e.g. Zhang, 2008).

The ocean contribution to SST variability is often approximated by integrating temperatures from the surface to some characteristical depth (e.g. Yeager et al., 2012). The characteristical depth depends on the region of interest; in the North Atlantic, some previous studies have used 700m as lower bound (e.g. Zhang and Zhang, 2015). The integrated temperatures in this water volume are called upper ocean heat content (*UOHC*).

Changes in UOHC in the North Atlantic are controlled by convergence of heat in the ocean and vertical heat fluxes (surface heat fluxes, *SHFs*) at the ocean-atmosphere interface. The low-frequency changes of UOHC were linked to SST variability on the multi-decadal time scale under the term *Atlantic Multidecadal Oscillation* or, more recently because of its lack of clear periodicity, *Atlantic Multidecadal Variability* (*AMV*, e.g. Kerr, 2000; Ting et al., 2011; Peings et al., 2016). The connection of UOHC variability to the AMV indicates that SST variability on the decadal time scale and longer is in parts of the North Atlantic dominated by low-frequency ocean dynamics (e.g. Eden and Willebrand, 2001; Zhang et al., 2016). The idea of an ocean dominated AMV was contested recently by studies suggesting that the AMV is mainly the result of stochastic atmospheric forcing (Clement et al., 2015; Bellomo et al., 2016; Cane et al., 2017). However, several papers attributing SST variability in the North Atlantic to changes in ocean overturning and explaining a physical mechanism seem to support the hypothesis that the ocean dominates AMV variability (e.g. Zhang, 2008; Zhang and Zhang, 2015). Zhang (2008) and Zhang and Zhang (2015) described a mechanism connecting strong phases of AMOC-related OHT to ocean heat convergence and consequently UOHC variability in the North Atlantic (fig. 1.2). Strong phases of the AMOC and associated OHT in the northern North Atlantic were shown to originate from surface density anomalies in the subpolar North Atlantic and to propagate southward at a slow advection speed. The authors showed that this slow southward propagation of strong OHT phases led to a heat convergence anomaly North of the OHT anomaly and a heat divergence anomaly South of the OHT anomaly, constituting an UOHC anomaly dipole between positive anomalies in the North Atlantic Subpolar Gyre and negative anomalies in the Gulf Stream region, the so-called *AMOC Fingerprint* (Zhang and Zhang, 2015, figs. 1.2a,b). Because of the slow propagation of OHT phases, the time lag between the initial AMOC and OHT anomaly at 50°N and the AMOC Fingerprint was found to be between 2 and 12 years, depending on the model setup (fig. 1.2c; Zhang and Zhang, 2015). The meridional coherence of AMOC anomalies is a necessary prerequisite for studying the propagation of AMOC anomalies across latitudes. Zhang (2010) showed that AMOC anomalies were particularly meridionally coherent when calculated in density coordinates.

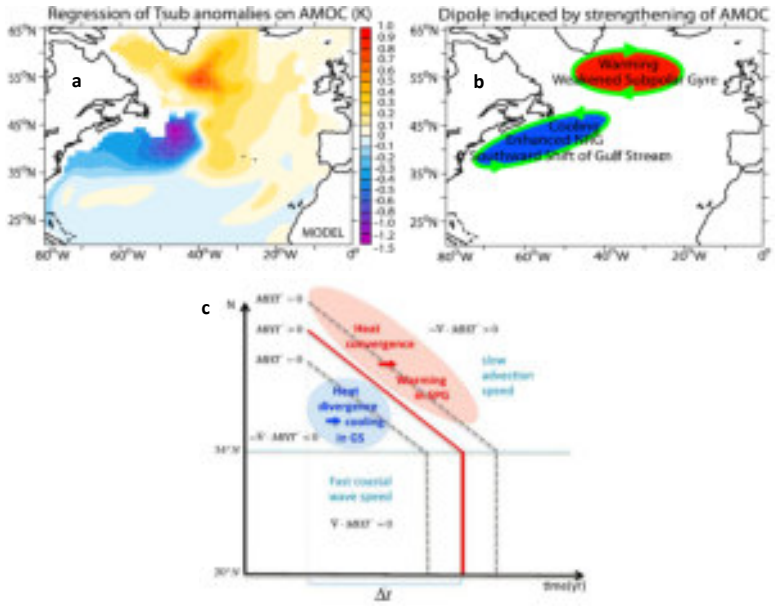


Figure 1.2: An illustration of the physical mechanism connecting AMOC variability in the subpolar North Atlantic to SST variability proposed by Zhang (2008) and Zhang and Zhang (2015). (a) shows the regression of temperature anomalies in the North Atlantic on AMOC in the GFDL-CM2.1 (Zhang, 2008). In (b), this temperature pattern is conceptualized by Zhang (2008). (c) shows a conceptual depiction of the AMOC-SST mechanism presented in Zhang and Zhang (2015), illustrating the southward propagation of positive AMOC anomalies (here: MHT = Meridional Heat Transport) over time and associated heat convergences and divergences.

The AMOC Fingerprint was connected to the AMV and thus hypothesized to influence the skill of decadal UOHC and SST hindcasts (Zhang, 2008; Zhang and Zhang, 2015). In this dissertation I use the Max-Planck-Institute Earth System Model (*MPI-ESM*) to examine this hypothesis. However, the mechanism connecting ocean circulation changes to SST variability on decadal time scales suggested by Zhang and Zhang (2015) was shown in only one climate model, the *GFDL Climate Model 1.2*, so far. For an application to a hindcast study in the MPI-ESM, this mechanism therefore has to be understood in the MPI-ESM prior to analyzing hindcasts. I thus derive the following first central research question:

⇒ **Can the mechanism leading to the AMOC Fingerprint proposed by Zhang and Zhang (2015) be found in the MPI-ESM-LR, and how is this mechanism characterized?**

I will in this chapter establish a mechanism in the MPI-ESM connecting AMOC and OHT to SST variability on the decadal time scale. Going further, I will examine the effect of this mechanism on surface air temperatures (*SATs*). I will then proceed to examine how this mechanism affects temperature predictability on the same time scale.

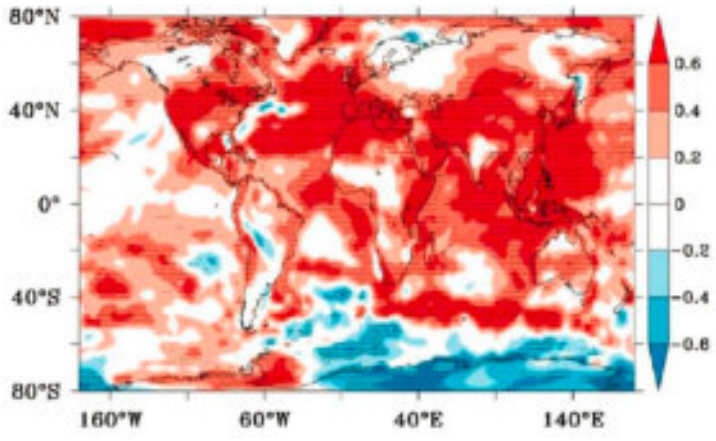


Figure 1.3: A typical map of surface temperature hindcast skill 2-5 years into the future, evaluated as anomaly correlation coefficients (from Müller et al., 2014). The North Atlantic region shows high skill.

1.3 Decadal Hindcasts in the North Atlantic Region

Climate prediction several years into the future has received increasing scientific attention recently (e.g. Palmer et al., 2004; Smith et al., 2007; Keenlyside et al., 2008; Doblas-Reyes et al., 2013; Müller et al., 2014). Good hindcast skill was found in the North Atlantic region and in surface temperatures over Europe on time scales of 2-8 years (fig. 1.3, e.g. Smith et al., 2007; Yeager et al., 2012; Müller et al., 2012). While hindcast skill for surface temperatures was found to have a tendency to be weaker over continental regions, it was found to be mostly high over the ocean. This stems from the inertia of the ocean, a *memory* in the climate system, which preconditions decadal temperature predictability in the North Atlantic region (e.g. Branstator and Teng, 2010; Matei et al., 2012; Collow et al., 2015). Recent studies suggested that phases of strong OHT in the subpolar Atlantic Ocean linked to AMOC variability influence North Atlantic UOHC (defined as the heat contained in the top 700m of the ocean, e.g. in Zhang and Zhang, 2015), which might improve surface temperature predictability on a time scale of 2 to 12 years (Zhang, 2008; Zhang and Zhang, 2015). However, the specific advantage of knowing the OHT-strength at the start of a prediction on the quality of the predictability estimate of North Atlantic UOHC and surface temperatures was not shown so far. Here, I analyze the dependency of the decadal predictability of North Atlantic UOHC, SSTs, and SATs on the state of the OHT at 50°N at the start of the prediction.

Decadal SST and SAT hindcasts were demonstrated to be particularly skillful in the North Atlantic in several recent studies (e.g. Matei et al., 2012; Yeager et al., 2012; Klöwer et al., 2014; Müller et al., 2014; Robson et al., 2017; Yeager and Robson, 2017). Since surface temperature hindcast skill is often found to be stronger over the ocean than over land, it is commonly assumed that the source of decadal temperature prediction

skill in the North Atlantic resides in the ocean (Yeager and Robson, 2017). Specifically, high decadal surface temperature hindcast skill was found in regions where ocean dynamics influence surface temperature variability more strongly than atmospheric dynamics in many studies (e.g. Yeager and Robson, 2017). There are two current hypotheses concerning the origin of the oceanic memory that preconditions hindcast skill in the North Atlantic region: that decadal surface temperature hindcast skill originates from the persistence of heat content in the ocean (e.g. Meehl et al., 2009), and that skill originates from low-frequency ocean circulation (e.g. Robson et al., 2013, 2014).
Case studies showed that low-frequency variability of the oceanic circulation can precondition hindcast skill of UOHC and SSTs on the decadal time scale. Specifically, the 1960s cooling of the North Atlantic Subpolar Gyre (Robson et al., 2014), and the 1920s (Müller et al., 2014) and 1990s (Yeager et al., 2012; Robson et al., 2013) warmings were shown to be predictable several years in advance when the state of the ocean was initialized in the respective model simulation. These studies showed that individual events of strongly anomalous UOHC in the North Atlantic could have been predicted in the past. Found predictability might even be connected to decadal predictability for SATs over parts of Europe as well (Robson et al., 2012). While the connection of subpolar AMOC and OHT variability on UOHC predictability was shown in these case studies, there was no systematic evaluation yet. The physical mechanism connecting OHT to UOHC variability in the North Atlantic region suggested by Zhang and Zhang (2015) provides a framework to test the influence of OHT dynamics on UOHC predictability more generally.
In this dissertation, I go beyond previous studies and systematically identify the specific influence of variability of subpolar AMOC and OHT on decadal hindcast skill of surface temperatures in the North Atlantic. Based on the findings from chapter 3, I assess hindcast skill in the 20^{th} century for strong and weak subpolar OHT phases separately. I then compare the influences of OHT variability and UOHC persistence on decadal surface temperature hindcast skill. To reconcile the hypothesis brought up in previous studies that ocean overtuning dynamics dominate decadal surface temperature predictability, I ask:

⇒ **How strongly do ocean overturning dynamics influence the skill of SST hindcasts in the North Atlantic region?**

Following the findings from chapters 3 and 4, I will show some of their implications for the evaluation of decadal hindcast skill. Specifically, I will first place these findings in the context of other model simulations with the same climate model and in the context of observations to see how representative they are of different modes of climate variability. I will then discuss how representative skill estimates found for any period of the past can be for actual forecasts.

1.4 Understanding North Atlantic Climate Variability in the MPI-ESM-LR

In assimilation model experiments, observational data is used to constrain the variability the model produces on its own to stay within the bounds of observed variability (e.g. Palmer et al., 2004). The hope of applying this technique in the context of decadal hindcast studies is to generate a four-dimensional climate state estimate that is similar to the observed climate (e.g. Keenlyside et al., 2008). This state estimate is then used to start hindcast runs with the free model from a relatively realistic climate state to produce skillful hindcasts (e.g. Yeager et al., 2012; Robson et al., 2013). In the past, this approach has proved to be very effective in improving decadal hindcast skill. However, there is an issue connected to the assimilation of observations into climate models: it is from the assimilation model itself unclear how representative dynamical features found within it are with respect to observations and the underlying model. In this dissertation, I present an approach to place findings from the first two chapters in the context of observations and the MPI-ESM-LR.

Numerous studies reported differences between assimilation experiments, non-initialized model simulations, and observations (e.g. Balsameda and Anderson, 2009; Pohlmann et al., 2017). The relative contribution of model and observed climate variability to climate variability in assimilation experiments varies in space and time (Servonnat et al., 2015). It is therefore inherently unclear how the dynamics found in assimilation model experiments need to be interpreted: as observed variability, as model variability, or as something else.

In the latter case - the interpretation of assimilation model simulations as their own domain of model simulation with its very own mode of climate variability - findings concerning climate variability and predictability obtained from these model simulations need an extra step of analysis to be fully understood. Because while it is relatively clear what climate variability found in model studies or observational studies represents (model variability and observed climate variability, respectively), the respective roles of observed and modeled climate variability in producing climate variability in an assimilation run are unclear for the aforementioned reasons.

I will in chapter 5 examine findings from chapters 3 and 4 in more detail, placing them in the broader climatic context of model-based and observed climate variability. I will put particular focus on North Atlantic OHT and SST variability. Specifically, I will answer the following question:

⇒ **Is the previously discussed climate variability reasonable with respect to both model variability produced by the MPI-ESM-LR and observations?**

After placing the findings concerning decadal temperature variability and hindcast skill

in the North Atlantic in the context of overall climate variability, I move on to show how hindcast skill can be expected to change over time. This will not only help to understand climate hindcasts of the past, but also suggest how to translate hindcast skill found for the past into the credibility of forecasts.

1.5 Non-Stationary North Atlantic Surface Temperature Prediction Skill

In this dissertation I shed light on the time-dependence, or non-stationarity, of decadal climate hindcasts. I also suggest an approach to accurately estimate the credibility of individual forecasts using hindcast simulations.
Decadal climate hindcasts are conventionally evaluated over a certain period in the past. The mean skill found for these hindcasts is then assumed to reflect the credibility of an individual forecast. However, several studies showed recently that hindcast skill estimates can differ substantially depending on the period that the hindcasts are evaluated for. This was shown for seasonal hindcasts of the winter North Atlantic Oscillation (Weisheimer et al., 2017; O'Reilly et al., 2017) and for decadal hindcasts of the North Atlantic Subpolar Gyre region (Brune et al., 2017). The skill estimates derived for the past by conventional hindcast studies are therefore not representative for the whole period they cover. Hindcast skill estimates are thus likely not directly applicable for individual forecasts.
To examine the applicability of current measures of decadal hindcast skill for the estimation of the credibility of any individual forecast, I ask:

⇒ **Are mean hindcast skill estimates appropriate to estimate the credibility of a single temperature forecast in the North Atlantic region?**

This chapter will provide some exciting insights into where decadal climate prediction research might, or should, develop in the future. It will also draw on the findings from chapters 3 and 4 and highlight their implications.

In this dissertation, I examine the influence of low-frequency ocean variability on the predictability of surface temperatures, studying the North Atlantic region. Chapter 2 presents the methods I will use in this dissertation. In chapter 3, I use an initialized version of the MPI-ESM to examine a physical mechanism suggested by Zhang and Zhang (2015) that connects variability of the heat transported by the AMOC into the subpolar North Atlantic to North Atlantic SST variability on the decadal time scale. I discuss the influence of this mechanism on decadal surface temperature predictability in chapter 4. A paper that was published in the Journal of Climate summarizes findings

presented in chapters 3 and 4.
In chapter 5, I elaborate on the findings from chapter 3 and investigate which mode of climate variability (model-based or observed) is found in the assimilation experiment. This will help to place the findings from chapters 3 and 4 in the wider climatic context. In chapter 6, I examine how representative hindcast skill estimates for the past can be considered for individual decadal climate forecasts. This will allow me to integrate my findings into the scientific context and discuss implications of these findings for future decadal climate hindcast and predictability studies in the concluding chapter. Findings presented in chapter 6 are the subject of a paper that is currently in preparation for submission.

2 | Model and Methods

To answer the questions I bring forward in this dissertation, I use the fully coupled Max Planck Institute Earth System Model (*MPI-ESM*). Specifically, I use the low-resolution (*LR*) model version that was used in the 5^{th} phase of the Coupled Model Intercomparison Project. This model version uses the ocean model MPIOM (Jungclaus et al., 2013) at a nominal horizontal resolution of 1.5° and 40 vertical levels which is interactively coupled to the atmospheric model ECHAM6 (Stevens et al., 2013) of the horizontal resolution T63 with 47 vertical levels with the top at 0.1 hPa. The model used here has a curvilinear grid with three poles, one over Greenland, one over Sibiria, and one over Antarctica. Therefore, the actual model resolution in the North Atlantic region is considerably higher than the average resolution, which improves the representation of climate variability in the North Atlantic region. I use several configurations of the MPI-ESM-LR that I will outline in the following.

2.1 piControl, HIST and RCP

To understand how ocean overturning influences surface temperatures in the MPI-ESM-LR, I use a pre-industrial control (*piControl*) simulation of 1000 years, and a historical simulation for 1896-2005 (*HIST*). Atmospheric greenhouse gas emissions are kept constant at the pre-industrial level in piControl. Therefore, the 1000 years of the piControl simulation represent an estimate of the internal variability created by the MPI-ESM-LR. In HIST, atmospheric greenhouse gas concentrations follow the observed trend for that time period. HIST can therefore be used to examine the reaction of the MPI-ESM-LR to greenhouse gas warming until the end of the 20^{th} century. I subtract the linear trend from piControl and HIST before performing any analysis.
In addition to the piControl and HIST model simulations, I use two greenhouse gas emission scenario projections with the MPI-ESM-LR to assess the change of physical mechanisms identified in this thesis with global warming. Specifically, I consider the time period 2191-2300 in the two warming scenarios *RCP4.5* and *RCP8.5*. These correspond to a moderate warming scenario (RCP4.5) and a business-as-usual scenario of strong warming (RCP8.5). To assess climate variability in the 22^{nd} century beyond the

global warming trend within that century, I subtract the linear trend from the 110-year-long time series of RCP4.5 and RCP8.5.

2.2 The MPI-ESM-LR Decadal Prediction System

In this dissertation, I use the decadal prediction system from Müller et al. (2014) with the fully coupled MPI-ESM-LR to explore decadal climate variability and predictability in the North Atlantic. The setup consists of an assimilation experiment covering the period 1901-2010, and ten-year-long hindcast runs with the fully coupled MPI-ESM-LR, which are started (*initialized*) from the assimilation experiment at the beginning of every year, and are after that only subject to observed greenhouse gas forcing.
The assimilation experiment consists of three ensemble members that are constructed by forcing the MPIOM at the surface with fluxes of momentum, energy, and freshwater (Müller et al., 2015). These fluxes are obtained for the period 1872-2010 from three different randomly selected realizations of the 20^{th} century reanalysis (Compo et al., 2011). Four-dimensional salinity and temperature fields from the resulting ocean states are then nudged into the fully coupled MPI-ESM-LR for the period 1901-2010 (Müller et al., 2014). I use the ensemble mean of the three resulting climate states as an estimate of climate variability in the 20^{th} century.
I remove the mean seasonal cycle from each ensemble member of the assimilation experiment separately to ensure that my analyses are not dominated by the signal of a seasonal cycle. I then form an ensemble mean of the three realizations of the assimilation run. I detrend the ensemble mean and form anomalies against its mean state. Detrending ensures that the variability and predictions that I examine in this dissertation refer to the internal variability of the climate system, and are not contaminated by the long-term trend. The detrended ensemble mean of anomalies will henceforth be referred to as *ASSIM*.
From every individual ensemble member, Müller et al. (2014) started ten year long simulations with the free fully coupled MPI-ESM-LR at the beginning of every year. I evaluate the ensemble mean of these hindcasts against the assimilation experiment to reconstruct the skill of decadal predictions for the 20^{th} century.
I remove the mean seasonal cycle from every ensemble member of the hindcasts individually. The ensemble members of the hindcast experiments are bias corrected against the corresponding realization of the MPIOM simulation from Müller et al. (2015), by setting the mean climate state in the hindcast experiments to that of the corresponding MPIOM simulation. Then, I construct an ensemble mean from the hindcast ensemble members. I detrend the ensemble mean hindcasts, and form anomalies against the mean state of the ensemble mean. The resulting detrended and bias corrected ensemble mean

hindcasts will be referred to as *HC* throughout this dissertation.
Starting coupled model simulations from assimilation experiments can lead to a strong drift where model physics and *real* physics disagree (e.g. Smith et al., 2013; Sanchez-Gomez et al., 2016; Pohlmann et al., 2017). Because the ocean state estimate used here for assimilation was based on the same model version as the ocean component of the coupled model it was assimilated into, and because it was forced exclusively at the ocean surface, data assimilation in this simulation represents a relatively *soft* approach to data assimilation. This potentially reduces the amount of drift that can be expected from this assimilation model simulation. Moreover, the interior ocean can freely adjust to the surface forcing. This enables an examination of the temporal and spatial development of surface-induced changes in the ocean state and dynamics, like the mechanism leading to the AMOC Fingerprint (Zhang and Zhang, 2015), with minimal perturbance from data assimilation. On the other hand, a surface-forced ocean state estimate might represent the three-dimensional ocean state only to a limited degree. However, the ocean state estimate that was used to produce ASSIM was previously found to produce reasonable climate variability in the North Atlantic region that is in agreement with observations and reanalyses in the atmosphere as well as the ocean for the entire 20^{th} century, like the 1960 cooling and the 1920 and 1990 warmings (Müller et al., 2015). A drawback of the HC simulations is the limited ensemble size of three members. Regardless, the HC ensemble mean was shown to reproduce climate variability in the North Atlantic region robustly for the entire time series 1901-2010, which led to an increase in North Atlantic surface temperature prediction skill when using 1901-2010 for the evaluation of skill compared to 1960-2010 (Müller et al., 2014). Because of the aforementioned reasons, ASSIM and HC are appropriate tools for the examination of the research questions brought forward in this dissertation.

2.3 Observations

Wherever possible, I place my findings in the context of temperature observations. I use in this dissertation the Hadley Center Sea Ice and Sea Surface Temperature data set (*HadISST*) from Rayner et al. (2003). This SST state-estimate is an ocean reanalysis based on SST and sea ice concentration data that is interpolated onto a grid using a reduced scale optimal interpolation procedure (Rayner et al., 2003). HadISST provides an SST and sea ice reanalysis for the period 1870 to 2018 (as of April 2018) that is widely used to verify model simulations and hindcasts (e.g. Brune et al., 2017). In this dissertation I use the period 1901-2010 of HadISST data. To assess internal climate variability beyond linear warming, I subtract the linear trend for 1901-2010 from HadISST.

2.4 Post-Processing and Methods

In this dissertation, I use AMOC anomalies in density coordinates, calculated from vertical diapycnal transports that are calculated from divergences of horizontal transports. I define total OHT in time ($Q(t)$) as the depth- and longitude- integrated product of three-dimensional y-velocity (v) and potential temperature (Θ) fields as in Jayne and Marotzke (2001). This is formulated as

$$Q(t) = \rho_0 c_p \int \int_{-H}^{0} v\Theta dz dx,$$

with ρ_0: density of sea water ($1025\, kg \cdot m^{-3}$), and c_p: specific heat of sea water ($3994\, J \cdot kg^{-1} \cdot^{\circ} C^{-1}$) integrated over depth z, up to maximum depth level H, and longitudinal extent of the Atlantic basin x (Jayne and Marotzke, 2001).
Upper ocean heat content (*UOHC(t)*) at every horizontal grid point is calculated by integrating potential temperature Θ in the upper 700 m of the ocean (the upper 20 layers in the MPI-ESM-LR):

$$UOHC(t) = \rho_0 c_p \int_{700m}^{0m} \Theta dz.$$

In this dissertation, I examine the heat exchange between the ocean and the atmosphere using surface heat fluxes (*SHF*). These are here defined as the total surface heat fluxes over sea. This includes shortwave, longwave, sensible and latent heat fluxes. These fluxes are defined positive downward.
Climate co-variability is assessed using Pearson correlation coefficients. These correlation analyses sometimes involve lagging one time series with respect to the other to find temporal shifts between variability patterns. The statistical significance of these correlations is assessed using a Monte-Carlo procedure. This procedure shows the likelyhood that the correlation values occur by chance. The amount of total climate variability explained by certain features of the climate system is analyzed using an Empirical Orthogonal functions (*EOF*) analysis (Storch and Zwiers, 1999). Wherever I compare different physical modes of the climate system, like strong and weak phases of ocean heat transport, I use a composite mean analysis, i.e. I compare the mean states of the different physical modes. These composite means are lagged against each other where appropriate. To test whether composite mean climate states related to different climate modes are statistically different from the mean climate variability, I use a two-sided t-test.
I assess the hindcast skill of HC using Anomaly Correlation Coefficients (*ACCs*). ACCs

are formulated as

$$ACC = \frac{\sum_{i=1}^{n} f_i' a_i'}{\sqrt{\sum_{i=1}^{n} f_i'^2 \sum_{i=1}^{n} a_i'^2}},$$

with n: number of samples, f': anomaly of the forecast value, and a': anomaly of the verifying value (Jolliffe and Stephenson, 2012). ACCs thus provide an estimate of the co-variability of HC and ASSIM for a given period of time in the past. The statistical significance of the skill estimates of the hindcasts presented in this dissertation is assessed using a Monte-Carlo procedure. This procedure shows the likelyhood that HC produces 'observed' climate variability by chance. The predictability I discuss here is tested using a leave-one-out cross-validation (Arlot and Celisse, 2010) to ensure that hindcast skill that I diagnose is not dominated by individual years.

I form annual mean anomalies of OHT, AMOC, UOHC and SSTs. From SATs, I construct annual and seasonal mean anomalies for winter (January, February, March; *JFM*), spring (April, May, June; *AMJ*), summer (July, August, September; *JAS*), and autumn (October, November, December; *OND*) from HC and ASSIM. The definition of these seasons is chosen to avoid averaging over model initialization in HC. I use full values of both annual and seasonal mean SHFs, not anomalies, to understand the full energy exchange between the ocean and the atmosphere. Analyses in piControl, HIST, the RCP scenarios, and HadISST only use annual means.

3 Ocean Overturning and North Atlantic Temperatures

3.1 Introduction

The objective of this chapter is to identify the specific influence of strong and weak phases of the AMOC and associated ocean heat transport on SSTs in the North Atlantic up to a decade into the future. Zhang and Zhang (2015) used simulations with the GFDL-CM2.1 model to present a physical mechanism that connects ocean overturning variability in the subpolar North Atlantic with a particular SST pattern, the *AMOC Fingerprint*, about a decade later. The AMOC Fingerprint features a SST dipole of a positive SST anomaly in the North Atlantic subpolar gyre and a negative SST anomaly in the Gulf Stream region.

In their study, Zhang and Zhang (2015) showed that the evolution of the AMOC Fingerprint depends on a slowly southward propagating AMOC and OHT anomaly that originates from a surface density anomaly in the North Atlantic. While propagating southward slowly, this AMOC and OHT anomaly constantly transports heat northward, leading to a heat convergence anomaly North of the AMOC and OHT anomaly and a heat divergence anomaly South of it (cf. fig. 1.2c; Zhang and Zhang, 2015). At 35°N, the AMOC and OHT anomaly breaks down and continues to travel southward at the speed of a coastal Kelvin wave, which inhibits the formation of the AMOC Fingerprint South of that latitude.

The evolution of the AMOC Fingerprint has so far only been examined in the GFDL model. I therefore use in this chapter the ASSIM simulation to examine the robustness of the evolution of the AMOC Fingerprint in the MPI-ESM-LR. In the process, I identify possible model-specific aspects in the formation of the AMOC Fingerprint. I begin by replicating the study by Zhang and Zhang (2015) in the ASSIM simulation. Subsequently, I address some questions that the study by Zhang and Zhang (2015) left open: I investigate in composite mean upper ocean heat content and SST anomalies to determine whether there is a difference in the influences of strong and weak phases of AMOC and OHT on the formation of the AMOC Fingerprint. I then examine the influence of the AMOC Fingerprint on surface air temperatures.

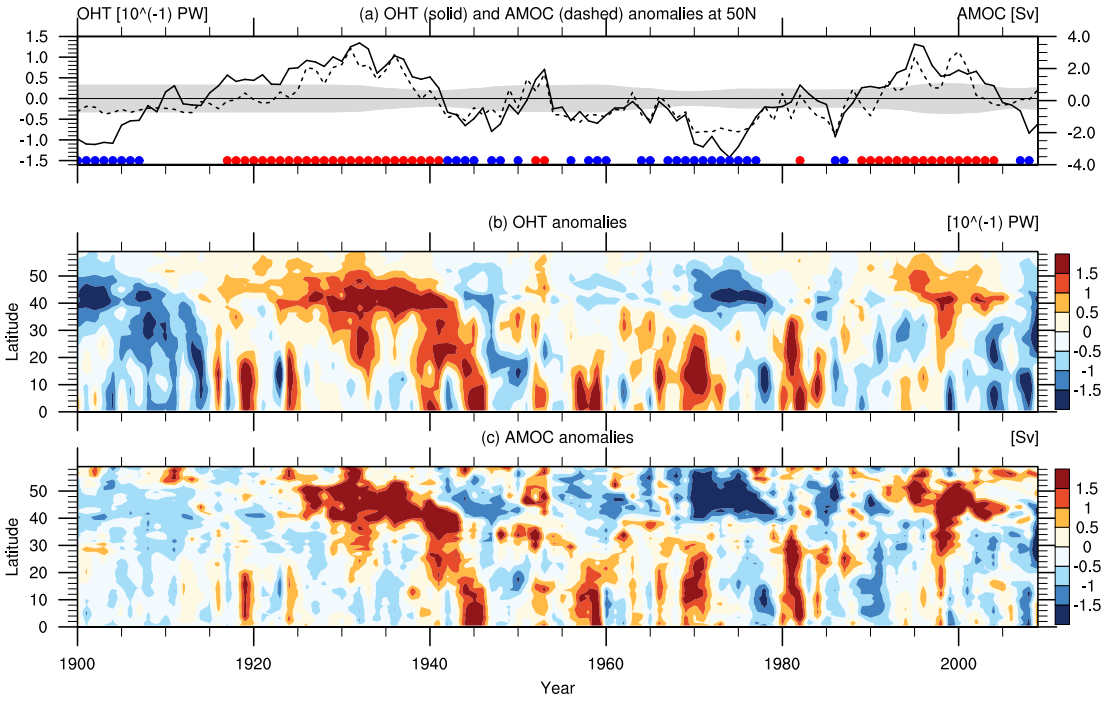


Figure 3.1: (a) Detrended anomalies of AMOC maximum at 50°N (dashed line, [Sv]) and total OHT at 50°N (solid line, [10^{-1} PW]) in the ASSIM-simulation. The grey area denotes a half standard deviation above and below the mean of the previous 30 years. Strong and weak OHT phases, i.e. years where the solid line lies outside the grey area, are marked with red and blue dots at the bottom, respectively. Hovmöller Diagrams of OHT anomalies (b) and AMOC maximum anomalies (c) illustrate the development of strong and weak anomalies of OHT and AMOC in space (y-axis, [°latitiude]) and time (x-axis [yrs]). OHT and AMOC time series are detrended at each latitude.

3.2 Meridional Coherence of AMOC and OHT in the North Atlantic

The key feature of the findings by Zhang and Zhang (2015) is the slow southward propagation of OHT phases in the North Atlantic. I therefore first analyze the meridional coherence of AMOC maximum and total OHT in the assimilation model experiment ASSIM, and test whether AMOC variability and heat transport variability are linked in this model. Annual mean anomalies of AMOC maximum and OHT are largely coherent at 50°N ($corr = 0.84$, fig. 3.1a), and seem closely connected across the entire North Atlantic at both decadal and longer time scales (fig. 3.1b,c). In ASSIM, the 20^{th} century is characterized by substantial multidecadal variability in both AMOC and OHT with stable strong anomalies in the 1920s and 1990s and an episode of less stable weak anomalies in between. This variability is similar to previously published estimates (e.g. Robson et al., 2013, 2014; Müller et al., 2014).

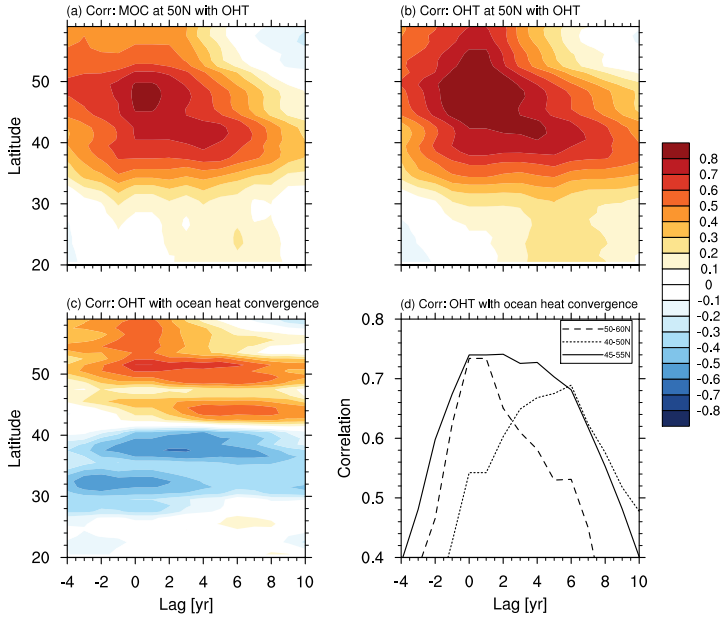


Figure 3.2: ASSIM-based lead-lag correlations of OHT at 50°N with AMOC maximum anomalies (a), OHT anomalies (b), and ocean heat convergence (c). Lead-lag correlations of OHT at 45, 50, and 55°N with oceanic heat convergence between 40-50°N, 45-55°N, and 50-60°N, respectively, are shown in (d). Positive lags indicate that OHT leads and vice versa.

As suggested by Zhang and Zhang (2015), OHT anomalies in ASSIM generally originate in the North Atlantic between 50-60°N and propagate southward slowly; this propagation is closely linked to AMOC dynamics (fig. 3.2a,b). South of 35°N, both AMOC and OHT anomalies show only limited correlation to OHT anomalies at 50°N, and the correlation that can be seen shows the same lag to OHT at 50°N. The latter indicates that South of 35°N, AMOC and OHT anomalies propagate southward at the speed of a coastal Kelvin wave as described in Zhang and Zhang (2015).

OHT anomalies at 50°N (henceforth OHT_{50N}) and ocean heat convergence anomalies between 55-40°N several years later are highly correlated (fig. 3.2c). This effect results from the slow southward propagation of the OHT anomaly North of 35°N, and is in line with the findings of Zhang and Zhang (2015). Because the OHT anomaly propagates more slowly southward than it advects heat northward, an ocean heat convergence anomaly arises just North of the OHT anomaly, and an ocean divergence anomaly arises just South of the OHT anomaly. The heat convergence signal is particularly strong and long-lasting between 45-55°N, where I find a high positive correlation when OHT_{50N} leads ocean heat convergence by 0 to 8 years (fig. 3.2c,d). The time lags of maximum correlation between OHT_{50N} and ocean heat convergence anomalies decrease with increasing latitude of the area affected by ocean heat convergence (fig. 3.2d). Because of the faster propagation of the OHT anomalies South of 35°N, I do not find a ocean heat convergence signal South of 40°N (fig. 3.2c). The strong and long-lasting ocean heat convergence anomalies North of 40°N accumulate heat and potentially lead

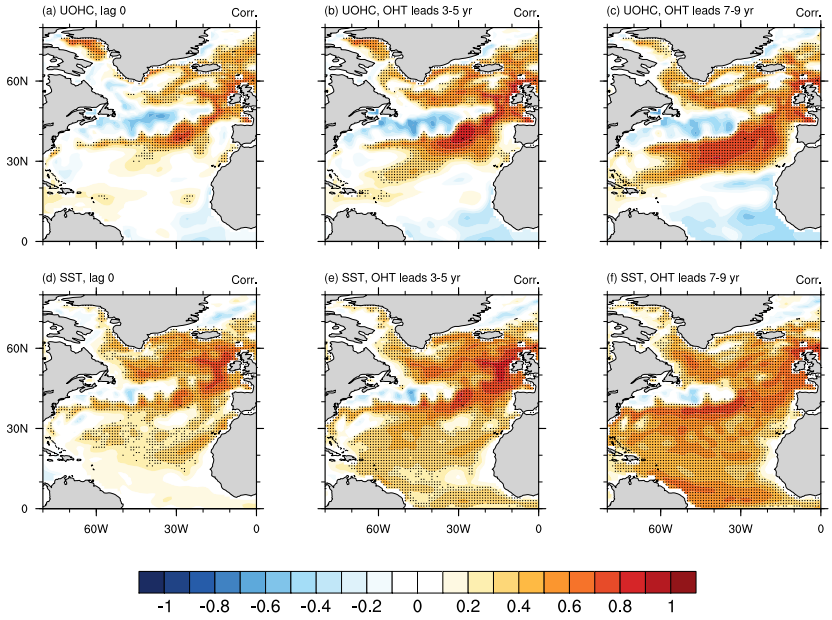


Figure 3.3: Point-by-point correlation of OHT_{50N} with upper ocean heat content of the upper 700m in the North Atlantic in ASSIM at lag 0 (a) and when OHT_{50N} leads by 3-5 years (b) and 7-9 years (c), and with SSTs at lag 0 (d) and when OHT_{50N} leads by 3-5 years (e) and 7-9 years (f). Stippling indicates significant correlations at the 99% level.

to persistent UOHC (Dong et al., 2007; Zhang and Zhang, 2015).

In ASSIM, I find an area of UOHC in the North Atlantic to be significantly correlated with OHT_{50N} variability at lag 0, as well as when OHT_{50N} leads by 3-5 and 7-9 years (fig. 3.3a-c). The highest correlations are located in the northeast North Atlantic and form a crescent shape around a strong negative correlation in the Gulf Stream region. The overall connection of UOHC to OHT_{50N} increases with increasing years that OHT_{50N} leads, but the pattern stays the same. SST anomalies (fig. 3.3d-f) show largely the same shape of correlation to OHT_{50N} anomalies as UOHC anomalies at all time lags (fig. 3.3a-c). Like with UOHC, the connection of SSTs to OHT_{50N} increases with increasing years that OHT_{50N} leads. Most of this connection is found North of 40°N.

Correlation patterns of sea surface height (SSH) anomalies to OHT_{50N} (fig. 3.4a-c) are very similar to those of UOHC. SSH anomalies primarily originate from changes in the ocean and are barely affected by atmospheric processes. Thus, UOHC anomalies can be assumed to show the oceanic contribution to temperature changes in the surface ocean in regions where they coincide with SSH anomalies. This is a connection that was also shown by Zhang and Zhang (2015).

At the ocean surface, the influence of OHT_{50N} variability on sea surface salinity is closely connected to the OHT-influence on SSTs as shown by correlation maps in figure 3.4d-f. This supports that SSTs are a surface signal that arises from a combination of oceanic and atmospheric forcing. In conjunction with the similar time lags I find in

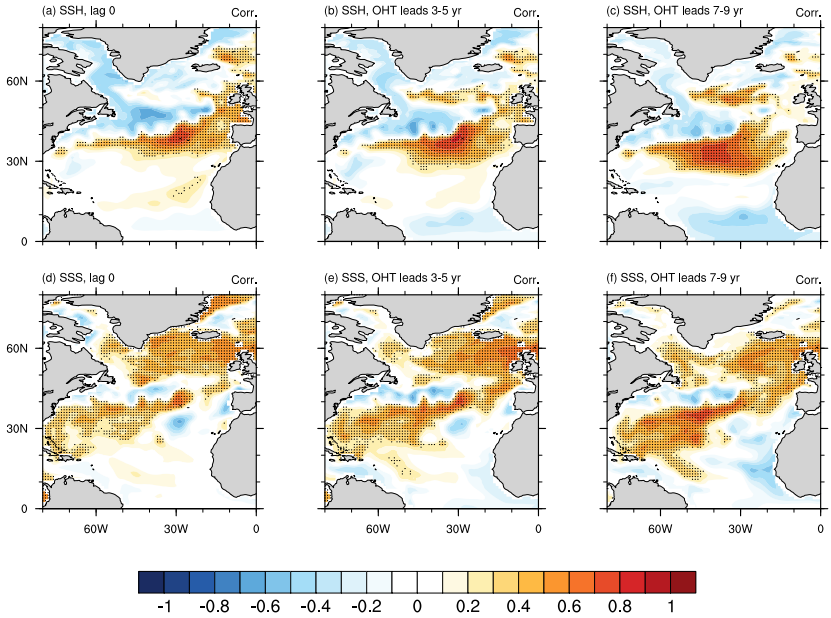


Figure 3.4: Point-by-point correlation of OHT_{50N} with sea surface height in the North Atlantic in ASSIM at lag 0 (a) and when OHT_{50N} leads by 3-5 years (b) and 7-9 years (c), and with sea surface salinity at lag 0 (d) and when OHT_{50N} leads by 3-5 years (e) and 7-9 years (f). Stippling indicates significant correlations at the 99% level.

this analysis and in the ocean heat convergence correlation (cf. fig. 3.2c,d), I conclude that the mechanism described by Zhang and Zhang (2015) as leading to the AMOC Fingerprint is responsible for the OHT-UOHC/SST correlation pattern.

The correlation patterns found for UOHC and SSTs correspond well with the first principal component of their respective empirical orthogonal functions (*EOFs*, fig. 3.5). In addition, the timeseries corresponding to these EOFs are highly correlated with OHT_{50N} when OHT leads by 8 years (UOHC: 0.76, SST: 0.68). This underlines that OHT_{50N} variability and the mechanism leading to the AMOC Fingerprint are responsible for much of the UOHC and SST variability in the extratropical North Atlantic.

The UOHC and SST pattern I find differs from the pattern Zhang and Zhang (2015) described. Specifically, I find the strongest positive anomaly in the northeast Atlantic, while Zhang and Zhang (2015) find this positive anomaly in the central subpolar gyre (cf. fig 1.2a). Because of the different UOHC and SST pattern in the MPI-ESM, I will henceforth refrain from calling this pattern the *AMOC Fingerprint* to avoid confusion, but will instead refer to the MPI-ESM-specific AMOC Fingerprint as the *characteristic SST pattern*.

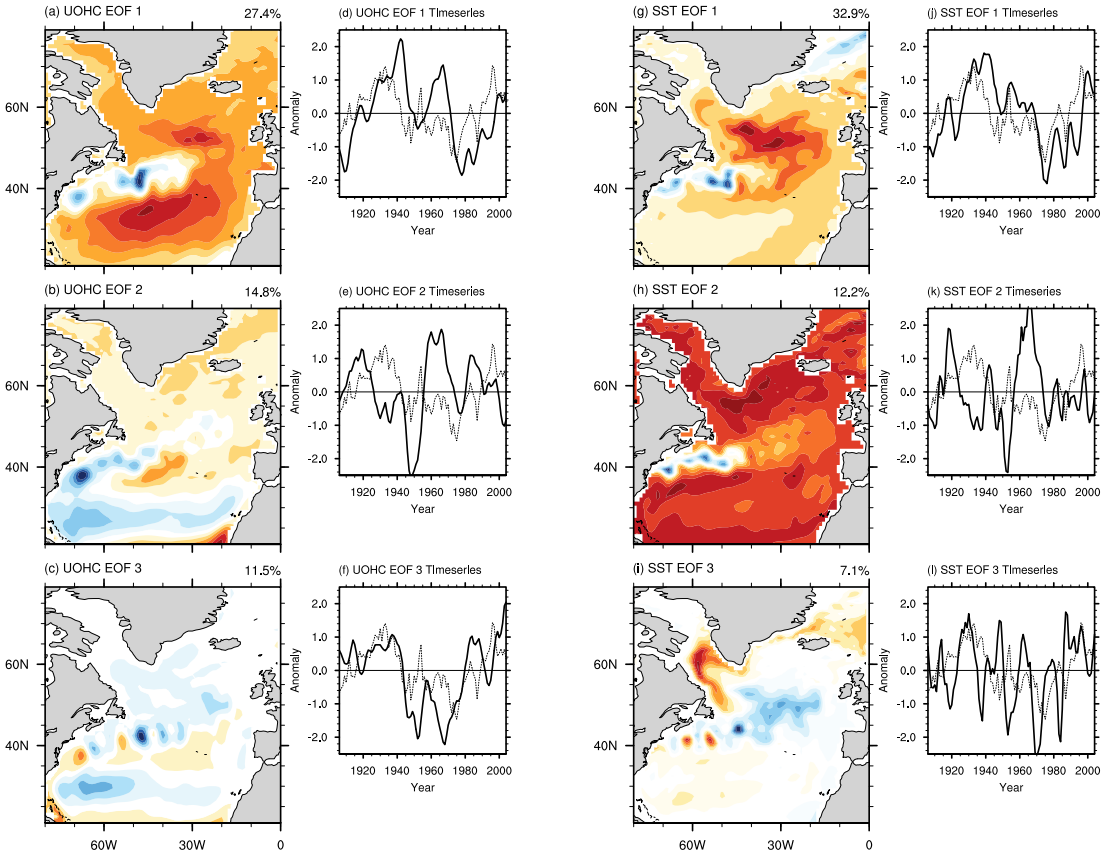


Figure 3.5: The first three Empirical Orthogonal Functions (EOFs; arbitrary units) of UOHC (a-c) and SSTs (g-i) in the North Atlantic between 20 and 80°N in ASSIM. Corresponding normalized time series are shown as thick solid lines in (d-f, in PW) and (j-l, in K), respectively. In the time series graphs, the dashed line shows OHT variability at 50°N [10^{-1} PW].

3.3 Subsampling Overturning States by Strong and Weak Ocean Heat Transport

Going beyond Zhang and Zhang (2015), I examine the emergence of the characteristic SST pattern and the role of UOHC in more detail with particular attention to strong and weak phases of OHT_{50N}. Composite mean OHT anomalies in the North Atlantic before, during, and after strong, weak, and neutral anomalies of OHT_{50N} in ASSIM reveal the respective influences of these OHT phases on SST anomalies separately. For that, overturning states are subsampled for years in which the OHT_{50N} is at least half a standard deviation above or below its mean of the preceding 30 years (fig. 3.1a). I define years that are not indentified as strong or weak OHT phases as *neutral* to understand climate variability that is not connected to particularly strong OHT anomalies for comparison. No conclusions presented in this dissertation change substantially, though, if other possible criteria are used to select strong, weak, and neutral phases of OHT_{50N}

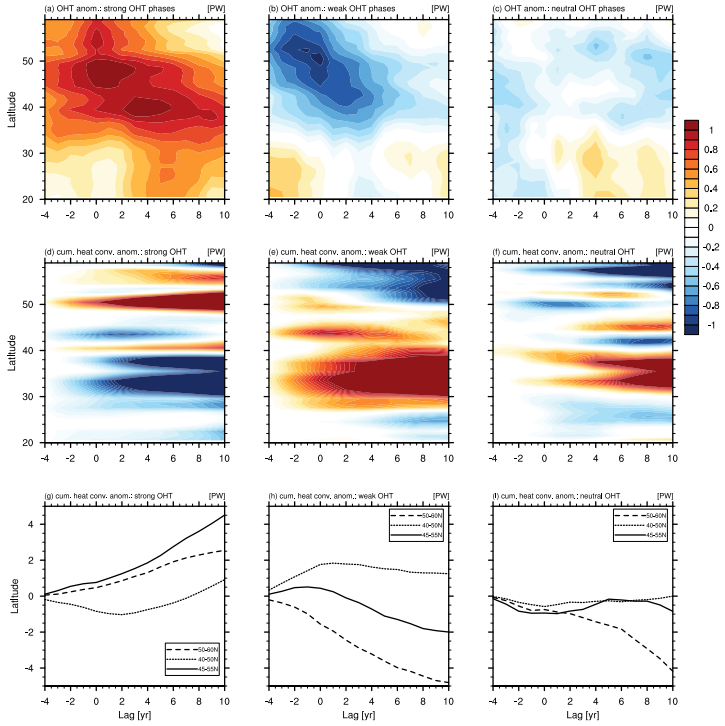


Figure 3.6: Composite means before, during, and after strong (left column), weak (middle column), and neutral (right column) OHT_{50N} phases in ASSIM. I show mean OHT anomalies (a-c) and cumulative ocean heat convergence relative to lag 0 (d-f) against latitude. (g-i) show composite mean cumulative heat convergence between 40-50°N, 45-55°N, and 50-60°N for strong (g), weak (h), and neutral (i) OHT phases at 45, 50, and 55°N, respectively. Positive lags indicate that OHT leads and vice versa.

(e.g. a full standard deviation above or below the mean, or above or below the mean of the previous 30 years).

Composite mean OHT anomalies associated with strong OHT_{50N} phases show a very pronounced and long-lived positive signal that propagates southward between 50 and 35°N. This OHT propagation persists for up to 9 years after a strong OHT_{50N} anomaly (fig. 3.6a). For weak OHT_{50N} phases, the negative southward propagating OHT anomaly disappears almost completely 6 years after the OHT_{50N} anomaly (fig. 3.6b). No distinct OHT propagation signal is connected to neutral OHT_{50N} phases (fig. 3.6c). This leads to different ocean heat convergence signals between phases of strong, weak and neutral OHT_{50N}.

Cumulative heat convergence anomalies illustrate the influence of strong and weak OHT_{50N} on the heat that accumulates in the ocean: cumulative heat convergence is calculated by integrating ocean heat convergence anomalies at every latitude between the OHT_{50N} anomaly and time lags between -4 and 10 (fig. 3.6d-f). Strong OHT_{50N} phases are followed by a strong and long-lived cumulative heat convergence anomaly that extends approximately from 45-55°N after 5 years (fig. 3.6d). This cumulative heat convergence anomaly is strongest between 45-55°N compared to other latitudinal bands

(fig. 3.6g). Weak OHT_{50N} phases are followed by a weak cumulative heat divergence anomaly that emerges at longer lag of 6 or more years and extends from 50°N northward (fig. 3.6e). Here, a positive cumulative heat convergence anomaly South of 50°N and a negative cumulative heat convergence anomaly North of 50°N are stronger than the heat convergence anomaly between 45-55°N (fig. 3.6h). The cumulative ocean heat convergence signal following neutral phases is largely similar to that following weak OHT_{50N} phases (fig. 3.6f), which does not lead to a distinct ocean heat convergence signal North of 45°N (fig. 3.6i). The strongest ocean heat convergence signal that cannot be found after all OHT_{50N} phases is the strong ocean heat convergence anomaly between 45-55°N after strong OHT_{50N} phases. This indicates that the effect of OHT_{50N} anomalies on the characteristic SST pattern is only strong between 45-55°N after strong OHT_{50N} phases, which would indicate an asymmetric response of SSTs to subpolar OHT. Other heat convergence signals are symmetric between all different OHT phases which indicates an influence of OHT variability on surface temperatures in these latitudinal bands. However, no large differences of temperature signals between different OHT_{50N} phases can be expected there.

Composite mean upper ocean heat content anomalies during and after phases of strong and weak OHT_{50N} (fig. 3.7) show a similar shape as the UOHC correlation maps in figures 3.3a-c (keeping in mind that positive correlations correspond to positive composite means for strong OHT phases and negative composite means for weak OHT phases). Following strong OHT_{50N} phases, a strong positive UOHC anomaly between 45-55°N appears in the northeast Atlantic and increases in strength over time (fig. 3.7a-c). By contrast, the UOHC anomaly is only weakly negative in this area after weak OHT_{50N} phases at all time lags (fig. 3.7d-f). Neutral phases are, particularly North of 40°N, not obviously connected to the correlation patterns I previously found (fig. 3.7g-i). This is in line with the findings based on the composite mean cumulative heat convergence anomalies (cf. fig. 3.6d-i). The role of the ocean in the formation of this UOHC anomaly is underpinned by the fact that I find a SSH anomaly of very similar shape at the same time (fig. 3.8).

Another signal I find in the composite mean UOHC anomalies is a dipole between the Gulf Stream region and its South-Eastern edge (fig. 3.7). This pattern can also be attributed to composite mean cumulative ocean heat convergence anomalies (cf. fig. 3.6d-i). However, it appears to be a symmetric response as this pattern arises after neutral, weak and strong OHT_{50N} phases at similar strengths. UOHC in that area is therefore likely to arise from physical processes that are not connected to the southward propagation of OHT phases in the North Atlantic region and associated ocean heat convergence. I also find this response in SSH anomalies (fig. 3.8).

Composite mean SST anomalies compared to the composite mean UOHC anomalies described above show a similar pattern North of 40°N (fig. 3.9). I find a strong signal

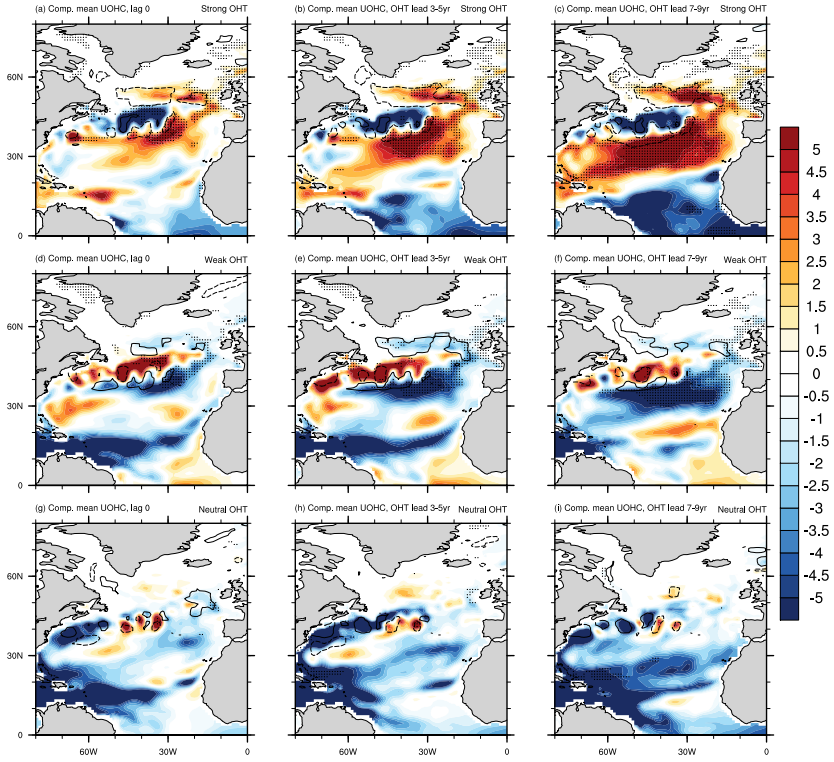


Figure 3.7: Composite mean upper ocean heat content anomalies [PW] in ASSIM, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean UOHC 3-5 (b,e,h) and 7-9 (c,f,i) years after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level. Contours show significant (at the 99% level) net ocean-atmosphere surface heat fluxes into the ocean (solid contours) and out of the ocean (dashed contours).

in composite mean SST anomalies in the northeast Atlantic that propagates westward over time after both strong and weak OHT_{50N} phases. Again, neutral OHT_{50N} phases are not strongly connected to SST anomalies North of 40°N. After strong OHT_{50N} phases, composite mean SST anomalies grow continuously stronger over time (fig. 3.9a-c), whereas composite mean SST anomalies after weak OHT_{50N} phases become very weak in the northeast Atlantic after 7-9 years (fig. 3.9f). I therefore conclude that ocean heat convergence influences SSTs most strongly in the northeast Atlantic several years after phases of strong OHT_{50N}. Moreover, the asymmetric response of SST anomalies in the northeast subpolar North Atlantic that originates from asymmetric OHT dynamics in the North Atlantic is one important finding from this chapter.

3.4 The Influence on Surface Air Temperatures

I now examine the influence of the mechanism leading to the characteristic SST pattern on ocean-atmosphere surface heat fluxes and subsequently surface air temperatures over Europe. This analysis will give an indication of the effect that changes in subpolar North Atlantic ocean circulation can have on European temperatures up to a decade

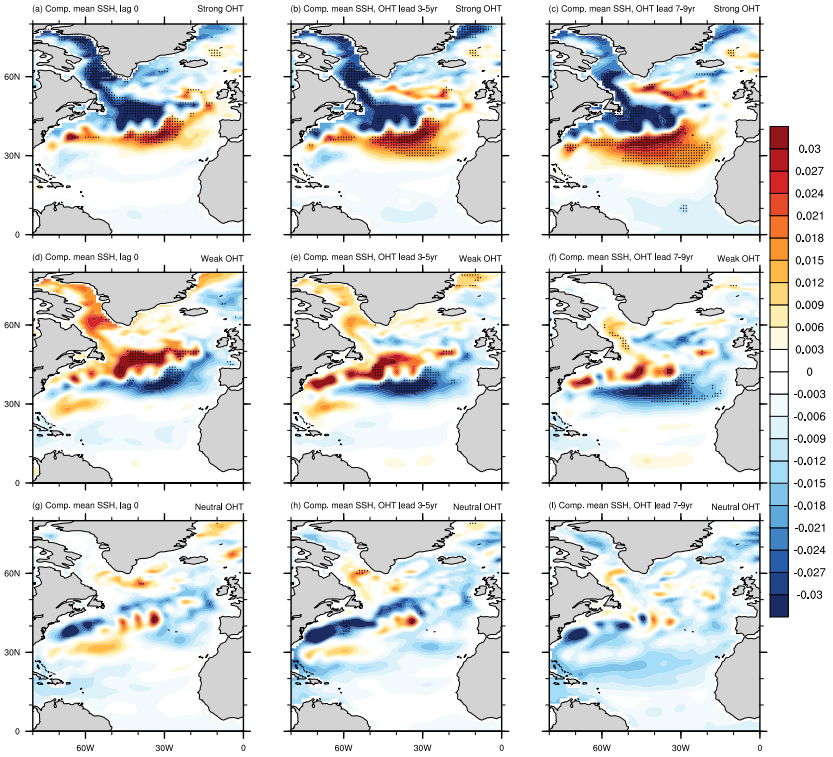


Figure 3.8: Composite mean sea surface height anomalies [m] in ASSIM, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSH 3-5 (b,e,h) and 7-9 (c,f,i) years after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level.

in advance, which might indicate high decadal surface temperature prediction skill over land.

SHFs are correlated with OHT_{50N} in similar areas as SSTs: in the subpolar gyre region and South of the gulf stream front (fig. 3.10a-c, cf. fig. 3.3). Unlike with SSTs, SHFs are negatively correlated to OHT_{50N} variability: positive phases of ocean heat transport are associated with later upward surface heat fluxes. As positive SST anomalies largely coincide with upward SHFs, it can be concluded that the ocean dominates surface temperature changes in these areas on decadal time scales. Interestingly, I find very little dependence of the correlations of SHFs with OHT_{50N} on the amount of years by which OHT leads. This indicates that the flux of heat from the ocean into the atmosphere is relatively constant. However, I do find a slight increase in SHF anomalies with lead time, particularly in the subpolar gyre region.

SAT anomalies in the North Atlantic region are also highly correlated to OHT_{50N} variability. With increasing lag (OHT_{50N} anomalies lead), the area of significant positive correlations increases strongly (fig. 3.10d-f). SAT variability is significantly correlated to OHT_{50N} phases over the northeast Atlantic, Scandinavia, the Iberian Peninsula, and large parts of northern Africa at lag 0 (fig. 3.10d). With increasing years that OHT_{50N} anomalies lead, significant SAT correlations propagate inland, covering much of Europe

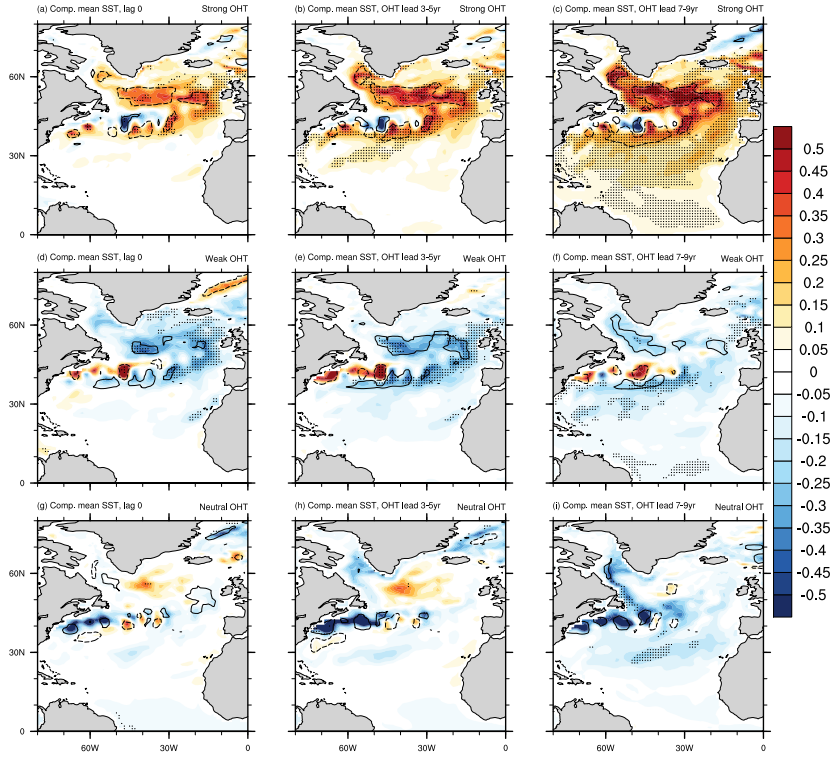


Figure 3.9: Composite mean SST anomalies [K] in ASSIM, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSTs 3-5 (b,e,h) and 7-9 (c,f,i) years after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level. Contours show significant (at the 99% level) net ocean-atmosphere surface heat fluxes into the ocean (solid contours) and out of the ocean (dashed contours).

and the Arabian Peninsula when OHT_{50N} leads by 3-5 years, after which correlations stay largely constant (fig. 3.10e,f). This shows a robust statistical relationship between OHT_{50N} variability and annual mean surface temperature variability over Europe on the decadal time scale. To better understand the specific influence of OHT_{50N} phases on European SATs, I now examine the mean SHF and SAT state connected to strong and weak phases of OHT_{50N}.

Significant composite mean SHFs (contours in fig. 3.9; upward SHFs generally correspond to positive SST anomalies and vice versa) show some asymmetries between strong and weak OHT_{50N} phases, which highlights the impact of OHT dynamics on surface heat fluxes. Note that I use total values and not anomalies for SHFs. Specifically, a zonal asymmetry appears after 7-9 years with upward SHFs across most of the North Atlantic between 50-55°N and 20-50°W where SST composites are high after strong OHT_{50N} phases (fig. 3.9c). I find downward SHFs in the Labrador Sea and in the northeast Atlantic between 50-55°N and 10-20°W where SST composites are weak after weak OHT_{50N} phases (fig. 3.9f). These areas can be regarded as areas where SHFs contribute strongly to SST variability alongside OHT. This finding indicates that the ocean influences the atmosphere more strongly after strong OHT_{50N} phases than after

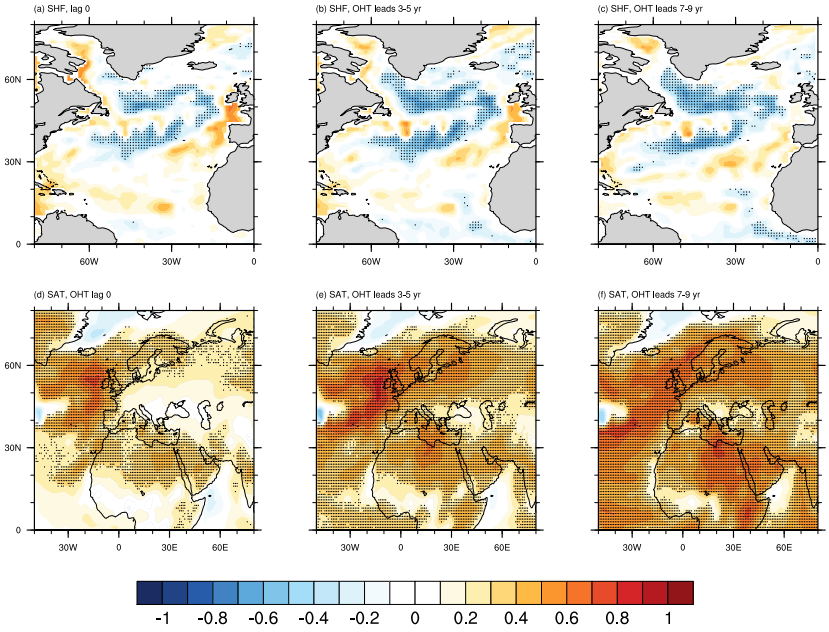


Figure 3.10: Point-by-point correlation of OHT_{50N} with ocean-atmosphere surface heat fluxes in ASSIM at lag 0 (a) and when OHT_{50N} leads by 3-5 years (b) and 7-9 years (c), and with surface air temperatures at lag 0 (d) and when OHT_{50N} leads by 3-5 years (e) and 7-9 years (f). Stippling indicates significant correlations at the 99% level.

weak OHT_{50N} phases, which might lead to different SAT signatures related to strong and weak OHT_{50N} phases.

I find distinct signals of strong and weak OHT_{50N} phases in SATs over Europe. Specifically, strong OHT_{50N} phases are followed by positive temperature anomalies over the subpolar gyre and western and northern Europe that increase with increasing lag (fig. 3.11a-c), whereas weak OHT_{50N} phases are followed by significant SAT anomalies over the subpolar gyre and Europe in general with the strongest anomaly over eastern Europe (fig. 3.11d-f). This could indicate a northward shift of the jet stream with stronger warming in the northeast Atlantic after strong OHT_{50N} phases as indicated by the composite mean SST analysis shown in figure 3.9. Interestingly, there is almost no significant signal in composite mean SATs following neutral OHT_{50N} phases (fig. 3.11g-i), which indicates that, on decadal time scales, the characteristic SST pattern and the physical mechanism preconditioning its formation - for both strong and weak OHT_{50N} phases - play an important role in modulating European SAT anomalies. However, SATs vary strongly on the seasonal time scale. I will therefore now examine the connection of seasonal SATs to OHT_{50N} variability in the subpolar North Atlantic on the decadal time scale. Because of the limited influence of neutral OHT_{50N} phases on European SATs I find here, I will henceforth focus exclusively on strong and weak OHT_{50N} phases.

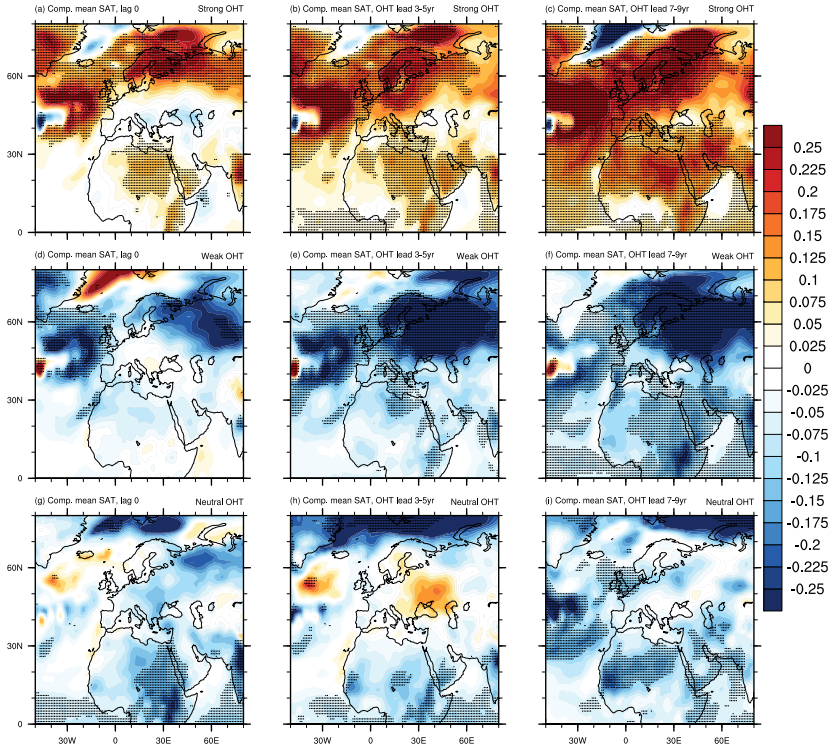


Figure 3.11: Composite mean SAT anomalies [K] in ASSIM, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSTs 3-5 (b,e,h) and 7-9 (c,f,i) years after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level.

3.5 Seasonal Impact of Ocean Dynamics on the Atmosphere

The analysis of seasonal mean SHFs and SATs allows me to assess the oceanic influence on SATs at the seasonal level as suggested by e.g. Arthun et al. (2017). Because of the strong SST signal on the decadal time scale (cf. fig. 3.9), I focus on composite mean SHFs and SATs in ASSIM 7-9 years after strong and weak OHT_{50N} phases.

I generally find upward SHFs after strong OHT_{50N} phases and downward SHFs after weak OHT_{50N} phases (fig. 3.12a-j). SHFs are strong in winter and fall after both strong and weak OHT_{50N} phases, and after strong OHT_{50N} phases in spring (fig. 3.12b,c,e,g,j). Winter and spring SHF, too, show the zonal asymmetry I found on annual mean SHFs (cf. fig. 3.9) with strong SHFs towards the western North Atlantic after strong OHT_{50N} phases and towards the eastern North Atlantic after weak OHT_{50N} phases (fig. 3.12b,c,g,h). These findings indicate that low-frequency variability from the ocean is transported into the atmosphere most strongly in winter, spring and fall, influencing SATs. They further indicate an asymmetric influence of strong and weak phases of subpolar OHT on seasonal SATs.

The response of seasonal SATs is generally symmetric between strong and weak OHT_{50N} phases over land (fig. 3.12l-o,q-t). There are some exceptions: in winter, SATs over Scan-

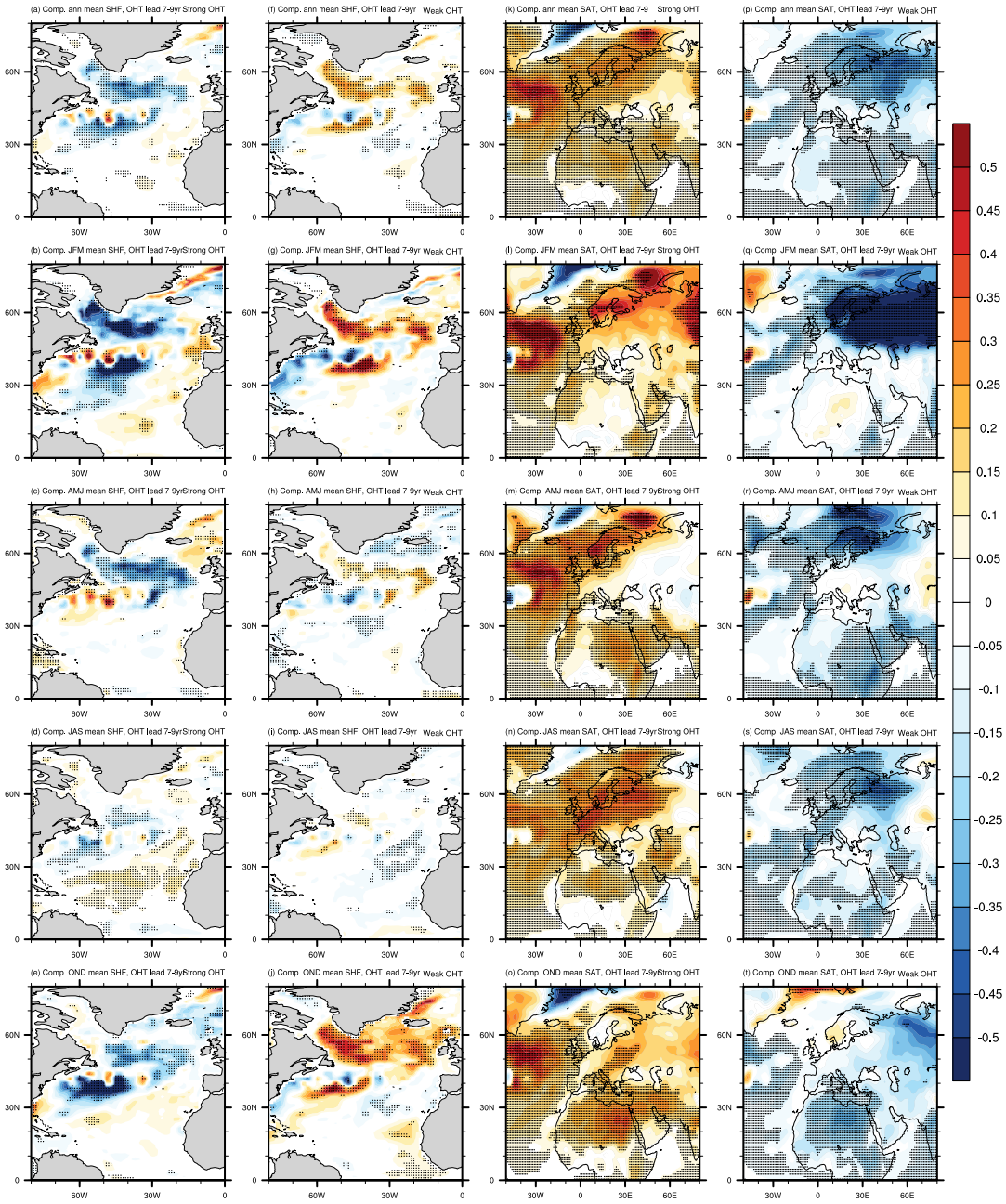


Figure 3.12: Composite mean SHFs $[W/m^2]$ in ASSIM when strong (a-e) and weak (f-j) OHT_{50N} phases lead by 7-9 years, and composite mean SAT anomalies [K] in ASSIM when strong (k-o) and weak (p-t) OHT_{50N} phases lead by 7-9 years. I show composites for annual (first row), JFM (second row), AMJ (third row), JAS (fourth row) and OND (fifth row) means. Stippling indicates significance at the 90% level.

dinavia show strong anomalies after both strong and weak OHT_{50N} phases. However, the anomalies only differ strongly from the mean variability after weak OHT_{50N} phases. The large areas of strong but insignificant wintertime SAT anomaly around Scandinavia indicate a high overall temperature variability in that season and area - this is likely related to the impact of the winter North Atlantic Oscillation (e.g. Visbeck et al., 2001). In summer, SAT anomalies are significant over Scandinavia after weak OHT_{50N} phases, and significant over large parts of western Europe and the UK after strong phases of OHT_{50N}. As I find little connection of SHF and SAT anomalies of any same season, there likely is a seasonal-scale lag in the influence of SHF anomalies on SAT variability over Europe (as suggested by e.g. Czaja and Frankignoul, 2002). On the decadal time scale time, any SAT signal is likely to be modulated by the ocean. Further, a connection of SAT anomalies over Scandinavia to subpolar OHT variability is in line with the findings from Arthun et al. (2017), which indicates that the mechanism leading to the characteristic SST pattern might play a role in modulating the decadal SAT prediction skill they find over Scandinavia.

3.6 Discussion

In this chapter, I show in an initialized version of the MPI-ESM-LR that the suggestion of Zhang and Zhang (2015) that the ocean heat transport in the North Atlantic region modulates SST variability has merit. However, some of my findings require discussion, like the role of the AMV in my findings, the model specificity of my findings, the contribution of strong and weak OHT_{50N} phases to SST variability, and the role of ocean-atmosphere surface heat fluxes.

The Atlantic Multidecadal Variability (AMV) is a major driver of North Atlantic temperature variability (e.g. Clement et al., 2015; Zhang et al., 2016). The mechanism I describe here, as well as the characteristic SST pattern, seem to be closely related to the AMV in the MPI-ESM-LR (fig. 3.13). A connection of a similar SST shape to the AMV was also shown in Delworth et al. (2017). In my study, OHT_{50N} leads the AMV by 8 years (max. correlation: 0.9), which indicates that the ocean at least contributes to the formation of the AMV. As a result, the physical mechanism invoked in this study can be seen not only as a contributor to decadal SST variability in the northeast Atlantic, but also as a contributor to AMV variability on the same time scale. Note that the AMV time series in ASSIM is also closely connected with the AMV in HadISST (fig. 3.13), which is an indication that ASSIM reflects *real* climate variability in the North Atlantic reasonably well.

A major caveat of the results I present here using ASSIM is the use of three ensemble members of one model. The small ensemble size was in the past shown to be appropriate

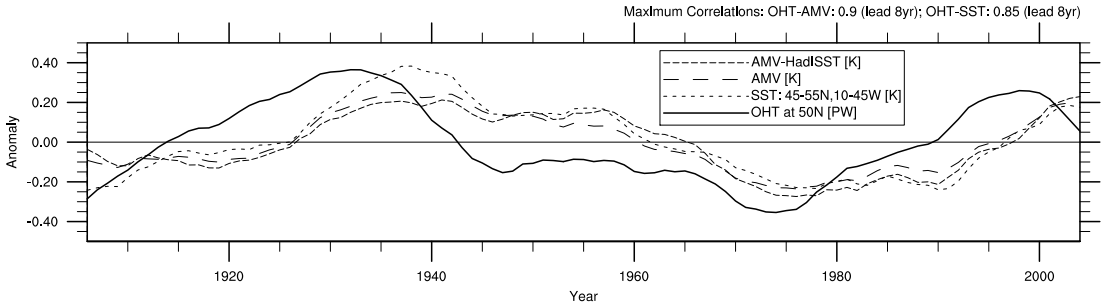


Figure 3.13: Time series from ASSIM of OHT_{50N} anomalies (solid line, [PW]), mean temperature anomalies in the box indicated in figs. 4.5, 4.6 (stippled line, [K]), and AMV as defined as average SST anomalies between 0 and 60°N in the North Atlantic (dashed line, [K]). I also show the AMV time series from HadISST observations [K] in narrow dashes. Time series are low-pass filtered with an 11-year running mean.

to study North Atlantic climate variability (Müller et al., 2014, 2015). To some degree, my findings have to be interpreted as model-specific. In particular, I find an eastward displacement of the characteristic SST pattern which connecteds to the OHT_{50N} variability compared to Zhang and Zhang (2015), while finding a very similar dynamical explanation for the origin of the characteristic SST pattern. This indicates that the MPI-ESM-LR reacts similarly to OHT anomalies in the North Atlantic as the GFDL CM2.1 used in Zhang and Zhang (2015), but these anomalies affect a different area in the North Atlantic. The more eastward characteristic SST pattern could indicate a more zonal Gulf Stream in the MPI-ESM compared to the GFDL model - in that case, the use of a higher resolution model might alleviate this problem (see e.g. Drews and Greatbatch, 2017). However, as the physical mechanism leading to the variability patterns identified here is consistent with previous publications, the conclusion that OHT in the subpolar North Atlantic influences SST variability as described above - probably somewhat further to the west - is valid.

An important finding from this research is the asymmetric response of North Atlantic SSTs to strong and weak phases of subpolar OHT. As a result, stronger and more persistent temperature anomalies can be expected in the North Atlantic after strong than after weak OHT_{50N} phases. An implication of this finding is of methodological nature: it suggests that simple correlation analyses are not enough to understand the dynamics of a system. While this finding is not necessarily new, this study emphasizes again that conclusions drawn from a correlation study should be taken with a pinch of salt, and deeper analyses of different possible states of a system are needed to fully comprehend its dynamics.

I find different lengths of strong and weak OHT_{50N} phases in ASSIM (cf. fig. 3.1). Although the numbers of years of strong and weak OHT_{50N} as identified by the criterion I define are similar (weak OHT_{50N}: 36 years; strong OHT_{50N}: 40 years), strong OHT_{50N}

phases appear more coherently than weak ones. This might affect the composite mean analysis I conduct, as longer OHT phases might also have a more pronounced impact on SSTs. I will take up this issue in chapter 5 of this dissertation, attempting to resolve it.

This study generally supports the suggestion by earlier studies that temperatures in the eastern North Atlantic are controlled by oceanic heat advection, while temperatures in the western North Atlantic are controlled by surface heat fluxes (e.g. Robson et al., 2017). This study suggests, however, that this notion depends on the strength of ocean heat transport. When OHT_{50N} is strong, I find in ASSIM that the northeast Atlantic is strongly influenced by ocean heat advection. During and after weak OHT_{50N}, SHFs become more important in the modulation of surface temperatures in the northeast Atlantic (fig. 3.9). This might be related to the weaker underlying UOHC anomaly following a weak OHT_{50N} phase.

I show that decadal variability of North Atlantic SSTs is strongly influenced by low-frequency ocean heat transport variability in the MPI-ESM-LR. As ocean variability is thought to modulate climate predictability on decadal time scales, I will analyze in the upcoming chapter the influence of the dynamical mechanism presented above on the skill of decadal surface temperature hindcasts in the North Atlantic region.

4 | Decadal Hindcasts in the North Atlantic Region

4.1 Introduction

In the previous chapter I show that the suggestion by Zhang and Zhang (2015) that ocean overturning variability in the subpolar North Atlantic influences SST variability in the North Atlantic region for up to a decade ahead has merit. Besides describing the physical mechanism leading to the characteristic SST pattern, Zhang and Zhang (2015) hypothesized that this physical mechanism influences surface temperature prediction skill in the North Atlantic on the same time scale. I use this chapter to reconcile the findings from chapter 3 with Zhang's hypothesis.

In the past, case studies showed a connection of particularly strong and weak phases of ocean overturning to high skill in SST hindcasts (e.g. Yeager et al., 2012; Robson et al., 2013, 2014). The mechanism shown by Zhang and Zhang (2015) suggests that findings from these case studies are generally applicable. However, the influence of the physical mechanism leading to the characteristic SST pattern on decadal surface temperature hindcast skill was not assessed so far.

This lack of systematic assessment of the influence of ocean overturning on decadal SST hindcast skill is mainly due to methodological limitations. For such an assessment, a long time series with many individual hindcast simulations is necessary to ensure statistically robust results. The HC hindcast experiments for the period 1901-2010 that I introduced earlier represent such a set of hindcast simulations.

This chapter builds on the connection between OHT_{50N}, UOHC, SSTs, SHFs and SATs described in chapter 3. I investigate systematically whether and how predictability of these parameters on the decadal time scale (3-5 years and 7-9 years ahead: at *lead years* 3-5 and 7-9), depends on the strength of the OHT_{50N} at the initialization of the hindcast. I evaluate HC against ASSIM using anomaly correlation coefficients (ACCs), for the entire time series, after strong, weak, and neutral OHT_{50N} phases to connect predictability to the physical mechanism leading to the characteristic SST pattern. Comparing the influences of OHT variability and UOHC persistence on the skill of decadal surface temperature hindcasts diagnoses the relative contributions of strong

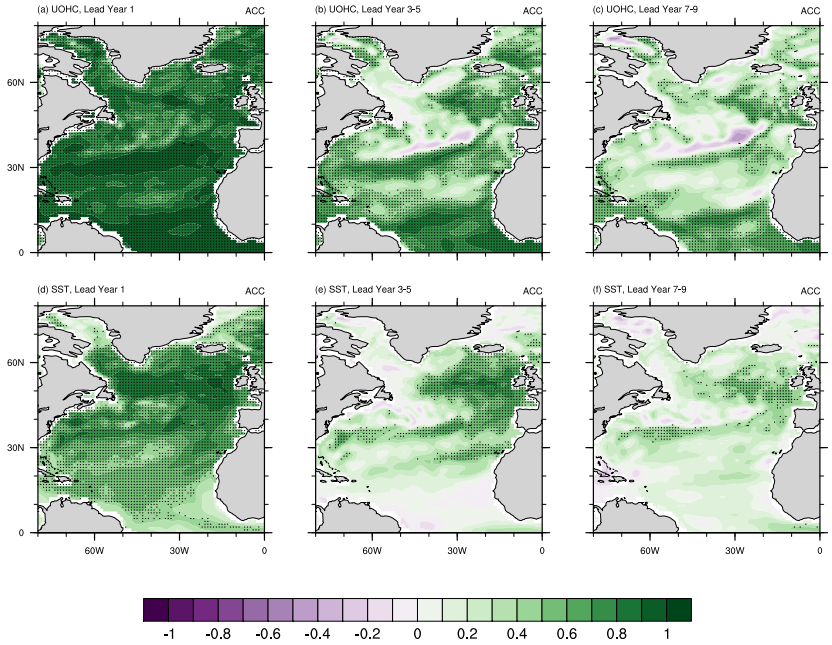


Figure 4.1: Anomaly Correlation Coefficients (ACCs) of UOHC (upper 700m) from the HC experiments against ASSIM at lead years 1 (a), 3-5 (b) and 7-9 (c). I show ACCs for SSTs at lead years 1 (d), 3-5 (e) and 7-9 (f). Stippling indicates significant ACCs at the 99% level.

and weak phases of OHT_{50N} and UOHC persistence to hindcast skill separately. I then examine the influence of skillful SST hindcasts on the skill of decadal hindcasts of seasonal surface air temperatures.

4.2 Skillful Hindcasts of North Atlantic Surface Temperatures

For 1901-2010, I find upper ocean heat content to be significantly predictable in the entire North Atlantic at lead year 1 with particularly high ACCs in the northeast Atlantic (fig. 4.1a). At lead years 3-5 and 7-9, I find significant ACCs for UOHC in the northeast Atlantic as well (figs. 4.1b,c). Similarly, SSTs show significant and high ACCs in the northeast Atlantic at lead years 1, 3-5 and 7-9 (fig. 4.1d-f).
High ACCs for annual mean surface heat fluxes and surface air temperatures (fig. 4.2) are mostly confined to regions of strong SST anomalies in the composite means (cf. fig. 3.3). This indicates a strong influence of the ocean on the skill of atmospheric hindcasts on decadal time scales. Like ACCs of SST hindcasts, ACCs of both SHF and SAT hindcasts are highest at lead year 1 and decrease afterwards. However, for SHFs and SATs, ACCs decrease much faster than for both UOHC and SSTs which is due to a comparatively strong influence of predictable low-frequency ocean variability on UOHC and SSTs. The shape of high decadal SHF ACCs appears more closely connected to

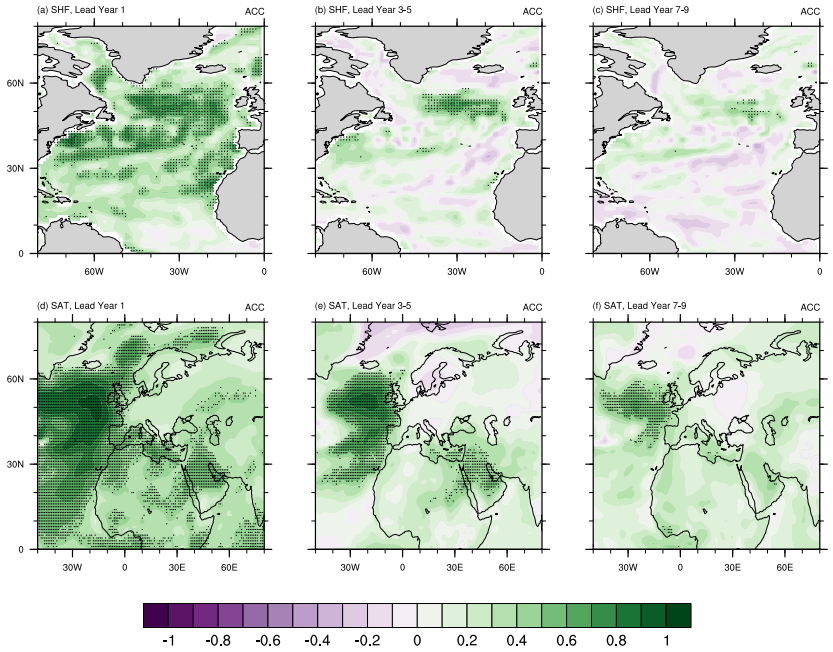


Figure 4.2: Anomaly Correlation Coefficients (ACCs) of SHFs from the HC experiments against ASSIM at lead years 1 (a), 3-5 (b) and 7-9 (c). I show ACCs for SATs at lead years 1 (d), 3-5 (e) and 7-9 (f). Stippling indicates significant ACCs at the 99% level.

SST variability than to SHF variability (cf. figs. 3.3, 3.10).

Unlike in the correlation analysis (cf. fig. 3.10), ACCs for annual mean SATs are insignificant over Europe on the decadal time scale (fig. 4.2). This is an indication that HC is not capable of capturing SAT variability accurately. However, the strong seasonality of SHF and SAT variability (cf. fig. 3.12) prevents any final conclusions at this point.

SHF ACCs show strong seasonality at all lead times (fig. 4.3). SHFs are mostly predictable in winter (JFM) and spring (AMJ) on all time scales, however, the spring predictability is clearly the strongest among the seasons on decadal time scales. Incidentally, spring is the season that shows one of the strongest SHF anomalies at the decadal time scale (cf. fig. 3.12). Further, this analysis indicates that SHFs are strongest and most predictable in spring. This could lead to significant ACCs of spring SATs on the decadal time scale.

ACCs for SATs show less seasonality than those of SHFs (fig. 4.4). However, some features stand out: ACCs at lead year 1 are highest in winter and spring (fig. 4.4d,g), SAT-ACCs at lead years 3-5 are almost independent of the season, and ACCs 7-9 years into the future are highest in winter, spring and fall. These features are not obviously connected to SHF predictability, but rather to the very little seasonal difference in hindcast skill that SSTs exhibit (not shown). This indicates that surface heat flux predictability does not have to be high to precondition good ACCs for SAT, but that SAT predictability is more directly connected to SST predictability. I do not find

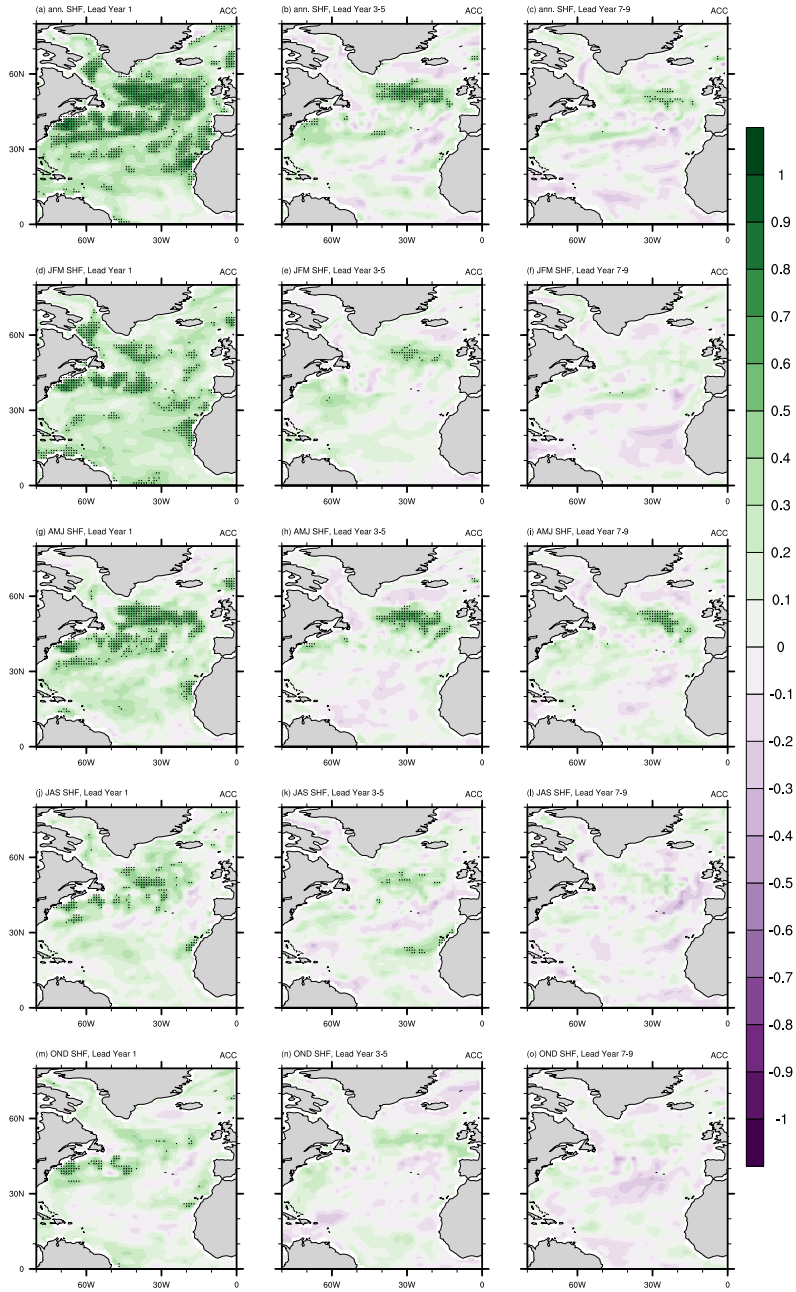


Figure 4.3: Anomaly Correlation Coefficients (ACCs) of SHFs from the HC experiments against ASSIM at lead years 1 (left column), 3-5 (middle column) and 7-9 (right column). Lines show ACCs for annual means (a-c), JFM means (d-f), AMJ means (g-i), JAS means (j-l), and OND means (m-o). Stippling indicates significant ACCs at the 99% level.

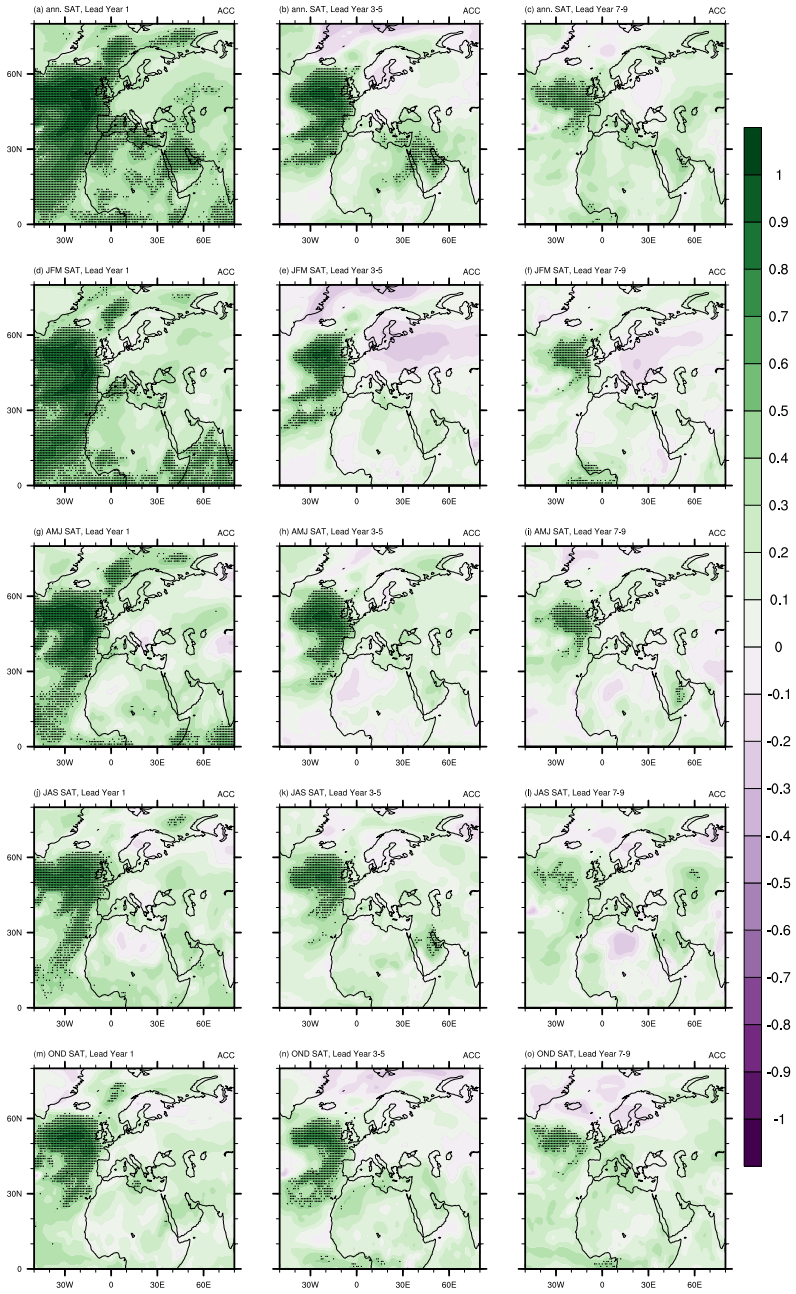


Figure 4.4: Anomaly Correlation Coefficients (ACCs) of SATs from the HC experiments against ASSIM at lead years 1 (left column), 3-5 (middle column) and 7-9 (right column). Lines show ACCs for annual means (a-c), JFM means (d-f), AMJ means (g-i), JAS means (j-l), and OND means (m-o). Stippling indicates significant ACCs at the 99% level.

significant ACCs in decadal hindcasts for seasonal SATs over land, not even in spring: significant ACCs are confined to the ocean areas.
The general structure of both UOHC and SST ACCs at lead years 3-5 and 7-9 resembles the shape of the characteristic SST pattern (cf. fig. 3.3) and the composite mean UOHC and SST patterns I discuss above (cf. figs. 3.7, 3.9). This indicates that OHT_{50N} anomalies play an important role in modulating UOHC and SST predictability. Because I find that SAT predictability is closely connected to SST predictability, ACCs of SAT hindcasts over land could also be connected to OHT_{50N} variability. I will examine this hypothesis in the upcoming section by looking at ACCs after strong and weak OHT_{50N} phases separately.

4.3 Hindcast Skill after Subsampling OHT Phases

4.3.1 Influence of Ocean Heat Transport on Decadal Predictability in the Ocean

I analyze ACCs of UOHC in the North Atlantic for hindcasts started in years of strong (fig. 4.5a-c), weak (fig. 4.5d-f) and neutral (fig. 4.5g-i) OHT_{50N} phases separately to identify the specific influence of the phase of OHT_{50N} on hindcast skill in the North Atlantic. I find significant UOHC ACCs at all lead years after both strong and weak OHT_{50N} anomalies. At lead year 1, ACCs are very similar after strong and weak OHT_{50N} phases, and for the entire time series (figs. 4.5a,d; 4.1a). At lead years 3-5 and 7-9, ACCs are generally higher after both strong and weak phases of OHT_{50N} than for the entire time series. Large areas of significant UOHC ACCs, like the area South of the Gulf Stream front and the Tropical Atlantic, show almost no asymmetry between strong and weak OHT_{50N} phases. However, I find significant ACCs in large parts of the northeast Atlantic after strong OHT_{50N} phases (fig. 4.5b,c), while ACCs are significant in the central North Atlantic after weak OHT_{50N} phases (fig. 4.5e,f).
After neutral OHT_{50N} phases, ACCs for UOHC in the North Atlantic are only high at lead year 1 (fig. 4.5g). At lead years 3-5, the ACC pattern for UOHC after neutral OHT_{50N} phases resembles that following weak OHT_{50N} phases, although it is slightly weaker after neutral phases (fig. 4.5h). This indicates that, on this time scale, OHT_{50N} variability influences ACCs of UOHC hindcasts only weakly. A possible explanation for the emergence of high UOHC ACCs on this time scale could be the persistence of UOHC. At lead years 7-9, there is no considerable ACC pattern in UOHC hindcasts following neutral OHT_{50N} phases (fig. 4.5i), underscoring the high impact of OHT_{50N} variability on hindcast skill on this time scale.
ACCs of SSTs behave very similar to ACCs of UOHC. SST ACCs are significant at all lead years after both strong and weak OHT_{50N} phases (fig. 4.6). At lead year 1 they are similar to ACCs diagnosed for the entire time series (figs. 4.6a,d; 4.1d). At lead years

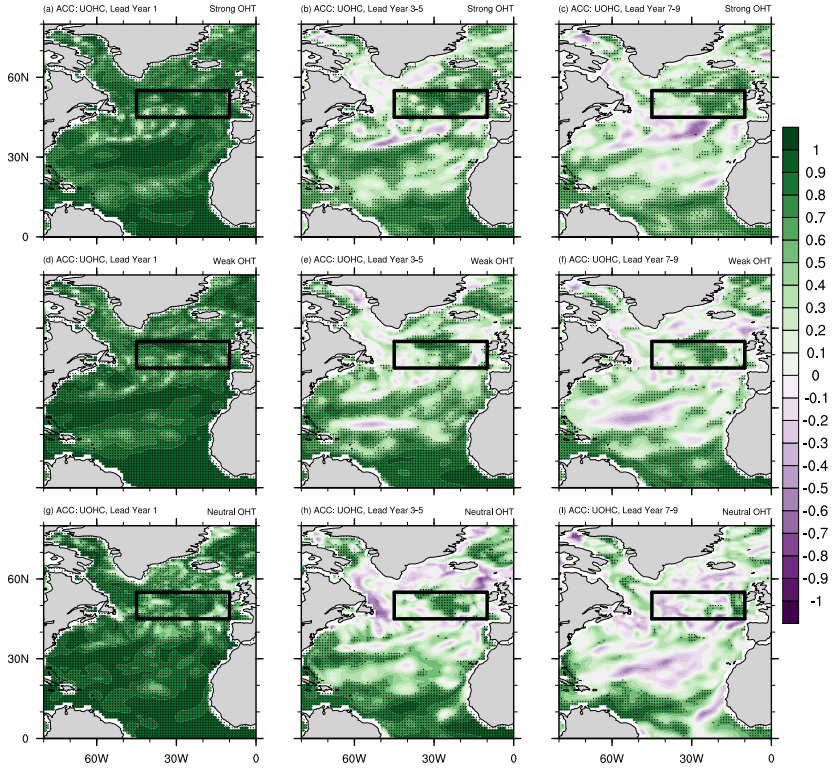


Figure 4.5: Anomaly Correlation Coefficients (ACCs) Upper ocean heat content (upper 700m) from the HC experiments against ASSIM after strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show ACCs at lead years 1 (a,d,g), 3-5 (b,e,h) and 7-9 (c,f,i). Stippled areas indicate significant ACCs at the 99% level. The black box is the area that UOHC is averaged over for figure 4.7a.

3-5 and 7-9, I find significant ACCs of SSTs to be higher and to cover larger areas after strong than after weak OHT_{50N} phases (fig. 4.6b,c,e,f). The conclusions concerning neutral OHT_{50N} phases are the same for SST as they are for UOHC (fig. 4.6g-i).
Many areas of significant ACCs for SSTs South of 40°N do not show much asymmetry between strong and weak OHT_{50N} phases. This agrees with my findings from the UOHC predictability study and suggests that the influence of OHT_{50N} on asymmetric predictability of UOHC and SSTs in those areas is limited. However, I find a zonal difference in ACCs between strong and weak OHT_{50N} phases with significant ACCs in the northeast Atlantic after strong OHT_{50N} phases, and more centrally located significant ACCs after weak OHT_{50N} phases.
I attribute the zonal asymmetry I find between the ACC patterns of both UOHC and SSTs at lead years 7-9 after strong and weak OHT_{50N} phases to the zonally asymmetric significant composite mean SHFs I show in figures 3.7 and 3.9. After both strong and weak OHT_{50N} phases, I find predictable UOHC and SSTs in areas that are characterized by little heat exchange with the atmosphere. This is an indication that significant ACCs are indeed connected to low-frequency ocean dynamics and that ACCs decrease strongly where the atmosphere contributes strongly to surface temperature variability.

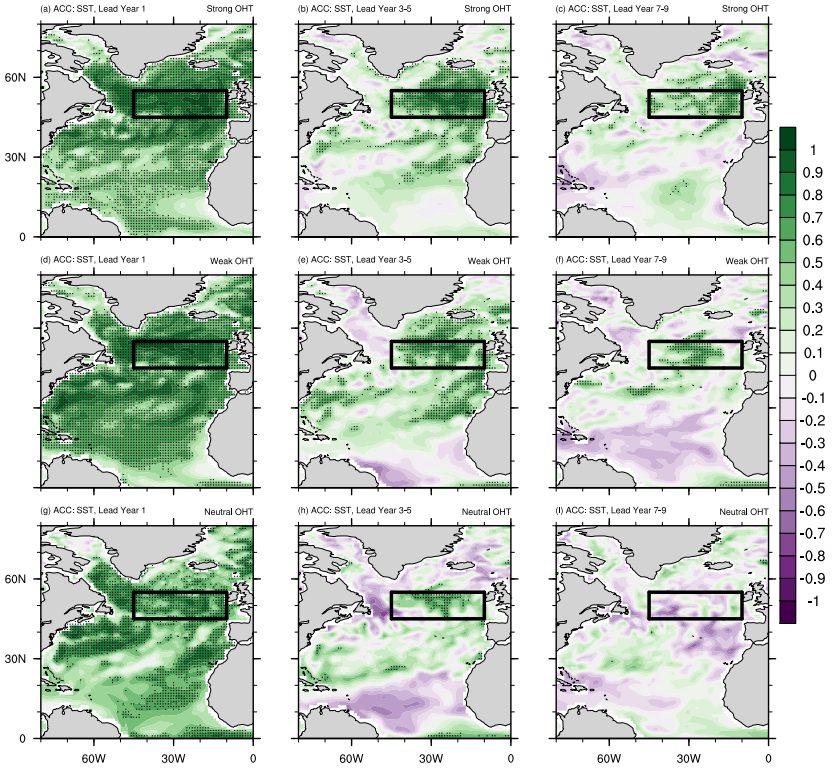


Figure 4.6: Anomaly Correlation Coefficients (ACCs) of SSTs from HC against ASSIM after strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show ACCs at lead years 1 (a,d,g), 3-5 (b,e,h) and 7-9 (c,f,i). Stippled areas indicate significant ACCs at the 99% level. The black box is the area that SSTs are averaged over for figure 4.7b.

ACCs for average UOHC and SSTs in the northeast Atlantic (45-55°N, 45-10°W, cf. black box in figs. 4.5 and 4.6) further illustrate the overall effect of strong and weak OHT_{50N} phases on decadal predictability of UOHC and SSTs (fig. 4.7). Note that the chosen box covers the area where I find significant predictability of UOHC and SSTs after both strong and weak OHT_{50N} phases. ACCs evaluated over the entire time series for both UOHC and SSTs are generally high in the northeast Atlantic for up to 9 lead years and outperform persistence forecast at lead years > 1 (fig. 4.7).

After strong OHT_{50N} phases, ACCs in the northeast Atlantic are significantly higher than after weak OHT_{50N} phases and for the entire time series at lead years 2-7 for UOHC (fig. 4.7a). Similarly, I find ACCs for UOHC after weak OHT_{50N} phases that are significantly lower than ACCs after strong OHT_{50N} phases and for the entire time series for lead years 2-6. UOHC ACCs after weak OHT_{50N} phases re-emerge after lead year 7, which might be connected to trends in UOHC anomalies.

SSTs show significantly higher ACCs after strong than after weak OHT_{50N} phases and for the entire time series at lead years 2-9 (fig. 4.7b). After weak OHT_{50N} phases, ACCs for SSTs largely follow ACCs for the entire time series. This analysis provides further indication that the mechanism leading to the characteristic SST pattern influ-

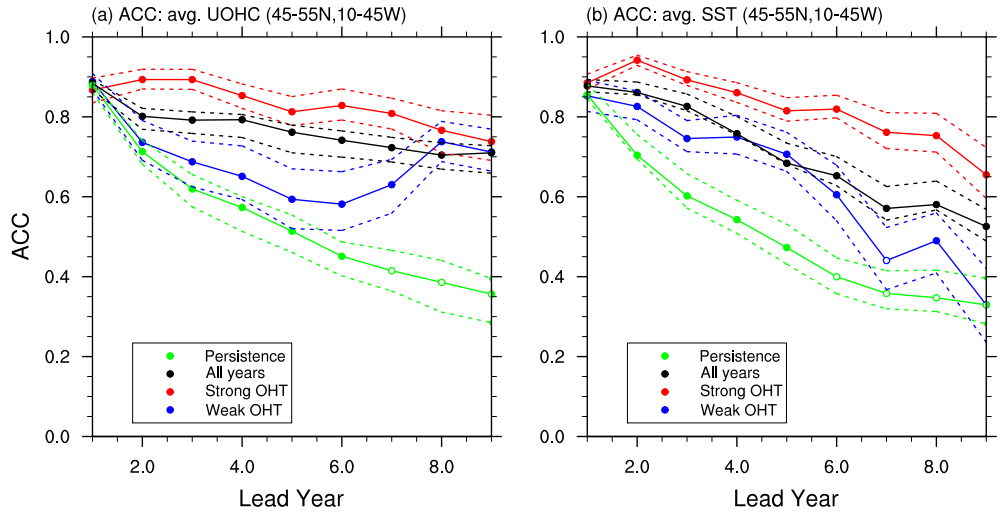


Figure 4.7: Anomaly Correlation Coefficients (ACCs) for average UOHC (a), and SSTs (b) over the box indicated in figs. 4.5,4.6 against lead time. I use ASSIM as a reference. The black line shows ACCs over the entire time series, green shows persistence forecast, the blue line shows predictability after weak, and the red line predictability after strong OHT_{50N} phases. The dashed lines indicate the interquartile ranges around the mean predictability for each color. Wherever these dashed lines do not overlap, two lines can be considered statistically significantly different. Solid points represent ACCs significant at the 99% level.

ences UOHC and SST predictability on the decadal time scale. Strong OHT_{50N} phases are particularly important in modulating ACCs for both UOHC and SSTs in the North Atlantic, as they lead to higher predictability of SSTs on the decadal time scale. I explore next whether a similar conclusion can be drawn for dynamical hindcasts of annual and seasonal mean SHFs and SATs.

SHF ACCs after strong, weak and neutral OHT_{50N} phases (fig. 4.8a-f) largely resemble the ACC pattern found for the entire time series for SHFs at all lags (cf. fig. 4.2). Asymmetries between strong, weak and neutral OHT_{50N} phases are similar to those found for SSTs. This underpins that ACCs of SHF hindcasts are more closely connected to SST-ACCs and SST variability than to SHF variability. As ACCs of SST hindcasts are generally high where SHFs are low, this shows that SHFs can be predicted skillfully in areas where they are low. This limits the impact on SAT predictability that can be expected from predictable surface heat fluxes.

Subsampling for different phases of subpolar OHT does not improve ACCs of decadal surface temperature hindcasts over Europe (fig. 4.9). In general, ACCs of decadal SAT hindcasts are closely related to those of SSTs: there are high ACCs over the ocean, particularly in the characteristic SST pattern region. Consequently, ACCs of SATs over the ocean are on the decadal time scale highest after strong OHT_{50N} phases, intermediately high after weak OHT_{50N} phases, and virtually absent after neutral OHT_{50N} phases (fig. 4.9c,f,i). This leads me to conclude that there is an influence of the OHT_{50N} phase on ACCs of SAT hindcasts, and that this influence is particularly strong on the

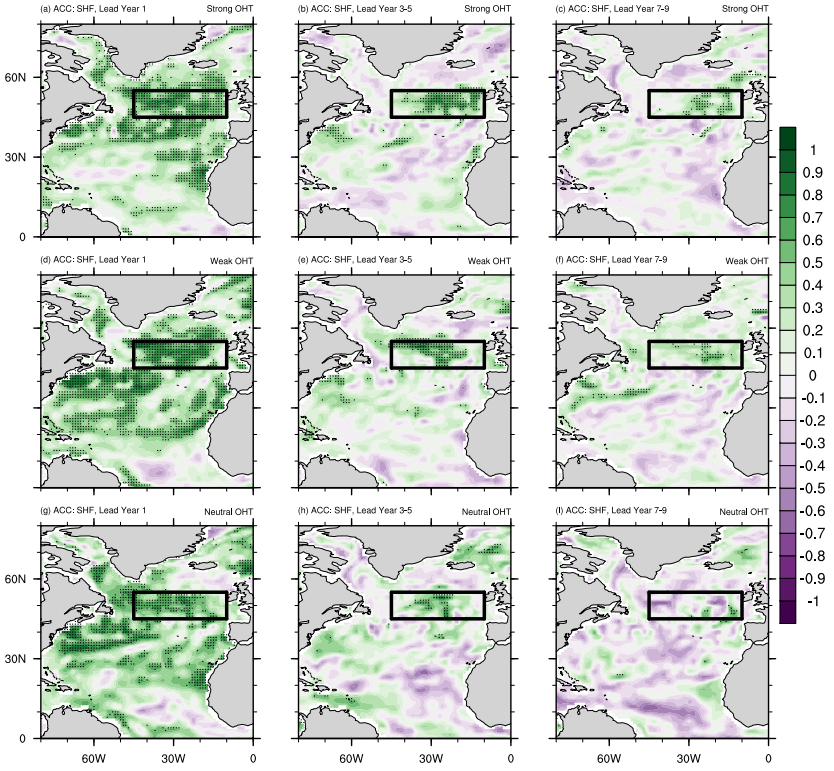


Figure 4.8: Anomaly Correlation Coefficients (ACCs) for SHFs from HC against ASSIM after strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show ACCs at lead years 1 (a,d,g), 3-5 (b,e,h) and 7-9 (c,f,i). Stippled areas indicate significant ACCs at the 99% level.

decadal time scale. However, HC and ASSIM do not show an influence of OHT_{50N} on SAT-ACCs for hindcasts over land. This might change when evaluating SAT hindcasts for different seasons instead of annual means.

4.3.2 Influence of Ocean Heat Transport on Hindcasts of Seasonal SATs

ACCs are high for seasonal mean SHFs at lead year 7-9 in areas where ACCs are high for SSTs at the same lead time after both strong and weak OHT_{50N} phases (fig. 4.10a-j; cf. fig. 4.6c,f). Specifically, SHF ACCs are high in the northeast Atlantic after strong, and high in the central North Atlantic after weak OHT_{50N} phases. ACCs for SHFs are significant in the North Atlantic in spring after strong OHT_{50N} (fig. 4.10c). In all other seasons, ACCs for SHFs are largely insignificant following both strong and weak OHT_{50N}. Following the assumption that SAT ACCs of decadal hindcasts are modulated by the ocean, this suggests that there could be limited ACCs for SAT hindcasts at seasonal means in HC, because the predictable SST signal is only in spring transported into the atmosphere.

ACCs are only marginally significant for seasonal mean SATs over land at lead year 7-9 after strong and weak OHT_{50N} phases (fig. 4.10k-t). The shape of these ACCs also bears

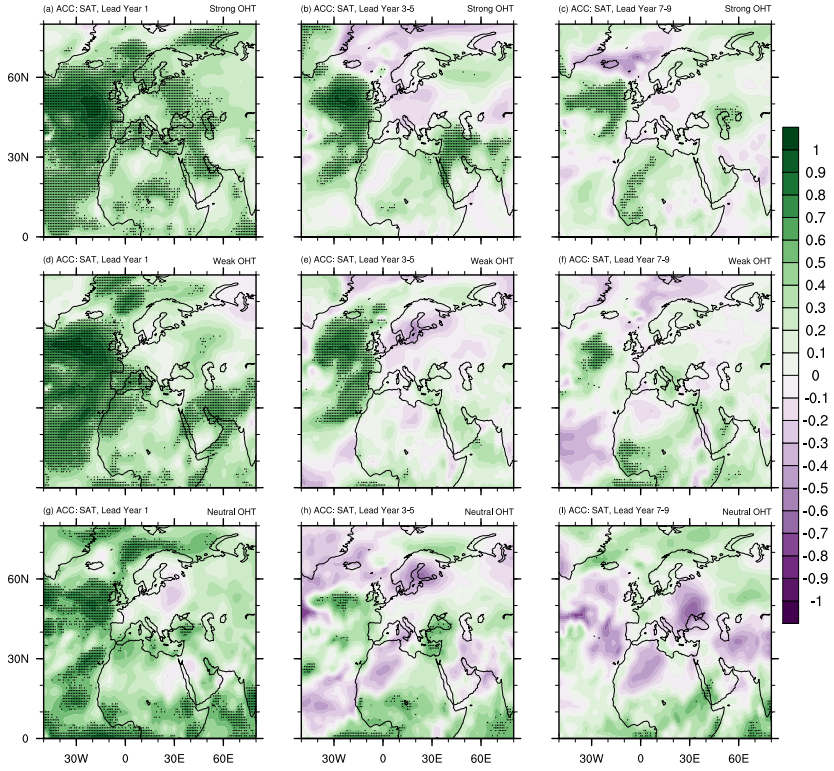


Figure 4.9: Anomaly Correlation Coefficients (ACCs) for SATs from HC against ASSIM after strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show ACCs at lead years 1 (a,d,g), 3-5 (b,e,h) and 7-9 (c,f,i). Stippled areas indicate significant ACCs at the 99% level.

little resemblence to the composite mean SAT patterns I have identified before (cf. fig 3.12). There is some seasonal variability in SAT ACCs, but this appears to not be linked to SHF predictability. Specifically, I find high ACCs for SATs over the ocean across all seasons, while ACCs are generally low over land. The shape of these ACCs generally follows that of predictable SSTs. This illustrates that, in HC, the influence of North Atlantic OHT variability on SAT hindcasts is limited to the areas that are immediately influenced by the ocean. However, the predictability of SHFs doew not play a role in the prediction of SATs. HC does apparently not properly represent teleconnection mechanisms that could transport SST anomalies to Europe at both annual and seasonal means.

4.4 Discussion

This chapter demonstrates that the mechanism suggested by Zhang and Zhang (2015) and discussed in chapter 3 modulates SST predictability on decadal time scales. The results presented here are generally in line with previous studies of decadal hindcasts of North Atlantic surface temperatures (e.g. Matei et al., 2012; Robson et al., 2013, 2014; Müller et al., 2014). Here, I discuss some of my findings, like the model specificity of

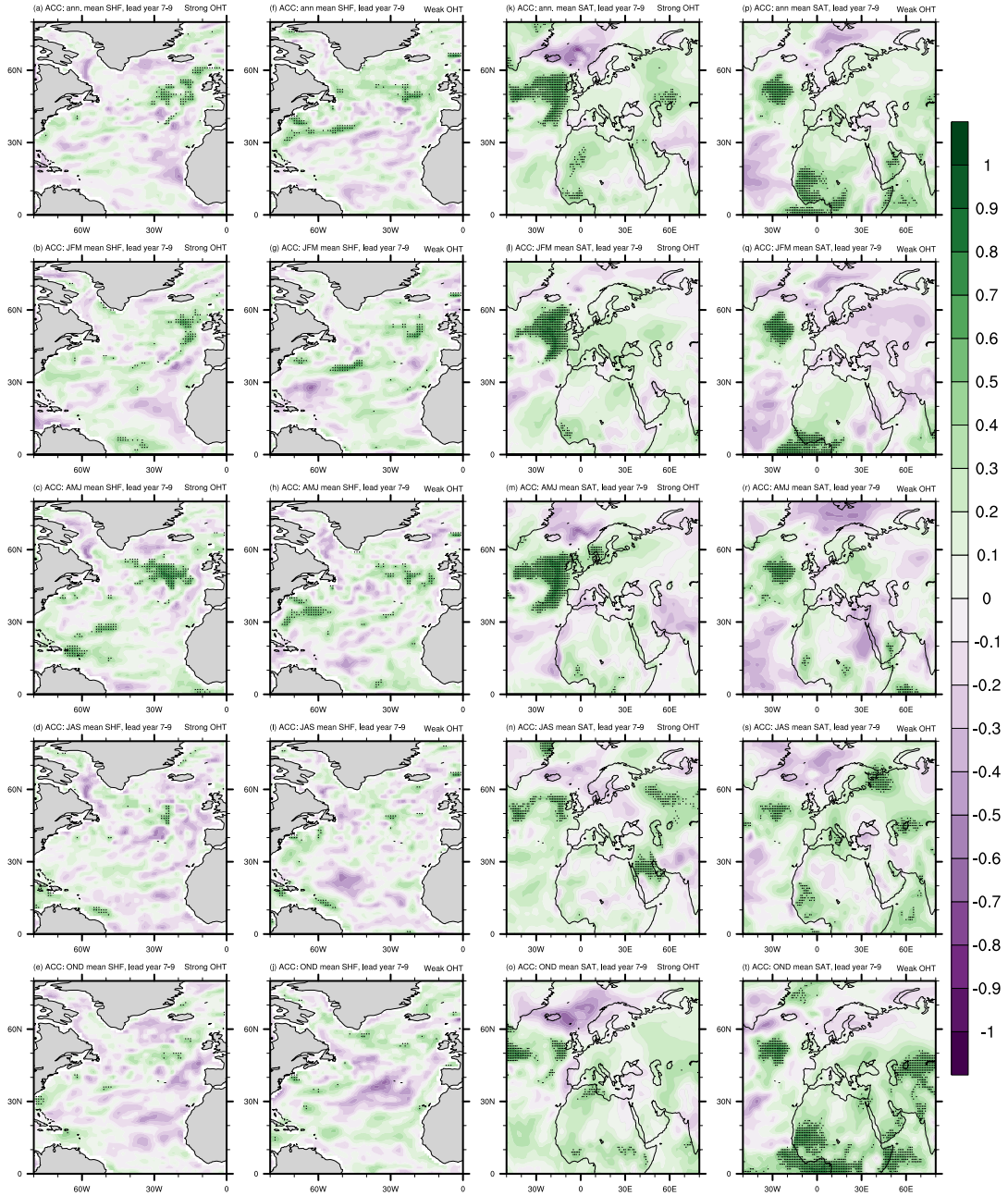


Figure 4.10: Anomaly Correlation Coefficients (ACCs) at lead year 7-9 from HC against ASSIM for SHFs after strong (a-e) and weak (f-j) OHT_{50N} phases, and SAT ACCs after strong (k-o) and weak (p-t) OHT_{50N} phases. I use annual (first row), JFM (second row), AMJ (third row), JAS (fourth row) and OND (fifth row) means. Stippling indicates significance at the 99% level.

my findings, ACCs (or lack thereof) following neutral OHT_{50N} phases, and the missing ACCs of dynamical surface air temperature hindcasts. I will also discuss the hindcast analyses on subsampled time series using OHT_{50N} variability, and give an indication of how the influence of OHT_{50N} dynamics on ACCs for hindcasts compares to that of UOHC persistence. I will then highlight some implications of this study that require further analysis.

In this hindcast study, I use one model with three ensemble members, and evaluate hindcasts against the assimilation experiment, which might well introduce some model-specificity into my findings. This is underlined by the fact that I find only little signal for SST hindcasts in an evaluation of ACCs against HadISST observations (not shown). My findings from the previous chapter, however, show that the North Atlantic AMV variability in HadISST is closely connected to the AMV in the model. This study therefore supports the notion brought forward frequently (e.g. Boer et al., 2016; Yeager and Robson, 2017), that the AMV is a driver of North Atlantic temperature predictability. Meanwhile, it is likely that a lack of ACCs of hindcasts evaluated against observations originates from the zonal displacement of the characteristic SST pattern due to a too zonal Gulf Stream in ASSIM and HC, and that both ASSIM and HC in fact get SST variability generally right. In that case, the influence of different phases of subpolar OHT on the skill of SST hindcasts would be expected further west than shown in this study. The finding of Müller et al. (2014), showing that ASSIM and HC produce reasonable climate variability, supports this statement.

I find that for both UOHC and SSTs, decadal predictability is connected to the physical mechanism leading to the characteristic SST pattern. The general absence of high ACCs in hindcasts started in neutral OHT_{50N} phases supports this claim. The Labrador Sea plays a peculiar role in this context, as it is covered by the characteristic SST pattern but I do not find SST predictability there. This can be attributed to the incapability of the non-initialized MPI-ESM-LR to represent temperature variability in the Labrador Sea (e.g. Brune et al., 2017) and stronger surface heat fluxes in that area, which overwrite the predictable temperature signal from the ocean (cf. fig. 3.9). Therefore, predictability is lost faster in that area than elsewhere in the North Atlantic region.

The influence of the characteristic SST pattern on ACCs of dynamical surface air temperature hindcasts on the decadal time scale is found to be limited in this dissertation. Arthun et al. (2017) showed skill for decadal winter SATs over Scandinavia that is connected to ocean heat transport using observations and a statistical model. I find no predictability in dynamical hindcasts of that area, except for coastal zones (cf. fig. 4.10). This lack of ACCs is likely an artifact of the limited resolution, non-stratosphere resolving nature (this would be necessary for the representation of teleconnections; e.g. Hoskins and Karoly, 1981), or few ensemble members of the model setup I use (for a recent discussion of dynamic multiyear hindcasts over land, see e.g. Sheen et al., 2017).

However, my results support Arthun et al. (2017) in that there appears to be a connection of OHT_{50N} to Scandinavian SATs (cf. fig. 3.12). This could be invoked to improve decadal SAT predictability over Scandinavia in dynamical models using a larger ensemble, more frequent hindcast initialization, or a higher resolution model that better represents atmospheric variability at all levels.

I find a strong influence of phases of subpolar OHT on the skill of decadal hindcasts for SSTs in the North Atlantic with a possibility for good ACCs of atmospheric surface temperature hindcasts over land in future studies. The findings presented so far do not, however, reconcile the debate in previous studies on the relative effects of UOHC persistence and dynamical ocean processes (i.e. overturning or OHT) on hindcast skill on these time scales yet (e.g. Müller et al., 2014; Yeager et al., 2012; Robson et al., 2013). In figure 4.11, an analogue to figure 4.7 is shown, but instead of OHT_{50N} strength, I use UOHC in the box indicated in figures 4.5 and 4.6 for subsampling to exemplify the influence of UOHC persistence on ACCs of UOHC and SSTs in hindcasts. On short time scales of up to 4 years, UOHC persistence influences the skill of temperature hindcasts about as strongly as OHT dynamics. This is coincident with the time scale on which I find high ACCs for UOHC and SSTs after all OHT_{50N} strengths including neutral phases, which underpins my hypothesis that those ACCs arise from UOHC persistence. After lead year 5, hindcasts subsampled for strong OHT_{50N} phases consistenly outperform those subsampled for any phase of UOHC persistence. This shows that, while both UOHC persistence and OHT dynamics are equally important for temperature hindcasts 3-5 years into the future and on shorter time scales, subpolar OHT dynamics are more important than UOHC persistence in shaping ACCs at lead years 7-9.

ASSIM produces more coherent strong OHT_{50N} phases than weak ones. As in the previous chapter, this asymmetric coherence of OHT_{50N} phases could impact the findings of this chapter by making strong OHT_{50N} phases appear more impactful than they are. Unfortunately, this problem cannot be approached in this study directly, but will have to be studied in more detail in a more conceptual model setup. I will take up this issue in the next chapter.

Brune et al. (2017) show that the estimated skill of decadal surface temperature hindcasts in the North Atlantic varies depending on the time that the skill is evaluated for. This finding implies that the use of any decadal hindcast skill estimate - produced for the past - is limited for the estimation of the credibility of an actual forecast. This study suggests ocean heat transport in the subpolar North Atlantic as a possible indicator of credibility of an actual forecast. As there is currently no long-term observational data set of OHT in the subpolar North Atlantic, this suggestion can currently not be verified against observations. The OSNAP project (Lozier et al., 2017) aims to construct such an observational data set. As it could be used as an indicator for the credibility of an actual forecast, the data set produced by the OSNAP campaign holds exciting oppor-

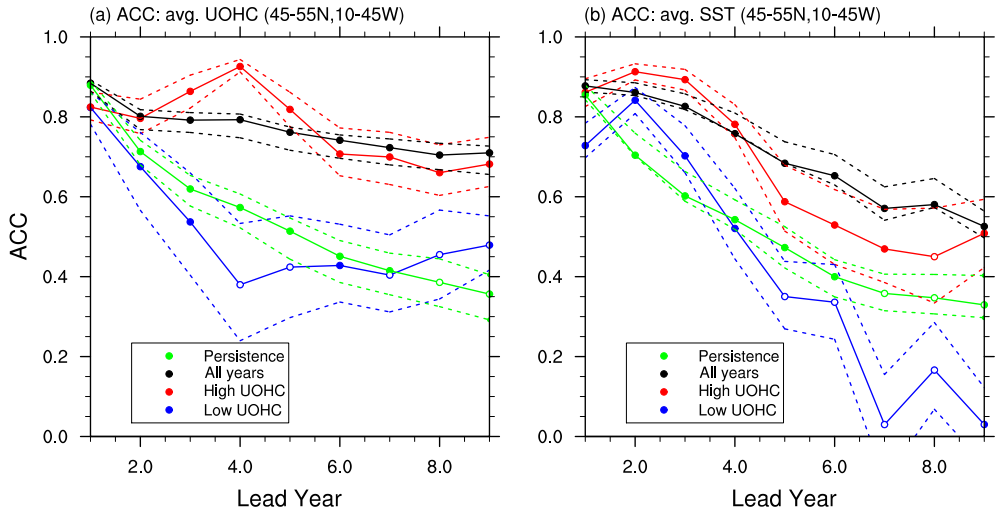


Figure 4.11: Anomaly Correlation Coefficients (ACCs) for average UOHC (a), and SSTs (b) over the box indicated in figs. 4.5,4.6 against lead time. I use ASSIM as a reference. The black line shows ACCs over the entire time series, green shows persistence forecast, the blue line shows predictability after weak, and the red line predictability after strong phases of upper ocean heat content in the black box. This figure illustrates the influence of UOHC persistence on the skill of UOHC and SST hindcasts. The dashed lines indicate the interquartile ranges around the mean predictability for each color. Solid points represent ACCs significant at the 99% level.

tunities for the suggestions I bring forward in this study.

4.5 Implications of this Research

Chapters 3 and 4 of this dissertation suggest that the skill of decadal surface temperature predictions in the North Atlantic depends on the strength of subpolar ocean heat transport in the Atlantic ocean at the start of the prediction. This finding could be used to judge the credibility that would be expected from an actual temperature forecast. These results provoke two thoughts in particular:

(i) Is the climate variability found in ASSIM representative of other modeled or observed climate variability? Assimlation model experiments combine a model's climate dynamics with observed dynamics. It is therefore inherently unclear from the analysis of ASSIM presented here - and in fact for the interpretation of all assimilation-based hindcast studies - whether its dynamics should be interpreted as model or observed dynamics. It is therefore important to place the climate variability produced in the assimilation model experiment in the broader context of climate variability produced by the climate model that is used, and of observed climate variability. Specifically, I discuss in chapter 3 that ASSIM shows peculiarly coherent strong OHT_{50N} phases. This coherence needs closer examination concerning its origins.

(ii) As ACCs of SSTs in the North Atlantic are higher after phases of strong ocean heat

transport than after weak phases, I conclude that hindcast skill changes over time. This indicates that no study of decadal climate hindcast skill performed for a fixed period in the past enables conclusions for the credibility of an actual forecast. Chapter 4 suggests that the credibility of a decadal forecast of North Atlantic surface temperatures could be judged using the phase of OHT_{50N} at the start of the forecast. Before concrete solutions to this problem can be evaluated, though, the non-stationarity of decadal surface temperature hindcast skill has to be understood better. Moreover, the influence of the phase of subpolar OHT at the beginning of an individual prediction on the quality of the prediction has to be evaluated for individual forecasts, not a time-average, to make the prediction quality found for the past transferable to the credibility of individual forecasts.

In the next chapter, I will take up thought (i) and place climate variability found in the ASSIM simulation in the context of the MPI-ESM-LR and observations, and discuss how these findings help to interpret the results presented in the first two chapters. Subsequently, addressing thought (ii), I will explore the time-dependence (or non-stationarity) of decadal SST hindcast skill in the 20^{th} century using HC and HadISST observations. This will advance our understanding of the role that physical processes play in the modulation of decadal climate predictability. Moreover, this chapter will fundamentally question the way that decadal climate hindcast studies are interpreted, and suggests improvements that should be made to enable a translation of hindcast skill estimates to forecasts.

5 | Understanding North Atlantic Climate Variability in the MPI-ESM-LR

5.1 Introduction

A fundamental problem with modelling studies is that it is from the model itself unclear how representative dynamical features found in models are with respect to observations. In assimilation experiments, this problem gets an additional dimension as the mixture of observed and model dynamics prevents a clear attribution of found dynamical processes to either model or observed dynamics. This makes it difficult to judge the representativeness of assimilation model dynamics with respect to observations. In the previous chapters, I find peculiar subpolar North Atlantic ocean heat transport dynamics in ASSIM: strong phases of OHT_{50N} appear much more coherently than weak ones. Here, I present an attempt to resolve the issue of placing dynamical variability found in assimilation model experiments between model variability and observed variability using the North Atlantic ocean heat transport variability found in ASSIM as a case study.
In chapter 3, I show that in ASSIM there is a large influence of subsurface ocean overturning variability, i.e. subpolar ocean heat transport, on SST variability in the North Atlantic. This physical mechanism also shows a pronounced influence on the skill of decadal surface temperature hindcasts in the North Atlantic region. In particular, I find very coherent strong OHT_{50N} phases with a strong influence on North Atlantic SSTs and incoherent weak OHT_{50N} phases with a weaker influence on North Atlantic SSTs. However, from the presented analyses alone, it cannot be understood how realistic this link between OHT, SSTs and surface temperature hindcast skill is, and where this finding has to be placed in terms of model and observed variability.
I will in this chapter attempt to resolve the question of robustness of the previously presented physical mechanism leading to the characteristic SST pattern in the face of similarly coherent strong and weak OHT_{50N} phases. This chapter further addresses the question whether particularly coherent OHT_{50N} phases are likely to be found in non-initialized model simulations or observations. I approach this problem by comparing ASSIM to uninitialized simulations with the same version of the MPI-ESM-LR, and to HadISST observations. The 1000-year-long piControl simulation of constant

pre-industrial greenhouse gas concentrations (piControl) can be used to approximate internal variability produced by the MPI-ESM-LR (e.g. Olonscheck and Notz, 2017). A historical simulation forced with the historical greenhouse gas concentrations from 1896 to 2005 (HIST) enables a direct comparison of ASSIM to a non-initialized version of the MPI-ESM-LR for the same time period. An assessment of two climate change scenarios (RCP4.5 and RCP8.5) in the MPI-ESM-LR for the period 2191-2300 points out possible changes of the physical mechanism leading to the characteristic SST pattern with global warming.

This chapter will predominantly focus on the asymmetric influence of strong and weak OHT_{50N} phases on North Atlantic SSTs I find in the previous chapters. I will then proceed and examine whether the model on its own is able to reproduce the coherent strong OHT_{50N} phases found in ASSIM, and assess the influence of the coherence of OHT_{50N} phases on SST variability. In the process, I will also look at HadISST observations of sea surface temperatures to estimate the representativeness of my findings for observed climate variability. This will help to place the findings from the last two chapters in the broader climatic context, and to judge their representativeness.

5.2 Ocean Overturning Dynamics in the Non-Initialized MPI-ESM-LR

In the piControl model simulation, annual mean AMOC maximum and total OHT_{50N} are correlated 63% for the full 1000 year time series (fig. 5.1a). While this is much lower than the correlation in ASSIM, it is still significant at the 99% confidence level. Moreover, AMOC maximum and total OHT variability in the North Atlantic appear very similar in their meridional variabilities (fig. 5.1b,c). Therefore, the piControl simulation produces ocean overturning variability that supports my finding that OHT and AMOC dynamics are connected in the North Atlantic. However, the low correlation between AMOC and OHT variability at 50°N in piControl might indicate that the very high correlation I find in ASSIM is somewhat peculiar. Unlike in ASSIM, strong and weak OHT_{50N} phases selected by the criterion of half a standard deviation above or below the mean state, are in piControl both similarly long and appear similarly coherent (dots in fig. 5.1a).

The piControl simulation alone is not sufficient to place findings from ASSIM in the context of the MPI-ESM-LR variability as there is no greenhouse gas forcing in piControl while there is such forcing in ASSIM. I therefore compare OHT and OHT-SST dynamics from ASSIM not only to piControl, but also to a historical simulation without data assimilation. To further understand the influence of greenhouse gas forcing, I also compare ocean overturning dynamics found in ASSIM to a simulation with the MPI-

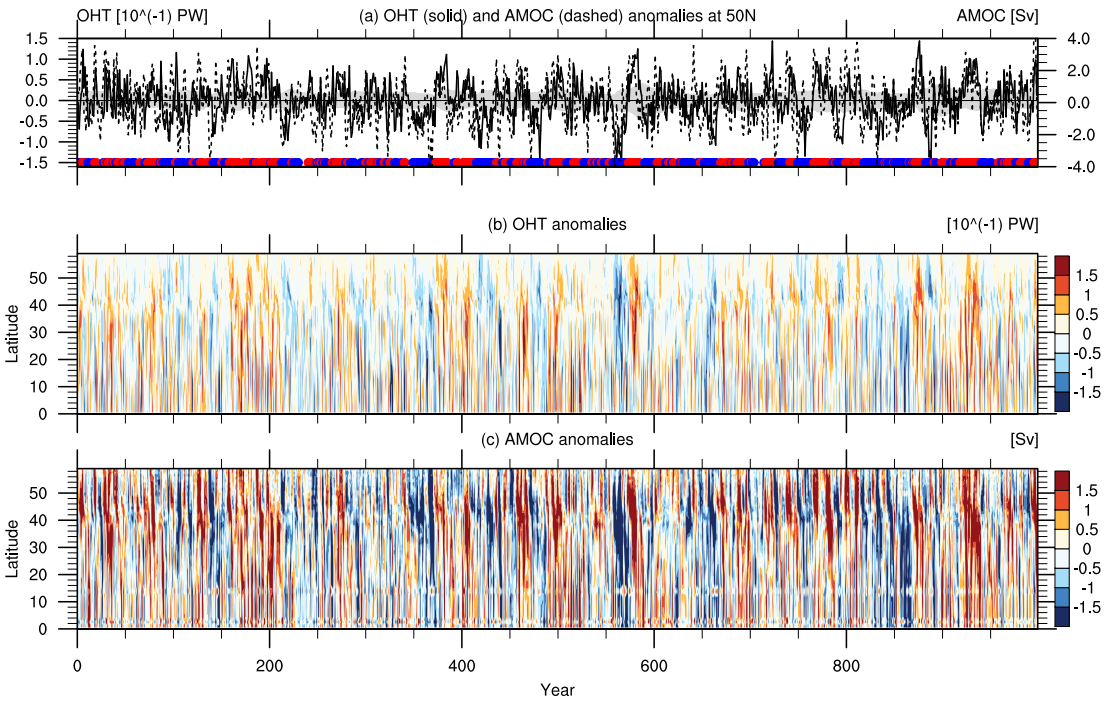


Figure 5.1: (a) Detrended anomalies of AMOC maximum at 50°N (dashed line, [Sv]) and total OHT_{50N} (solid line, [10^{-1} PW]) in the piControl-simulation. The grey area denotes a half standard deviation above and below the mean of the previous 30 years. Strong and weak OHT_{50N} phases, i.e. years where the solid line lies outside the grey area, are marked with red and blue dots at the bottom, respectively. Hovmöller Diagrams of OHT anomalies (b) and AMOC maximum anomalies (c) illustrates the development of strong and weak anomalies of OHT and AMOC in space (y-axis, [°latitude]) and time (x-axis [yrs]). OHT and AMOC time series are detrended at each latitude.

ESM-LR forced by the RCP4.5 scenario. While I use the full 1000 years in piControl to include all modes of variability that the model produces on its own, I use the 110 year period between 1896 and 2005 from HIST to make the results as comparable as possible to the 1901 to 2010 period used in ASSIM. In the RCP4.5 experiment, I use the last 110 years of the simulation, 2191 to 2300, to capture the maximum greenhouse gas concentration influence on ocean dynamics. To better understand possible AMOC-OHT variability the MPI-ESM-LR can produce, I will reproduce figures from the beginning of this dissertation using piControl, HIST, and RCP4.5 model runs.

There is a tendency to a southward propagation of OHT anomalies that originate at 50°N in all model simulations (fig. 5.2). However, in none of the simulations is this southward propagation even closely as pronounced as in ASSIM. The HIST simulation shows subdecadal OHT dynamics relatively similar to those found in ASSIM, while the southward propagation of OHT phases in the North Atlantic is very weak in piControl and virtually non-existent in RCP4.5. Following the hypothesis of Zhang and Zhang (2015) that OHT dynamics influence ocean heat convergence and that the southward propagation plays a crucial role in this, subpolar OHT phases should be followed by

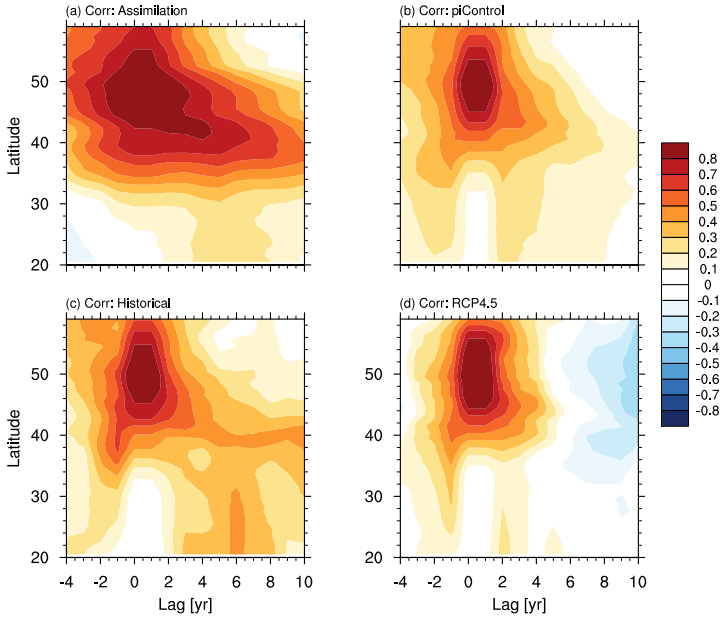


Figure 5.2: Lead-lag correlations of OHT at different latitudes with OHT_{50N} found in (a) ASSIM, (b) piControl, (c) HIST, and (d) RCP4.5. Time series are detrended at every latitude individually. Positive lags indicate that OHT leads and vice versa.

a weak ocean heat convergence signal in the North Atlantic in piControl, HIST and RCP4.5.

Indeed, ocean heat convergence in the North Atlantic shows less connection to OHT_{50N} variability in piControl, HIST and RCP4.5 than in ASSIM (fig. 5.3). However, all simulations show a (predominantly weak) dipole ocean heat convergence anomaly signal between a positive correlation to OHT_{50N} North of roughly 40°N when OHT_{50N} leads by 2-8 years, and an instant negative correlation South of 40°N. This effect is most pronounced in HIST and piControl, and less strong in RCP4.5. These findings support the role of southward propagation of OHT phases in the formation of ocean heat convergence in the subpolar North Atlantic ocean several years later. Do these OHT_{50N} phases influence North Atlantic SSTs, as shown in ASSIM, in piControl as well?

Both UOHC (fig. 5.4a-c) and SSTs (fig. 5.4d-f) in the subpolar North Atlantic appear to be modulated by OHT_{50N} up to 9 years earlier in piControl. The shape of these correlations resembles that found in the same analysis in ASSIM (cf. fig. 3.3). This indicates that a similar mechanism gives rise to surface temperature anomalies of this shape in piControl and ASSIM, which is also in line with the ocean heat convergence analysis presented above. Particularly at long lags, i.e. 7-9 years after OHT_{50N} phases, the UOHC and SST correlations to OHT_{50N} are much lower in piControl than in ASSIM. This relatively shortly-lived influence of OHT_{50N} on SSTs, too, is in line with the diagnose of southward propagating OHT anomalies and associated ocean heat convergence anomalies that are shorter live in piControl than in ASSIM. It therefore appears that

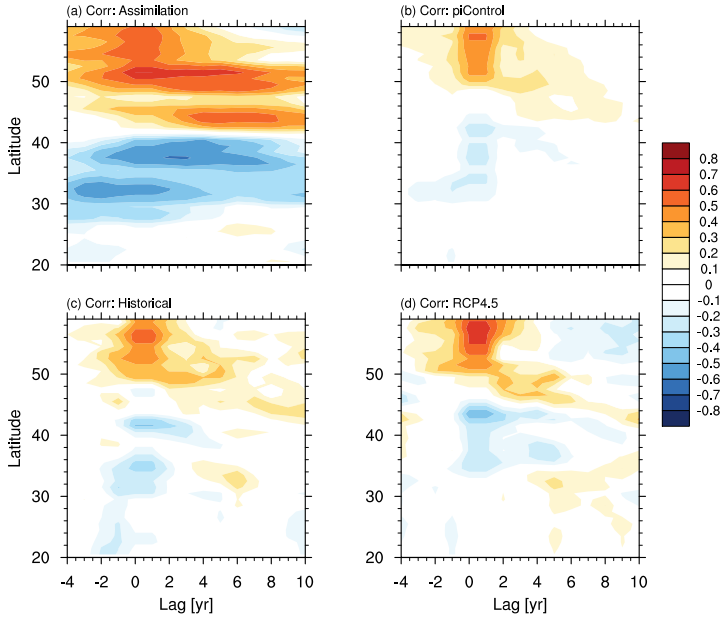


Figure 5.3: Lead-lag correlations of ocean heat convergence at different latitudes with OHT_{50N} found in (a) ASSIM, (b) piControl, (c) HIST, and (d) RCP4.5. Time series are detrended at every latitude individually. Positive lags indicate that OHT leads and vice versa.

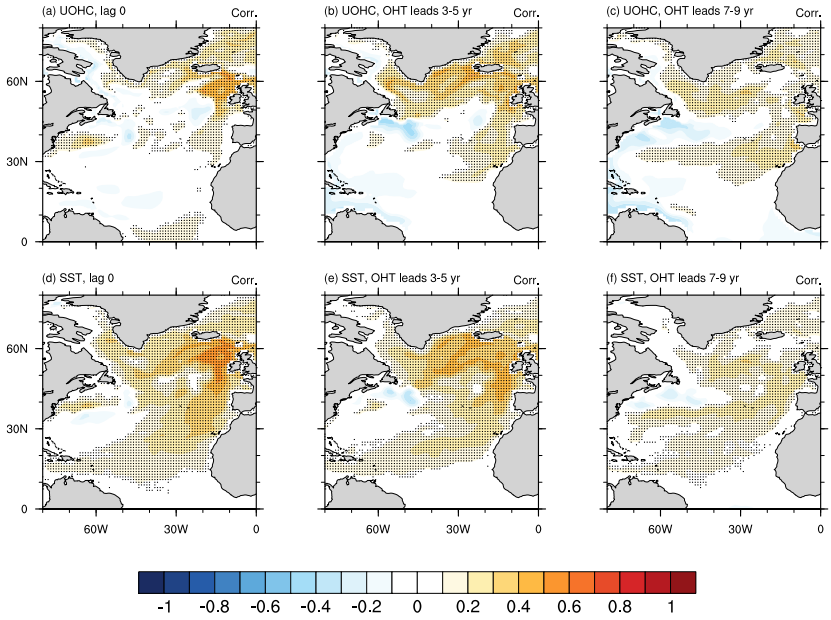


Figure 5.4: Point-by-point correlation of OHT_{50N} with upper ocean heat content of the upper 700m in the North Atlantic in piControl at lag 0 (a) and when OHT_{50N} leads by 3-5 years (b) and 7-9 years (c), and with SSTs at lag 0 (d) and when OHT_{50N} leads by 3-5 years (e) and 7-9 years (f). Stippling indicates significant correlations at the 99% level.

the mechanism leading to the characteristic SST pattern in ASSIM works in piControl as well. However, this finding does not address the 'problem' of differently coherent OHT_{50N} phases in ASSIM, as described above. To better understand OHT_{50N} and

SST dynamics in piControl, HIST and RCP4.5, I will therefore now examine strong and weak phases of OHT_{50N} separately.

5.3 Subsampled Overturning States in the MPI-ESM-LR

In each model simulation, I subsample ocean states in which OHT_{50N} is above or below the mean of the preceding 30 years. I then construct composite mean OHT anomalies across latitudes after strong and weak OHT_{50N} phases to examine strong and weak OHT_{50N} phases separately. This enables me to place the findings from the variability analysis in ASSIM in the context of overall OHT variability in the MPI-ESM-LR.
In piControl, the southward propagation of OHT phases in the subpolar North Atlantic is comparable to ASSIM (fig. 5.5a-c). However, the OHT dynamics show a more shorter-lived signal in piControl than in ASSIM, and after 5 years, there is no difference in the OHT anomaly North of 40°N between strong and weak OHT_{50N} phases in piControl. The HIST simulation shows a much stronger OHT propagation signal after weak OHT_{50N} phases than after strong OHT_{50N} phases, while the overall OHT propagation signal is clearly southward with time in both cases (fig. 5.5d,e). The strong anomaly after weak OHT_{50N} phases could indicate a stronger influence of OHT_{50N} variability on surface temperatures after weak OHT_{50N} phases than after strong OHT_{50N} phases in HIST. In the RCP4.5 simulation, there is a very short-lived southward OHT_{50N} propagation signal after strong OHT_{50N} phases for up to 4 years. However, I find almost no southward propagation of OHT_{50N} phases after weak OHT_{50N} phases (fig. 5.5g,h). This indicates a weak influence of weak OHT_{50N} phases on SSTs, and a strong, but short-lived, influence of strong OHT_{50N} phases on SSTs for up to 4 years in RCP4.5. After neutral OHT_{50N} phases, composite mean OHT anomalies are negligible in all model runs (fig. 5.5c,f,i). The weak influence of weak OHT_{50N} phases and stronger influence of strong OHT_{50N} phases on ocean heat convergence can be assessed by comparing composite mean cumulative ocean heat convergence signals in the three MPI-ESM-LR simulations.
Cumulative ocean heat convergence signals after strong and weak OHT_{50N} phases in the subpolar North Atlantic are largely symmetric in piControl (fig. 5.6a,b). There appears, however, to be a slight asymmetry in the ocean heat convergence signals North of 50°N with a stronger convergence signal after strong OHT_{50N} phases, which is consistent with my findings from the OHT propagation signals. Similarly, the ocean heat convergence signal South of 40°N is stronger after weak than after strong OHT_{50N} phases, indicating a seesaw-like mechanism connecting OHT anomalies to ocean heat convergence similar to that found for seasonal AMOC dynamics in the subtropical Atlantic in Duchez et al. (2016). By contrast, neutral phases show no distinctive ocean heat convergence signal (fig. 5.6c).

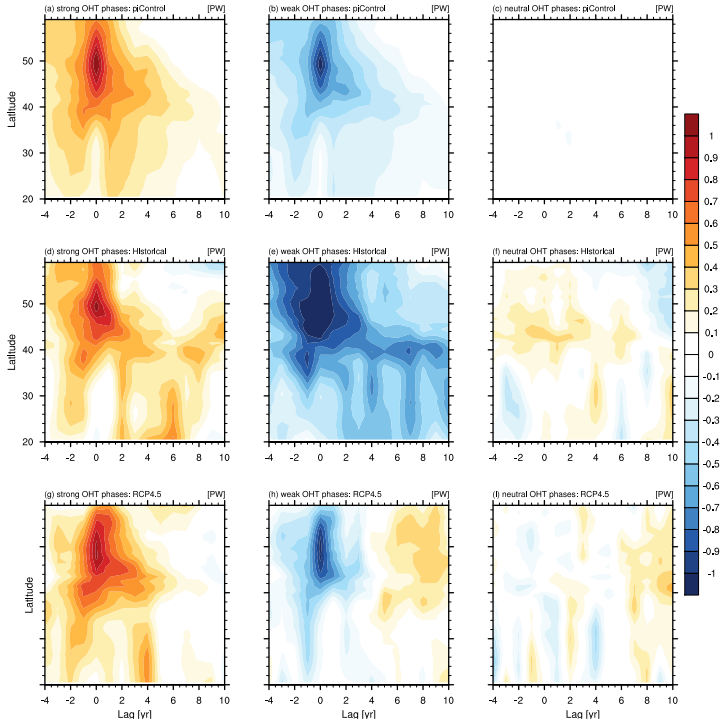


Figure 5.5: Composite mean OHT anomalies [10^{-1}PW] before, during, and after strong (left column), weak (middle column), and neutral (right column) OHT_{50N} anomalies shown in latitude against lag. (a-c) show piControl, (d-f) show HIST, and (g-i) show RCP4.5. Time series are detrended at every latitude individually. Positive lags indicate that OHT leads and vice versa.

In HIST, there is a pronounced asymmetry in cumulative ocean heat convergence signal in the subpolar North Atlantic between strong and weak OHT_{50N} phases with a much stronger heat convergence signal after strong OHT_{50N} phases than after weak ones (fig. 5.6d,e). This is counter-intuitive to the findings from the OHT propagation analysis that suggested a strong influence from weak OHT_{50N} phases. Additionally, neutral OHT_{50N} phases that are not connected to distinct composite mean OHT anomalies in HIST, are connected to a quite strong ocean heat convergence anomaly (fig. 5.6f). This indicates that other mechanisms than the mechanism leading to the characteristic SST pattern are at work in the North Atlantic in HIST. On the other hand, this result is in line with the findings from piControl and ASSIM, that both show a stronger influence of strong than of weak OHT_{50N} phases on ocean heat convergence. Therefore, an alternative interpretation of this finding is that the strong influence of strong OHT_{50N} phases on subpolar ocean heat convergence in the North Atlantic is less sensitive to the strength of the southward propagating OHT anomaly than it is to the sign of the anomaly. Consistently with this hypothesis, strong OHT_{50N} phases influence subpolar North Atlantic ocean heat convergence more strongly than weak OHT_{50N} phases in RCP4.5 (fig. 5.6g,h).

Composite mean SSTs in piControl after strong and weak OHT_{50N} phases show some

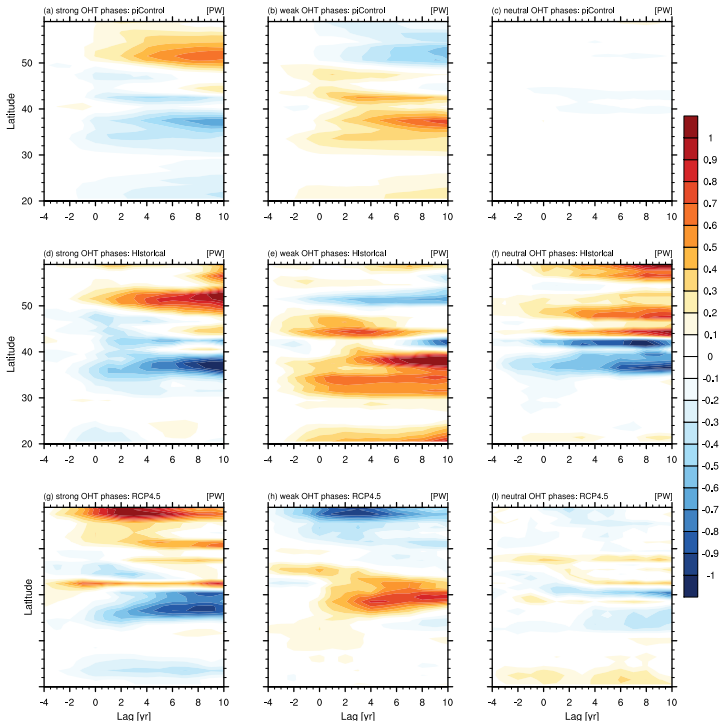


Figure 5.6: Composite mean cumulative heat convergence anomalies [PW] before, during, and after strong (left column), weak (middle column), and neutral (right column) OHT_{50N} anomalies shown in latitude against lag. Heat convergence anomalies are calculated cumulatively with respect to strong and weak OHT_{50N} phases. (a-c) show piControl, (d-f) show HIST, and (g-i) show RCP4.5. Time series are detrended at every latitude individually. Positive lags indicate that OHT leads and vice versa.

asymmetry (fig. 5.7). While composites are similarly significant in their difference from the mean variability across all lags and after both strong and weak OHT_{50N} phases, composite mean SST anomalies are stronger 3-5 years after strong OHT_{50N} phases than after weak OHT_{50N} phases. At longer lags, i.e. 7-9 years, there are much weaker composite mean SST anomalies in piControl than in ASSIM (cf. fig. 3.9), which is related to the shorter-lived OHT and ocean heat convergence anomalies in piControl. After neutral OHT_{50N} phases, there is no distinct composite mean SST pattern at all lags (fig. 5.7g-i). Remarkably, unlike in ASSIM, composite mean ocean-atmosphere surface heat fluxes show almost no asymmetry between strong and weak OHT_{50N} phases in piControl. However, as heat fluxes are assimilated in ASSIM, they are directly related to climate variability in the twentieth century reanalysis (Compo et al., 2011) - it is therefore likely that the asymmetry of SHFs found in ASSIM is more realistic than symmteric SHFs in piControl with respect to observations.

In the historical and the RCP4.5 simulation, a conclusive SST pattern is difficult to find at any decadal lag after both strong and weak OHT_{50N} phases (figs. 5.8, 5.9). While a coherent SST pattern appears in both simulations at lag 0, SSTs in the subpolar North Atlantic show largely insignificant anomalies at longer lags. An exception is the northeast Atlantic 7-9 years after strong OHT_{50N} phases in HIST, where I find a

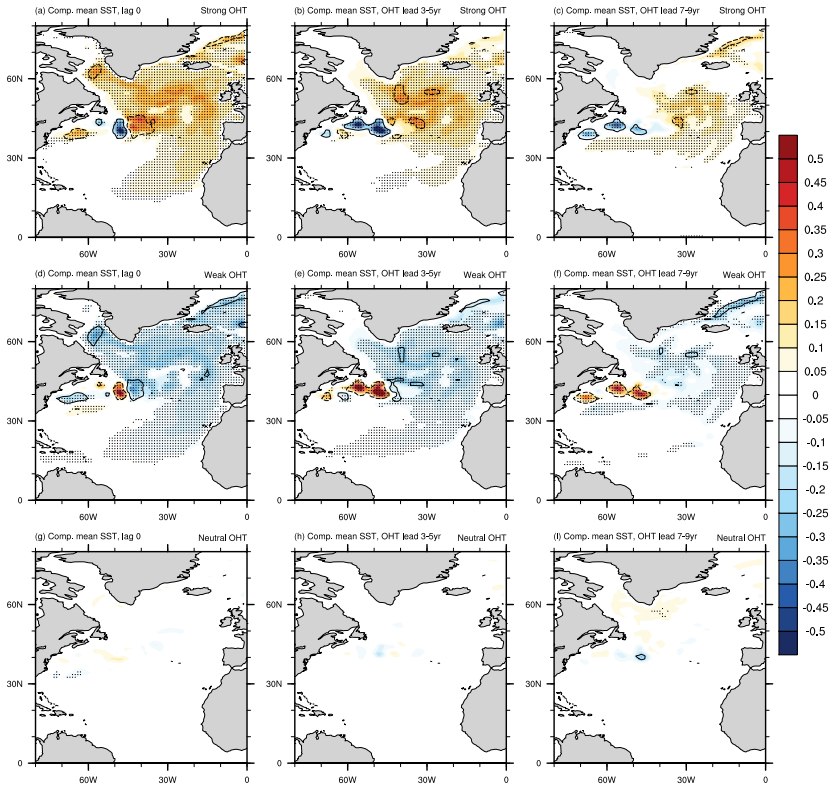


Figure 5.7: Composite mean SST anomalies [K] in piControl, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSH 3-5 (b,e,h) and 7-9 (c,f,i) after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level. Contours show significant (at the 99% level) net ocean-atmosphere surface heat fluxes into the ocean (solid contours) and out of the ocean (dashed contours).

scattered area of barely significant SST anomalies without a strong signal after weak OHT_{50N} phases (fig. 5.8b,e). As there is no consistent signal in that area throughout lead times shorter than 7-9 years, the physical consistency of this SST anomaly is debatable. In both model simulations, no clear SST signal can be identified following neutral OHT_{50N} phases.

SHFs show largely inconclusive patterns in both simulations at long time lags, although strong SHFs appear to be linked to places with relatively strong SST anomalies. These findings show that ocean heat convergence anomalies that persist for only a few years indeed have a weak influence on SST patterns, and if there is an influence on SSTs, I find it on short lags of only a few years. This analysis therefore shows that the mechanism leading to the characteristic SST pattern in ASSIM is robust in other simulations with the MPI-ESM-LR, however, the uninitialized model seems to generally generate shorter OHT_{50N} phases than ASSIM that are characterized by an influence of OHT_{50N} phases on SST anomalies on time scales no longer than a few years. Besides ASSIM, only the piControl simulation is characterized by an influence of subpolar OHT variability on SST anomalies on the decadal time scale.

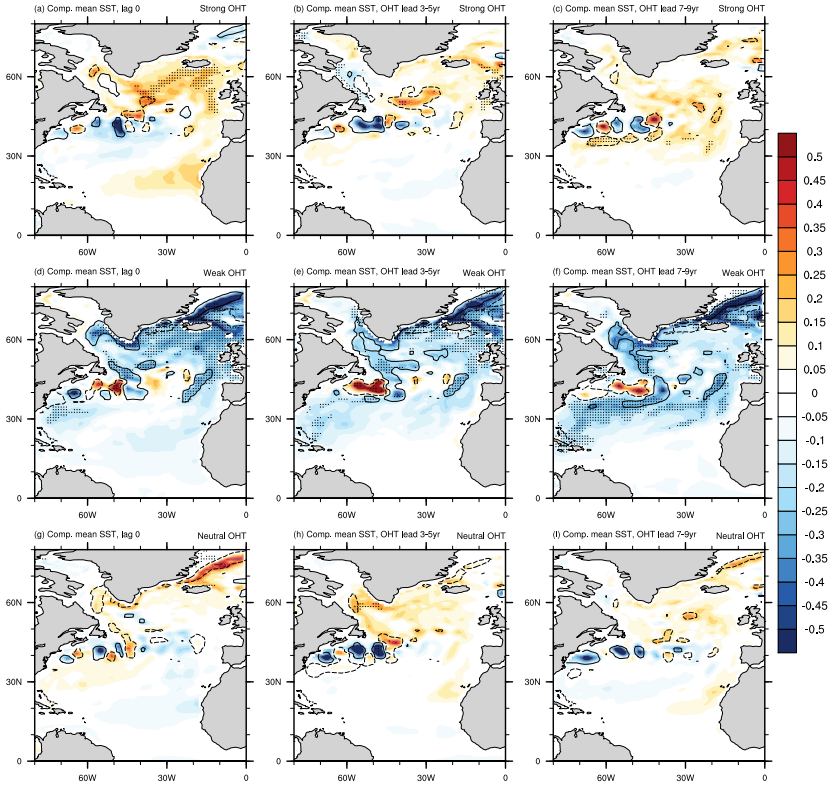


Figure 5.8: Composite mean SST anomalies [K] in HIST, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSH 3-5 (b,e,h) and 7-9 (c,f,i) after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level. Contours show significant (at the 99% level) net ocean-atmosphere surface heat fluxes into the ocean (solid contours) and out of the ocean (dashed contours).

5.4 Possible Reasons for Differently Coherent OHT Phases

Attempts to reproduce the characteristic SST pattern in other simulations with the MPI-ESM-LR were relatively successful, which inspires confidence that the non-linear influence of strong and weak OHT_{50N} phases on North Atlantic SSTs found in ASSIM is robust. The characteristic SST pattern is particularly strong in the piControl simulation, which indicates that the model alone is able to produce this SST pattern without external forcing. The characteristic SST pattern is therefore unlikely to originate exclusively from global warming. The decadal OHT_{50N} variability I find in ASSIM in the subpolar North Atlantic appears, however, somewhat peculiar concerning the coherence of OHT_{50N} phases.

An important finding from this analysis is that the time scale on which OHT_{50N} phases influence SST variability in the North Atlantic depends on the coherence of OHT_{50N} phases. More coherent OHT_{50N} phases lead to more long-lived SST anomalies in the

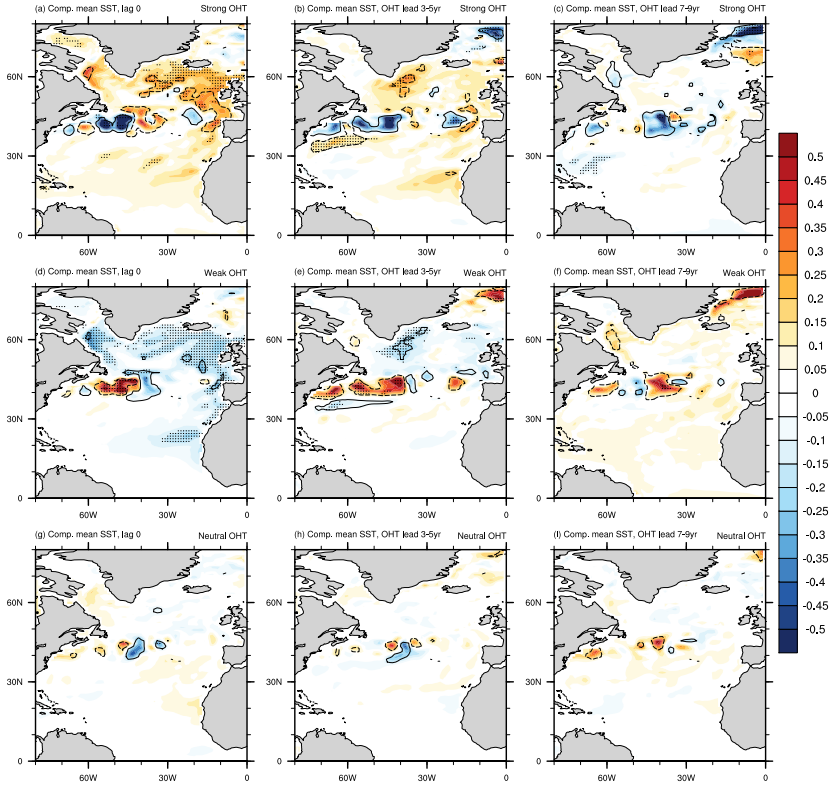


Figure 5.9: Composite mean SST anomalies [K] in RCP4.5, related to strong (a-c), weak (d-f), and neutral (g-i) OHT_{50N} phases. I show composites at lag 0 (a,d,g) and composite mean SSH 3-5 (b,e,h) and 7-9 (c,f,i) after strong and weak OHT_{50N} phases. Stippling indicates significance at the 99% level. Contours show significant (at the 99% level) net ocean-atmosphere surface heat fluxes into the ocean (solid contours) and out of the ocean (dashed contours).

North Atlantic. However, none of the analyses performed above give an indication if and by how much ASSIM over- or underrepresents the coherence of strong and weak OHT_{50N} phases in the subpolar North Atlantic.

The previous findings show that there is a difference in the length of subpolar OHT phases between ASSIM and piControl: strong OHT_{50N} phases are coherent in both ASSIM and piControl, while weak OHT_{50N} phases are similarly coherent in piControl, but relatively incoherent in ASSIM. There are three possible reasons for these different coherences, and understanding which one is responsible for these differences will help to place the findings from the analysis of ASSIM in a wider context. The different coherences of strong and weak OHT_{50N} phases in ASSIM could: reflect an artifact of 'subsampling' in the sense that the piControl simulation encompasses 1000 years, whereas ASSIM is 110 years long and more coherent strong and more incoherent weak OHT_{50N} phases might occur naturally; be related to the data assimilation that is applied in ASSIM which might perturb the model, leading to less coherent weak or more coherent strong OHT_{50N} phases; or be an effect of global warming, as there is no change of mean greenhouse gas concentration in piControl whereas greenhouse gases

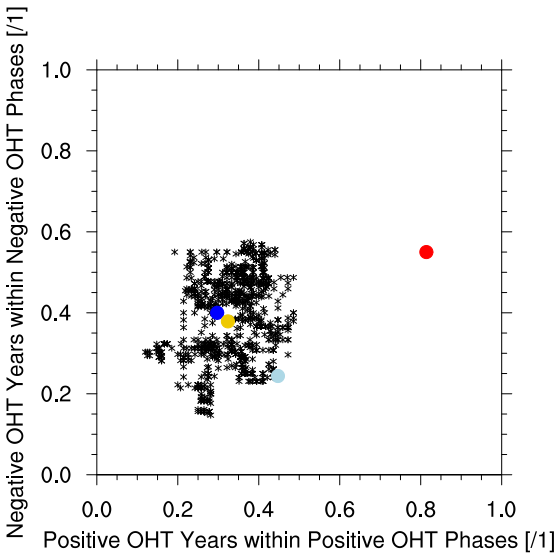


Figure 5.10: The ratio of years of strong OHT_{50N} that are framed by years of strong OHT_{50N} with respect to total years of strong OHT_{50N} on the x-axis, shown against the ratio of years of weak OHT_{50N} that are framed by years of weak OHT_{50N} with respect to total years of weak OHT_{50N} on the y-axis. The red dot shows this ratio in ASSIM (1901-2010), the yellow one in piControl (1000 years), the dark blue one in HIST (1896-2005), and the light blue one in RCP4.5 (2191-2300). Black crosses denote ratio combinations calculated from all 110 year long subperiods of piControl.

increase in ASSIM. I will in the following use the piControl, HIST, RCP4.5 simulations alongside an RCP8.5 simulation and HadISST observations (Rayner et al., 2003) and assess which of these processes produces which coherence of subpolar OHT phases.

Strong OHT_{50N} phases in ASSIM are much more coherent than strong OHT_{50N} phases produced by the MPI-ESM-LR as internal variability (fig. 5.10). I compare the coherences of strong and weak OHT_{50N} phases in ASSIM, HIST, and RCP4.5 to all possible 110-year-long segments of the 1000-year-long piControl simulation to estimate the coherences for subpolar OHT variability generated as internal variability by the MPI-ESM-LR. Only the coherences of strong and weak OHT_{50N} phases in HIST and RCP4.5, and the coherence of weak OHT_{50N} phases in ASSIM lie within the range of variability that is produced by the MPI-ESM-LR. This indicates that the assimilation applied to the MPI-ESM-LR to produce ASSIM generates overly coherent strong OHT phases in the subpolar North Atlantic. An implication of this finding is that ASSIM might over-estimate the influence of strong OHT_{50N} phases on northeast Atlantic SST anomalies.

To get a better indication of how to place the coherence of not only OHT_{50N} phases, but also SST variability in ASSIM in the wider context, I also show coherences of strong and weak phases of the Atlantic Multidecadal Variability (*AMV*, fig. 5.11). The AMV is an important mode of temperature and climate variability in the North Atlantic that may be linked to AMOC and OHT variability (e.g. Delworth et al., 2017). I here define the AMV as average SSTs in the North Atlantic between the equator and 60°N; the

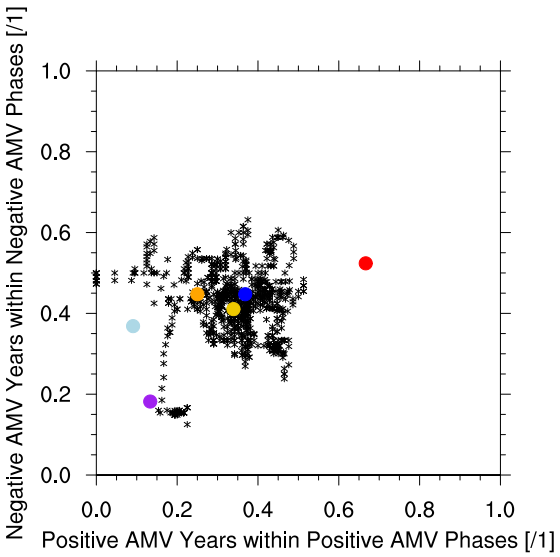


Figure 5.11: The ratio of years of strong AMV that are framed by years of strong AMV with respect to total years of strong AMV on the x-axis, shown against the ratio of years of weak AMV that are framed by years of weak OHT_{50N} at AMV with respect to total years of AMV on the y-axis. The red dot shows this ratio in ASSIM (1901-2010), the yellow one in piControl (1000 years), the dark blue one in HIST (1896-2005), the light blue one in RCP4.5 (2191-2300), the purple one in RCP8.5 (2191-2300), and the orange one in HadISST observations (1901-2010). Black crosses denote ratio combinations calculated from all 110 year long subperiods of piControl.

AMV index in ASSIM is highly correlated to the characteristic SST pattern index at no lag (0.96, cf. fig. 3.13), and therefore also highly correlated to OHT_{50N}. It is worth noting that, unlike previous studies, I do not smooth SSTs with a low-pass filter in the calculation of the AMV index for this chapter.

Coherences for the AMV index from the same model experiments as before, as well as the RCP8.5 scenario and HadISST observations for 1901-2010, show that positive AMV phases produced by ASSIM are more coherent than others (fig. 5.11). Similar to my previous findings concerning subpolar OHT variability, both positive and negative AMV indices in HIST, RCP4.5 and RCP8.5, and in HadISST, and negative AMV indices in ASSIM show coherence that lies within the internal variability produced by piControl. These findings appear to confirm my previous conclusion concerning the reliability of the variability produced by ASSIM.

As mentioned before, I do not smooth SSTs in the calculation of the AMV index for this analysis. However, for the correlation and composite mean analyses performed throughout this dissertation, I use three-year average SSTs to filter out year-to-year variability. I show coherences for the AMV index smoothed with a three-year running mean in figure 5.12.

When a 3-year running mean is applied, the coherences of positive and negative AMV phases of all model simulations (including ASSIM) and HadISST lie within the spread of coherences generated by piControl (fig 5.12). This effect is robust for OHT_{50N} phases

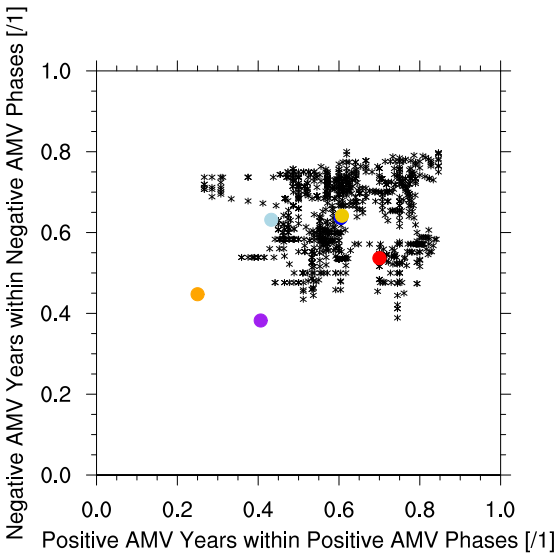


Figure 5.12: The ratio of years of strong AMV that are framed by years of strong AMV with respect to total years of strong AMV on the x-axis, shown against the ratio of years of weak AMV that are framed by years of weak AMV with respect to total years of weak AMV on the y-axis. Time series were smoothed with a three-year running mean prior to the definition of strong and weak phases. The red dot shows this ratio in ASSIM (1901-2010), the yellow one in piControl (1000 years), the dark blue one in HIST (1896-2005), the light blue one in RCP4.5 (2191-2300), the purple one in RCP8.5 (2191-2300), and the orange one in HadISST observations (1901-2010). Black crosses denote ratio combinations calculated from all 110 year long subperiods of piControl.

(not shown). These findings indicate that the coherence of phases of North Atlantic climate variability in ASSIM is accurate in the realm of model variability of the MPI-ESM-LR on a frequency lower than the year-to-year variability. Accordingly, ASSIM can be used to study representative SST dynamics when at least a 3-year-smoothing is applied. However, there is still a pronounced discrepancy between the coherence ratios in the model simulations and in HadISST, which indicates that the year-to-year AMV variability in the MPI-ESM-LR does not accurately reproduce observed AMV variability. This discrepancy, however, diminishes when a low-pass filter of more than 10 years is applied (not shown).

The coherences of OHT_{50N} and AMV phases are consistently lower for RCP4.5 and RCP8.5 than for the other model simulations and observations. Moreover, RCP8.5 shows lower coherence than RCP4.5. This indicates that with progressing greenhouse gas forcing, i.e. progressing global warming, the coherence of OHT_{50N} and AMV phases can be expected to decrease. This effect is particularly strong at the end of the 22^{nd} century.

5.5 Discussion

Here I present some evidence that the physical mechanism leading to the characteristic SST pattern in the North Atlantic is robust within the MPI-ESM-LR. Strong OHT_{50N} phases consistently lead to stronger SST anomalies up to a decade later than weak OHT_{50N} phases. The coherence of OHT_{50N} and AMV phases found in ASSIM can be produced by the MPI-ESM-LR if a running mean of three or more years is applied. This illustrates that the mode of climate variability found in chapter 3 does likely not exaggerate strongly the influence of OHT variability on SSTs. I will in this section discuss the limitations of this finding.

I show that strong OHT phases in the subpolar North Atlantic consistently influence SSTs stronger than weak OHT_{50N} phases in simulations with the MPI-ESM-LR that are characterized by similarly coherent strong and weak OHT_{50N} phases. However, the physical reason for the more pronounced influence of strong phases of subpolar OHT on North Atlantic SSTs compared to weak cases remains unexplored in this research. A deeper physical understanding of why this asymmetry exists - and in fact whether it can also be found in other global climate models - would advance our understanding of North Atlantic climate variability and should therefore be subject to future research. While a more in-depth examination on this effect lies outside the scope of this dissertation, I recommend a timely study here.

A comparison to the piControl, HIST and RCP model simulations and HadISST observations shows that the coherence of annual mean OHT_{50N} phases in the ASSIM simulation is unlikely to stem from internal model variability or greenhouse gas forcing. Therefore, the strong coherence of strong OHT_{50N} phases in ASSIM likely originates from the data assimilation that is applied. However, when using a three-year running mean to smooth OHT and SST variability, the coherence of strong and weak phases of OHT_{50N} and North Atlantic SSTs in ASSIM lies well within the model variability produced by piControl. The data assimilation might therefore act to smoothen year-to-year ocean variability in ASSIM. In the smoothed case, differently coherent OHT_{50N} and SST phases in ASSIM and piControl, HIST and RCP can therefore be attributed to internal variability; they occur naturally in the MPI-ESM-LR. This indicates that the mechanism leading to the characteristic SST pattern that I present here does not change strongly with greenhouse gas forcing by the 20^{th} century. I do however find a tendency for weaker coherence of both strong and weak OHT_{50N} phases in the last century of the RCP8.5 scenario. This indicates that the influence of strong and weak OHT_{50N} phases on North Atlantic surface temperatures might become shorter-lived with climate change.

These findings have an important implication for the research presented in chapter 4 of this dissertation: as the stronger influence of strong phases of subpolar OHT on North

Atlantic surface temperatures compared to weak phases can be shown to be robust, the higher ACCs for SSTs after strong than after weak OHT_{50N} phases in the northeast Atlantic up to a decade into the future are likely to be robust as well.
After showing that changes in decadal surface temperatures are robustly modulated by subpolar OHT, and therefore the skill of decadal SST hindcasts likely depends on OHT_{50N} variability, I will now discuss an important implication of this finding: the non-stationarity of hindcast skill. Because the findings presented up to now do leave us with a problem: if decadal hindcast skill varied in the past, how can we know how credible individual forecasts will be?

6 | Non-Stationary North Atlantic Surface Temperature Prediction Skill

6.1 Introduction

Conventional studies of decadal climate hindcasts result in one estimate of hindcast skill for the period in the past that is analyzed (e.g. Boer et al., 2016; Yeager and Robson, 2017). This skill estimate is often assumed to reflect the credibility of any individual forecast that would be conducted using the same prediction system, independent of the climate state from which this forecast is started. However, for the 20^{th} century the skill of SST hindcasts in the North Atlantic region depends on the strength of subpolar ocean heat transport at the start of the hindcast (see chapter 4). The skill of hindcasts is therefore time-dependent - or *non-stationary* (as also suggested by Weisheimer et al., 2017; Brune et al., 2017; O'Reilly et al., 2017). As a consequence, hindcast skill estimates derived for one period in the past do not represent the skill of all individual hindcasts within that period. Previous research suggested that this issue might be addressed by using longer and hence statistically more robust estimates of hindcast skill (e.g. Müller et al., 2014). While the estimate of mean hindcast skill would indeed be on better statistical grounds when a longer hindcast period is used, these hindcast skill estimates are unlikely to be appropriate to judge the credibility of individual decadal climate forecasts.

Here, I suggest the opposite: instead of using longer hindcast experiments to estimate skill in a representative way, I suggest to focus on shorter periods of time, but to take the climate state in which these hindcasts are started into account. This analysis, performed for the 20^{th} century, will indicate a way to estimate the credibility that can be expected from an individual decadal climate forecast based on the climate state at the start of that forecast.

ACCs of decadal North Atlantic SST hindcasts on average depend on the strength of subpolar OHT at the beginning of the hindcast (see chapter 4). OHT_{50N} variability is therefore likely a good indicator of decadal SST hindcast skill. This finding, however, is based on the average skill of approximately 40 individual decadal SST hindcasts after each strong and weak OHT_{50N} phases. The specific influence of the phase of subpolar

North Atlantic OHT at the start of one individual hindcast on the skill of that hindcast to predict North Atlantic SSTs on the decadal time scale is therefore still unclear. This problem, however, cannot be solved using conventional metrics of hindcast skill, as all conventional hindcast skill metrics are based on an assessment of multiple years. In this chapter, I propose an approach based on ACCs that connects the skill of individual hindcasts to the climatic state at the beginning of each hindcast.

Here, I test my approach to estimating the skill of individual hindcasts on hindcasts of the AMV in the 20^{th} century. I show in chapter 3 that the variability of the AMV is closely connected to that of the characteristic SST pattern, i.e. its evolution can be expected to be connected to the physical mechanism leading to the characteristic SST pattern and therefore variability of OHT_{50N}. Chapter 3 also suggests that AMV variability in ASSIM is similar to observed AMV variability. Hindcast skill for the AMV can therefore be evaluated against observations, which enhances the conclusions that can be drawn from this analysis.

I first show for the period 1901-2010 the non-stationarity of AMV hindcast skill in HC against HadISST observations to illustrate the spread of hindcast skills that is averaged over by estimating average hindcast skill for a long time series. I then propose a new metric that connects the skill of individual AMV hindcasts with the strength of subpolar North Atlantic OHT at the beginning of each individual hindcast. Findings from this chapter enable a direct translation of hindcast skill to the credibility of an individual decadal SST forecast in the North Atlantic region, using the strength of subpolar OHT at the start of the forecast as an indicator.

6.2 Non-Stationary Decadal AMV Hindcast Skill

An analysis of the non-stationarity of AMV hindcast skill for the 20^{th} century illustrates the range of skill that conventional hindcast skill estimates average over. As an example, AMV hindcast skill in the 20^{th} century at lead years 3-5 ranges between ACCs of 0 and 0.8 for a running 40-year window (fig. 6.1). One assessment of AMV hindcast skill for 40 years in the 20^{th} century is therefore not representative of the hindcast skill that would be found in a different 40-year-period. Figure 6.1 therefore supports the findings by Brune et al. (2017) and illustrates that SST hindcast skill in the North Atlantic is non-stationary. A hindcast skill estimate derived for the entire 20^{th} century thus averages over many differently predictable climate states, and is therefore not representative of the skill of every individual hindcast within that period.

AMV hindcast skill in the 20^{th} century strongly depends on the number of years that it is evaluated for (fig. 6.2). This dependence of AMV hindcast skill on the length of the evaluation period illustrates the extent to which hindcast skill estimates over a certain

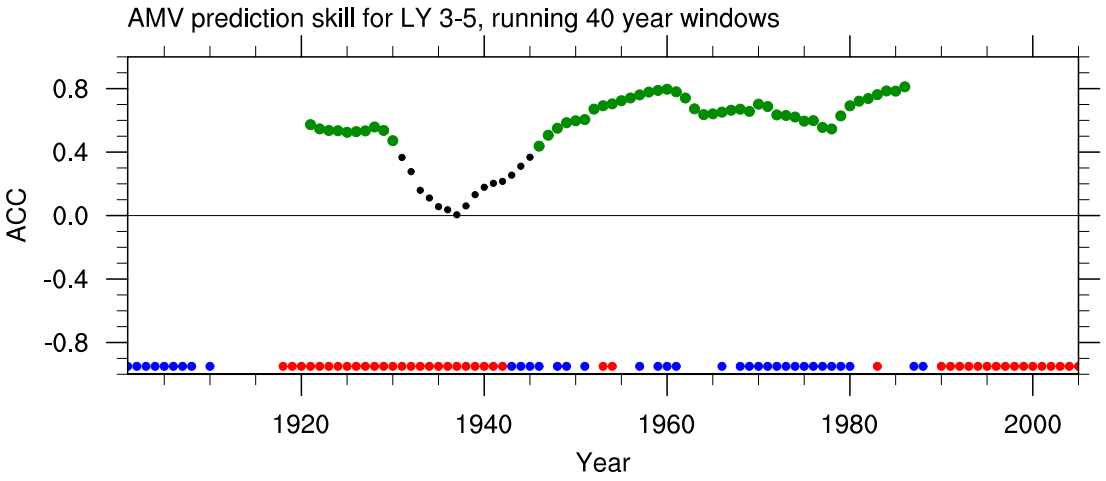


Figure 6.1: ACCs for the AMV (integrated SSTs in the North Atlantic between 0 and 60°N) evaluated against HadISST at lead years 3-5. Black dots indicate ACCs for a 40-year-long window around every dot. Green dots indicate significant ACCs at the 99% level. Red and blue dots at the bottom of the plot indicate strong and weak OHT_{50N} phases in ASSIM as identified in fig. 3.1.

period misrepresent the skill of individual hindcasts within that period. Figure 6.2 shows that differently long windows can lead to substantially different hindcast skill estimates for the same time period. Hindcast skill estimates derived for short time periods are sometimes as high as skill estimates for a longer time period, and sometimes lower. This illustrates that assessing hindcast skill for a long period likely results in a high skill estimate (as was shown by e.g. Müller et al., 2014), but that skill estimate does not reflect the hindcast skill of every shorter time period within that period. Thinking one step further, this analysis suggests that a hindcast skill estimate derived for several years, however many, is unlikely to reflect the hindcast skill of all individual hindcasts within that period.

The analysis presented above highlights that conventional measures of hindcast skill likely misrepresent the skill of individual hindcasts, because conventional skill measures rely on an analysis of many years. Because conventional hindcast skill estimates are not representative of the skill of individual hindcasts within the period they cover, these skill estimates cannot be directly translated into the credibility of an individual forecast. Estimating the credibility of an individual forecast therefore requires an estimate of skill for individual hindcasts. Showing that individual hindcasts are systematically more or less skillful when they are started in certain climate states would enable translating hindcast skill into forecast credibility by using the state of the climate system a the start of a forecast as an indicator.

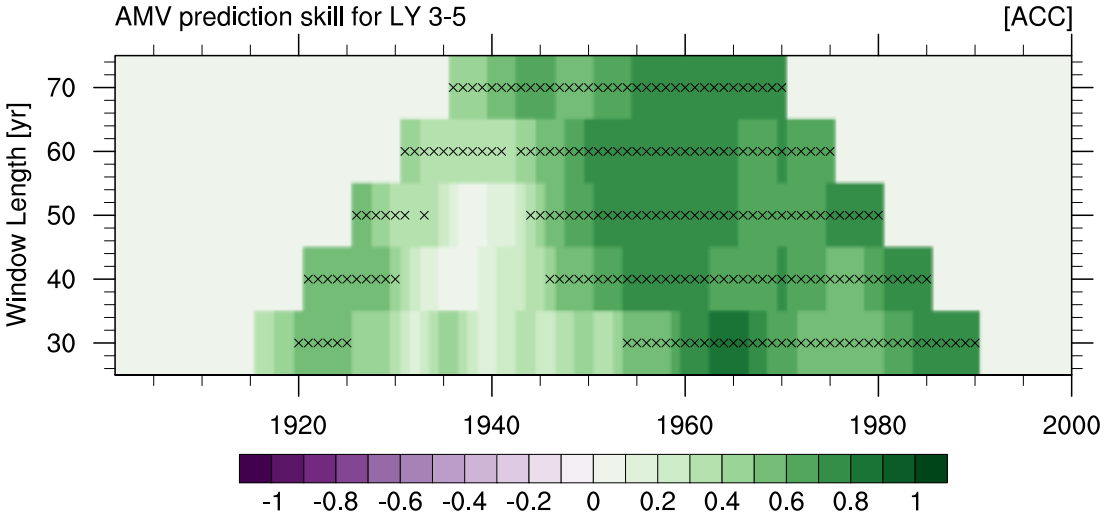


Figure 6.2: ACCs for the AMV evaluated against HadISST for lead years 3-5 as in figure 6.1, but for different running evaluation window lengths on the y-axis. Crosses indicate significant ACCs at the 99% level.

6.3 Towards Expected Prediction Skill

In chapter 4, I demonstrate the average influence of strong, weak, and neutral OHT_{50N} phases on decadal SST hindcast skill in the North Atlantic region. Could the phase of subpolar OHT at the start of a single hindcast therefore be used as an indicator of the decadal AMV skill of individual hindcasts? Figure 6.1 reveals no obvious connection between OHT_{50N} phases and non-stationary AMV hindcast skill. I will thus approach this issue by comparing predicted AMV anomalies to observed AMV anomalies for the period 1901-2010 individually, taking the strength of OHT_{50N} at the beginning of the prediction into account.

Individual predicted AMV anomalies at lead years 3-5 match reasonably well to observed AMV anomalies for 1901-2010 (fig 6.3): The majority of predicted AMV anomalies fall within one standard deviation of the corresponding observed AMV anomalies. This good overall match between individual hindcasts and the corresponding observations explains the high hindcast skill I find for the AMV in HC (cf. fig 6.1).

In some cases, the predicted AMV anomaly does not match the observed AMV anomaly: the predicted AMV anomaly lies more than one standard deviation away from the observed value (fig. 6.3). When OHT_{50N} is weak at the start of a hindcast, predicted AMV anomalies do not match the observed AMV more often than when OHT_{50N} is strong or neutral at the start of a hindcast (33%/25%/23% hindcasts following weak/neutral/strong OHT_{50N} phases do not match the observations). Individual AMV hindcasts are therefore more likely to match the observed AMV anomaly after strong and neutral phases of subpolar OHT than after weak OHT phases.

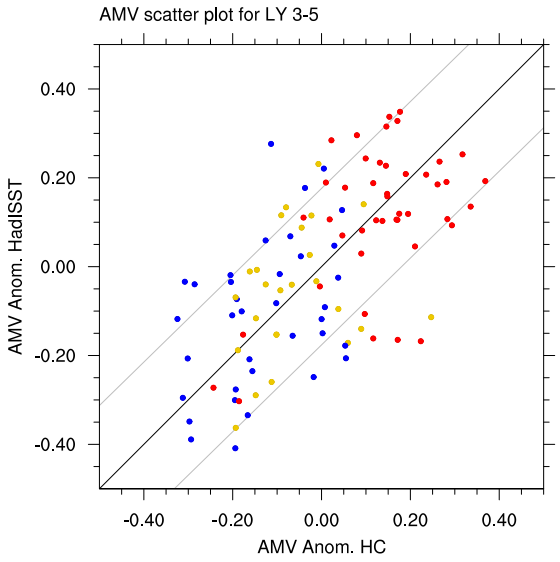


Figure 6.3: AMV anomalies calculated by integrating North Atlantic SSTs between the equator and 60°N [K] produced by HC against those observed in HadISST for lead years 3-5. Red dots indicate a strong OHT_{50N} phase in ASSIM at the beginning of the hindcast according to the criterion of half a standard deviation above or below the mean, while blue dots indicate weak, and yellow dots indicate neutral OHT_{50N} phases. The black line shows the 1:1 line, the grey lines a full sample standard deviation from the 1:1 line.

Indicating on a time series of OHT_{50N} individual hindcasts whose AMV hindcasts do not match the observed anomaly reveals that these years predominantly show OHT_{50N} phases that fall within one standard deviation of the mean state (fig. 6.4). Only two AMV hindcasts that were started when OHT_{50N} is more than one full standard deviation above the mean, and four AMV hindcasts started when OHT_{50N} is more than one standard deviation below the mean, do not match the observed AMV anomaly. Therefore, a single predicted AMV anomaly can be expected to lie within one standard deviation from the observed AMV anomaly when the prediction is initialized in a year in which OHT_{50N} is at least one standard deviation above or below the mean.
Analyses shown in chapter 4 indicate that the connection of strong OHT_{50N} phases to North Atlantic SST hindcast skill is stronger at lead years 7-9 than at lead years 3-5. This finding appears to be robust when analyzing whether AMV hindcasts at lead years 7-9 lie within one standard deviation of the observed AMV anomaly: Only few AMV hindcasts started in years of strong subpolar OHT do not match the observed AMV anomaly (31%/48%/18% hindcasts following weak/neutral/strong OHT_{50N} phases do not match the observations; fig. 6.5). It is noteworthy that almost half of the hindcasts following neutral OHT_{50N} phases predict an AMV anomaly that is more than one standard deviation off the observed value. This finding supports earlier results that attributed individual predicted AMV anomalies matching observations to either strong or weak phases of subpolar OHT at the initialization of the hindcast.
Predicted AMV anomalies, which do not match the observed value at lead year 7-9,

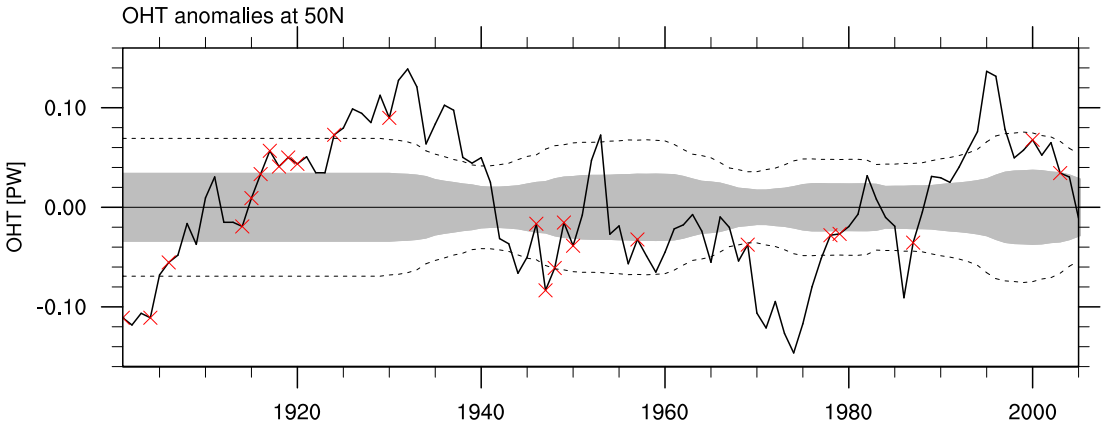


Figure 6.4: OHT_{50N} anomalies in ASSIM. The grey area denotes half a standard deviation of the OHT_{50N} of the preceding 30 years. Dashed lines show a full standard deviation of the preceding 30 years around the mean. Red crosses show years after which predicted AMV anomalies at lead years 3-5 lie more than one standard deviation away from the observed AMV anomaly (figure 6.3).

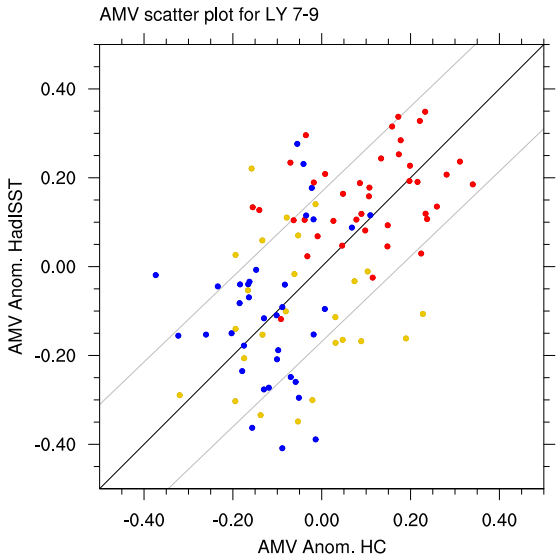


Figure 6.5: AMV anomalies calculated by integrating North Atlantic SSTs between the equator and 60°N [K] produced by HC against those observed in HadISST for lead years 7-9. Red dots indicate a strong OHT_{50N} phase in ASSIM at the beginning of the hindcast according to the criterion of half a standard deviation above or below the mean, while blue dots indicate weak, and yellow dots indicate neutral OHT_{50N} phases. The black line shows the 1:1 line, the grey lines a full sample standard deviation from the 1:1 line.

predominantly correspond to OHT_{50N} anomalies within one standard deviation from the mean state (fig. 6.6). If the OHT_{50N} anomaly is more than one standard deviation from the mean at the beginning of an individual decadal AMV hindcast, the predicted AMV anomaly can be expected to lie within one standard deviation of the observed value. Exceptions are two cases where AMV hindcasts 7-9 years ahead do not match observations after a strong OHT_{50N} phase, and five cases where AMV hindcasts do not

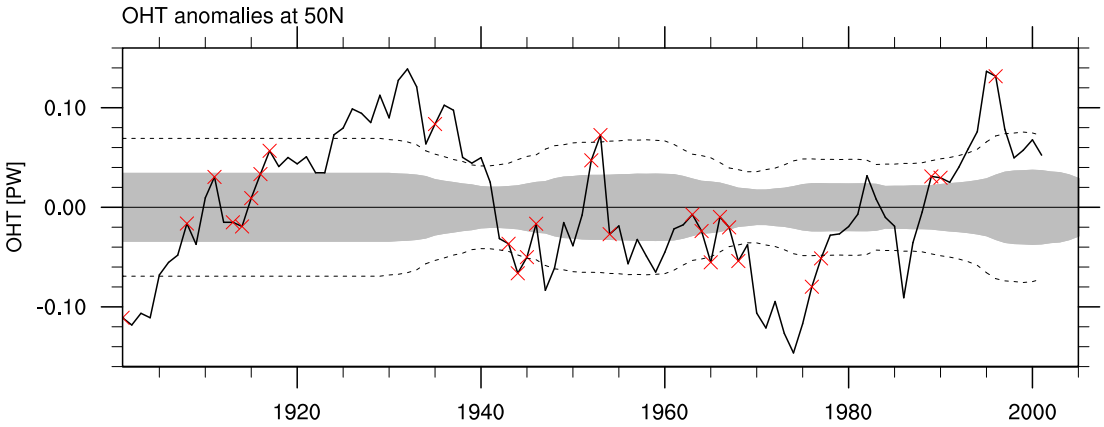


Figure 6.6: OHT_{50N} anomalies in ASSIM. The grey area denotes half a standard deviation of the OHT_{50N} of the preceding 30 years. Dashed lines show a full standard deviation of the preceding 30 years around the mean. Red crosses show years after which predicted AMV anomalies at lead years 7-9 lie more than one standard deviation away from the observed AMV anomaly (figure 6.5).

match the observed AMV anomaly after a weak OHT_{50N} phase. The connection of strong and weak OHT_{50N} anomalies to the skill of decadal SST hindcasts in the North Atlantic, found in chapter 4, is therefore largely valid for individual hindcasts.

6.4 Ocean Heat Transport as an Indicator of Prediction Skill: Limitations and Implications

Findings presented in this chapter show that it is possible to connect the strength of OHT_{50N} at the start of an individual decadal AMV-hindcast to the extent to which the predicted AMV matches the observations. When subpolar OHT shows an anomaly of more than one standard deviation above or below the mean at the beginning of a hindcast, AMV hindcasts predominantly fall within one standard deviation of the observed value. Strong phases of OHT_{50N} consistently lead to AMV hindcasts that fall within one standard deviation of the observed value. This effect likely arises from the physical mechanism leading to the characteristic SST pattern. AMV hindcast skill therefore depends on the climate state at the initialization of a hindcast. From this knowledge it can be inferred that observing the strength of subpolar OHT at the start of an individual SST forecast in the North Atlantic gives a good indication of the credibility of that forecast.

The analysis presented here comes, like most analyses presented in this dissertation, with the important caveat of the use of just one decadal prediction system. Although I showed in chapter 5 that the variability produced by ASSIM and HC is representative of model variability of the MPI-ESM-LR, the findings presented here are the product of only one model and will have to be replicated with other prediction systems to assess

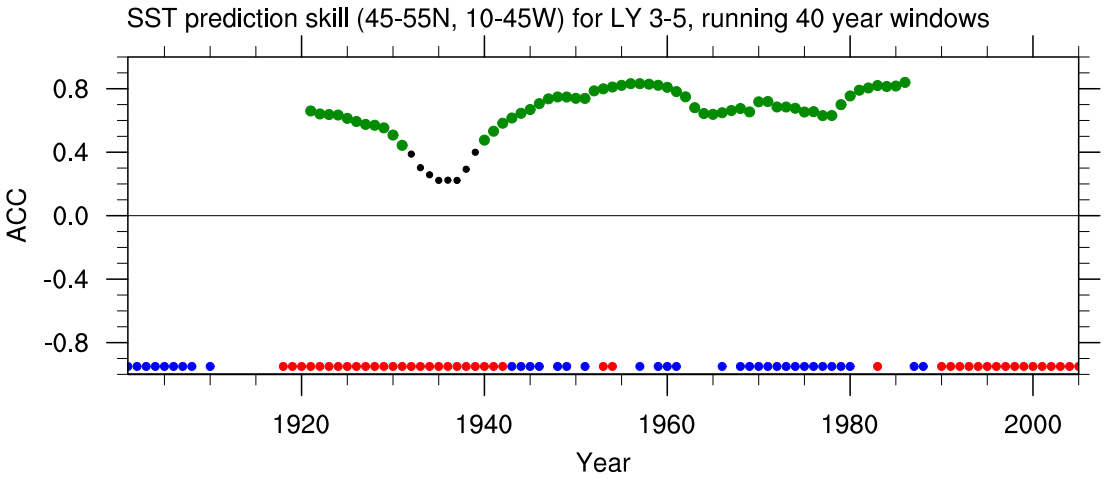


Figure 6.7: ACCs for the AMV (integrated SSTs in the North Atlantic between 0 and 60°N) evaluated against ASSIM at lead years 3-5. Black dots indicate ACCs for a 40-year-long window around every dot. Green dots indicate significant ACCs at the 99% level. Red and blue dots at the bottom of the plot indicate strong and weak OHT_{50N} phases in ASSIM as identified in fig. 3.1.

their robustness. In this chapter, there is one more issue of model specificity that needs to be discussed.

The strong non-stationarity found in this section could well be a model-artifact: maybe both ASSIM and HC produce unrealistic AMV variability around the 1940s that decreases hindcast skill around this time. An evaluation of AMV hindcasts in HC against ASSIM analogue to figure 6.1 shows, however, that the strong decrease in ACCs around the 1940s is robust (fig. 6.7). Moreover, the high hindcast skill found prior to 1930 suggests that the relative sparsity of observations in that time period does not systematically decrease the skill of hindcasts for that period. The non-stationary AMV hindcast skill I find in this section is therefore likely to be connected to physical mechanisms rather than a model artifact or connected to sparse observations.

The many years of neutral OHT_{50N} phases around 1920 and 1940 that are connected to predicted AMV anomalies that do not match the observed AMV (cf. fig. 6.4) explain the decrease in non-stationary AMV hindcast skill around the 1940s (cf. fig. 6.1). At that time, the time windows used for evaluating hindcast skill include many individual AMV hindcasts that do not match observations. The persistent phases of neutral OHT_{50N} in the 1920s and 1940s are therefore likely the reason for the decreased AMV hindcast skill estimates in the mid-1940s.

This chapter also addresses the common assumption in hindcast studies that the robustness of the hindcast skill estimate increases with an increasing time window for which hindcasts are evaluated. My findings show that this assumption generally holds for the statistical robustness of hindcast skill. However, results presented here also illustrate that a hindcast skill estimate calculated for a long period of time tends to over- or (in

rare cases) underestimate the skill of individual hindcasts within that period. With growing length of the evaluation window of hindcasts, this error increases. Long hindcast time series are thus, despite their strong statistical rigour, less useful to estimate the credibility of an individual forecast, than short hindcast time series.

The analysis presented here suggests that the skill of individual hindcasts can be immediately translated into the credibility of an individual forecast. An implicit assumption of this is that the physical mechanism used as an indicator for prediction skill does not change between the period that it is tested for (here: the 20^{th} century), and the forecasted period. As I show in chapter 5 of this thesis that the physical mechanism leading to the characteristic SST pattern is likely to only change with climate change by the end of the 22^{nd} century, it is unlikely that this physical mechanism will change on the short term.

I demonstrate in this chapter that conventional hindcast skill estimates are rarely suitable to estimate the credibility of a single forecast when that forecast is started. Without taking physical mechanisms into account, the skill derived by conventional hindcast studies cannot be understood as representative for any individual forecast. If, however, physical mechanisms like OHT_{50N} variability and the mechanism leading to the characteristic SST pattern are considered, the difference between hindcast skill and the credibility of a forecast can be estimated, and hindcast skill can be transferred to forecast credibility. This analysis highlights that skill estimates derived from hindcast analysis have to be based on individual years to be translateable to individual forecasts. New skill measures going beyond the approach presented in this chapter, will therefore have to be developed in the future and combined with conventional hindcast skill estimates to truly estimate the credibility of individual forecasts.

7 | Conclusions

This dissertation explores the skill of decadal temperature predictions in the North Atlantic region. In the introduction I ask four questions concerning the role of subpolar AMOC and ocean heat transport for temperature variability and predictability in the North Atlantic region. In this chapter I will first answer these questions one by one, then draw final conclusions concerning the findings presented in this dissertation.

7.1 Ocean Overturning and North Atlantic Temperatures

A study by Zhang and Zhang (2015) showed that variability of ocean overturning in the subpolar North Atlantic influences SSTs in the North Atlantic up to about a decade ahead. The resulting SST pattern is characterized by a positive anomaly in the North Atlantic subpolar gyre and a negative anomaly in the Gulf Stream region. The authors hypothesized that the physical mechanism they proposed modulates the skill of decadal SST predictions. Because Zhang and Zhang (2015) only show the AMOC Fingerprint in the GFDL climate model, I examine the evolution of Zhang's mechanism in the MPI-ESM-LR in this chapter, and answer the guiding question:

⇒ **Can the mechanism leading to the AMOC Fingerprint proposed by Zhang and Zhang (2015) be found in the MPI-ESM-LR, and how is this mechanism characterized?**

Decadal North Atlantic climate variability in an initialized simulation with the MPI-ESM-LR features the mechanism leading to the AMOC Fingerprint. Annual mean AMOC and OHT are highly correlated in this ASSIM model experiment. ASSIM further shows southward propagating AMOC and OHT phases in the North Atlantic on the time scale of up to a decade. This gives rise to the convergence of heat in the North Atlantic subpolar gyre region, and is restricted to the Atlantic North of 40°N. These AMOC and OHT dynamics are very similar to those found in Zhang and Zhang (2015).

Ocean heat content of the upper 700 m (UOHC) and SSTs in the North Atlantic are influenced in ASSIM for up to a decade ahead by the mechanism

described by Zhang and Zhang (2015). The most prominent SST anomaly connected to OHT phases at 50°N (OHT_{50N}) is found in the northeast Atlantic, whereas Zhang and Zhang (2015) find it to be more centrally located. The mechanism leading to the AMOC Fingerprint is therefore robust in at least two climate models, but its exact location is model-dependent. In the MPI-ESM-LR, its eastward displacement is likely due to an excessively zonal Gulf Stream path. The model-specific AMOC Fingerprint is in this dissertation called the *characteristic SST pattern*. In ASSIM, the characteristic SST pattern explains in ASSIM most UOHC and SST variability in the North Atlantic, indicating it is the dominant mode of decadal temperature variability there.

Neutral OHT_{50N} phases are only for half a decade connected to the characteristic SST pattern. Going beyond the original study by Zhang and Zhang (2015), I conduct a separate analysis of the influence of strong, weak and neutral phases of OHT_{50N} on North Atlantic SSTs. SSTs in the northeast Atlantic after neutral OHT_{50N} phases show a weak characteristic pattern up to 5 years into the future. This pattern is likely connected to persistent UOHC phases, supporting Zhang and Zhang (2015) as well as findings from this chapter that anomalous subpolar ocean overturning is particularly important in shaping North Atlantic SSTs on the decadal time scale.

OHT variability in the subpolar North Atlantic influences AMV variability at least partially. AMV variability is in ASSIM closely related to the characteristic SST pattern, and subpolar North Atlantic OHT consistently leads the AMV by 8 years. Meanwhile, the physical mechanism leading to the characteristic SST pattern starts from ocean surface density anomalies in the Labrador Sea. These could result from atmospheric forcing. This could reconcile the disagreement by Clement et al. (2015) and Zhang et al. (2016): this dissertation suggests the AMV as a signal of ocean-filtered stochastic atmospheric variability with a crucial part for both elements of the climate system.

The influence of strong and weak phases of OHT_{50N} on North Atlantic SSTs is asymmetric. Strong phases of OHT in the North Atlantic propagate southward more consistently than weak ones. As a result, strong OHT_{50N} phases lead to stronger anomalies of ocean heat convergence North of 40°N than weak ones. There is a strong ocean-driven asymmetry between strong and weak overturning phases in the North Atlantic. As a result, both UOHC and SSTs are strongly influenced by this asymmetry between strong and weak OHT phases in the subpolar North Atlantic. North of 40°N, a strong characteristic SST pattern arises for several years after strong phases of OHT_{50N}. This pattern is much weaker after weak OHT_{50N} phases and disappears by the seventh year. The physical reason for this asymmetry cannot be fully explained in this dissertation.

Surface heat fluxes dampen asymmetric decadal SST variability in the North Atlantic, which arises from strong and weak ocean overturning phases. Strong

phases of OHT_{50N} are connected to strong surface heat fluxes (SHFs) from the ocean into the atmosphere in the northwest Atlantic. After weak OHT_{50N} phases, a bipolar pattern of heat fluxes into the ocean arises with peaks in the Labrador Sea and the eastern North Atlantic.

The mechanism leading to the characteristic SST pattern influences seasonal surface air temperatures over Europe on the decadal time scale. The surface heat fluxes connected to OHT_{50N} variability indicate that subpolar SST variability influences atmospheric temperatures (SATs). OHT_{50N} variability is connected to annual and seasonal mean SAT variability up to 10 years ahead. This influence is particularly strong in western, northern, and eastern Europe.

The asymmetric connection of strong and weak phases of OHT_{50N} to North Atlantic SSTs can also be found in SATs. Strong OHT_{50N} phases lead to strong positive annual mean SAT anomalies in northern Europe around the UK and Scandinavia. After weak OHT_{50N} phases, a strong negative SAT anomaly arises in southern Scandinavia, the Baltics, and eastern Europe. This connection is particularly strong in winter and spring. While strong OHT_{50N} phases are connected to SAT anomalies in all seasons, weak OHT_{50N} phases show a distinct pattern in winter and spring, and almost no SAT pattern in the other seasons. This analysis illustrates the strong influence of the phase of subpolar OHT on European SAT variability.

7.2 Decadal Hindcasts in the North Atlantic Region

Studies showed a connection of strong AMOC anomalies to the skill of decadal North Atlantic SST hindcasts, based on case studies (Yeager et al., 2012; Robson et al., 2013, 2014). Here, I examine the speculation by Zhang and Zhang (2015) that the physical mechanism leading to the characteristic SST pattern modulates decadal SST hindcast skill. I evaluate systematically the specific influence of the variability of subpolar OHT on decadal North Atlantic SST prediction skill. Findings from this chapter are very indicative of the credibility of actual decadal SST predictions in the North Atlantic region. I draw on the findings from the previous chapter and answer my guiding research question:

⇒ **How strongly do ocean overturning dynamics influence the skill of SST hindcasts in the North Atlantic region?**

The physical mechanism leading to the characteristic SST pattern, i.e. ocean overturning dynamics, play an important role in modulating the skill of decadal SST predictions. SST hindcasts show high anomaly correlation coefficients (ACCs), a measure for the skill of hindcasts, in the area that is dominated by the characteristic SST pattern on the decadal time scale. When OHT_{50N} is in a neutral phase

at the beginning of a hindcast, ACCs for UOHC and SSTs are much lower at lead years 3-5 and 7-9 than when OHT_{50N} is either particularly strong or weak at the beginning of a hindcast. Low ACCs can be found at lead years 3-5 for UOHC and SSTs after neutral OHT_{50N} phases, which is connected to UOHC persistence. At lead years 7-9, there are no significant ACCs for UOHC and SST hindcasts on the decadal time scale in the North Atlantic after neutral OHT_{50N} phases, underscoring that overturning dynamics are essential in modulating the skill of surface temperature predictions on the decadal time scale.

The influence of OHT_{50N} variability on ACCs of North Atlantic SST hindcasts shows an asymmetry similar to the OHT_{50N} influence on decadal SST variability. UOHC and SST hindcasts show consistently higher ACCs after strong OHT_{50N} phases than after weak OHT_{50N} phases. This difference is likely to be significant in the area of the characteristic SST pattern.

Predictable SST anomalies arising from ocean heat convergence are at all examined lead times balanced by dampening SHFs. The asymmetry of ACCs after strong and weak OHT_{50N} phases is governed by two factors: the asymmetric ocean heat convergence signal after strong and weak OHT_{50N} phases (which leads to differently strong UOHC and SST anomalies), and the asymmetric surface heat flux patterns following the strong and weak OHT_{50N} phases.

ACCs are high in the northeast Atlantic after strong OHT_{50N} phases, and high in the central North Atlantic after weak OHT_{50N} phases. On the decadal time scale, the effect of SHFs dampening SST anomalies and consequently ACCs is particularly pronounced. After weak OHT_{50N} phases, strong SHFs in the eastern part of the characteristic SST pattern mask much of the SST anomaly arising from ocean variability, while strong SHFs in the western North Atlantic mask some SST anomaly arising there after strong OHT_{50N} phases. This leads to a zonal asymmetry in ACCs after strong and weak OHT_{50N} phases on the decadal time scale.

The model simulations analyzed here show very little significant ACCs in decadal SAT hindcasts over Europe. However, ACCs for SATs on the decadal time scale are high in the region of the characteristic SST pattern. This illustrates that OHT dynamics modulate the skill of SAT hindcasts at least in that region. While SAT-ACCs are generally limited over land in this study, some significant ACCs are found over the UK and western Scandinavia in spring after strong OHT_{50N} phases, and over eastern Scandinavia in summer after weak OHT_{50N} phases. These are areas in which SATs are affected at the same time by the characteristic SST pattern, which illustrates that there is potential for improvement of the skill of decadal predictions of seasonal SATs by using subpolar North Atlantic OHT as an indicator in a model of higher resolution, more ensemble members, or a more frequent initialization.

7.3 Understanding North Atlantic Climate Variability in the MPI-ESM-LR

Assimilation model experiments are a combination of a climate model and observational data. It is therefore from an assimilation model experiment alone unclear whether the detected climate variability has to be interpreted as model variability, observed climate variability, or something else entirely. To better understand the physical mechanism derived in chapter 3 and its influence on decadal SST prediction skill shown in chapter 4, I compare in chapter 5 the climate variability found in ASSIM to that produced by the non-initialized MPI-ESM-LR, and to the SST observations data set HadISST. I thus answer the guiding reserach question of this chapter:

⇒ **Is the previously discussed climate variability reasonable with respect to both model variability produced by the MPI-ESM-LR and observations?**

The physical mechanism leading to the characteristic SST pattern is robust within the MPI-ESM-LR. piControl, HIST and RCP4.5 show a similar mechanism leading to the characteristic SST pattern as ASSIM. The physical mechanism I find in ASSIM that modulates decadal SST variability and predictability is therefore not exclusively an artifact of model initialization.

Climate variability in ASSIM has to be smoothed with a 3-year running mean to reflect climate variability in the MPI-ESM-LR. The coherence of annual mean strong OHT_{50N} phases lies well outside the spread produced by piControl, HIST, HadISST, and both examined RCP scenarios. ASSIM is therefore unlikely to produce reasonably asymmetric influences of strong and weak phases of OHT_{50N} on annual mean SST variability in the North Atlantic. However, smoothing the OHT_{50N} and/or SST time series with a running mean ≥ 3 years places the OHT_{50N} and SST variability produced by ASSIM within reasonable coherence of OHT_{50N} and SST phases. This indicates that the findings presented in this dissertation are within the variability produced by the MPI-ESM-LR, as I use SST filtered with a 3-year running mean throughout this dissertation.

The time scale at which subpolar OHT variability influences SSTs will likely decrease with global warming. HadISST shows lower coherence of SST phases in the North Atlantic than the MPI-ESM-LR. The mechanism leading to the characteristic SST pattern is therefore likely to have a slightly shorter-lived influence on SSTs in reality than in the MPI-ESM-LR. Both RCP4.5 and RCP8.5 show shorter coherence of OHT_{50N} and SST phases than piControl, HIST and ASSIM. In a changing climate, the characteristic SST pattern will therefore likely vary at a higher frequency. This might also be true for overall climate variability in the North Atlantic.

Strong phases of subpolar OHT robustly influence ACCs for SSTs in the

North Atlantic more strongly than weak OHT phases. According to findings presented in this chapter, the overall conclusions concerning the influence of OHT_{50N} variability on the skill of SST hindcast hold. However, the asymmetry between strong and weak OHT_{50N} phases is likely to be exaggerated in this study. Nonetheless, the conclusion that strong OHT_{50N} phases influence ACCs for SSTs in the North Atlantic more favorably than weak OHT_{50N} phases is likely robust.

7.4 Non-Stationary North Atlantic Surface Temperature Prediction Skill

Conventional decadal climate prediction studies produce one skill estimate for the past (e.g. Boer et al., 2016). This skill estimate is then understood to reflect the credibility of any forecast with the examined prediction system. In chapter 4, however, I show that SST hindcast skill changes with changing OHT_{50N} in the North Atlantic. These findings support previous studies (e.g. Brune et al., 2017) in that decadal SST prediction skill in the North Atlantic can be assumed to be time-dependent, or non-stationary. This, in turn, limits the applicability of hindcast estimates produced for a fixed period in the past for the credibility of individual forecasts. In chapter 6, I propose a new approach to the estimation of the skill of hindcasts that is more applicable to forecasts, answering the guiding question:

⇒ **Are mean hindcast skill estimates appropriate to estimate the credibility of a single temperature forecast in the North Atlantic region?**

The applicability of hindcast skill estimates for individual forecasts is limited. AMV hindcast skill varies between ACCs of 0 and 0.8, depending on the time-horizon that it is evaluated for, the length of the evaluation time window, and lead time. Hindcast skill is generally high for long evaluation time windows and decreases with shorter evaluation time windows. I find high skill in the early 20^{th} century, around the 1950s, and towards the end of the century. Hindcast skill is low around the 1940s. The estimates of hindcast skill depend therefore on the point in time at which they are issued - this presents a substantial limitation to all climate prediction studies that do not account for this issue.

Extending the length of the period over which hindcast skill is evaluated does not facilitate the translation of hindcast skill into the credibility of individual forecasts. The changing skill estimate with length of the evaluation time window show that longer periods of observational records are needed to produce robust estimates of hindcast skill. This finding highlights, however, that skill estimates obtained for a long period of time do not reflect the skill of individual hindcasts within

this time period, but rather an overall estimate. Shorter evaluation periods have to be considered to produce a conclusive estimate of hindcast skill that is representative for all predictions issued in this period.

Individual decadal forecasts of North Atlantic SSTs are likely credible whenever OHT_{50N} is more than one standard deviation higher than the mean state at the beginning of the forecast. Strong phases of subpolar OHT in the North Atlantic consistently lead to skillful individual hindcasts of the AMV anomaly. This is not the case for neutral or weak phases of OHT_{50N}, which sometimes lead to skillful AMV hindcasts and sometimes not.

Other physical mechanisms can likely be used to judge the credibility of forecasts of other variables in other regions and on different time scales. Findings presented in this chapter indicate that the mechanism leading to the characteristic SST pattern can be used to judge the credibility of individual decadal SST forecasts in the North Atlantic. This implies that other physical mechanisms can be used to judge the credibility of forecasts of other variables in other regions and on different time scales. This implication needs further investigation in the future, highlighting other physical mechanisms, and developing measures of hindcast skill that allow for the evaluation of skill for individual years.

7.5 Hindcast Skill versus Forecast Skill: A Shift of Paradigm?

I opened this thesis by quoting Klaus Schulze, one of the pioneers of electronic music: *Everything changes, permanently. How boring if it wouldn't.* I would like to end this dissertation by looking back and reconciling this quote with the findings I presented in this dissertation: What is changing?

In recent years, an increasing number of studies showed events of strong climate variability to be well predictable on the decadal time scale (Yeager et al., 2012; Robson et al., 2013; Müller et al., 2014; Robson et al., 2014). While these case studies implied a change in hindcast skill over time, the actual time dependence of the skill of hindcasts more than a year ahead was first shown by Brune et al. (2017) for North Atlantic SSTs. However, the study by Brune et al. (2017) lacked a specific physical explanation for the non-stationarity of hindcast skill.

By systematically combining both - the time-dependence of prediction skill with a physical mechanism that explains how this time-dependence arises - this dissertation shows for the first time that there are times at which prediction systems that indicate a high overall hindcast skill show no skill whatsoever. This is an important finding for the interpretation of existing hindcast studies: any hindcast skill that is found for a long time period should not be expected from an individual climate forecast in the absence

of an indication that that prediction skill is actually constant over time - which, as this work indicates, is unlikely.

Finally, this work presents an important step towards the estimation of the credibility of a single decadal climate forecast. When it cannot be known whether hindcast skill estimates produced for a long time period are representative of the skill of any hindcast within this period, these hindcasts skill estimates cannot be used to judge the credibility of a single climate forecast. Physical mechanisms can help judging whether the credibility of a forecast is high or low. For North Atlantic SSTs, subpolar ocean heat transport at the start of the forecast is likely a good indicator of the credibility of a decadal temperature forecast. More research needs to be done on other regions and other time scales to verify these results, identifying other physical mechanisms for other regions.

Klaus Schulze's quote is therefore applicable to this dissertation in many ways. It does here not only reference changing climate and climate variability, but also changes in hindcast and forecast skill, and the interplay between variability and predictability. I also present evidence that the connection between variability and predictability should be expected to change over time. Finally, findings from this dissertation suggest that the research on hindcasts itself will have to change in the future to produce hindcast skill estimates that are translateable into the credibility of forecasts. Change is therefore the central topic of this work. The physical mechanisms that connect climate variability to decadal prediction skill are not trivial to identify, though, and this dissertation only represents the first attempt to systematically connect the two. There is a lot of work left to be done, and our understanding of the connection of climate variability and its prediction is almost certainly going to change in the future. Which is a good thing - after all, it would be *boring if it wouldn't*.

Bibliography

Arlot, S., and A. Celisse, 2010: A survey of cross-validation procedures for model selection. *Statistics Surveys*, **4 (0)**, 40–79, doi:10.1214/09-SS054, URL http://projecteuclid.org/euclid.ssu/1268143839.

Arthun, M., T. Eldevik, E. Viste, H. Drange, T. Furevik, H. L. Johnson, and N. S. Keenlyside, 2017: Skillful prediction of northern climate provided by the ocean. *Nature Communications*, **8**, 15 875, doi:10.1038/ncomms15875, URL http://www.nature.com/doifinder/10.1038/ncomms15875.

Balsameda, M., and D. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters*, **36 (L01701)**, L01 701, doi:10.1029/2008GL035561.

Bellomo, K., A. Clement, L. Murphy, L. Polvani, and M. Cane, 2016: New observational evidence for a positive cloud feedback that amplifies the atlantic multidecadal oscillation. *Geophys. Res. Lett.*, **43 (18)**, 9852–9859, doi:10.1002/2016GL069961.

Bjerknes, J., 1964: Atlantic air-sea interaction. *Advances in Geophysics*, **10**, 1–82.

Boer, G., and Coauthors, 2016: The decadal climate prediction project (dcpp). *Geoscientific Model Development*, **9**, 3751–3777, doi:10.5194/gmd-9-3751-2016.

Branstator, G., and H. Teng, 2010: Two Limits of Initial-Value Decadal Predictability in a CGCM. *Journal of Climate*, **23 (23)**, 6292–6311, doi:10.1175/2010JCLI3678.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2010JCLI3678.1.

Brune, S., A. Düsterhus, H. Pohlmann, W. A. Muller, and J. Baehr, 2017: Time dependency of the prediction skill for the north atlantic subpolar gyre in initialized decadal hindcasts. *Climate Dynamics*, **4 (0)**, doi:10.1007/s00382-017-3991-4.

Cane, M. A., A. C. Clement, L. N. Murphy, and K. Bellomo, 2017: Low-pass filtering, heat flux, and atlantic multidecadal variability. *Journal of Climate*, **30 (18)**, 7529–7553, doi:10.1175/JCLI-D-16-0810.1.

Clement, A., K. Bellomo, L. N. Murphy, M. A. Cane, T. Mauritsen, G. Radel, and B. Stevens, 2015: The Atlantic Multidecadal Oscillation without a role for ocean

circulation. *Science*, **350 (6258)**, 320–324, doi:10.1126/science.aab3980, URL http://www.sciencemag.org/cgi/doi/10.1126/science.aab3980.

Collow, T. W., W. Wang, A. Kumar, and J. Zhang, 2015: Improving Arctic Sea Ice Prediction Using PIOMAS Initial Sea Ice Thickness in a Coupled Ocean–Atmosphere Model. *Monthly Weather Review*, **143 (11)**, 4618–4630, doi:10.1175/MWR-D-15-0097.1, URL http://journals.ametsoc.org/doi/10.1175/MWR-D-15-0097.1.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Quarterly Journal of the Royal Meteorological Society*, **137 (654)**, 1–28, doi:10.1002/qj.776, URL http://doi.wiley.com/10.1002/qj.776.

Czaja, A., and C. Frankignoul, 2002: Observed impact of atlantic sst anomalies on the north atlantic oscillation. *Journal of Climate*, **15 (6)**, 606–623, doi:10.1175/1520-0442(2002)015.

Delworth, T., F. Zeng, L. Zhang, R. Zhang, G. Vecchi, and X. Yang, 2017: The central role of ocean dynamics in connecting the north atlantic oscillation to the extratropical componen of the atlantic multidecadal oscillation. *Journal of Climate*, **30 (16)**, 3789–3805, doi:10.1175/JCLI-D-16-0358.1.

Doblas-Reyes, F. J., and Coauthors, 2013: Initialized near-term regional climate change prediction. *Nature Communications*, **4**, 1715, doi:10.1038/ncomms2704, URL http://www.nature.com/doifinder/10.1038/ncomms2704.

Dong, S., S. L. Hautala, and K. A. Kelly, 2007: Interannual Variations in Upper-Ocean Heat Content and Heat Transport Convergence in the Western North Atlantic. *Journal of Physical Oceanography*, **37 (11)**, 2682–2697, doi:10.1175/2007JPO3645.1, URL http://journals.ametsoc.org/doi/abs/10.1175/2007JPO3645.1.

Drews, A., and R. Greatbatch, 2017: Evolution of the atlantic multidecadal variability in a model with an improved north atlantic current. *Journal of Climate*, **30 (14)**, 5491–5512, doi:10.1175/JCLI-D-16-0790.1.

Duchez, A., P. Courtois, E. Harris, S. Josey, T. Kanzow, R. Marsh, D. Smeed, and J.-M. Hirschi, 2016: Potential for seasonal prediction of atlantic sea surface temperatures using the rapid array at 26n. *Climate Dynamics*, **46 (15)**, 3351–3370, doi:10.1007/s00382-015-2918-1.

Eden, C., and J. Willebrand, 2001: Mechanism of interannual to decadal variability of the north atlantic circulation. *Journal of Climate*, **14**, 2266–2280.

Gaetani, M., and E. Mohino, 2013: Decadal prediction of the sahelian precipitation in cmip5 simulations. *Journal of Climate*, **49 (9-10)**, 7708–7719, doi:10.1175/JCLI-D-12-00635.1.

Gulev, S. K., M. Latif, N. Keenlyside, W. Park, and K. P. Koltermann, 2013: North Atlantic Ocean control on surface heat flux on multidecadal timescales. *Nature*, **499 (7459)**, 464–467, doi:10.1038/nature12268, URL http://www.nature.com/doifinder/10.1038/nature12268.

Hoskins, B. J., and D. J. Karoly, 1981: The steady linear response of a spherical atmosphere to thermal and orographic forcing. *Journal of Atmospheric Sciences*, **38**, 1179–1196, doi:10.1175/1520-0469(1981)038<1179:TSLROA>2.0.CO;2.

Jayne, S. R., and J. Marotzke, 2001: The dynamics of ocean heat transport variability. *Reviews of Geophysics*, **39 (3)**, 385–411, doi:10.1029/2000RG000084, URL http://doi.wiley.com/10.1029/2000RG000084.

Jolliffe, I. T., and D. B. Stephenson, 2012: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, Vol. 2. 2nd ed., Wiley and Sons, Ltd.

Jungclaus, J. H., and Coauthors, 2013: Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *Journal of Advances in Modeling Earth Systems*, **5 (2)**, 422–446, doi:10.1002/jame.20023, URL http://doi.wiley.com/10.1002/jame.20023.

Keenlyside, N. S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner, 2008: Advancing decadal-scale climate prediction in the North Atlantic sector. *Nature*, **453 (7191)**, 84–88, doi:10.1038/nature06921, URL http://www.nature.com/doifinder/10.1038/nature06921.

Kerr, R., 2000: A north atlantic climate pacemaker for the centuries. *Science*, **288 (5473)**.

Klöwer, M., M. Latif, H. Ding, R. Greatbatch, and W. Park, 2014: Atlantic meridional overturning circulation and the prediction of North Atlantic sea surface temperature. *Earth and Planetary Science Letters*, **406**, 1–6, doi:10.1016/j.epsl.2014.09.001, URL http://linkinghub.elsevier.com/retrieve/pii/S0012821X1400541X.

Lozier, S. M., and Coauthors, 2017: Overturning in the Subpolar North Atlantic Program: A New International Ocean Observing System. *Bulletin of the American Meteorological Society*, **98 (4)**, 737–752, doi:10.1175/BAMS-D-16-0057.1, URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0057.1.

Marotzke, J., and Coauthors, 2016: Miklip: A national research project on decadal climate prediction. *BAMS*, **12**, 2379–2394, doi:10.1175/BAMS-D-15-00184.1.

Marshall, J., H. Johnson, and J. Goodman, 2001: A study of the interaction of the north atlantic oscillation with ocean circulation. *Journal of Climate*, **14 (2)**, 1399–1421, doi:10.1175/1520-0442(2001)014<1399:ASOTIO>2.0.CO;2.

Matei, D., H. Pohlmann, J. Jungclaus, W. Müller, H. Haak, and J. Marotzke, 2012: Two Tales of Initializing Decadal Climate Prediction Experiments with the ECHAM5/MPI-OM Model. *Journal of Climate*, **25 (24)**, 8502–8523, doi:10.1175/JCLI-D-11-00633.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00633.1.

Meehl, G., and Coauthors, 2009: Decadal prediction: Can it be skillful? *Bulletin of the American Meteorological Society*, **90 (10)**, 1467–1486, doi:10.1175/2009BAMS2778.1.

Monerie, P.-A., L. Coquart, E. Maisonnave, M.-P. Moine, L. Terray, and S. Valcke, 2017: Decadal prediction skill using a high-resolution climate model. *Climate Dynamics*, **1 (1)**, 7708–7719, doi:10.1007/s00382-017-3528-x.

Müller, W. A., H. Pohlmann, F. Sienz, and D. Smith, 2014: Decadal climate predictions for the period 1901-2010 with a coupled climate model. *Geophysical Research Letters*, **41 (6)**, 2100–2107, doi:10.1002/2014GL059259, URL http://doi.wiley.com/10.1002/2014GL059259.

Müller, W. A., and Coauthors, 2012: Forecast skill of multi-year seasonal means in the decadal prediction system of the Max Planck Institute for Meteorology. *Geophysical Research Letters*, **39 (22)**, L22 707, doi:10.1029/2012GL053326, URL http://doi.wiley.com/10.1029/2012GL053326.

Müller, W. A., and Coauthors, 2015: A twentieth-century reanalysis forced ocean model to reconstruct the North Atlantic climate variation during the 1920s. *Climate Dynamics*, **44 (7-8)**, 1935–1955, doi:10.1007/s00382-014-2267-5, URL http://link.springer.com/10.1007/s00382-014-2267-5.

Olonscheck, D., and D. Notz, 2017: Consistently estimating internal climate variability from climate model simulations. *Journal of Climate*, **30 (23)**, 9555–9573, doi:10.1175/JCLI-D-16-0428.1.

O'Reilly, C., J. Heatley, D. MacLeod, A. Weisheimer, T. Palmer, N. Schaller, and T. Wollings, 2017: Variability in seasonal forecast skill of northern hemisphere winters over the twentieth century. *Geophys. Res. Lett.*, **44 (11)**, 5729–5738, doi:10.1002/2017GL073736.

Palmer, T. N., and Coauthors, 2004: Development of a european multimodel ensemble system for seasonal-to-interannual prediction (demeter). *Bulletin of the American*

Meteorological Society, **85 (6)**, 853–872, doi:10.1175/BAMS-85-6-853, URL http://journals.ametsoc.org/doi/abs/10.1175/BAMS-85-6-853.

Peings, Y., G. Simpkins, and G. Magnusdottir, 2016: Multidecadal fluctuations of the north atlantic ocean and feedback on the winter climate in cmip5 control simulations. *J. Geophys. Res. Atmos.*, 120, doi:10.1002/2015JD024107.

Pielke Sr., R., G. Liston, J. Eastman, and L. Lu, 1999: Seasonal weather prediction as an initial value problem. *Journal of Geophysical Research*, **104 (D16)**, 19 463–19 479, doi:10.1.1.167.9777.

Pohlmann, H., J. Kröger, R. J. Greatbatch, and W. A. Müller, 2017: Initialization shock in decadal hindcasts due to errors in wind stress over the tropical pacific. *Climate Dynamics*, **49 (1)**, 2685–2693, doi:10.1007/s00382-016-3486-8, URL https://link.springer.com/article/10.1007/s00382-016-3486-8.

Rayner, N., D. Parker, E. Horton, C. Folland, L. Alexander, D. Rowell, E. Kent, and A. Kaplan, 2003: Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *J. Geophys. Res.*, **108 (D14)**, 4407, doi:10.1029/2002JD002670.

Robson, J., I. Polo, D. L. Hodson, D. P. Stevens, and L. C. Shaffrey, 2017: Decadal prediction of the north atlantic subpolar gyre in the higem high-resolution climate model. *Climate Dynamics*, **2 (0)**, 1–17, doi:10.1007/s00382-017-3649-2.

Robson, J., R. Sutton, and D. Smith, 2013: Predictable Climate Impacts of the Decadal Changes in the Ocean in the 1990s. *Journal of Climate*, **26 (17)**, 6329–6339, doi:10.1175/JCLI-D-12-00827.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-12-00827.1.

Robson, J., R. Sutton, and D. Smith, 2014: Decadal predictions of the cooling and freshening of the North Atlantic in the 1960s and the role of ocean circulation. *Climate Dynamics*, **42 (9-10)**, 2353–2365, doi:10.1007/s00382-014-2115-7, URL http://link.springer.com/10.1007/s00382-014-2115-7.

Robson, J. I., R. T. Sutton, and D. M. Smith, 2012: Initialized decadal predictions of the rapid warming of the North Atlantic Ocean in the mid 1990s. *Geophysical Research Letters*, **39 (19)**, L19 713, doi:10.1029/2012GL053370, URL http://doi.wiley.com/10.1029/2012GL053370.

Sanchez-Gomez, E., C. Cassou, Y. Ruprich-Robert, E. Fernandez, and L. Terray, 2016: Drift dynamics in a coupled model initialized for decadal forecasts. *Climate Dynamics*, **46 (1)**, 1819–1840, doi:10.1007/s00382-015-2678-y, URL https://link.springer.com/article/10.1007/s00382-015-2678-y.

Servonnat, J., J. Mignot, D. Guilyardi, E. Swingedouw, R. Séférian, and S. Labetoulle, 2015: Reconstructing the subsurface ocean decadal variability using surface nudging in a perfect model framework. *Climate Dynamics*, **44 (14)**, 315–338, doi:10.1007/s00382-014-2184-7.

Sheen, K., D. Smith, N. Dunstone, R. Eade, D. Rowell, and M. Vellinga, 2017: Skilful prediction of sahel summer rainfall on inter-annual and multi-year time scales. *Nature Communications*, **8**, 14 966, doi:10.1038/ncomms14966.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, and J. M. Murphy, 2007: Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model. *Science*, **317 (5839)**, 796–799, doi:10.1126/science.1139540, URL http://www.sciencemag.org/cgi/doi/10.1126/science.1139540.

Smith, D. M., R. Eade, and H. Pohlmann, 2013: A comparison of full-field and anomaly initialization for seasonal to decadal climate prediction. *Climate Dynamics*, **41 (11-12)**, 3325–3338, doi:10.1007/s00382-013-1683-2, URL http://link.springer.com/10.1007/s00382-013-1683-2.

Stevens, B., and Coauthors, 2013: Atmospheric component of the MPI-M Earth System Model echam6. *Journal of Advances in Modeling Earth Systems*, **5 (2)**, 146–172, doi:10.1002/jame.20015, URL http://doi.wiley.com/10.1002/jame.20015.

Storch, H. v., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*, Vol. 1. 1st ed., Cambridge University Press.

Timmermann, A., M. Latif, R. Voss, and A. Grötzner, 1998: Northern hemispheric interdecadal variability: a coupled air-sea mode. *Journal of Climate*, **11 (2)**, 1906–1931, doi:10.1175/1520-0442(1998)011<1906:NHIVAC>2.0.CO;2.

Ting, M., Y. Kushnir, R. Seager, and C. Li, 2011: Robust features of atlantic multi-decadal variability and its climate impacts. *Geophys. Res. Lett.*, **38**, L17 705, doi:10.1029/2011GL048712.

Visbeck, M., J. Hurrell, L. Polvani, and H. Cullen, 2001: The north atlantic oscillation: Past, present, and future. *Proceedings of the National Academy of Sciences*, **98 (23)**, 12 876–12 877, doi:10.1073/pnas.231391598.

Weisheimer, A., N. Schaller, C. O'Reilly, D. MacLeod, and T. Palmer, 2017: Atmospheric seasonal forecasts of the twentieth century: multi-decadal variability in predictive skill of the winter north atlantic oscillation (nao) and their potential value for extreme event attribution. *Quarterly Journal of the Royal Meteorological Society*, **143 (703)**, 917–926, doi:10.1002/qj.2976.

Yeager, S., A. Karspeck, G. Danabasoglu, J. Tribbia, and H. Teng, 2012: A Decadal Prediction Case Study: Late Twentieth-Century North Atlantic Ocean Heat Content. *Journal of Climate*, **25 (15)**, 5173–5189, doi:10.1175/JCLI-D-11-00595.1, URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI-D-11-00595.1.

Yeager, S., and J. Robson, 2017: Recent progress in understanding and predicting decadal climate variability. *Current Climate Change Reports*, **3 (3)**, 112–127, doi:10.1007/s40641-017-0064-z.

Zhang, J., and R. Zhang, 2015: On the evolution of Atlantic Meridional Overturning Circulation Fingerprint and implications for decadal predictability in the North Atlantic. *Geophysical Research Letters*, **42 (13)**, 5419–5426, doi:10.1002/2015GL064596, URL http://doi.wiley.com/10.1002/2015GL064596.

Zhang, R., 2008: Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophysical Research Letters*, **35 (20)**, doi:10.1029/2008GL035463, URL http://doi.wiley.com/10.1029/2008GL035463.

Zhang, R., 2010: Latitudinal dependence of atlantic meridional overturning circulation (amoc) variations. *Geophysical Research Letters*, **37 (0)**, L16 703, doi:10.1029/2010GL044474.

Zhang, R., R. Sutton, G. Danabasoglu, T. L. Delworth, W. M. Kim, J. Robson, and S. G. Yeager, 2016: Comment on "The Atlantic Multidecadal Oscillation without a role for ocean circulation". *Science*, **352 (6293)**, 1527–1527, doi:10.1126/science.aaf1660, URL http://www.sciencemag.org/cgi/doi/10.1126/science.aaf1660.