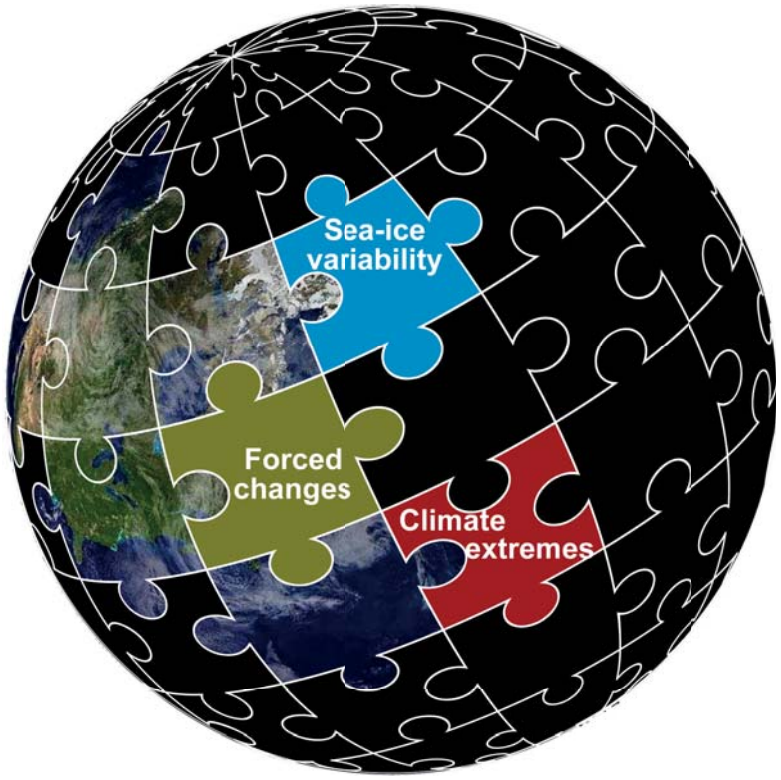
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

# Understanding internal variability of sea ice and surface air temperature




Dirk Olonscheck

Hamburg 2018

Berichte zur Erdsystemforschung
*Reports on Earth System Science*

209
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



# Understanding internal variability of sea ice and surface air temperature


Quantifying, unravelling and coping with chaos on a changing planet

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

Dirk Olonscheck

Hamburg 2018

Dirk Olonscheck

Max-Planck-Institut für Meteorologie
Bundesstrasse 53
20146 Hamburg

Tag der Disputation: 04.07.2018

Folgende Gutachter empfehlen die Annahme der Dissertation:

Dr. Dirk Notz
Prof. Dr. Martin Claußen

Titlepage2 figure credit:
https://sciencesprings.wordpress.com/2018/02/21/

Typeset using the classicthesis template developed by André Miede, available at:
https://bitbucket.org/amiede/classicthesis/

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 209
*Reports on Earth System Science / Max Planck Institute for Meteorology* 2018

ISSN 1614-1199

For my family.

## ABSTRACT

In this dissertation I investigate three aspects of internal climate variability: the quantification of changes in internal climate variability through time, the attribution of the future probability of climate extremes, and the understanding of the driving mechanisms for sea-ice variability.

I first introduce a new method that shifts our way of thinking about variability from the time domain towards the ensemble domain. From this method, I infer consistent estimates of internal climate variability and its change through time across models. The multi-model view allows me to provide robust evidence that the internal variability of annual near-surface air temperature will remain unchanged on a global average, but will likely increase in many tropical, subtropical, and polar regions and likely decrease in mid to high latitudes under large $CO_2$ forcing. The internal variability of Arctic and Antarctic sea-ice volume and Antarctic sea-ice area will in all likelihood decrease proportionally to the mean sea-ice state. I further find that the simulated trends in sea-ice volume and area are mostly plausible with respect to observed trends when model-specific internal variability is taken into account.

Secondly, I attribute the future probability of temperature and precipitation extremes to changes caused by a shift in the mean and changes caused by a change in higher-order moments. Based on an empirical threshold approach for multiple models, I show that the increased probability of hot extremes and the largely vanishing probability of cold extremes under large $CO_2$ forcing is mainly determined by the shift in the mean. In contrast, the changed probability of heavy precipitation extremes can be mainly attributed to the projected change in higher-order moments.

Finally, I challenge the wide-spread belief that intricate atmospheric or oceanic effects and feedbacks are important drivers of the substantial year-to-year variability of Arctic sea-ice area. I instead provide robust evidence that most sea-ice variability is directly driven by atmospheric temperature fluctuations. This implies that possible feedback-driven tipping points in the sea-ice system are unlikely to exist and sets a natural limit to seasonal predictions of sea ice.

My findings open up pathways for an explicit quantitative consideration of internal climate variability in climate studies and for the reduction of uncertainties in future climate projections. The methods used here could be applied to any climate variable and hence be used to further unravel mysteries of internal climate variability in related fields such as paleoclimatology, or Southern Ocean variability. Enabled by the combination of new ways of thinking and growing computational power, my findings foster our understanding of internal climate variability on a changing planet.

# ZUSAMMENFASSUNG

In dieser Dissertation untersuche ich drei verschiedene Aspekte der internen Klimavariabilität: die Quantifizierung von zeitlichen Änderungen der internen Klimavariabilität, die Ursachenaufteilung von zukünftigen Auftretenswahrscheinlichkeiten von Klimaextremen, und das Verständnis der Antriebsmechanismen für Schwankungen in der Meereisfläche.

Im ersten Teil der Dissertation stelle ich eine neue Methode vor, die unsere Denkweise über Klimavariabilität von der zeitlichen Dimension zur Ensemble-Dimension lenkt. Ich nutze diese Methode, um konsistente Abschätzungen der Klimavariabilität und ihrer zeitlichen Änderung von mehreren Modellen abzuleiten. Die gleichzeitige Abschätzung für mehrere Modelle erlaubt mir robust nachzuweisen, dass die interne Variabilität von jährlicher Oberflächentemperatur unter starkem $CO_2$-Antrieb im globalen Mittel nahezu unverändert bleibt, in tropischen, subtropischen und polaren Regionen wahrscheinlich zunimmt, und in mittleren bis hohen Breiten wahrscheinlich abnimmt. Die interne Variabilität des arktischen und antarktischen Meereisvolumens und der antarktischen Meereisfläche nimmt sehr wahrscheinlich und proportional zum mittleren Zustand ab. Zudem zeige ich, dass die von Modellen simulierten Trends in Bezug auf die beobachteten Trends in der Meereisfläche und dem Meereisvolumen größtenteils plausibel sind, wenn die modellspezifische interne Klimavariabilität berücksichtigt wird.

Im zweiten Teil der Dissertation quantifiziere ich, wie stark die zukünftig veränderte Wahrscheinlichkeit von Temperatur- und Niederschlagsextremen durch die Verschiebung im Mittelwert und die Änderung der höheren Momente der Verteilungen verursacht wird. Mit Hilfe eines empirischen Grenzwertes und täglichen Daten von mehreren Modellen zeige ich, dass die zunehmende Wahrscheinlichkeit von Hitzeextremen und die zumeist verschwindende Wahrscheinlichkeit von Kälteextremen unter starkem $CO_2$-Antrieb hauptsächlich von der Verschiebung im Mittelwert bestimmt ist. Die veränderte Auftretenswahrscheinlichkeit von Starkniederschlagsextremen ist hingegen hauptsächlich mit der Änderung der höheren Momente der Verteilung zu erklären.

Im dritten Teil der Dissertation stelle ich die weit verbreitete Annahme infrage, dass komplizierte atmosphärische Effekte und Rückkopplungen wichtige Antriebe für die erheblichen zwischenjährlichen Schwankungen der arktischen Meereisfläche sind. Stattdessen weise ich robust nach, dass der Großteil der Meereisvariabilität direkt durch Schwankungen der atmosphärischen Temperatur verursacht ist. Dies macht mögliche, durch Rückkopplungen getriebene, Kipppunkte im System Meereis unwahrscheinlich und setzt eine natürliche Grenze für saisonale Vorhersagen von Meereis.

Die Erkenntnisse meiner Dissertation zeigen neue Wege auf, interne Klimavariabilität in Klimastudien explizit quantitativ zu berücksichtigen, und Unsicherheiten in zukünftigen Klimaprojektionen zu verringern. Die hier genutz-

ten Methoden können für jede Klimavariable angewandt und somit dafür verwendet werden, um neues Wissen zur internen Klimavariabilität in benachbarten Forschungsfeldern, wie der Paleoklimatologie oder der Variabilität im Südlichen Ozean, zu gewinnen. Die hier gewonnenen Erkenntnisse wurden durch die Kombination aus neuen Denkweisen und wachsender Rechenleistung ermöglicht und erhöhen unser Verständnis der internen Klimavariabilität auf einem sich verändernden Planeten.

# PRE-PUBLISHED WORK RELATED TO THIS DISSERTATION

Olonscheck, D. and Notz, D. (2017), "Consistently Estimating Internal Climate Variability from Climate Model Simulations", *Journal of Climate* 30.23, pp. 9555-9573. DOI: 10.1175/JCLI-D-16-0428.1.

Olonscheck, D., Mauritsen, T., and Notz, D., "Arctic sea-ice variability is primarily driven by atmospheric temperature fluctuations", *in preparation for resubmission to Nature Geoscience*.

# ACKNOWLEDGMENTS

I could only write this dissertation because of the many inspiring, supportive and lovely people that surrounded me in the last three and a half years. First of all, I like to thank my unique family - my sister, parents and grandparents - for the unconditional love and everlasting support that I received whenever needed. You are the best family I can imagine.

I am extremely grateful to Dirk Notz and Thorsten Mauritsen. Your brilliant supervision, continuous inspiration and focused guidance shaped my research and also myself as a scientist. Thanks for your high expectations and for providing the scientific freedom that allowed me to pursue my own research interests. I also owe to Jochem Marotzke and Lars Kaleschke. Thank you for your great advice and for keeping your eyes on my progress. I further like to acknowledge Martin Claußen, Inga Hense and Lars Kutzbach for serving as commission members.

Writing this dissertation was largely facilitated thanks to the smooth and reliable infrastructure at the Max Planck Institute for Meteorology, the trouble-free data management by DKRZ and the patient support by two of the greatest scientific programmers I have ever met, Karl-Hermann Wieners and Helmut Haak. Many thanks also to Antje Weitz and the IMPRS office for their excellent administrative support.

Special thanks go to very likeable colleagues: Clara Burgard, Laura Niederdrenk, Felix Bunzel and Niels Fuchs for the great time each day in our office and the many relaxing (coffee) breaks, and to the whole sea-ice group for the pleasant atmosphere, funny board game and christmas events. It is a pleasure to be part of this diverse and vibrant working group. I especially thank my PhD colleagues - Sabine Egerer, Vivienne Groner, Jessica Engels, Johannes Winckler and Leonard Borchert - to name just a few. I won't forget the amazing cycling tours, hikes and short trips that we have done in all seasons. Many thanks also to the loyal table tennis players of this institute.

I further like to hug all my friends for the deep and long-standing friendship that also enriched the past years and will continue to do so. Many thanks also to my flatmates Leonie and Vroni. You were such an important part of my life.

Finally, I am very grateful to the unique nature of the Earth that makes my life so diverse, peaceful and worth living. I very much hope that my work helps to protect this beauty.

CONTENTS

# LIST OF FIGURES

## LIST OF TABLES

# ACRONYMS

| | |
|---|---|
| CMIP5 | Coupled Model Intercomparison Project Phase 5 |
| GMST | global mean surface temperature |
| ICV | internal climate variability |
| IPCC | Intergovernmental Panel on Climate Change |
| PDF | probability distribution function |
| RCP | Representative Concentration Pathway |
| SAT | near-surface air temperature |

# 1 EXPLORING INTERNAL CLIMATE VARIABILITY

*Climate change and climate variability*

Climate science is not just an endeavor born out of scientists' curiosity but serves humankind to understand a basis for its own existence. Scientists therefore intensively study the mean climate state of the Earth and its long-term evolution (IPCC 2014a). However, rather than by the mean climate state, our everyday life is affected by climate variability. This makes us aware that climate variability is at least as important for the habitability on the Earth as the mean climate state and thus fundamentally relevant to understand in addition to the changes in the mean climate (Katz and Brown 1992). These "variations in the mean state on all spatial and temporal scales beyond that of individual weather events" (IPCC 2014b) stem from the redistribution and changes in the amount of energy around the globe. Climate variability results from different sources: either natural or anthropogenic forcing *external* to the Earth's system, or processes *internal* to the Earth's system. While we understand many aspects of the external sources such as variations in solar activity and volcanic eruptions, and the climate variability caused by changes of the orbital parameters of the Earth (e.g., Shindell et al. 2003), we lack fundamental knowledge of the functioning of the internally generated climate variability. The importance of internally generated variability within the climate system has been acknowledged in recent years (e.g., Deser et al. 2012, 2014), but the quantitative role and the physical mechanisms of the interactions of the atmosphere with the ocean and land surfaces are not well understood. In this dissertation, I specifically address the quantitative role and the physical mechanisms of the internal climate variability and its chaotic nature as a key characteristic of the Earth's climate system.

*Sources of climate variability*

*Internal climate variability*

## 1.1 CHAOS IN THE EARTH SYSTEM

*Historical context*

55 years ago, Edward N. Lorenz showed that the climate system is chaotic (Figure 1.1). Based on a system of ordinary differential equations, he was the first who recognized that small differences in a dynamical system could trigger very different results (Lorenz 1963). 42 years ago, Klaus Hasselmann elaborated on the finding by Lorenz. Based on a Statistical Dynamical Model, he showed that climate variability can be clearly separated by timescales into a rapidly varying weather component and a slowly responding climate component (Hasselmann 1976). Together with Claude Frankignoul, Hasselmann further showed in a 2-layer atmosphere-ocean-model that the short-timescale atmospheric forcing can produce large-scale, low-period climate variability (Frankignoul and Hasselmann 1977). Today, we know from a number of other landmark papers that in global climate models these small differences in initial conditions are not only decisive for weather predictions, but also persist

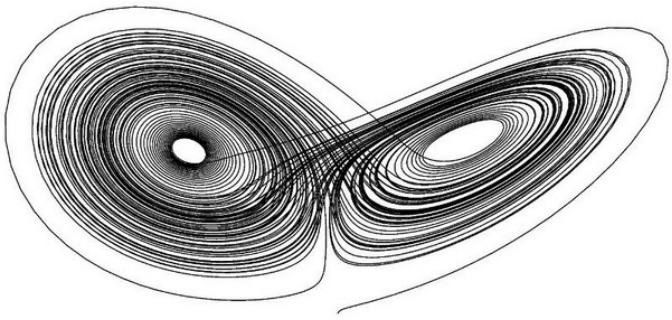


Figure 1.1: The Lorenz attractor, which became famous for the "Butterfly effect". The result differs dependent on the initial conditions.

on longer timescales (e.g., Deser et al. 2012, 2014; Hawkins and Sutton 2009; Knutti and Sedláček 2013; Marotzke and Forster 2015). The relative importance of uncertainty that arises from the initial conditions, from external forcing, and from the representation of the behavior of the climate system differs for different spatial and time scale averages and for different variables (Hawkins and Sutton 2009). Although the sampling uncertainty from initial conditions is especially important on shorter time scales and for smaller scale variables, the projection of any climate variable derived from a single simulation of a climate model is affected by its initial conditions (Collins et al. 2013).

*Chaotic but not random*

I here follow the spirit of Lorenz (1993) who defined chaos by referring to "variations that are not random but look random". Randomness is identical to the absence of determinism, which allows in a sequence of events that everything that can ever happen can happen next. In contrast, the chaotic climate system is a dynamical system that is deterministic but sensitively dependent on the initial conditions. Internal climate variability (ICV), the focus of this dissertation, simply results from this chaotic nature of the climate system with its largest contribution from variations in the atmospheric circulation (e.g., Deser et al. 2014; Wallace et al. 1995). Although chaos is deterministic and hence potentially predictable, the uncertainty from ICV arises because of our incomplete knowledge of the full state of the climate system from which we initialise a model simulation. The climate-modelling community accounts for the uncertainty in the initial conditions by running multiple simulations with a single model with tiny, random perturbations in the initial conditions, or by initialising a single model from atmospheric or oceanic states displaced in time. In the light of these efforts, I outline three scientific obstacles that so far prevented and in part still impede a better understanding of ICV.

## 1.2 OBSTACLES TO OUR UNDERSTANDING OF INTERNAL CLIMATE VARIABILITY

Although the understanding of climate variability is a primary goal of climate science (Hasselmann 1976), the functioning of ICV and its quantitative role within the climate system are still insufficiently understood. I consider the following three obstacles as key components that explain why major knowledge

gaps persist and conflicting findings exist in the three research areas that I address in this dissertation.

### 1.2.1 *Obstacle 1: A lack of robust estimates*

*... from observations*

The first obstacle arises from our inability to robustly infer ICV from observations or from climate-model simulations. At least three reasons inhibit inferring the Earth system's ICV directly from observations. First, observational records are often too short to allow for an estimate that includes variability on decadal or longer timescales. Second, the evolution of any observable is a combination of ICV and an externally forced signal. Properly disentangling the two is challenging especially when both components interact (Kirtman et al. 2013) and when the forced signal changes at a different time frequency than the ICV. Third, observations are not the entire truth. They are often incomplete (e.g., Huang et al. 2017; Karl et al. 2015), of insufficient quality and hence less reliable the further they reach back in time (e.g., Flato et al. 2013; Walsh et al. 2017). Further, the products from observations may differ substantially because of the data post-processing that relies on modelling (e.g., Bunzel et al. 2016). Although these limitations can be reduced, they cannot be entirely eliminated (e.g., Flato et al. 2013).

*... from climate-model simulations*

In contrast to the limited possibility to assess the climate system's internal variability directly from observations, climate models are powerful tools to study ICV and its change through time. Their controlled setup outperforms observations with respect to possible temporal length, spatial coverage, different prescribed forcings and the separation of internally generated and externally forced signals. However, we also struggle to consistently estimate ICV from climate-model simulations. The CMIP5 models simulate a broad range of ICV, which points to their differing representation of ICV (Flato et al. 2013). Consistent estimates of modelled ICV are only available for an unforced state that is commonly estimated from a preindustrial control simulation with constant external forcing (Figure 1.2a for GMST). However, preindustrial control simulations are unsuitable to represent the ICV of simulations with a changing external forcing (e.g., Brown et al. 2017; Kay et al. 2015). Modelled ICV under changing external forcing can now be estimated from multiple ensemble simulations of a single model, enabled by growing computational capacity largely due to parallel computing. Although running multiple simulations of individual models became common in the CMIP5 framework, the low numbers of ensemble simulations for each model are considered too small for a proper assessment of forced changes in ICV over time. Some modelling centers now run large ensembles with a single model, which allows one to clearly separate ICV from the forced signal for different forcing periods (e.g., Banerjee et al. 2017; Frankignoul et al. 2017; Hedemann et al. 2017; Känel et al. 2017; Kay et al. 2015). However, the results from single-model large ensembles are prone to the structural uncertainty of the specific model. While a multi-model approach is desired to increase the sampling of both initial conditions and model properties (Kirtman et al. 2013), running additional large ensem-

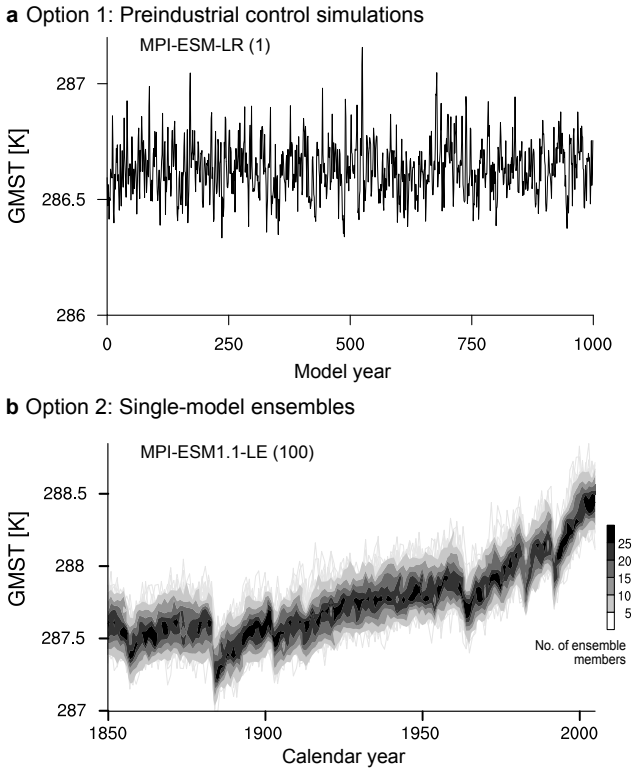



Figure 1.2: Existing options to estimate ICV. **a** A 1000-year long preindustrial control simulation for annual global mean surface temperature (GMST) from MPI-ESM-LR. **b** 100 ensemble simulations for historical annual GMST from MPI-ESM1.1-LE. The temperature range in 1850 in **b** differs from the range in **a** because of the different model versions.

bles with other models consumes substantial computational resources and is hence limited.

*Existing knowledge gaps*

Despite the two complementary options to estimate ICV from either preindustrial control simulations or single-model ensemble simulations, a full quantification of ICV has not been possible. For instance, so far one could not consistently estimate forced changes of internal variability over time, nor robustly evaluate climate-model simulations. I postulate that answering these questions was largely inhibited by the lack of new approaches to sensibly combine the existing information from climate-model simulations.

*A way forward*

In Chapter 2 of this dissertation, I introduce a new method to overcome our inability to infer robust estimates of ICV and its changes through time from climate-model simulations. I argue that multimodel-consensus estimates for different forcing periods can be derived with the tools we have at hand, but require a new way of thinking about variability both in the time domain and in the ensemble domain. I use these estimates for assessing changes of ICV over time and for a robust evaluation of climate-model simulations, but the method could be used more widely for other climate variables and different applications.

### 1.2.2 *Obstacle 2: A lack of attribution of changes*

The second obstacle concerns our inability to robustly assess how and why the probability of climate extremes changes under climate change. Despite the tremendous societal interest and the urgent need from adaptation and planning, robust regional information on changes in the future probability of climate extremes is lacking.

*... with climate-model simulations*

I consider the combination of the following three aspects as the major limitation for a robust assessment and a better understanding of the origin of future climate extremes. First, a precise detrending of the considered timeseries is required, but this is challenging and is usually performed by removing a linear or higher-order polynomial fit (e.g., Lewis and King 2017; Ylhäisi and Räisänen 2014). The presence of trends in the mean of temperature and precipitation however sensitively increases the variance (Huntingford et al. 2013; McKinnon et al. 2016; Rhines and Huybers 2013) and consequently impacts estimates of skewness and kurtosis. The difficulty to robustly separate signal from noise explains why the possible changes in higher-order moments of a distribution and their impact on the probability of climate extremes are rarely examined in the literature. Second, many approaches to investigate changes in climate extremes require assumptions about the underlying distribution and its changes (e.g., Lewis and King 2017; Rahmstorf et al. 2015). The adequacy of such assumptions is, however, questionable especially when applied uniformly to every grid cell (e.g., Cavanaugh et al. 2015; Lewis and King 2017). Third, the small amount of daily model output compared to monthly model output in general, and from high-resolution global climate models in particular impedes robust multi-model analyses. While climate models are our only tool to project climate extremes, the model resolution, process representation and parameterisations of current global climate models are widely considered insufficient for simulating many processes relevant for climate extremes. Therefore, an assessment of future climate extremes should at least not rest on a single model.

*Existing knowledge gaps*

I postulate that the described challenges so far prevented us to make substantial progress in estimating the future probability of climate extremes and in attributing the origin of the projected changes. As a consequence, we lack global estimates on how much the probability of climate extremes is projected to change on a local level and which proportion of such changes can be attributed to a shift in the mean and to a change in higher-order moments such as ICV.

*A way forward*

In Chapter 3 of this dissertation, I address the poor understanding of the origin of the projected changes by quantifying the future probabilities of temperature and precipitation extremes and by attributing them to changes caused by a shift in the mean and changes caused by a change in higher-order moments. Based on a proper removal of trends and an empirical threshold approach that uses daily simulations from multiple models and requires no assumption on the underlying distribution, I present a method for obtaining robust in-

formation on the origin of the future probability of climate extremes within the limits set by the ability of our global climate models to represent such extremes.

### 1.2.3 *Obstacle 3: A lack of technical scope*

The third obstacle originates from our inability to fully grasp the causes of ICV. Simply stating "It's just internal variability" hardly satisfies scientific curiosity and builds barriers to our understanding. Explaining ICV instead requires the understanding of the mechanisms that cause it.

*The power of global climate models*

Global climate models are an indispensable tool for studying the mechanisms that cause ICV, because ICV largely manifests in the interactions between all components of the climate system (Kayano et al. 2005; Mitchell 1976). However, unravelling the physical mechanisms of ICV and their quantitative role in a chaotic system is challenging in global climate models, because the endeavor necessitates a broad technical scope. First, it requires long fully coupled climate-model simulations in high temporal resolution to properly account for the interactions of the climate components on all timescales up to centuries, which generates a tremendous amount of data to be stored and analysed. Second, it requires a sophisticated methodology to quantitatively separate the physical mechanisms that contribute to ICV. The growing computational power largely enabled by the parallel computing in supercomputers and the increased data-storage capacities now allow for unravelling the physical mechanisms of ICV.

*Existing knowledge gaps*

By using the growing computational power, I quantify the underlying mechanisms that cause the internal variability of Arctic sea-ice area. The scientific community commonly believes that feedback mechanisms are an essential cause for sea-ice variability. The ice-albedo feedback that describes the increased warming and hence ice melting caused by a lower reflectivity of open ocean compared to ice is considered as a key driver. Other proposed mechanisms that cause Arctic sea-ice variability are the cloud feedback, the water vapour feedback, the forcing by surface winds and the meridional oceanic heat transport. We are certain about the existence of such radiative effects, feedbacks and forcings, which is well reasoned by physical laws and conceptual understanding. However, we know much less about the quantitative relevance of such processes for Arctic sea-ice variability. Previous studies assessed individual drivers, in part limited to certain regions within the Arctic Ocean, and quantified their role based on correlations to the sea-ice area. I here propose that such regional viewpoints and the lack of powerful modelling tools inhibited us so far from obtaining a consistent picture on the driving mechanisms of Arctic sea-ice variability which prevents a more complete understanding of the Arctic climate system.

*A way forward*

In Chapter 4 of this dissertation, I quantify the effect of the known physical mechanisms on Arctic sea-ice variability to provide consistent insights into the processes that primarily cause the variability. In contrast to previous

studies, I choose an Arctic-wide perspective and quantify the sea-ice variability by switching the proposed mechanisms one-by-one non-interactive within the fully coupled climate system. I hereby aim to uncover the main mechanisms that cause the Arctic sea-ice area to vary in size from one year to the next.

In the following three chapters, I specifically address the outlined three obstacles and provide pathways to enable the quantification, the handling and the understanding of ICV on our changing planet.

## 1.3 RESEARCH QUESTIONS

*The important thing is not to stop questioning.*
*Curiosity has its own reason for existing.*

*— Albert Einstein*

---

I here outline eight guiding research questions to be answered in this dissertation. I focus on the internal variability of sea ice and SAT, but my methods can be applied widely to other climate variables.

*Research questions on changes in ICV*

In Chapter 2, I ask:

a. **How can ICV be consistently estimated for a changing forcing?**

b. **How does the internal variability of SAT, sea-ice area and sea-ice volume change under climate change?**

c. **How can climate-model simulations be robustly evaluated?**

Chapter 2 rests on work that I published jointly with Dirk Notz (Olonscheck and Notz 2017). The study has been slightly extended and adapted to fit the structure of this dissertation.

*Research questions on future climate extremes*

In Chapter 3, I ask:

a. **How are the background conditions for climate extremes, i.e. the mean and the internal variability of daily SAT and daily total precipitation, projected to change?**

b. **What is the future probability of climate extremes under strong global warming?**

c. **How much of the future probability of climate extremes can be attributed to a shift in the mean, and how much to changes in higher-order moments?**

Chapter 3 presents results from a paper in preparation.

*Research questions on the drivers of Arctic sea-ice variability*

In Chapter 4, I ask:

**a. Which mechanisms primarily drive the variability in Arctic sea-ice area?**

**b. What does the relative role of driving mechanisms imply for the functioning of the Arctic climate system?**

Chapter 4 is based on a study that I wrote with co-authors and that is currently under way for resubmission. The study has been adapted to fit the format of this dissertation.

In this chapter, I presented three obstacles that so far inhibited a better quantification and understanding of ICV within the climate system. To overcome the first obstacle, I develop a method for consistently estimating changes in ICV over time. I introduce and apply this method in the next chapter.

# 2

# CONSISTENTLY ESTIMATING INTERNAL CLIMATE VARIABILITY FROM CLIMATE-MODEL SIMULATIONS

## 2.1 SUMMARY

This chapter introduces and applies a new method to consistently estimate internal climate variability (ICV) for all models within a multi-model ensemble. The method regresses each model's estimate of ICV from the preindustrial control simulation on the variability derived from a model's ensemble simulations, thus providing practical evidence of the quasi-ergodic assumption. The method allows one to test in a multi-model consensus view how the internal variability of a variable changes for different forcing scenarios. Applying the method to the CMIP5 model ensemble shows that the internal variability of globally-averaged SAT remains largely unchanged for historical simulations and might decrease for future simulations with a large $CO_2$ forcing. Regionally, the projected changes reveal likely increases in temperature variability in the tropics, subtropics, and polar regions, and extremely likely decreases in mid latitudes. Applying the method to sea-ice volume and area shows that their respective internal variability likely or extremely likely decreases proportionally to their mean state, except for Arctic sea-ice area, which shows no consistent change across models. For the evaluation of CMIP5 simulations of Arctic and Antarctic sea ice, the method confirms that internal variability can explain most of the models' deviation from observed trends but often not the models' deviation from the observed mean states. The new method benefits from a large number of models and long preindustrial control simulations, but it requires only a small number of ensemble simulations. The method allows for consistent consideration of ICV in multi-model studies and thus fosters understanding of the role of ICV in a changing climate.

## 2.2 INTRODUCTION

*Relevance of internal climate variability*

Internal variability of the climate system, caused by the system's chaotic nature, limits the predictability of climate (e.g., Deser et al. 2014) and represents a major source of uncertainty for climate projections (e.g., Deser et al. 2012; Hawkins and Sutton 2009, 2011; Swart et al. 2015). Knowledge of ICV is a prerequisite for climate-change attribution (e.g., Marotzke and Forster 2015; Swanson et al. 2009; Trenberth 2011) and climate-model evaluation (e.g., Flato et al. 2013; Notz 2015; Stroeve et al. 2014). However, robustly quantifying ICV in climate studies remains challenging. Here I examine how ICV in global climate models estimated from preindustrial climate simulations relates to ICV estimated from the ensemble spread of historical and future cli-

mate simulations. This allows me to develop a new method to estimate ICV for individual model simulations, which I apply for assessing changes in ICV over time and for evaluating climate-model simulations.

*Control-simulation approach*

The magnitude of the ICV of climate-model simulations is usually estimated by using one of two different approaches (Collins et al. 2013). The first approach, here called "control-simulation approach," is based on the analysis of preindustrial control simulations with constant external forcing (for diverse applications see for example Huber and Knutti 2014; Palmer and McNeall 2014; Resplandy et al. 2015; Roberts et al. 2015; Schindler et al. 2015; Schneider and Kinter 1994; Swanson et al. 2009). Apart from model drift, any climate variability of the preindustrial control simulations is ICV. Preindustrial control simulations, typically spanning many centuries, are commonly sufficiently long to also include multidecadal and longer-term ICV. They are usually available for any climate model and are, for example, part of the entrance criteria for a model to participate in phase 6 of the Coupled Model Intercomparison Project (Meehl et al. 2014). However, the control-simulation approach is commonly considered unsuitable for representing the ICV of simulations with a different or changing external forcing (e.g., Kay et al. 2015).

*Ensemble-spread approach*

The second approach, here called "ensemble-spread approach," addresses this possible limitation (e.g., Deser et al. 2012, 2014; Wettstein and Deser 2014). The ensemble-spread approach is based on ensemble simulations with slightly different initial conditions, with each realisation subject to the same external forcing. The ensemble spread between these different realisations of a single model measures the ICV for different forcing scenarios. Some modelling groups run large ensembles of a single model to disentangle the internally and externally forced contributions in a simulation, because a sufficiently high number of ensemble simulations is required to estimate the model's total ensemble spread. However, running multiple realisations with any given global climate model consumes substantial computational power. As a consequence, many modelling groups provide only a single realisation or a small number of ensemble simulations. This inhibits a robust and consistent estimation of model-specific ICV for different forcing scenarios for a given multi-model ensemble.

*Combining both approaches*

I here address this common problem of multi-model studies by examining the relationship between the estimate of preindustrial ICV from the control-simulation approach and the estimate of historical or future ICV from the ensemble-spread approach. I expect similarities between both estimates of ICV following the quasi-ergodic assumption, which states that the variance of one sequence of events over time equals the ensemble variance at a given time (e.g., Hingray and Said 2014; Neumann 1932). However, changes in external forcing might alter the internal variability of a climate variable over time (Lu et al. 2014; Sutton et al. 2015). By relating both approaches across a multi-model ensemble, I derive estimates of ICV for different forcings of each given model.

*Application to CMIP5*

I apply my method to the model ensemble of phase 5 of the Coupled Model Intercomparison Project (CMIP5) for SAT, sea-ice volume and sea-ice area (i)

to investigate whether the magnitude of internal variability changes over time and (ii) to robustly evaluate CMIP5 sea-ice simulations.

*Assessing changes of internal variability over time*

Possible changes of internal variability over time were investigated by Huntingford et al. (2013), Thompson et al. (2015) and Holmes et al. (2016). Huntingford et al. (2013) used model output from 17 CMIP5 models to investigate the time-evolving global temperature variability. They examined 11- and 31-year detrended ensemble-mean historical and RCP8.5 simulations, and found that so far, the globally averaged temperature variability has been stable but is projected to decrease in future. Holmes et al. (2016) analysed future changes in winter and summer temperature variability in a 17-member ensemble from a global climate model forced by the SRES A1B emission scenario. They removed 40-year linear trends in ensemble-mean temperature and found strong regional changes in temperature variability. In contrast to these previous approaches, mine does not require the removal of any trend and thus allows for a clear separation of ICV from external forcing based on ensemble simulations of individual models. Thompson et al. (2015) estimated the uncertainty in projections of future climate trends arising from ICV. Using an analytic model that requires a time-stationary standard deviation for different forcing scenarios, they related the statistics of the preindustrial control simulation to the spread of trends in the 40-member ensemble of the global climate model CCSM3. They concluded that for most regions, the preindustrial control simulation is sufficient to represent the ICV derived from the CCSM3 large ensemble. Based on my multi-model approach, I derive estimates of ICV for different forcing scenarios for all CMIP5 models; my estimates are largely independent of the ensemble size of a single model. Further, my approach does not require the standard deviation of a variable to be stationary in time and thus can be used for variables whose standard deviation changes for different forcing scenarios.

*Evaluating sea-ice simulations*

For the evaluation of CMIP5 simulations, I limit my analysis to sea ice. Recent studies evaluated CMIP5 sea-ice simulations of Arctic sea-ice extent (e.g., Flato et al. 2013; Massonnet et al. 2012; Notz 2014; Stroeve et al. 2012), Antarctic sea-ice extent (e.g., Mahlstein et al. 2013; Zunz et al. 2013) as well as Arctic sea-ice thickness and volume (Shu et al. 2015; Stroeve et al. 2014). All of these studies stressed the large influence of internal variability. The influence of internal variability on Arctic sea-ice trends was specifically estimated by Swart et al. (2015), who investigated both the CMIP5 models and the 30-member ensemble of the global climate model CESM1. They concluded that internal variability must be carefully accounted for when evaluating sea-ice simulations, which was also spelled out in a dedicated study by Notz (2015). Nevertheless, model-specific estimates of internal variability for the satellite period (1979–today) across the CMIP5 model ensemble do not exist yet. My method now provides such estimates and consistently considers internal variability in an evaluation of the CMIP5 sea-ice simulations.

Section 2 introduces my method to robustly estimate ICV for different forcing scenarios across models. Section 3 presents the data used. Sections 4 and 5 demonstrate usage of this method by applying it to specific climate ob-

servables, namely, annual SAT, sea-ice volume, and sea-ice area. Section 6 summarizes my findings.

## 2.3 METHOD AND APPLICATIONS

### 2.3.1 *Method*

I regress the standard deviation from the control-simulation approach on that of the ensemble-spread approach. I specifically explain here how this regression allows me to derive a model-specific estimate of ICV across models for simulations with different forcing scenarios. My method consists of four steps:

1. I calculate the standard deviation of the preindustrial control simulation for each model.
2. I calculate the ensemble standard deviation for each model that provides ensemble simulations.
3. I regress the estimates from step 1 onto those from step 2 and fit a regression line through these estimates. I use this basic version of my method to assess changes of ICV over time.
4. If the regression can robustly be determined, I use it in a final step to translate the estimates of the preindustrial standard deviations to estimates corresponding to different forcing scenarios for models with a single simulation. This extended version of my method gives us consistent estimates of ICV for all models.

I now explain these four steps in detail.

#### 2.3.1.1 *Control-simulation approach*

I calculate the standard deviation $\sigma_{piC}$ for a variable $x$ from the preindustrial control simulation of length $T$ of each model across each output interval of interest $t$ (e.g., year):

$$\sigma_{piC}(T) = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(\overline{x} - x_t)^2} \tag{2.1}$$

*Removing model drift*

However, for two reasons I cannot obtain a reliable estimate of ICV from directly applying this control-simulation approach to all simulations. First, some preindustrial control simulations still drift substantially, likely because the models are at the beginning of the control simulation not yet in equilibrium with the preindustrial forcing (e.g., Frankcombe et al. 2015; Knutson et al. 2013). I hence remove the least squares linear trend from each model's preindustrial control simulation to minimize model drift. Second, because of multidecadal and longer-term ICV, the standard deviation also depends on

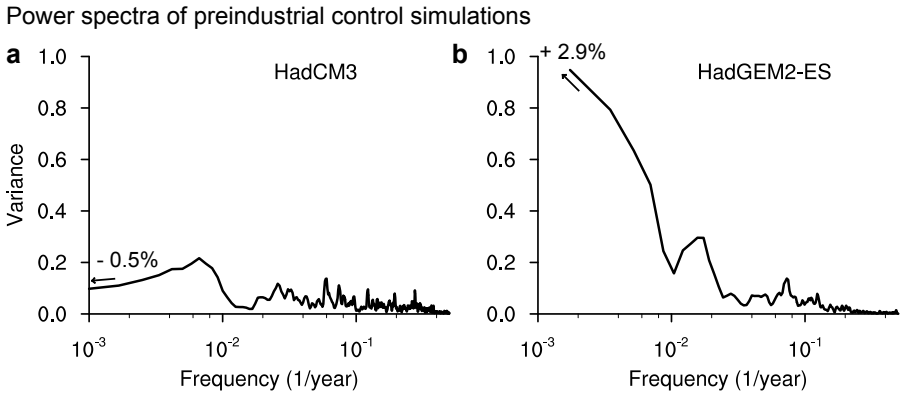



Figure 2.1: Logarithmic power spectra for **a** a preindustrial control simulation that is sufficiently long (HadCM3) and **b** a preindustrial control simulation that is too short (HadGEM2-ES) to represent the model's total internal variability of annual GMST. The number indicates the change in variance at the last time-equivalent time step of the spectral distribution of variability. Note that preindustrial control simulations shorter than the one of HadGEM2-ES can be fully adequate because the suitability depends on the manifestation of ICV in a model and on the variable analysed.

*Testing for the sufficiency of the simulation length*

the length T of the control simulation, which strongly varies among models (cf. Table 2.1). To test whether the control simulations of all models are sufficiently long to largely cover the model's total ICV, I analyse their spectral distribution of variability. The logarithmic power spectra reveal that for some models the maximum spectral power of the control simulation does still considerably increase at the time scale provided by the length of the control simulation (see Figure 2.1 for HadCM3 compared to HadGEM2-ES). I here consider an increase in the spectral power at the time scale of the length T of the control simulation as considerable when the control simulation's variance increases by more than 1% of the total variance for the last time-equivalent time step $\Delta(T-(T-1_{equiv}))$ of the spectral distribution of variability. I choose this criterion that still allows for slight increases in variance to not exclude too many models. When a model provides too short a control simulation, I assume that the ICV as given by the control simulation is not the total ICV of this particular model. I here disregard such models.

#### 2.3.1.2 *Ensemble-spread approach*

I calculate the ensemble standard deviation $\sigma_{ens}$ for a variable $x$ as the square root of the ensemble variance across the different ensemble simulations $n$ of a model with N ensemble simulations for each output interval of $t$ (e.g., each year) averaged over the simulation length T:

$$\sigma_{ens}(N,T) = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left[\frac{1}{N-1}\sum_{n=1}^{N}(\overline{x_t}-x_{n,t})^2\right]} \tag{2.2}$$

*Advantages of the ensemble-spread approach*

The ensemble-spread approach also includes long-term ICV, because calculating the ensemble standard deviation does not require one to remove any model trend. The approach avoids possible underestimates of a model's total ICV due to small numbers of ensemble simulations because it benefits

from a largely increased sample size caused by considering the estimates of ensemble variance at every output interval. For example, N = 3 ensemble simulations with T = 150 years correspond to a simulation with T = 450 years. Hence, the ensemble-spread approach then relies on a sample size of 450, which is similar to the standard length of a control simulation. To nevertheless test the sensitivity of the ensemble standard deviations to very small numbers of available ensemble simulations, I assess the range of ensemble standard deviations from any pairwise combination of a model's ensemble simulations (i.e., N = 2; see Section 2.5). The upper bound of this range is given by the two ensemble members with the most contrarily temporal evolution, while the lower bound results from the two ensemble members that are most similar. I use this range to estimate the uncertainty of the ensemble standard deviation.

*Testing for the sufficiency of small ensemble sizes*

To further assess the suitability of the approach for small numbers of ensemble simulations, I use the 100-member ensemble of the Max Planck Institute Earth System Model (MPI-ESM1.1-LE) in low resolution covering the period 1850–2005 (see Section 2.4). To examine the trade-off between the required number of ensemble members and the necessary time-averaging length for robust estimates of GMST variability, I calculate the respective ensemble standard deviation from all possible consecutive combinations of the 100 ensemble simulations for varying ensemble sizes and time-averaging lengths (Figure 2.2). I find that ensemble sizes as small as three are representative for the best estimate of the model-simulated variability (horizontal line) at least for

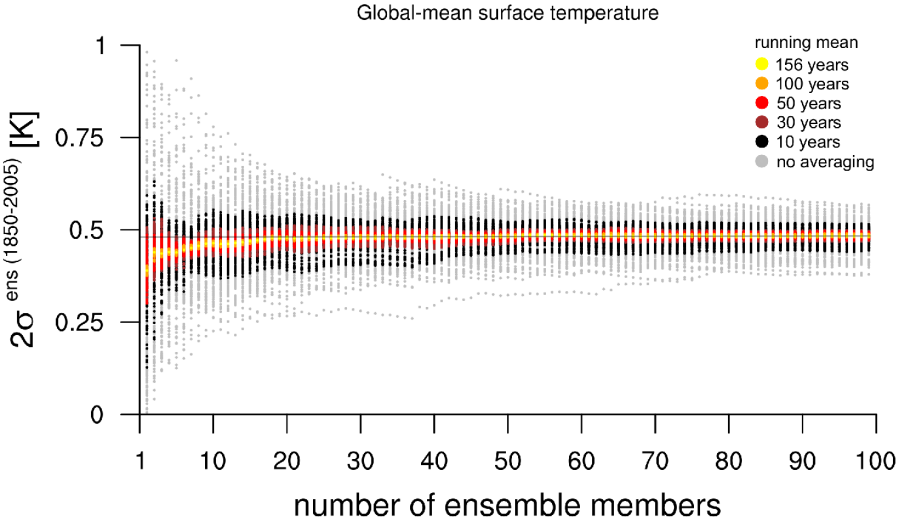



Figure 2.2: Ensemble standard deviation of GMST calculated for varying ensemble size and time-averaging length from all possible consecutive combinations of the 100 historical simulations from MPI-ESM1.1-LE. The horizontal line marks the best estimate of simulated variability inferred from all 100 ensemble simulations and the full time-averaging length. The spread of estimates on the right tail is conservative because the larger the ensemble sizes, the more the combinations suffer from resampling of ensemble members. The estimates differ from the one in Figure 2.4a because I here show the GMST variability instead of the globally-averaged SAT variability at every grid cell.

averaging periods of 100 years or longer. I further find that this representativity of small ensemble sizes for long averaging periods is not only true for GMST variability, but also holds on a grid-cell level (see Figure A.1 for the grid cell Hamburg).

*Role of volcanic forcing*

The ensemble-spread approach purely samples ICV and circumvents uncertainties due to inconsistent natural external forcings over time in CMIP5. While preindustrial and future RCP8.5 simulations generally do not include natural external forcing from volcanic eruptions, historical simulations do contain volcanic eruptions (e.g., Santer et al. 2014). Since volcanic eruptions considerably contribute to the variability of climate variables, such as GMST (e.g., Bradley and Jones 1992; Briffa et al. 1998) and sea-ice area (Rosenblum and Eisenman 2016), this would bias estimates of preindustrial and future natural variability low compared to estimates of historical natural variability. However, for my estimates of historical ICV derived from the ensemble-spread approach, each ensemble member has experienced the same volcanic forcing. Apart from the minor effects of synchronized ensemble spread in years of volcanic eruptions, my ensemble standard deviation is thus independent from volcanic forcing.

#### 2.3.1.3 *Basic version of the method*

Depending on the specific application, my method can be used in two versions, here called "basic version" and "extended version." To present both versions of the method, I use a fictitious example as sketched in Figure 2.3.

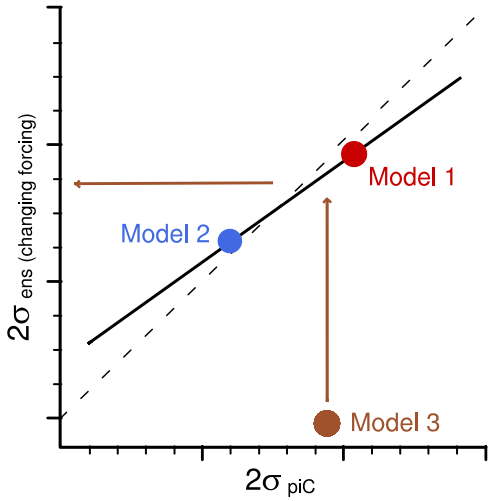



Figure 2.3: Schematic view of the method for estimating ICV for different forcing scenarios. The basic version of the method regresses the estimate of ICV derived from the preindustrial control simulation of a model (x axis) on the ensemble standard deviation of models with ensemble simulations such as models 1 and 2 (y axis). The unity line as a reference is indicated by the dashed black line. For the extended version, a constructed ensemble standard deviation can be derived for models with a single simulation (model 3) using the regression line through models 1 and 2. The extended version requires a consistent response of the models with ensemble simulations.

*... for models with ensemble simulations*

For the basic version, I simply relate the estimate of ICV from the control-simulation approach ($2\sigma_{\mathrm{piC}}$, x axis) to that of the ensemble-spread approach ($2\sigma_{\mathrm{ens}}$, y axis). This can be done for a single model that provides ensemble simulations, such as model 1, but also across all models of a given multi-model ensemble that provide ensemble simulations by linear least squares regression, such as model 1 and model 2. This basic version allows one to directly compare the preindustrial ICV to the ICV for different forcing scenarios for models with ensemble simulations. I exploit the multi-model relationship of the basic version to test for a change of ICV over time in application 1.

#### 2.3.1.4 *Extended version of the method*

*... for models with a single simulation only*

The extended version builds on the basic version. By using the regression line through the estimates of ICV of model 1 and model 2, I derive an ensemble standard deviation for models with a single simulation from the model's preindustrial estimate of ICV (model 3, situated at the zero line of the ensemble standard deviation). This procedure is based on the key assumption that a relationship found for many models is valid for other models as well. I justify the assumption by the underlying theory of quasi ergodicity that I show to hold for the models with ensemble simulations. To be applicable to a variable, the procedure requires a similar response from models with ensemble simulations for this variable. If this prerequisite is fulfilled, then the procedure allows one to circumvent limitations associated with models that have a single realisation only. I keep the original estimates for models with ensemble simulations and do not adjust their standard deviations to the regression line. I use these original estimates of ICV for models with ensemble simulations and the derived estimates from the extended version of my method for models with a single simulation to consistently evaluate sea-ice simulations from all models in application 2.

### 2.3.2 *Applications*

*Defining the likelihood of change in ICV*

To obtain a consensus estimate of the projected direction and magnitude of a possible change in ICV over time (application 1), I evaluate the regression line obtained in the simple version of my method at the location of the multi-model mean. To examine the likelihood of an identified change, I test whether the confidence interval of the regression line includes the unity line. Following the IPCC terminology, I define a change as likely when the 66% confidence interval of the regression line does not include the unity line and as extremely likely when the 95% confidence interval of the regression line does not include the unity line. In contrast, I define a change as possible when the 66% confidence interval of the regression line does include the unity line.

*Testing for the plausibility of simulations*

To evaluate the CMIP5 sea-ice simulations (application 2), I consider both the simulated ICV $\sigma_{\mathrm{mod}}$ that I derive from the extended version of my method and the observational or reanalysis uncertainty $\delta_{\mathrm{ref}}$ (see Appendix A). To combine both sources of uncertainty, I follow an approach of Santer et al.

(2008) that was adopted by Stroeve et al. (2012). This approach uses a plausibility variable as a measure of model fidelity,

$$\phi = \frac{\overline{\mathrm{mod}} - \overline{\mathrm{ref}}}{\sqrt{\sigma^2_{\mathrm{mod}} + \delta^2_{\mathrm{ref}}}} \tag{2.3}$$

that weights the distance between any time-averaged CMIP5 model simulation ($\overline{\mathrm{mod}}$) and the time-averaged reference data ($\overline{\mathrm{ref}}$) by the internal variability of the simulations and the observational or reanalysis uncertainty. The plausibility variable $\phi$ thus quantifies how far the model deviates from the reference data in units of the associated quantity- and model-specific uncertainty.

## 2.4 DATA

*CMIP5 simulations*

To demonstrate the usage of my method, I apply it to the internal variability of annual SAT and of Arctic and Antarctic sea-ice volume and sea-ice area as simulated by the models that took part in CMIP5 (Taylor et al. 2012). These simulations are available from the Earth System Grid data portal of the Earth System Grid Federation (http://esgf-node.llnl.gov/search/cmip5/) and from the data portal from the Centre for Environmental Data Analysis (ftp://ftp.ceda.ac.uk/badc/cmip5/). For both SAT and sea-ice volume and area, I analyse the preindustrial control simulation, the historical simulations (1850–2005), and future simulations (2006–2100) driven by the RCP8.5 scenario (Moss et al. 2010) of each CMIP5 model.

*SAT*

For SAT, I analyse gridded monthly mean data of 145 historical simulations from 44 different climate models and 87 RCP8.5 simulations from 40 different climate models (Table 2.1). To account for regional changes in temperature variability across models, I regrid all CMIP5 simulations by bilinear interpolation on a grid resolution of 1.8947° x 3.75°. I use the regridded data to calculate annual globally-averaged SAT by weighting the SAT with the area of the model grid cells and then averaging annually and globally. The globally-averaged SAT variability is derived by first calculating the ensemble standard deviation at every grid cell and by then averaging globally.

*Sea-ice concentration and thickness*

For sea ice, I analyse 145 historical simulations from 44 different climate models and 88 RCP8.5 simulations from 40 different climate models that provide gridded monthly mean data of sea-ice concentration and sea-ice thickness (Table 2.1). From these, sea-ice area is calculated by multiplying the area of the model grid cells with their sea-ice concentration, which is then added up over all grid cells for either the Northern or the Southern Hemisphere. The model output of sea-ice thickness is the equivalent thickness averaged over the grid cell assuming that the grid cell is entirely ice covered. Sea ice volume is calculated as the product of the area of the model grid cells and their equivalent sea-ice thickness, which again is added up over all grid cells for both hemispheres. For reasons explained by Notz (2014), such as differences in grid geometry and misleading results with respect to model quality as a result of synthetic biases in sea-ice extent, I focus on the more direct and more physical metric sea-ice area instead of sea-ice extent.

Table 2.1: CMIP5 simulations and single-model large ensembles used.

| Model name | Control run length T | # of SAT simulations | | # of sea-ice simulations | | |
|---|---|---|---|---|---|---|
| | [years] | historical | RCP8.5 | historical | RCP8.5 | extended to 2014 |
| ACCESS1-0 | 500 | 2 | 1 | 3 | 1 | 1 |
| ACCESS1-3 | 500 | 3 | 1 | 3 | 1 | 3 |
| bcc-csm1-1 | 500 | 3 | 1 | 3 | 1 | 1 |
| bcc-csm1-1-m | 500 | 3 | 1 | 3 | 1 | 1 |
| BNU-ESM | 559 | 1 | 1 | 1 | 1 | 1 |
| CanESM2 | 1096 | 5 | 5 | 5 | 5 | 5 |
| CCSM4 | 501 | 6 | 6 | 6 | 6 | 6 |
| CESM1-BGC | 500 | 1 | 1 | 1 | 1 | 1 |
| CESM1-CAM5 | 319 | 3 | 3 | 3 | 3 | 3 |
| CESM1-FASTCHEM | 222 | 3 | - | 3 | - | - |
| CESM1-WACCM | 200 | 1 | 3 | 1 | 3 | - |
| CMCC-CESM | 277 | 1 | 1 | 1 | 1 | 1 |
| CMCC-CM | 330 | 1 | 1 | 1 | 1 | 1 |
| CMCC-CMS | 500 | 1 | 1 | 1 | 1 | 1 |
| CNRM-CM5 | 850 | 10 | 5 | 10 | 5 | 5 |
| CNRM-CM5-2 | 150 | 1 | - | 1 | - | - |
| CSIRO-Mk3-6-0 | 500 | 10 | 10 | 10 | 10 | 10 |
| EC-EARTH | 452 | 8 | 8 | 10 | 10 | 10 |
| FGOALS-g2 | 700 | 4 | 1 | 4 | 1 | 1 |
| FIO-ESM | 800 | 3 | 3 | 3 | 3 | 3 |
| GFDL-CM3 | 500 | 5* | 1 | 5* | 1 | 5* |
| GFDL-ESM2G | 500 | 3* | 1 | 1* | 1 | 1* |
| GFDL-ESM2M | 500 | 1* | 1 | 1* | 1 | 1* |
| GISS-E2-H | 780 | 6 | 1 | 6 | 1 | 5 |
| GISS-E2-H-CC | 251 | 1 | 1 | 1 | 1 | 1 |
| GISS-E2-R | 850 | 6 | 1 | 6 | 1 | 6 |
| GISS-E2-R-CC | 251 | 1 | 1 | 1 | 1 | 1 |
| HadCM3 | 1200 | 10* | - | 10* | - | 10* |
| HadGEM2-CC | 240 | 1* | 3 | 1* | 3 | 1* |
| HadGEM2-ES | 576 | 5* | 4 | 4* | 4 | 4* |
| inmcm4 | 500 | 1 | 1 | 1 | 1 | 1 |
| IPSL-CM5A-LR | 1000 | 6 | 4 | 6 | 4 | 4 |
| IPSL-CM5A-MR | 300 | 3 | 1 | 3 | 1 | 3 |
| IPSL-CM5B-LR | 300 | 1 | 1 | 1 | 1 | 1 |
| MIROC5 | 670 | 5 | 3 | 5 | 3 | 5 |
| MIROC-ESM | 630 | 3 | 1 | 3 | 1 | 3 |
| MIROC-ESM-CHEM | 255 | 1 | 1 | 1 | 1 | 1 |
| MPI-ESM-LR | 1000 | 3 | 3 | 3 | 3 | 3 |
| MPI-ESM-MR | 1000 | 3 | 1 | 3 | 1 | 3 |
| MPI-ESM-P | 1156 | 2 | - | 2 | - | - |
| MRI-CGCM3 | 500 | 3 | 1 | 3 | 1 | 1 |
| MRI-ESM1 | 250 | 1 | 1 | 1 | 1 | 1 |
| NorESM1-M | 501 | 3 | 1 | 3 | 1 | 3 |
| NorESM1-ME | 252 | 1 | 1 | 1 | 1 | 1 |
| CESM1-CAM5-BGC-LE | 1800 | 35** | 35 | 35** | 35 | - |
| MPI-ESM1.1-LE | 2000 | 100 | - | 100 | - | - |

* historical simulations that start in year 1860 (**1920) only

*Single-model large ensembles*

To evaluate the applicability of small ensemble sizes of individual CMIP5 models, I additionally analyse single-model large ensemble simulations of updated versions of two models that are part of CMIP5: the 35-member ensemble of the NCAR Community Earth System Model (CESM1-CAM5-BGC) covering the period 1920–2100 with 2006–2100 forced by RCP8.5 (Kay et al. 2015) and the 100-member ensemble of the Max Planck Institute Earth System Model (MPI-ESM1.1-LE) in low resolution covering the period 1850–2005.

## 2.5 APPLICATION 1: ASSESSING CHANGES OF INTERNAL CLIMATE VARIABILITY OVER TIME

### 2.5.1 *Surface air temperature*

*Global-mean change*

I first apply the basic version of my method to annual SAT (Figure 2.4). For the historical simulations, plotting the global-mean estimates of internal variability from the control-simulation approach ($2\sigma_{\mathrm{piC}}$, x axis) against that from the ensemble-spread approach ($2\sigma_{\mathrm{ens}}$, y axis) across the CMIP5 models results in a linear one-to-one relationship (black regression line in Figure 2.4a). Model estimates of ensemble standard deviation that deviate from the one-to-one relationship are usually very uncertain as shown by the test on how

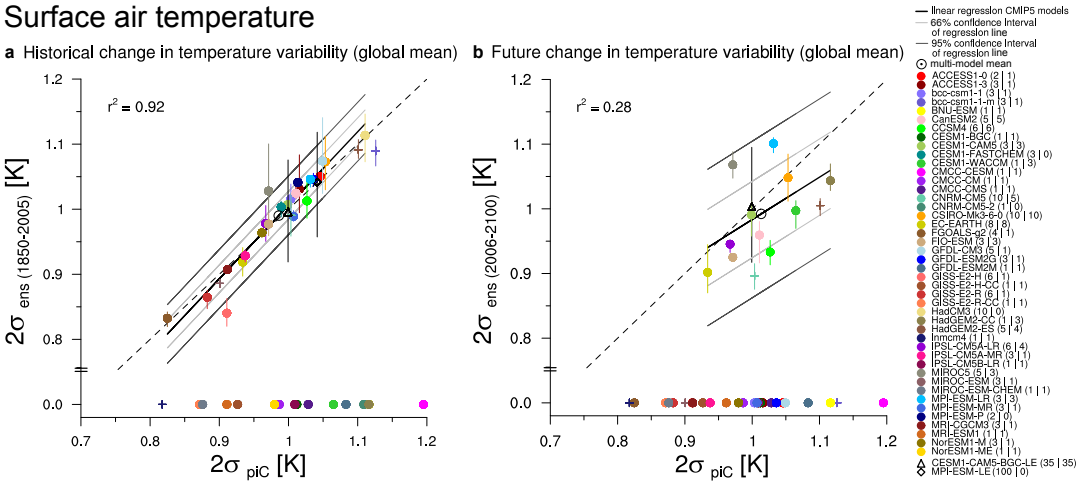



Figure 2.4: Relationship between the standard deviation of each CMIP5 model preindustrial control simulation (x axis) and the ensemble standard deviation of the corresponding **a** historical simulations and **b** RCP8.5 scenario runs (y axis) for annual globally-averaged SAT. CMIP5 models that provide ensemble simulations (colored filled dots with nonzero ensemble standard deviation) are considered for calculating the regression line, the $r^2$, and the 66% and the 95% confidence intervals, and for calculating the multi-model mean (black circled dot). CMIP5 models that have a too short preindustrial control-simulation length to cover their total ICV are shown as + symbols. Models with a single simulation are situated at the zero line of the ensemble standard deviation. Large ensembles are denoted by a triangle or a diamond. The number of ensemble simulations used to calculate the ensemble standard deviation is given in parentheses first for the historical simulations and second for the RCP8.5 scenario runs. Uncertainty as a result of the different numbers of ensemble simulations is displayed by assessing the ensemble standard deviation from any pairwise combination of a model's ensemble simulations (vertical bars).

representative these estimates are for a model's total ensemble standard deviation described in Section 2.3.1 (see vertical bars in Figure 2.4a,b). I have additional support for the one-to-one relationship from the two models that provide large ensemble simulations, CESM1-CAM5-BGC and MPI-ESM1.1-LE (black triangle and diamond, respectively). I thus detect no robust change in time-averaged internal variability of annual globally-averaged SAT between the preindustrial and the historical period across the CMIP5 models.
For the future simulations forced by RCP8.5, I detect a possible decrease in time-averaged internal variability of annual globally-averaged SAT compared to the preindustrial climate (Figure 2.4b).

*Regional changes*

To examine the direction and relative magnitude of regional changes in SAT variability, I use my method to calculate the ratio $\sigma_{ens}/\sigma_{piC}$ averaged for all models for every grid cell (colored patterns in Figure 2.5a,b). For the historical simulations, I find weak regional changes in SAT variability (Figure 2.5a). While tropical regions show a possibly (no stippling) increased SAT variability

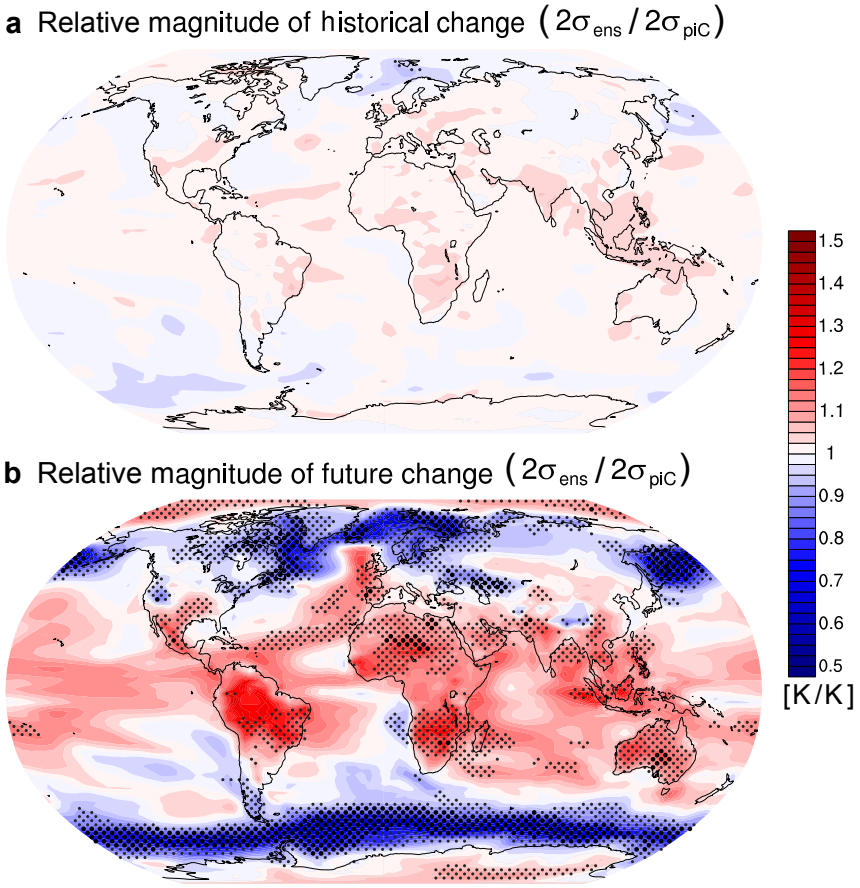



Figure 2.5: Regional changes in variability of SAT. **a** Magnitude of relative change between the preindustrial SAT variability and the historical SAT variability, and **b** the magnitude of relative change between the preindustrial SAT variability and the SAT variability of a future climate forced by the RCP8.5 emission scenario. A possible increase (red shades) and a possible decrease (blue shades) in SAT variability, and likely changes (light stippling) and extremely likely changes (strong stippling) are shown.

compared to the preindustrial climate (red), many mid-latitude regions show a possibly decreased SAT variability (blue). The pattern of SAT-variability change for the historical period already depicts the pattern of change projected for the future.

For the future simulations forced by RCP8.5, I detect strong regional changes in SAT variability (Figure 2.5b). Whereas many tropical, subtropical, and polar regions show a likely (light stippling) or possibly increased SAT variability, many mid-latitude regions show a likely decreased SAT variability. Many mid- to high-latitude oceans show extremely likely changes (strong stippling). The mid- to high-latitude oceans also show the strongest absolute changes in SAT variability (Figure A.2b). These pronounced decreases over mid- to high-latitude oceans can be explained by the strong sea-ice retreat under large $CO_2$ forcing because the past strong interannual variability in the sea-ice cover is replaced by open oceans every year.

*Comparison to previous studies*

My findings of a globally-averaged stable internal variability of SAT for the time-averaged historical period and a possibly decreased internal variability of SAT for the time-averaged future climate forced by RCP8.5 agree with the result by Huntingford et al. (2013) that so far the variability of GMST has not changed but is projected to decrease in the future. They are also in line with the single-model result by Hawkins et al. (2016), who found a decreased GMST variability of about 10% for idealized initial condition ensembles of a global climate model forced by a 1% $CO_2$ increase per year.

Regionally, my findings confirm the result by Huntingford et al. (2013) of increased variability in regions of low variation and decreased variability in regions of high variation. My results also agree with the pattern of future temperature-variability change found by Holmes et al. (2016, their Figures 2 c,d and 4 c,d) and are further consistent with Screen (2014) and Schneider et al. (2015). They all argue that polar amplification decreases the temperature variability in Northern Hemisphere mid latitudes, because northerly winds are warming more rapidly than southerly winds, especially in winter. However, my multi-model result often identifies likely and possible changes because the model estimates differ widely in the direction and/or magnitude of regional changes.

In contrast to these previous studies that show regional changes in future SAT variability, Thompson et al. (2015) found a time-stationary internal variability. They show that for most regions, the preindustrial control simulation is sufficient to represent the future internal variability derived from the CCSM3 large ensemble. Kay et al. (2015) support this independence of regional internal variability from external forcing based on 34-year trends in winter SAT from CESM1-CAM5-BGC-LE. These inconsistent results, which are mainly based on ensembles of individual models, demand a multi-model approach. My multi-model approach discloses often very different model estimates of projected SAT-variability change that hinders a robust projection of changes in many regions of the globe. My approach therefore cautions one to use only individual models for assessing future changes in SAT variability and now allows for the interpretation of single-model results in a multi-model context.

### 2.5.2 *Sea-ice metrics*

*No historical change in variability*

I now apply the basic version of my method to sea ice. In an analogy to SAT, I find remarkable similarity between the preindustrial and the historical internal variability for both Northern and Southern Hemisphere sea-ice volume and area (Figure 2.6). The single-model large ensembles of CESM1-CAM5-BGC and MPI-ESM1.1-LR (black triangle and diamond, respectively) confirm the one-to-one-relationship. When testing for a change in sea-ice internal variability with respect to a future RCP8.5-forced climate, I separately analyse winter sea ice (Figure 2.7) and summer sea ice (Figure 2.8). For winter Arctic and Antarctic sea-ice volume, I detect an extremely likely decreased internal variability compared to the preindustrial climate (Figure 2.7a,b), while for winter Antarctic sea-ice area I find a likely decreased internal variability (Figure 2.7d). For winter Arctic sea-ice area (Figure 2.7c), the model responses differ substantially and the future variability proves largely independent of a model's preindustrial variability. I suggest two counteracting effects that cause the models to disagree even on the direction of change. On the one hand, the variability of Arctic sea-ice area decreases because the mean sea-

*Future change in variability in winter*

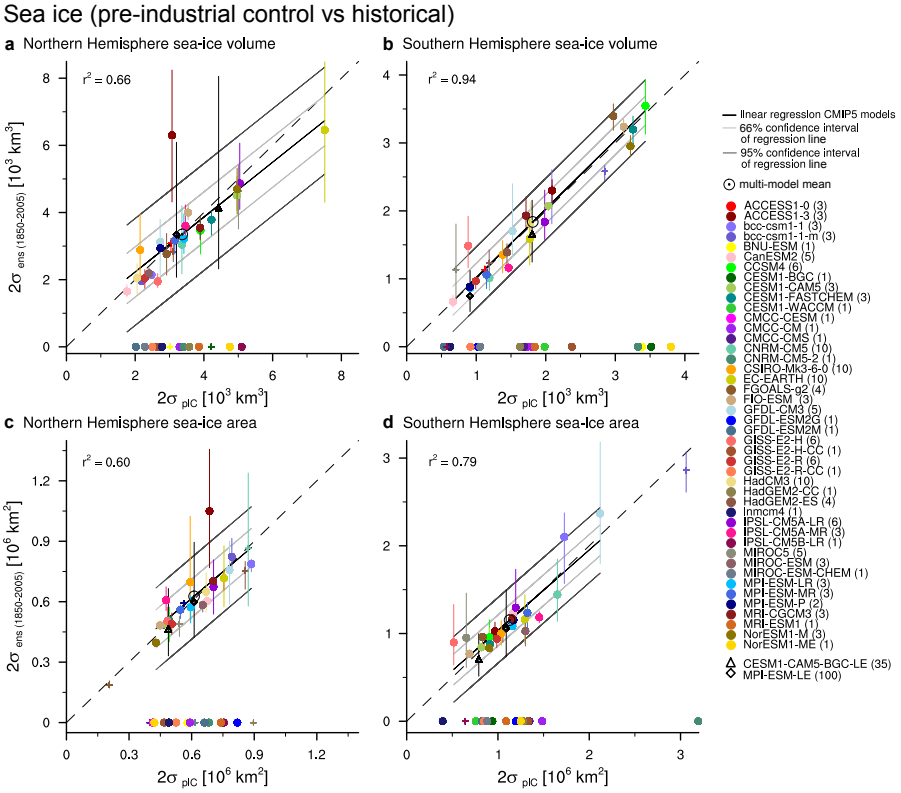



Figure 2.6: Relationship between the standard deviation of each CMIP5 model preindustrial control simulation (x-axis) and the ensemble standard deviation of the corresponding historical runs (y-axis) for annual **a** Northern and **b** Southern Hemisphere sea-ice volume, and **c** Northern and **d** Southern Hemisphere sea-ice area. CMIP5 models shown as colored filled dots are considered for calculating the regression line, the $r^2$, the 66 % and the 95 % confidence intervals when several simulations are provided.

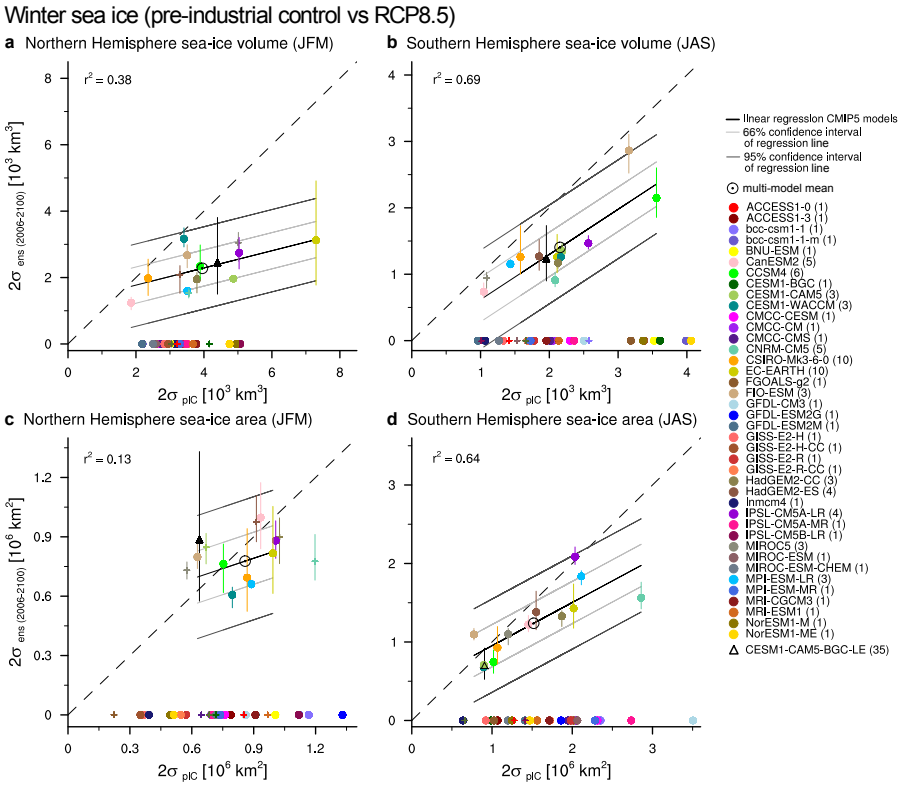



Figure 2.7: As in Figure 2.6, but for the ensemble standard deviation of the corresponding RCP8.5 scenario runs (y axis) for winter conditions.

ice area decreases (see Section 2.5.3). On the other hand, the variability of Arctic sea-ice area increases when the sea ice is detached from continental boundaries (Eisenman et al. 2011), when the sea ice becomes thinner (e.g., Bitz and Roe 2004; Notz 2009), and when the high-latitude temperature variability increases (cf. Figure 2.5b). The inconsistent model responses even on the direction of change might reflect the different manifestation and timing of these counteracting processes in each model.

*Future change in variability in summer*

For summer sea ice, I consider only sea-ice volume larger than $1 \cdot 10^3$ km$^3$ and sea-ice area larger than $1 \cdot 10^6$ km$^2$ to prevent artifacts arising from touching the lower bound of zero sea ice. As for winter, I find an extremely likely decreased internal variability of summer Arctic and Antarctic sea-ice volume and a likely decreased internal variability of summer Antarctic sea-ice area compared to preindustrial conditions (Figure 2.8a,b,d). My result for summer Antarctic sea-ice area is consistent with Goosse et al. (2009, their Figure 1b), who show a decreasing sea-ice variability with a decreasing mean state for March Antarctic sea-ice extent.

*Variability of Arctic sea-ice area in summer*

In contrast to these variables, the future internal variability of summer Arctic sea-ice area possibly increases and becomes largely independent of the different preindustrial manifestations of internal variability (Figure 2.8c). The possible increase agrees with the result for CCSM3 (Holland et al. 2008) and is consistent with the study by Goosse et al. (2009) on September Arctic sea-ice extent. Based on 14 global climate models, Goosse et al. (2009, their

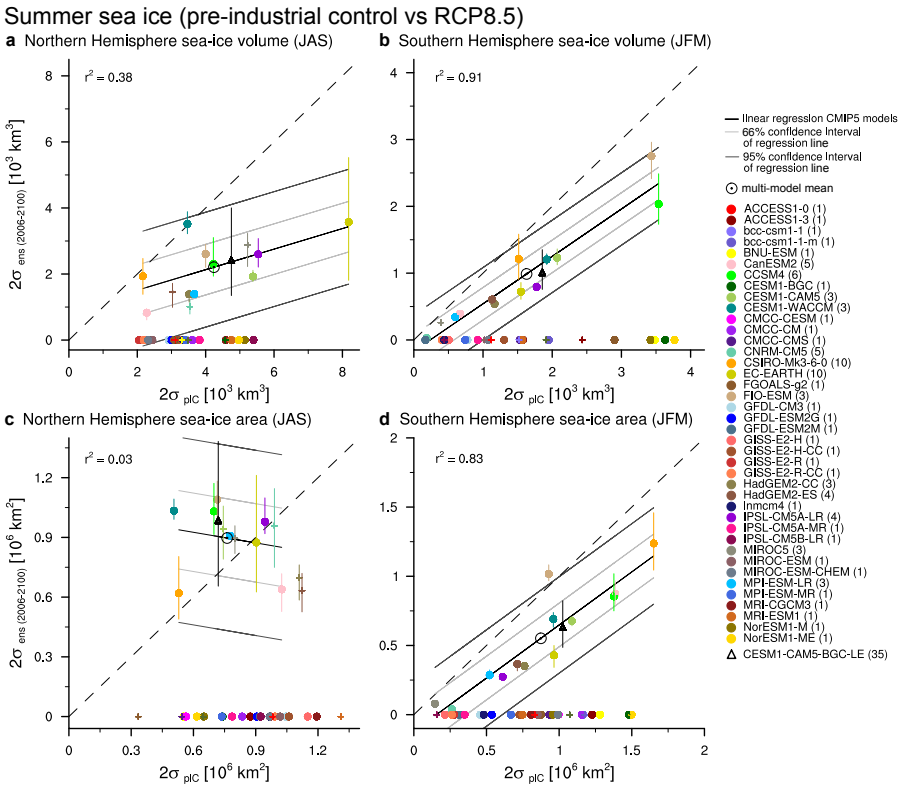



Figure 2.8: As in Figure 2.7, but for summer conditions. Note that only a sea-ice volume > 1 $\cdot$ $10^3$ km$^3$ and a sea-ice area > 1 $\cdot$ $10^6$ km$^2$ are considered when calculating the ensemble standard deviation.

Figure 1a) showed that the variability in September Arctic sea-ice extent increases until the mean sea-ice area has decreased to around 3 $\cdot$ $10^6$ km$^2$ and then decreases for a lower mean September sea-ice extent. Averaged over the period 2005–2100, I thus find increased variability in summer Arctic sea-ice area across all models (Figure 2.8c). As for winter, the variability in future summer Arctic sea-ice area is possibly increased because the sea ice becomes detached from continental boundaries, becomes thinner, and is vulnerable to increases in high-latitude temperature variability.

In summary, the magnitude of modeled annual mean internal variability of sea-ice volume and area remains largely unchanged for the historical period. In contrast, the magnitude of internal variability in winter and summer sea ice decreases in the RCP8.5 scenario, except for the variability of Arctic sea-ice area, which shows inconsistent model responses on the way to ice-free conditions. The inconsistent model responses for Arctic sea-ice area highlight the benefit of my multi-model approach compared to single-model studies, as it allows for the interpretation of single-model results in a multi-model context.

### 2.5.3 *Linking sea-ice variability to the mean sea-ice state*

Previous studies showed a linear relationship between mean temperature and temperature variability in the range of high-latitude annual mean tempera-

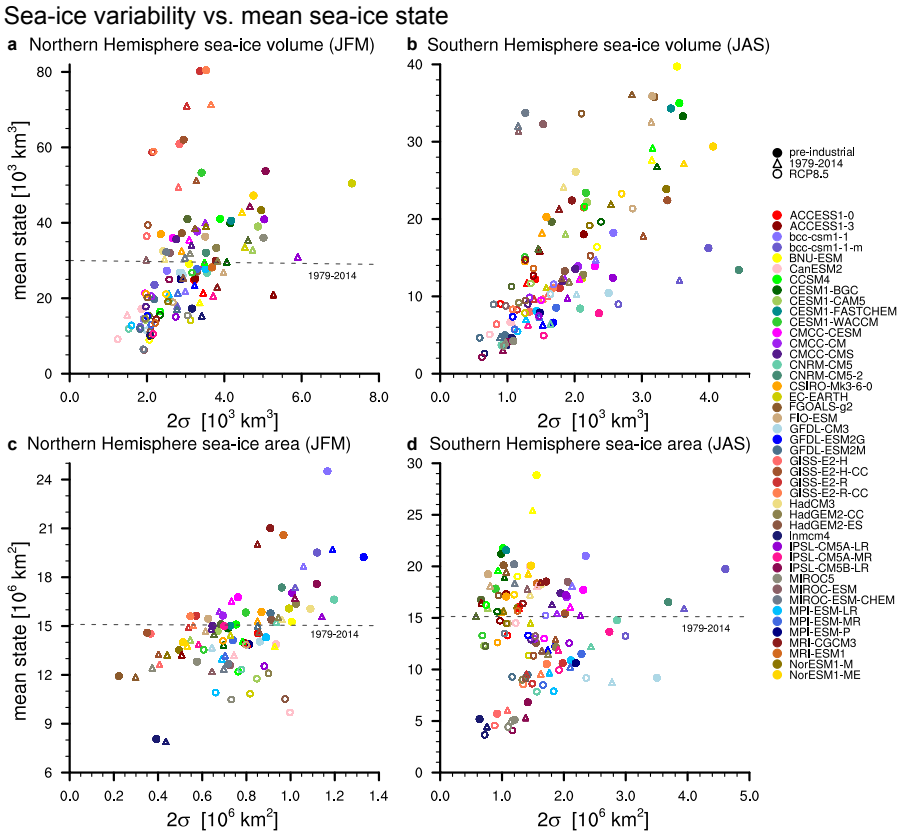



Figure 2.9: Relationship between the standard deviation and the mean state of the preindustrial (filled dots), reference data time period (triangles) and future RCP8.5-forced climate (circles) of winter **a** Northern and **b** Southern Hemisphere sea-ice volume, and **c** Northern and **d** Southern Hemisphere sea-ice area for the CMIP5 models. Symbols are shown each in the model-specific color. Dashed lines indicate the mean state of reference data from 1979 to 2014. Note that for future Arctic sea-ice area only the estimates from models with ensemble simulations are shown, because the regression line based on these models is highly uncertain and does not allow to robustly derive estimates for models with a single simulation (see Figure 2.7c).

tures (Esau et al. 2012, their Figure 3b) and a linear relationship between mean Arctic temperatures and mean Arctic sea-ice area (e.g. Gregory et al. 2002; Mahlstein and Knutti 2012) for global climate models. To test whether sea-ice variability is also linked to the mean sea-ice state in CMIP5 models, I apply the extended version of my method. The time-averaged evolution of internal variability for the sea-ice metrics in winter (Figure 2.9a–d) shows that models with a high preindustrial mean state simulate a higher preindustrial estimate of internal variability of the corresponding sea-ice variable than models with a low preindustrial mean state (filled dots). The cross-model relationship for the preindustrial state also holds for a single model over time, which one can infer from following individual models over time in Figure 2.9. If for the reference data period 1979–2014 (triangles) the modeled time-averaged mean sea-ice state is decreased compared to the preindustrial mean state (filled dots), then the internal variability is usually reduced sim-

ilarly. The decrease in internal variability continues for the RCP8.5-forced decrease in the mean sea-ice state (circles) for all metrics except for Arctic sea-ice area, whose future variability is largely independent of the future mean Arctic sea-ice area.

*Can the relationship between mean state and internal variability be used as an emergent constraint?*

Except for future Arctic sea-ice area, my analysis implies that the stronger the mean sea-ice state is altered, the more the internal sea-ice variability is reduced. From this relationship between the mean state and the variability of sea ice in CMIP5, I learn that when assuming that the models are realistic, knowledge about the mean state of the observable can be an emergent constraint for the system's internal variability, as also stated by Bathiany et al. (2016) based on two box models and a comprehensive Earth system model. However, because of the spread of modeled estimates, the relationship between the mean state and the variability permits only a rough estimate of the system's total variability from the mean state of a short observational time series. For comparison to observational or reanalysis data shown as dashed lines in Figure 2.9a,c,d, I show the model simulations for the reference data period 1979–2014 instead of those from the historical period.

## 2.6 APPLICATION 2: PLAUSIBILITY OF SEA-ICE SIMULATIONS

To present a second application of my method, I use its extended version for a robust and consistent assessment of the plausibility of sea-ice simulations from all CMIP5 models. I specifically stress here that only my robust estimation and consideration of modeled internal variability for evaluating sea-ice simulations is needed, since I do not know the system's true internal variability that otherwise could be used. I lack this knowledge because the observational record of sea-ice evolution is short and no robust estimate of the internal variability of the preindustrial sea-ice state exists. When applying the emergent constraint between internal variability and the mean state as discussed in Section 2.5.3, I face too broad a range of modeled estimates to derive a robust relationship that could be used to properly estimate the system's true internal variability based on the observed mean state.

*True sea-ice variability is unknown*

*Evaluating CMIP5 simulations of sea-ice area and volume*

The plausibility of the CMIP5 sea-ice simulations is tested for the metrics sea-ice volume and sea-ice area for the Northern Hemisphere and sea-ice area for the Southern Hemisphere both with respect to 30-year trends and the mean state. No evaluation of CMIP5 simulations of Southern Hemisphere sea-ice volume is provided because I lack a consistent long-term reanalysis data-set as a reference. Reference observational or reanalysis data for the considered measures are available from 1979 until today. To maximize overlap of this period with the CMIP5 simulations, I prolong the historical (1850–2005) sea-ice simulations until 2014 using simulations based on future RCP emission scenarios (Moss et al. 2010). Depending on availability and to maximize the number of models included in the analysis, I use RCP4.5 (preferred) or RCP8.5 for the extension of the historical sea-ice simulations. The choice of using either RCP4.5 or RCP8.5 for the extension of the historical runs until 2014 does not influence the evaluation result because both RCPs differ only slightly dur-

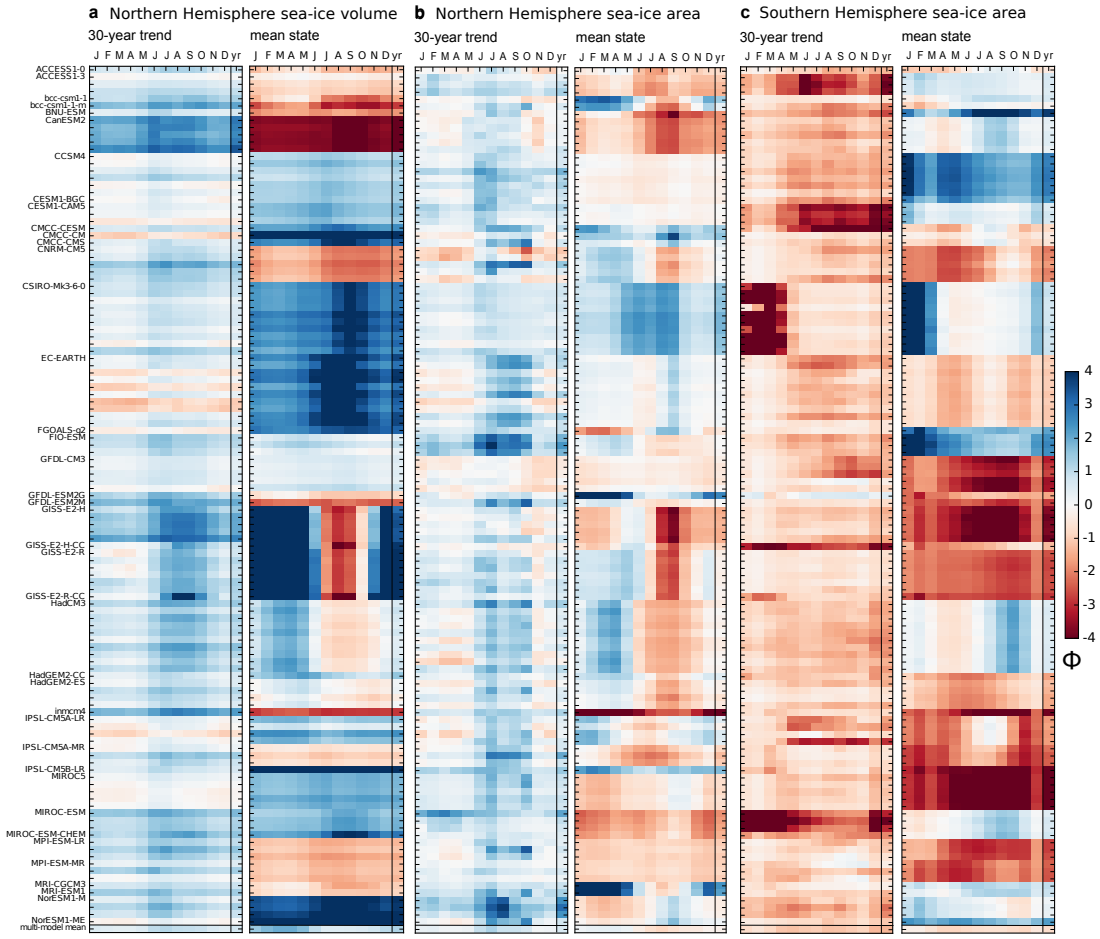



Figure 2.10: Portrait plot of the plausibility of CMIP5 sea-ice simulations for the 30-year trend and the mean state of **a** Northern Hemisphere sea-ice volume, **b** Northern Hemisphere sea-ice area, and **c** Southern Hemisphere sea-ice area based on the distance between each extended historical CMIP5 model simulation and reference data (PIOMAS for Northern Hemisphere sea-ice volume and the CDR satellite retrieval for sea-ice area). Deviations are shown in units of $\phi$, which combines $\delta_{ref}$ and $\sigma_{mod}$; a model's negative (red) and positive (blue) deviation with respect to reference data are indicated. Note that each model name is attached to the first ensemble simulation only.

ing this period. The extension of the historical sea-ice simulations reduces the total number of available simulations to 119 (see Table 2.1). For the evaluation of the mean sea-ice state, each model simulation is averaged over the period 1979–2014. The same is done for the reference data. In case of 30-year trends, both the model output and the reference data are averaged over the six 30-year linear trends obtained from the available 36-year-long time series from 1979 to 2014. CMIP5 model plausibility subdivided for each individual simulation and each month is presented as a portrait plot as introduced by Gleckler et al. (2008) (Figure 2.10). This is a condensed color-coded way to compare different variables of different model simulations to each other. The color indicates the likelihood of the model simulation to be consistent with the reference data. Red corresponds to a model's negative deviation with respect to the reference data, whereas blue indicates a model's positive devi-

ation. While $\phi = 0$ describes perfect agreement between the model output and the reference data, simulations are plausible at a likelihood of 95% when their deviations from the reference data result in $-2 < \phi < 2$. Deviations larger than $-3 < \phi < 3$ are plausible at a likelihood of 1%.

*Plausibility for 30-year trends versus the mean state*

By now consistently taking model-specific internal variability and reference data uncertainty into account, I find that internal variability can explain much of the differences between the models and the reference data for 30-year trends in sea-ice volume and sea-ice area (Figure 2.10a–c, left columns), whereas for some models it cannot explain the model biases of the mean state (Figure 2.10a–c, right columns). More specifically, my results reveal that for most models their internal variability is sufficiently high to explain the models' deviations from observed Northern Hemisphere trends in sea-ice volume and in sea-ice area, from Southern Hemisphere trends in sea-ice area, and the mean state of Northern Hemisphere sea-ice area. In contrast, for many models their internal variability cannot explain the model's deviation from the reanalysed mean state of Northern Hemisphere sea-ice volume and observed Southern Hemisphere sea-ice area.

*Comparison to previous studies*

My results confirm previous findings that modeled Northern Hemisphere trends in sea-ice area are generally less negative than the observed trends, especially in summer (e.g., Stroeve et al. 2012) and that modeled Southern Hemisphere trends in sea-ice area are more negative / less positive than the ones observed (e.g., Haumann et al. 2014; Mahlstein et al. 2013; Zunz et al. 2013). Nevertheless, the internal variability of most of the models can explain the annual mean deviations of the modeled trends in sea-ice area from the ones observed at a likelihood of 95%, in line with the discussion by Notz (2015).

Overall, the results show that the plausibility of models differs widely, both within one metric and across metrics. On the one hand, this variety is simply caused by the different model performances in simulating a sea-ice metric. On the other hand, this variety is also a result of the different model-specific internal variability taken into account for the model evaluation. For example, the simulations of two different models having the same distance to the reference data will have a different plausibility if one model is characterized by a different internal variability than the other. Evaluating the different model performances in the light of different model-specific internal variability is the strength of this evaluation approach.

## 2.7 CONCLUSIONS

*A new method to estimate ICV for different forcing scenarios*

I present a method that allows one to derive a robust estimate of ICV in climate-model simulations. I combine the control-simulation approach and the ensemble-spread approach that are commonly used for estimating ICV from model simulations. The method provides practical evidence that the quasi-ergodic assumption holds as long as the ICV is not changed by external forcing. Across different metrics, I find a linear relationship between the estimates from both approaches, which allows me to translate the modeled ICV of the preindustrial control climate to that of the historical and future

climate. This multi-model approach also allows for robust estimates of historical and future ICV for models with a single simulation. This new method for estimating ICV for different forcing scenarios is readily transferable to other variables and other applications in future multi-model studies. The applicability of my method is limited when the ensemble spread is averaged over much shorter periods than several decades, or when the ensemble standard deviations estimated for models with ensemble simulations have inconsistent directions of change, such as for future Arctic sea-ice area.

I present results from two applications of this method, namely, the assessment of changes of ICV over time and the evaluation of climate-model simulations. From applying my method to annual globally-averaged SAT and sea-ice volume and area for assessing large-scale changes of internal variability over time, I find

*... applied to examine changes of ICV over time*

1. a stable internal variability of annual globally-averaged SAT and sea-ice volume and area for the historical climate compared to the preindustrial climate,

2. a possibly decreased multi-model mean internal variability of annual globally-averaged SAT for the RCP8.5 scenario,

3. an extremely likely decreased internal variability of winter and summer Arctic sea-ice volume and winter and summer Antarctic sea-ice volume and a likely decreased internal variability of winter and summer Antarctic sea-ice area for a future climate forced by the RCP8.5 scenario, while winter and summer Arctic sea-ice area show inconsistent model responses,

4. changes in sea-ice variability to be largely controlled by changes in the mean sea-ice state, except for future Arctic sea-ice area, which gets detached from continental boundaries and is vulnerable to increased SAT variability.

On a regional scale, the method offers a multi-model consensus view on how and where internal temperature variability is projected to change in future. I find that

1. the pattern of possible SAT-variability change for the historical period already depicts the pattern of change projected for the future, in agreement with temperature extremes that occurred in the last decade (for an overview see Coumou and Rahmstorf 2012),

2. many mid latitudes are extremely likely to experience or will likely experience strong decreases in SAT variability, while many subtropical, tropical, and polar regions will likely experience strong increases in SAT variability under a future RCP8.5-forced climate. The multi-model consensus pattern of a projected increased temperature variability with possibly associated extreme events, especially on land, has major implications for society.

In the context of previous studies (Deser et al. 2000[a]; Holmes et al. 2016; Schneider et al. 2015; Screen 2014), my study suggests a close interplay between the mean state and the variability of SAT and sea-ice volume and area. Under global warming, the mean Arctic sea-ice cover decreases linearly with mean Arctic temperature (e.g., Gregory et al. 2002; Mahlstein and Knutti 2012). My results show that decreases in sea-ice variability are directly linked to the decrease in the mean sea-ice state as also found by Bathiany et al. (2016). This link between SAT variability and sea-ice variability via their mean states is enforced by the high sensitivity of thin sea ice to ocean and atmosphere temperatures (e.g., Bathiany et al. 2016; Bitz and Roe 2004; Notz 2009). For Arctic sea-ice area, the link between sea-ice variability and the mean sea-ice state is less prominent, likely because of the competing effect of increased SAT variability. When sea ice that keeps SAT close to the melting temperature is replaced by open ocean, the SAT variability can increase and thus allows for an increased variability of the remaining sea-ice area. For both annual globally-averaged SAT and sea-ice volume and area, applying my method reveals that the small CMIP5 ensemble size of a model is already representative of the model's total internal variability as shown for CESM1-CAM5 and MPI-ESM-LR based on their large ensembles. This representativeness of only few ensemble simulations of a model and the hugely different manifestation of ICV in CMIP5 models suggests that a multi-model approach offers more robust estimates for changes in ICV than results based on single-model large ensembles. The method proves powerful in addressing questions of regional temperature variability such as extreme events, which are commonly investigated using single-model large ensembles. Consequently, I consider the method as a useful tool for studies on ICV complementary to large ensemble simulations of a single model.

*... applied to evaluate climate-model simulations*

For the evaluation of climate-model simulations, the method permits a uniform consideration of model-specific ICV for all models independent from whether they provide several realisations or not. When applied to CMIP5 simulations of sea-ice volume and area, I conclude the following:

1. My multi-model approach discloses a highly variable model-specific internal variability of sea-ice volume and area. The different manifestation of internal variability in CMIP5 models hence must be considered in climate-model evaluation, as discussed previously by Stroeve et al. (2014), Notz (2014, 2015) and Swart et al. (2015).

2. The consideration of model-specific internal variability in evaluating CMIP5 sea-ice simulations is crucial for understanding the discrepancies between model output and reference data. The results allow for a distinction between model deviations that are plausible due to internal variability and reference data uncertainty and those that cannot be explained by these sources of uncertainty and thus point to model biases.

The applications discussed here show the potential of my simple method for estimating ICV for individual models and across multi-model ensembles. It allows me to gain both a robust assessment of temporal changes in variability

and a robust evaluation of model plausibility. In addition, my method allows me to directly quantify the agreement among models, which I often find to be quite low. I hence caution against the overinterpretation of possible changes in ICV obtained from single-model studies. Finally, while I limited myself to an assessment of SAT and sea-ice volume and area, my method is applicable for a wide range of climate variables and thus hopefully contributes to further understanding of ICV and its role for the climate evolution of our planet.

The multi-model approach that I introduced in this chapter complements our methodological toolset and enables us to understand changes in ICV over time. I now turn our attention to how such changes in ICV, specifically those in temperature and precipitation variability, are projected to impact the future probability of climate extremes. I investigate this question by attributing the projected change in probability to a shift in the mean and to changes in higher-order moments of the distributions of daily SAT and daily total precipitation. I quantify both contributions to the future probability of climate extremes in the next chapter.

# 3

# ATTRIBUTING FUTURE CLIMATE EXTREMES TO CHANGES IN MEAN AND HIGHER-ORDER MOMENTS

*Say you were standing with one foot in the oven*
*and one foot in an ice bucket.*
*According to the percentage people,*
*you would be perfectly comfortable.*

*— Bobby Bragan*

## 3.1 SUMMARY

Climate extremes can strongly affect human well-being, economy and ecosystems (IPCC 2012). Their extensive impacts demand for robust knowledge about future changes in climate extremes. However, quantifying the future probability of climate extremes under global warming is a scientific challenge, because the occurrence might not only be determined by the change in the mean, but also by respective changes in higher-order moments of the distribution. The relevance of changes in higher-order moments for the future probability of climate extremes is, however, difficult to assess and therefore debated (e.g., Fischer et al. 2013; McKinnon et al. 2016; Parey et al. 2013). Based on an empirical threshold approach, I here show that the increased future probability of hot extremes and the decreased future probability of cold extremes can be primarily attributed to the shift towards warmer temperatures rather than to changes in higher-order moments. Changes in higher-order moments are nevertheless important, as they counteract the increased future probability of hot extremes in high latitudes, strengthen the increased future probabilities of hot extremes in many mid latitudes, and reinforce the decreased future probability of cold extremes in mid- to high latitudes. Unlike for temperature extremes, the increased future probability of heavy precipitation extremes in most tropical and mid- to high-latitude regions and the decreased future probability mainly over subtropical oceans is primarily determined by changes in higher-order moments, while respective shifts in the mean precipitation amount are of secondary importance. By quantifying the global contributions from changes in mean and in higher-order moments to the future probability of climate extremes on the local scale, this study provides basic understanding to better inform society and adaptation planning.

## 3.2 INTRODUCTION

Global warming results in changes in the occurrence of climate extremes which impact society more than changes in the mean climate (e.g., Katz and

Brown 1992; Lewis and King 2017). Changes in climate extremes can be caused by a shift in the mean, by a changed variability and by a changed symmetry in distributions of temperature and precipitation (IPCC 2012), and they usually occur due to a combination of changes in these central statistical moments (e.g., Meehl et al. 2000; Trenberth et al. 2015). However, despite substantial research, the relative importance of changes in the central statistical moments for assessing the future probability of climate extremes is still unclear (e.g., Lewis and King 2017). To fill this gap, I here quantify the contributions from changes in the mean and from changes in higher-order moments to the future probability of temperature and precipitation extremes.

*Studies on temperature extremes*

*... sometimes consider changes in higher-order moments as unimportant*

Previous studies disagree on the importance of possible changes in higher-order moments for the observed and future probability of climate extremes. Whether past and future changes in temperature extremes can be approximated by the shift in the mean alone or are also strongly modulated by changes in higher-order moments is subject to considerable scientific debate. Some studies implicitly assumed unchanged variability (e.g., Zhou and Yu 2006) or argued that a shift in the mean accounts for much of the change in temperature extremes without requiring changes in higher-order moments (Donat and Alexander 2012; McKinnon et al. 2016; Rhines and Huybers 2013; Räisänen 2002; Simolo et al. 2011; Tingley and Huybers 2013). Specifically, Räisänen (2002) showed for 19 CMIP2 simulations with gradual doubling of $CO_2$ that future changes in the extremes of interannual variability of monthly temperature will be largely determined by the shift in the mean because the changes in variability are much smaller. Simolo et al. (2011) showed a dominant role of the shifted mean for observed changes in hot and cold extremes over Europe and ruled out contributions from changes in higher-order moments. Donat and Alexander (2012) reported that globally-observed changes in daily maximum and minimum temperature distributions are dominated by the shifted mean, with insignificant changes in variance and slight skewness towards hot temperatures. Based on instrumental data and proxy records, Tingley and Huybers (2013) showed that recent hot extremes can, to first order, be explained by the increased mean with constant variability in space and time. Based on quantile regression, McKinnon et al. (2016) quantified observed Northern Hemisphere mid-latitude changes in summer daily maximum and minimum temperature distributions and found that about 90% of the trends in temperature extremes can be explained by a shift in the mean with the remainder attributable to changes in the higher-order moments.

*... and sometimes consider changes in higher-order moments as important*

In contrast, other studies suggested that changes in higher-order moments, especially variance, are considerable and potentially more decisive for temperature extremes than the shift in the mean (e.g., Barrow and Hulme 1996; Mearns et al. 1984; Perkins-Kirkpatrick et al. 2017; Schär et al. 2004). Ylhäisi and Räisänen (2014) globally investigated future changes in skewness in a multi-model framework and showed that the assumption of a simple shift of the distribution does not hold for daily temperatures, because the temperature increase differs between the highest and the lowest percentiles. In line with their study, Lewis and King (2017) reported consistent skewness

in projected daily temperatures towards hot extremes for the 21$^{st}$ century and Lustenberger et al. (2014) found asymmetric changes in observed daily temperature variability for Europe. By fitting a normal distribution to Swiss summer temperatures, Schär et al. (2004) found that the 2003 heat wave in Europe cannot be explained only by the shift in the mean, but also requires an increased variability. Based on daily temperature time series for Eurasia and the United States, Parey et al. (2013) argued that trends in variance cannot be neglected when assessing changes in temperature extremes. Ballester et al. (2010) showed with an ensemble of regional climate simulations over Europe that deriving changes in temperature extremes requires knowledge of the mean, standard deviation and skewness of the distribution, and specifically point to the role of skewness.

*Studies on precipitation extremes*

*... consider changes in higher-order moments as important*

In contrast to the relative importance of changes in the mean and higher-order moments for temperature extremes in both observations and climate-model simulations, their relative importance for precipitation extremes is examined less in the literature. Karl and Knight (1998) found disproportionate changes across the precipitation distribution in the United States, with an increased proportion of total precipitation derived from heavy extremes at the expense of more moderate precipitation. Jakob and Walland (2016) examined the relative role of the shift in the mean over the 20$^{th}$ century and natural climate variability for the probability of climate extremes in Australia and showed that daily precipitation extremes rarely exhibit long-term change but are strongly modulated by the El Niño Southern Oscillation. Pendergrass et al. (2017) showed that global precipitation variability is projected to increase by at least as much as mean precipitation and less than heavy precipitation extremes. They argued that this muted increase in the mean precipitation relative to extreme precipitation requires a change in the shape of the distribution, specifically an increase in skewness, to break the symmetry between the changes in mean, variability, and extreme precipitation (see also Pendergrass and Gerber 2016).

Despite these research efforts, the relative contributions from changes in the mean and in higher-order moments to the future probability of climate extremes remain unclear. To better understand the origin of changes, I here quantify the future probability of temperature and precipitation extremes inferred from multiple models and systematically attribute the projected changed probability to a shift in the mean and to changes in higher-order moments.

## 3.3 DATA AND EMPIRICAL THRESHOLD APPROACH

*CMIP5 models with daily ensemble simulations*

To quantify the contributions from changes in the mean and changes in higher-order moments to the future probability of climate extremes, I use the CMIP5 model simulations of daily near-surface air temperature (SAT) and daily total precipitation from nine models that provide ensemble simulations for the historical period and for the future forced by the RCP8.5 emission scenario (Table 3.1). These simulations are available from the Earth System Grid data portal of the Earth System Grid Federation (http://esgf-node.llnl.gov/search/cmip5/).

Table 3.1: CMIP5 models with daily ensemble simulations used.

| Model name | Spatial resolution | # of simulations for SAT and total precipitation | |
|---|---|---|---|
| | latitude x longitude | historical | RCP8.5 |
| CanESM2 | 2.7906 x 2.8125 | 5 | 5 |
| CCSM4 | 0.9424 x 1.25 | 3 | 3 |
| CSIRO-Mk3-6-0 | 1.8653 x 1.875 | 10 | 10 |
| EC-EARTH | 1.1215 x 1.125 | 7 | 7 |
| HadGEM2-CC | 1.25 x 1.875 | 3* | 3** |
| HadGEM2-ES | 1.25 x 1.875 | 4 | 4** |
| IPSL-CM5A-LR | 1.8947 x 3.75 | 4 | 4 |
| MIROC5 | 1.4008 x 1.40625 | 3 | 3 |
| MPI-ESM-LR | 1.8653 x 1.875 | 3 | 3 |

* start in year 1960, ** end in year 2098

*Removal of trends*

I compare the probabilities of climate extremes for the future period 2034-2100 to those for the observational period 1950-2016. Given the high sensitivity of higher-order moments to the trends in both periods, the quantification of changes in higher-order moments requires a proper removal of trends. I detrend the observational-period and future-period daily temperature and precipitation timeseries of each ensemble member of a model in two steps. First, I calculate the ensemble mean of each model's ensemble simulations and apply a one-year running mean to the ensemble mean to remove its annual cycle. Second, I subtract this ensemble mean from each ensemble member to remove the forced signal on timescales longer than a year. Assuming an adequate removal of the forced signal, this isolates the internal climate variability (ICV) of each ensemble member which allows me to examine possible changes in higher-order moments in the distributions of SAT and total precipitation. To quantify the shift in the mean, I calculate the difference between the mean of both periods. To examine the future probability of climate extremes, I use the detrended first ensemble member of each model. The results are insensitive to the choice of the ensemble member.

*Empirical threshold approach*

I use an empirical threshold approach to quantify the contributions from a shift in the mean and a change in higher-order moments to the future probability of climate extremes (Figure 3.1). The approach considers the full distributions of SAT and total precipitation and does not rely on assumptions about the underlying distribution. I define higher-order moments as all statistical moments of a distribution higher than the first central moment, the mean. Higher-order moments thus include variance, skewness, kurtosis and all high-order moments. Following threshold indices for extremes (Zhang et al. 2011),

*Definition of extremes*

I define hot or cold extremes as the hottest or coldest 5% of days of the 66-year observational-period timeseries of SAT at every grid cell, which corresponds to all days that are warmer than the $95^{th}$ percentile or colder than the $5^{th}$ percentile of the observational-period PDF, respectively (see Figure 3.1a). Likewise I define heavy precipitation extremes as the wettest 5% of days of the 66-year observational-period timeseries of total precipitation at every grid

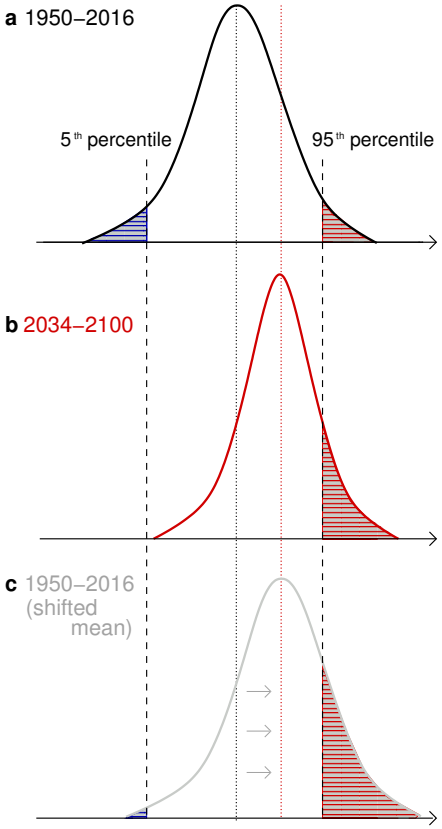



Figure 3.1: Schematic view of the empirical threshold approach to quantify the contributions to the future probability of climate extremes from a shift in the mean and from changes in higher-order moments for cold and hot extremes. **a** The 1950-2016 observational-period PDF that defines the 5% threshold for cold extremes and the 95% threshold for hot extremes. **b** The 2034-2100 future PDF that defines the future probability of cold and hot extremes caused by both a shift in the mean and the change in higher-order moments. **c** The observational-period PDF is shifted by its mean difference to the future PDF which isolates the respective changed probability that is attributable to the shift in the mean. The respective changed probability that can be attributed to changes in higher-order moments results from the difference in the areas with red lines between **b** and **c**.

cell, which corresponds to all days that are wetter than the $95^{\text{th}}$ percentile of the observational-period PDF. I use these observational-period thresholds for extremes, because humans are locally adapted to these extremes and a changed future probability with respect to the observational-period probability likely requires adaptation measures.

*Quantifying changes in probability*

To quantify the future probability of hot extremes, I use the future PDF and count the number of days that exceed the threshold criterion of the observational period (area with red lines in Figure 3.1b). To quantify the contribution to the future probability of hot extremes from a shift in the mean, I shift the otherwise unchanged observational-period PDF by adding the difference in the mean of both periods for every grid cell (Figure 3.1c). The number of days that exceed the $95^{\text{th}}$ percentile of the observational-period PDF defines the future probability of hot extremes that can be attributed to the shift in the mean (area with red lines in Figure 3.1c). To quantify the contribution to the

*... and attributing them to a shift in the mean*

*... and to changes in higher-order moments*

future probability of hot extremes from changes in higher-order moments, I subtract the change inferred from only the shift in the mean from the change caused by both the shift in the mean and the change in higher-order moments between both periods (area with red lines in Figure 3.1b minus that in Figure 3.1c). This difference quantifies the contribution from changes in higher-order moments that is possible with a concurrent shift in the mean. In addition, I use an alternative approach that quantifies the contribution from changes in higher-order moments without a concurrent shift in the mean. To do so, I shift the otherwise unchanged future PDF to the mean of the observational-period PDF and directly count the number of days that exceed the $95^{\text{th}}$ percentile of the observational-period PDF. For cold extremes and heavy precipitation extremes, I use the analogous procedure to quantify their future probabilities and the respective contributions from a shift in the mean and from changes in higher-order moments.

*Inter-model robustness*

To provide a multi-model mean estimate, I regrid the CMIP5 model estimates of the future probability of climate extremes on a 1°x1° grid by bilinear interpolation and average the estimates from all models. To quantify the inter-model robustness of the projected change in probability, I calculate the inter-model standard deviation normalised by the projected change of the multi-model mean. I consider a projected change as robust across models where the change of the multi-model mean is larger than the inter-model standard deviation.

This empirical threshold approach allows for a clear separation between robust changes in the future probability of climate extremes caused by a shift in the mean and those caused by changes in higher-order moments.

## 3.4 ORIGIN OF CHANGES IN THE FUTURE PROBABILITY OF CLIMATE EXTREMES

*Quantifying changes in mean and internal variability*

*... for SAT*

Before I attribute the future probabilities of climate extremes to a shift in the mean and to changes in higher-order moments, I first quantify how much the mean and the internal variability of daily SAT and daily total precipitation are projected to change between the observational and the future period (Figure 3.2 and Figure 3.3). As in Chapter 2 and Olonscheck and Notz (2017), I here use twice the ensemble standard deviation as a measure of ICV. By relating the estimates of globally-averaged SAT for all models between both periods (compare with Figure 2.4 and the method introduced in Section 2.3.1), I find that all nine CMIP5 models consistently project an increase in the mean and a decrease in the internal variability of SAT (Figure 3.2a,b). The projected decrease in the internal variability qualitatively differs from the finding in Figure 2.4b because of the use of daily data instead of annual data. For the observational period, the estimates from the three reanalysis products ERA-40/ERA-Interim (Dee et al. 2011), JRA-55 (Harada et al. 2016; Kobayashi et al. 2015) and NCEP/NCAR (Kalnay et al. 1996) agree with the magnitude of the simulated estimates of globally-averaged SAT. The pattern of the multi-model mean change reveals that the northern high latitudes contribute most

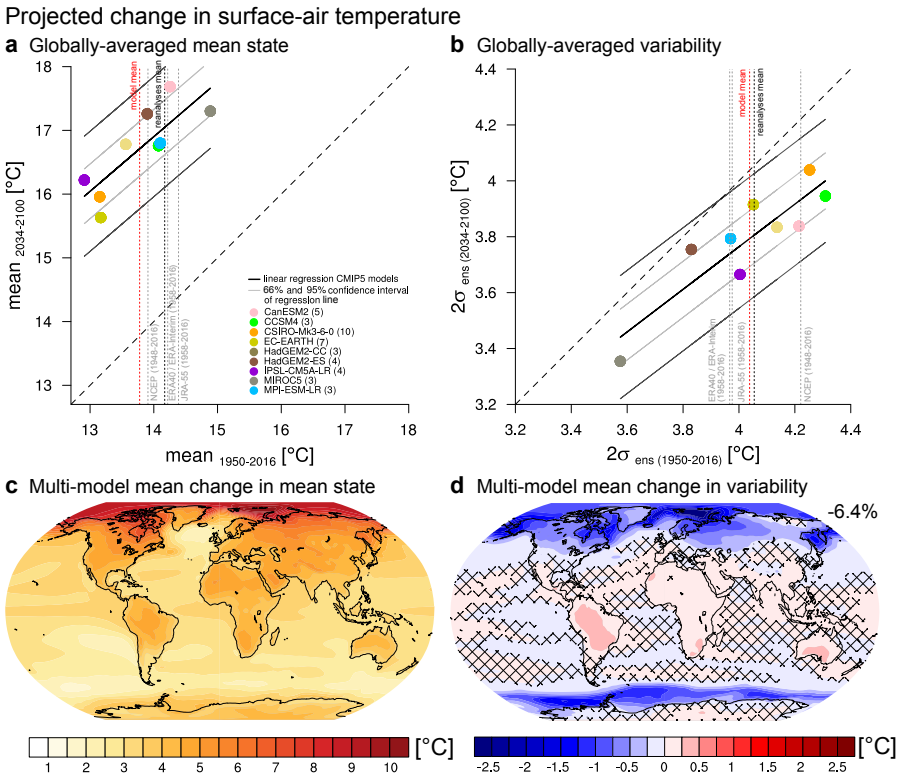



Figure 3.2: Projected change in daily near-surface air temperature (SAT) across models. **a** The globally-averaged mean state from 2034-2100 as a function of the globally-averaged mean state from 1950-2016, and **b** the globally-averaged internal variability from 2034-2100 as a function of the globally-averaged internal variability from 1950-2016 for nine CMIP5 models (filled dots). The estimates from three reanalysis products (vertical grey dotted lines and their mean in black) are compared to the multi-model mean estimate (vertical red dotted line) for the observational period. Multi-model mean change in **c** the mean state and **d** the internal variability between 1950-2016 and 2034-2100 at every grid cell. The absolute mean state and internal variability for both periods is shown in Figure B.1. Crossed areas mark insignificant changes at the 95% significance level.

to the globally-averaged increase in SAT (e.g., Pithan and Mauritsen 2014; Serreze and Barry 2011) and to the globally-averaged decrease in its internal variability (Figure 3.2c,d; compare Holmes et al. 2016; Schneider et al. 2015).

*... for total precipitation*

In contrast to daily SAT, I find that all nine CMIP5 models consistently project an increase in both the mean and the internal variability of globally-averaged daily total precipitation between both periods (Figure 3.3a,b), in line with Pendergrass et al. (2017) and Wentz et al. (2007). The three reanalysis products and the satellite and gauge measurements collected in the Global Precipitation Climatology Project (Huffman and Bolvin 2012) confirm the magnitude of the simulated estimates of total precipitation. The pattern of the multi-model mean change discloses that the globally-averaged increase in both the mean and the internal variability of total precipitation predominantly originates from the tropics and partly from the mid latitudes. The increases are opposed by decreases in particular in the dry eastern ocean basins that

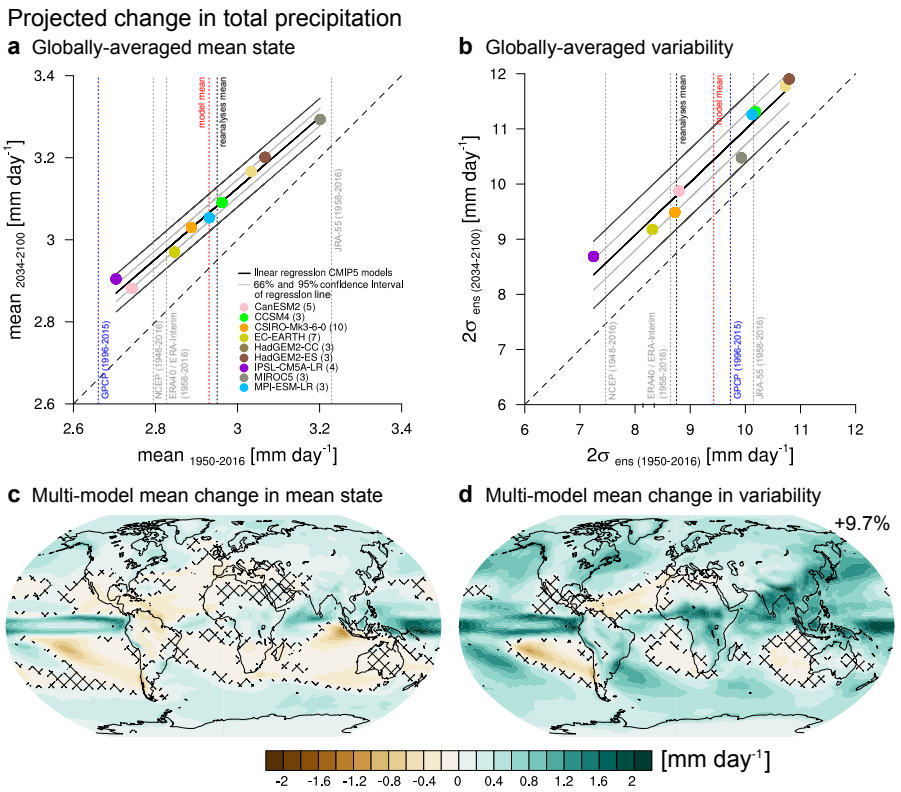



Figure 3.3: Projected change in daily total precipitation across models. **a** The globally-averaged mean state from 2034-2100 as a function of the globally-averaged mean state from 1950-2016, and **b** the globally-averaged internal variability from 2034-2100 as a function of the globally-averaged internal variability from 1950-2016 for nine CMIP5 models (filled dots). The estimates from three reanalysis products (vertical grey dotted lines and their mean in black) and observations from the Global Precipitation Climatology Project (vertical blue dotted line) are compared to the multi-model mean estimate (vertical red dotted line) for the observational period. Multi-model mean change in **c** the mean state and **d** the internal variability between 1950-2016 and 2034-2100 at every grid cell. The absolute mean state and internal variability for both periods is shown in Figure B.2. Crossed areas mark insignificant changes at the 95% significance level.

are projected to become even drier (Allan and Soden 2008) and less variable (Figure 3.3c,d).

The consistent findings of the projected decrease in internal variability of daily SAT and the projected predominant increase in the mean amount and internal variability of daily total precipitation under strong global warming raise the question how the projected changes impact the future probability of climate extremes. The projected changes suggest increases in hot extremes and decreases in cold extremes caused by the warming, decreases in temperature extremes in high latitudes caused by the decreased internal variability of SAT, and increases in precipitation extremes caused by both the increased precipitation amount and the increased internal variability of total precipitation in most regions. However, this assessment of the projected changes in temperature and precipitation does not account for changes in statistical moments higher than variance.

*Quantifying and attributing the future probability of climate extremes*

In light of the projected changes in the mean and the internal variability of SAT and total precipitation, I now quantify the future probability of temperature and precipitation extremes based on the empirical threshold approach that considers changes in all statistical moments and retains regional differences in seasonal variability. I attribute the projected changes in the probability of hot, cold and heavy precipitation extremes to a shift in the mean and to changes in higher-order moments. I do not examine the scarcity of precipitation because droughts are primarily determined by the duration of the lack of precipitation and related temperature and soil-moisture feedbacks (e.g., Seneviratne et al. 2014).

### 3.4.1 *Hot extremes*

*Future probability of hot extremes*

For a future forced with the RCP8.5 emission scenario, hot extremes are projected to become much more likely everywhere (Figure 3.4a). In some trop-

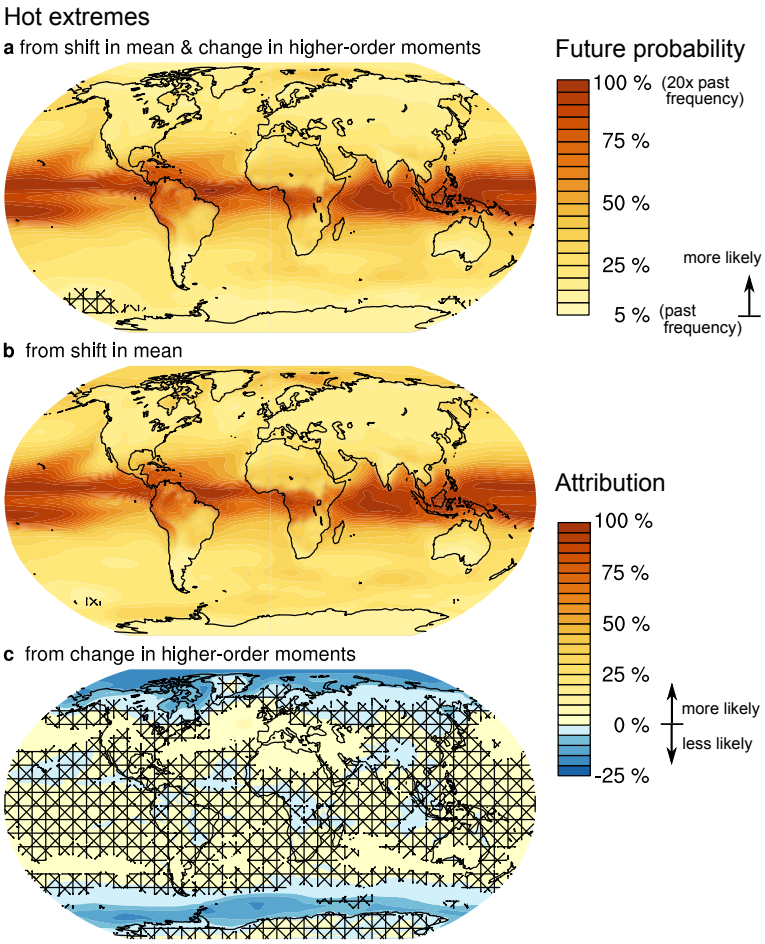



Figure 3.4: Attribution of the future probability of hot extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of hot extremes from the full change in the PDF of daily SAT between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

ical regions, all daily temperatures exceed in the future the temperature that defined a hot extreme in the past, representing a 20-fold increase in probability. In mid and high latitudes, the future probability of hot extremes is projected to increase substantially, but less compared to the tropics. Most of the increased future probability of hot extremes can be attributed to the shift to a warmer climate (Figure 3.4b). Despite the fact that global warming is most pronounced in mid and high latitudes (Figure 3.2c), the occurrence of local hot extremes increases less in mid and high latitudes than in the tropics. This counter-intuitive finding is caused by the small signal-to-noise ratio of SAT in mid and high latitudes compared to the tropics (e.g., King et al. 2015). In the tropics, the seasonal variability of SAT is small and the change in the mean shifts the full distribution to warmer temperatures partly without much overlap to the observational-period distribution (compare Figure 3.5c for Singapore). In contrast, the seasonal variability of SAT in mid and high latitudes is substantially larger than the change in the mean (compare Figure 3.5a-b,d).

*... caused by the shift in the mean*

*... caused by changes in higher-order moments*

The dominant role of changes in the mean state for tropical hot extremes overprints the small changes induced by higher-order moments in the tropics (Figure 3.4c). However, although being still smaller than changes induced by the warming shift, changes in higher-order moments modulate the future probability of hot extremes in mid and high latitudes. In line with the projected decreased internal variability of SAT especially in high latitudes (Figure 3.2d), the projected changes in higher-order moments in high latitudes counteract the increased future probability of hot extremes induced by the warming shift. In contrast, the changes in higher-order moments in many mid-latitude regions are projected to strengthen the increased probability induced by the warming shift, in line with previous findings (e.g., Schär et al. 2004). By using the alternative approach that quantifies the change from higher-order moments without a warming shift (Figure B.3a), I qualitatively find the same contribution from higher-order moments. However, the maximum possible decrease in the probability of hot extremes is, by definition, 5% since this indicates that temperatures representing past hot extremes no longer occur in the future. The potentially larger contribution from the change in higher-order moments becomes only effective for a concurrent warming shift (Figure 3.4c).

To illustrate the projected changes in the future probability of hot extremes and their attributable causes, I show the change in the PDFs for four individual cities around the globe (Figure 3.5, right tails). For Longyearbyen, the increased future probability of hot extremes, i.e. days warmer than the observational-period $95^{\text{th}}$ percentile, is caused by the shift in the mean, but partly compensated by the narrowing of the distribution. For Paris, the increased future probability of hot extremes is caused by both the shift in the mean and the widening of the distribution. For Singapore, the increased future probability of hot extremes is almost entirely caused by the shift towards warmer temperatures that overprints any changes in higher-order moments. For Buenos Aires, the increased future probability of hot extremes is largely caused by the shift in the mean and enhanced by a slight widening of the dis-

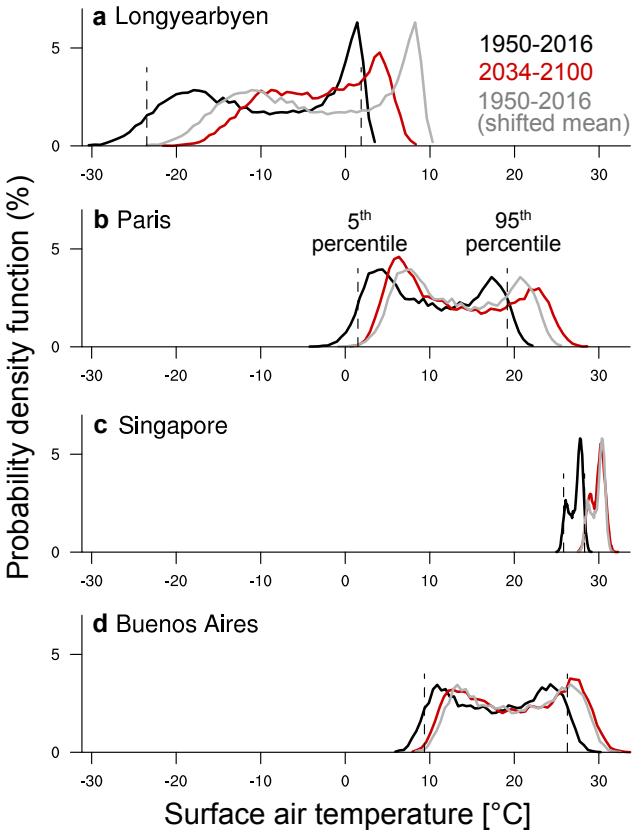



Figure 3.5: Multi-model mean change in the PDF of SAT in **a** Longyearbyen, **b** Paris, **c** Singapore and **d** Buenos Aires. The observational-period PDF (black) is compared to the future-period PDF (red) and the observational-period PDF shifted by the difference in the mean of both periods (grey). The $5^{th}$ and $95^{th}$ percentiles are shown as dashed lines.

tribution.

*Very hot extremes*

I further quantify the future probability of the 1% hottest extremes defined as the number of days warmer than the observational-period $99^{th}$ percentile (Figure B.4). Compared to the 5% hottest extremes examined before, I find that the pattern of the changed future probability of the 1% hottest extremes is very similar, but its magnitude of change is much higher. While the 5% hottest extremes show an up to 20-fold increase in frequency, the 1% hottest extremes show an up to 100-fold increase in frequency in some tropical regions. The relative contributions to this difference attributable to the shift in the mean and the change in higher-order moments are similar to the relative contributions for the 5% hottest extremes.

Overall, the increased future probability of hot extremes is primarily determined by the shift towards warmer temperatures. Projected changes in higher-order moments can counteract their increased future probability especially in high latitudes and strengthen their increased future probability in many mid latitudes.

### 3.4.2 *Cold extremes*

*Future probability of cold extremes*

For a future forced with the RCP8.5 emission scenario, cold extremes are projected to largely vanish in most regions (Figure 3.6a). However, in mid and high latitudes that are characterised by a substantial seasonal variability of SAT, cold extremes become less likely everywhere, but remain still possible. The projected decreased probability or vanishing of cold extremes can be largely attributed to the shift towards warmer temperatures (Figure 3.6b). Cold extremes remain similar likely as for the observational period where no or only slight warming is projected such as for the North Atlantic warming hole (e.g., Drijfhout et al. 2012; Menary and Wood 2018; Rahmstorf et al. 2015; Sgubin et al. 2017) and for parts of the Southern Ocean. However, the remaining high probabilities of cold extremes in these regions are not robust across models, but largely caused by two models – CSIRO-Mk3-6-0 and MIROC5 – that project an increased probability of cold extremes in these re-

*... caused by the shift in the mean*

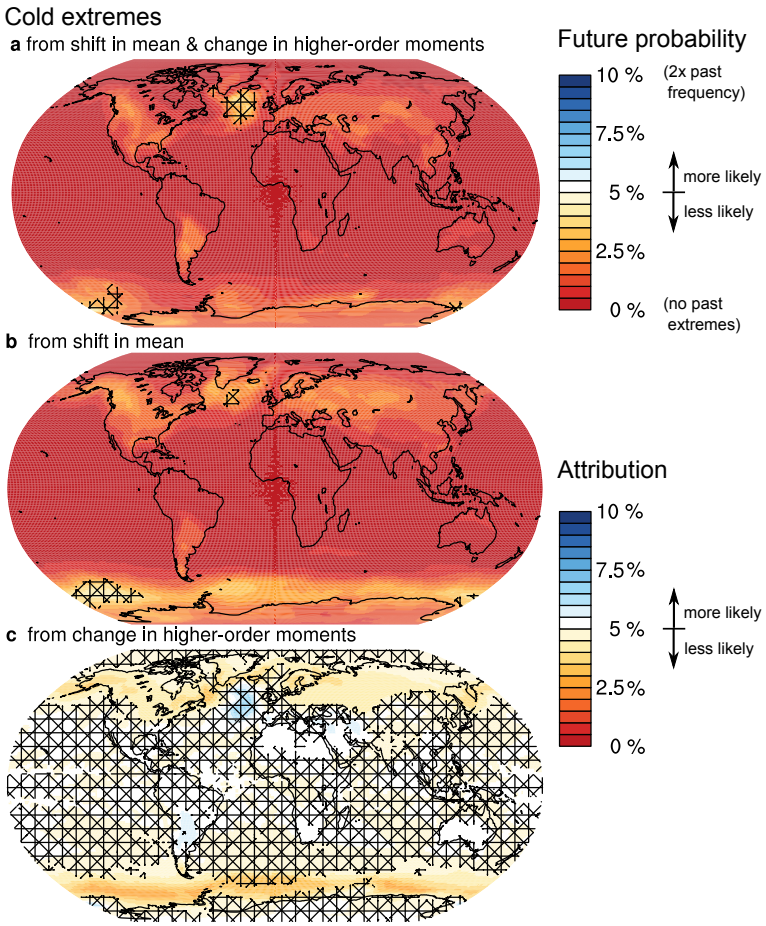



Figure 3.6: Attribution of the future probability of cold extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of cold extremes from the full change in the PDF of daily SAT between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

gions that are projected to warm least.

*... caused by changes in higher-order moments*

The remaining future probability of cold extremes from the shift in the mean in mid- to high latitudes becomes further reduced by a projected narrowing of the SAT distributions (Figure 3.6c), in line with the projected decreased internal variability of SAT in mid- to high latitudes (Figure 3.2d). Changes in higher-order moments are projected to counteract the decreased future probability caused by the shift in the mean in some Southern Hemisphere mid-latitude land regions, and – non-robustly across models – in the North Atlantic warming hole. From the alternative approach that quantifies the change from higher-order moments without a warming shift (Figure B.3b), I find that changes in higher-order moments would have the potential to substantially strengthen the decrease in the future probability of cold extremes in mid- to high latitudes. In contrast, but mostly not robust across models, changes in higher-order moments could partially compensate for the decreased future probability of cold extremes in tropical and subtropical regions. These findings of the potential effects from higher-order moments on the future probability of cold extremes are suggested by the projected decreased internal variability of SAT in mid- to high latitudes (Figure 3.2d)

The changes in the PDFs for the four cities illustrate the global patterns of the changed future probability of cold extremes (Figure 3.5, left tails). While cold extremes are projected to vanish in Longyearbyen and Singapore caused by the shift towards warmer temperatures without a remaining impact by higher-order moments, cold extremes are still possible in Paris and Buenos Aires and the probability for cold extremes in Buenos Aires is decreased less caused by the widening of the distribution.

*Very cold extremes*

Compared to the 5% coldest extremes examined before, I find the future probability of the 1% coldest extremes to decrease more strongly (Figure B.5), with similar relative contributions from the shift in the mean and the change in higher-order moments as for the 5% coldest extremes. Although not robust across models, the future probability of the 1% coldest extremes in the North Atlantic warming hole is decreased less than for the 5% coldest extremes because changes in the higher-order moments counteract the decreased probability caused by the warming shift and thereby retain a similar probability for cold extremes as for the observational period.

Similar to hot extremes, the decreased future probability or vanishing of cold extremes is primarily determined by the shift towards warmer temperatures. Projected changes in higher-order moments mostly further decrease the remaining future probability of cold extremes in mid- to high latitudes.

### 3.4.3 *Heavy precipitation extremes*

*Future probability of heavy precipitation extremes*

I now examine the future probability of heavy precipitation extremes. For a future forced with the RCP8.5 emission scenario, heavy precipitation extremes become more and up to twice as likely in many tropical, mid- and high-latitude regions, and less likely in many subtropical regions, particularly the eastern ocean basins (Figure 3.7a). This projected pattern of change

agrees with previous studies (Allan and Soden 2008; Fischer and Knutti 2015; Kharin et al. 2013; Pendergrass et al. 2017) and resembles both the projected change in the mean and in the internal variability of total precipitation (Figure 3.3c,d).

*... caused by the shift in the mean*

In contrast to the future probability of temperature extremes, the future probability of heavy precipitation extremes can be primarily attributed to changes in higher-order moments rather than to a shift in the mean precipitation. The shift in the mean precipitation contributes only to a minor part to the changed future probability of heavy precipitation extremes (Figure 3.7b). Whereas the increased amount of total precipitation in the high latitudes and parts of the tropics mainly along the Intertropical Convergence Zone slightly increases the future probability of heavy precipitation extremes in these regions, the decreased amount of total precipitation in many mid latitudes and the subtropics decreases the future probability of heavy precipitation extremes particularly

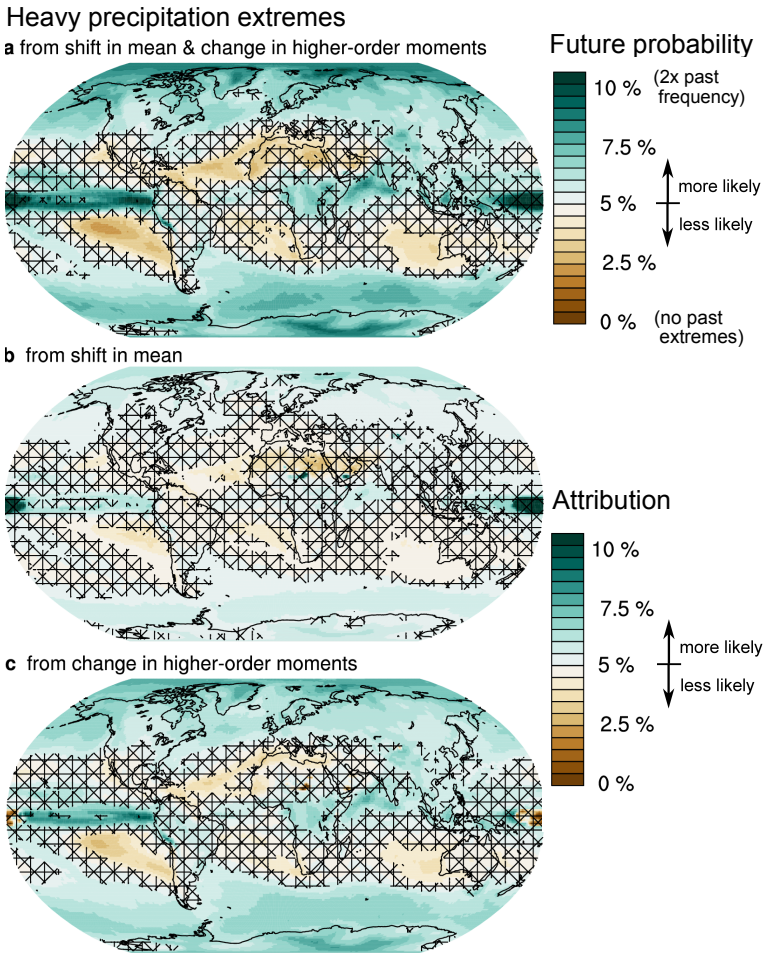



Figure 3.7: Attribution of the future probability of heavy precipitation extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of heavy precipitation extremes from the full change in the PDF of daily total precipitation between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

in the even drier eastern ocean basins. The future probability of heavy precipitation extremes is instead primarily determined by the change in higher-order moments in most regions (Figure 3.7c). The pattern of change caused by the change in higher-order moments largely reinforces the changes caused by the mean precipitation amount. Because of the generally small shift in the mean precipitation amount, I find a very similar result when quantifying the change from higher-order moments without a concurrent shift in the mean (Figure B.3c).

*... caused by changes in higher-order moments*

I further find that the future probability of the 1% heaviest precipitation extremes is changed much more strongly than the future probability of the 5% heaviest precipitation extremes, with more than a doubling of the past frequency in many tropical and high-latitude regions (Figure B.6). Compared to the attribution of changes in the 5% heaviest precipitation extremes, more of the strongly increased future probability of the 1% heaviest precipitation extremes is caused by the change in higher-order moments, with a relatively smaller contribution from the shift in the mean precipitation amount.

*Very heavy precipitation extremes*

Overall, the changed future probability of heavy precipitation extremes is primarily determined by the projected change in higher-order moments. The shift in the mean precipitation amount mainly reinforces the pattern of change caused by the change in higher-order moments.

In summary, the quantification and attribution of the future probability of climate extremes shows that future temperature extremes are dominated by the shifted mean, whereas future heavy precipitation extremes are primarily driven by changes in higher-order moments. This attribution of the projected change in climate extremes calls for considering both changes in the mean and changes in higher-order moments for a robust assessment of future climate extremes. However, the relative importance of the long-term shift in the mean and the change in higher-order moments strongly depends on the climate variable.

## 3.5 CONCLUSIONS

I quantify the future probabilities of climate extremes and attribute them to changes caused by a shift in the mean and changes caused by a change in higher-order moments. This attribution fosters our understanding of the relative role of climate change and climate variability for projected changes in climate extremes. I first quantify the projected change for the RCP8.5 emission scenario in both the mean and the internal variability of daily SAT and total precipitation as boundary conditions for the future probability of climate extremes. Based on consistent multi-model estimates, I find that

*Quantified changes in mean and internal variability of SAT and total precipitation*

1. the extremely likely future increase in the mean of daily globally-averaged SAT is accompanied by an extremely likely future decrease in its internal variability. The globally-averaged signal of the inverse relationship between the mean and the internal variability originates primarily from the northern mid and high latitudes,

2. the extremely likely future increase in the mean amount of daily globally-averaged total precipitation is accompanied by an extremely likely future increase in its internal variability,

3. the observational-period range of simulated globally-averaged SAT and total precipitation agrees with the range inferred from reanalysis products and observations.

*Quantified future probability of climate extremes attributed to its causes*

In light of the projected changes in the mean and the internal variability of daily SAT and total precipitation, I examine the future probability of temperature and precipitation extremes. Based on an empirical threshold approach, I quantify the respective contribution from a shift in the mean and a change in higher-order moments. For future temperature extremes, I find that

1. hot extremes become much more likely everywhere, but particularly in the tropics that show a high signal-to-noise ratio,

2. cold extremes are projected to vanish in most parts of the world, but remain possible on mid- and high-latitude land and in the Southern Ocean,

3. the future probability of hot and cold extremes can be primarily attributed to the shift in the mean towards warmer temperatures,

4. changes in higher-order moments are secondary, but can modulate the future probability of hot and cold extremes especially in mid and high latitudes.

For future precipitation extremes, I find that

1. heavy precipitation extremes become more likely in most tropical, mid- and high-latitude regions and less likely in the subtropics particularly in the eastern ocean basins,

2. the future probability of heavy precipitation extremes can be primarily attributed to changes in higher-order moments,

3. slight shifts in the mean towards higher precipitation amounts in the tropics and high latitudes and lower precipitation amounts in most subtropical and mid-latitude regions and particularly in the eastern ocean basins are secondary, but generally reinforce the changes caused by higher-order moments.

The presented findings reveal that the relative role of the shift in the mean and the change in higher-order moments for the future probability of climate extremes strongly differs for temperature and precipitation extremes. For temperature extremes, the strong warming signal largely overprints changes in higher-order moments especially where the seasonal temperature variability is low. The dominance of the warming signal further explains the strongly increased future probability of hot extremes on a planet with a future decreased temperature variability especially in high latitudes.

For heavy precipitation extremes, the warming-induced change in the mean amount of total precipitation is low compared to the warming-induced change in higher-order moments. Global warming causes the planet to become more extreme with respect to total precipitation because the changes in the mean and in higher-order moments cause wet regions to become even wetter and more variable, and dry regions to become even drier and less variable.

From this chapter, we learned that the strongly increased future probability of hot extremes and the strongly decreased future probability of cold extremes is dominated by the shifted mean, whereas the future probability of heavy precipitation extremes is primarily determined by changes in higher-order moments. A robust quantification of changes in the future probability of climate extremes hence requires the consideration of all statistical moments of the distributions of daily SAT and total precipitation from multiple models.
To improve our understanding of how ICV and its associated impacts arise, I will now dive into the mechanisms that cause its emergence. I focus on a variable that is both relevant to understand within the climate system and easy to grasp: the internal variability of the Arctic sea-ice area. I disclose the main cause why Arctic sea-ice area varies from one year to the next in the next chapter.

# 4

# ARCTIC SEA-ICE VARIABILITY IS PRIMARILY DRIVEN BY ATMOSPHERIC TEMPERATURE FLUCTUATIONS

Unwanted Feedback

*Sunbeams two-step over white blankets*
*Gliding painlessly between worlds,*
*Then stray too close to the edge*
*Where looming Arctic waves*
*Trap their latent heat;*
*Ending the dance,*
*And leaving*
*Only*
*Sea.*

*— Samuel Illingworth*

## 4.1 SUMMARY

The rapid decline of Arctic sea ice in the last decades is superimposed on strong interannual variability. The mechanisms that cause this variability are, however, still unclear (e.g., Årthun et al. 2012; Curry et al. 1995; Deser et al. 2000[b]; Ding et al. 2017; Fang and Wallace 1994; Hall 2004; Kapsch et al. 2013; Kashiwase et al. 2017; Letterly et al. 2016; Miles et al. 2014; Ogi et al. 2010; Park et al. 2015; Ukita et al. 2007; Woods and Caballero 2016; Zhang 2015) and their relative contributions are not quantified. Here I demonstrate that internal variability of sea ice is primarily caused by atmospheric temperature fluctuations. Other suggested drivers such as radiative effects and feedbacks related to surface albedo, clouds and water vapour, surface winds, and poleward oceanic heat transport together explain only 25% of the sea-ice variability. The dominating impact of atmospheric temperature fluctuations on sea ice is consistent with simulations from global climate models, reanalyses and observations. The direct response of sea ice to atmospheric temperature fluctuations explains why a feedback-driven tipping point is unlikely to exist in the sea-ice system (Bathiany et al. 2016; Tietsche et al. 2011) and limits prospects of interannual predictions of sea ice (Serreze and Stroeve 2015).

## 4.2 INTRODUCTION

*Arctic sea-ice area co-varies with temperature*

Arctic sea-ice area largely co-varies with Arctic mid-troposphere air temperature and Arctic sub-thermocline ocean temperature (Figure 4.1). However,

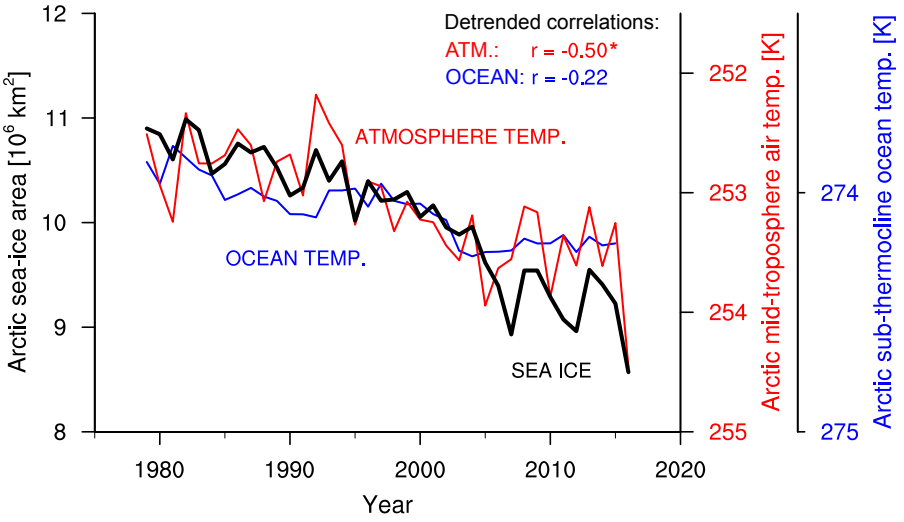



Figure 4.1: Evolution of Arctic sea-ice area, 60-90°N mid-troposphere air temperature and 60-90°N sub-thermocline ocean temperature from 1979 to 2016. The NSIDC passive microwave sea-ice concentration, the ERA-Interim reanalysis and the ORAS4 reanalysis are detrended to derive the indicated regression coefficients (see Section 4.3). Significant correlations are marked with an asterisk. The temperature axes are reversed and differently scaled for illustrative purposes.

*...but the link is modified by radiative effects, feedbacks, and forcings*

for the detrended timeseries a significant correlation only exists between Arctic sea-ice area and mid-troposphere air temperature (r = -0.50, $\alpha$ = 0.002, see Section 4.3), suggesting that variability in sea-ice area is mainly linked to fluctuations in atmospheric temperature (e.g., Deser et al. 2000[b]; Ding et al. 2017; Fang and Wallace 1994; Ukita et al. 2007). Understanding these links between Arctic sea-ice area and atmospheric and oceanic temperatures is difficult because the contributions from other suggested drivers, such as radiative effects and feedbacks related to surface albedo (e.g., Deser et al. 2000[b]; Hall 2004; Kashiwase et al. 2017), clouds (e.g., Letterly et al. 2016) and water vapour (e.g., Curry et al. 1995), and the forcings by surface wind (e.g., Ogi et al. 2010) and poleward atmospheric (e.g., Kapsch et al. 2013; Park et al. 2015; Woods and Caballero 2016) and oceanic energy transport (e.g., Årthun et al. 2012; Miles et al. 2014; Zhang 2015) are not well understood. Despite the extensive research fostered by scientific interest and societal needs, a complete Arctic-wide quantification of the drivers of sea-ice variability throughout a year is lacking. In this chapter, I quantify how much the suggested drivers contribute to sea-ice variability on annual to decadal timescales by systematically decoupling the relevant radiative effects, feedbacks and forcings in a state-of-the-art Earth system model.

*Quantifying suggested drivers*

## 4.3 METHODS

*Decoupling radiative effects, feedbacks, and forcings*

The method of decoupling radiative effects and feedbacks has been initially developed for ECHAM6 (Mauritsen et al. 2013) and elaborated to study the role of clouds for El Niño (Rädel et al. 2016). Here, I implement the method into the fully coupled Max Planck Institute Earth System Model (MPI-ESM1.2-

Table 4.1: List of experiments with their atmospheric $CO_2$ concentration, change in global-mean surface temperature (Δ GMST) and change in Arctic sea-ice area (Δ SIA) compared to the control experiment.

| Experiment | atm. $CO_2$ conc. [ppm] | Δ GMST [K] | Δ SIA [$10^6$ km$^2$] |
|---|---|---|---|
| Control (fully interactive) | 285 | 0 | 0 |
| Surface albedo non-interactive | 285 | -0.04 | 0.02 |
| Clouds non-interactive | 205 | 0 | -0.01 |
| Water vapour non-interactive | 165 | 0.04 | 0 |
| Surface wind non-interactive | 285 | -0.02 | 0.02 |
| Ocean heat transport non-interactive | 285 | -0.02 | 0.01 |
| All mechanisms non-interactive | 165 | 0.06 | -0.01 |

LR) to decouple the radiative effects and feedbacks within the atmosphere component ECHAM6.3 and the forcing by surface wind and poleward oceanic heat transport within the ocean component MPIOM1.6.

*... in the atmosphere*

The method consists of three steps: First, I perform a standard preindustrial control simulation run for 250 years with fully interactive feedbacks and forcings where I write out all relevant instantaneous fields at every 2-hourly radiation call. Second, I randomly shuffle the 2-hourly fields among the 250 years to eliminate auto-correlation but sustain the time of day and year. Third, the relevant time-shuffled fields are read into the models radiation calculations at every radiation call at the same time of day and year in 250-year long simulations with an otherwise identical setup. In the experiment with non-interactive surface albedo feedback, I prescribe the ice and land surface albedo in the visible and near infrared range. In the experiment with non-interactive cloud feedback, I prescribe the cloud liquid, ice, and cover fraction used in the atmospheric radiation calculations. In the experiment with non-interactive water vapour feedback, I prescribe the three-dimensional specific humidity field used in the atmospheric radiation calculations. In all three experiments, the respective fields are prescribed from the time-shuffled preindustrial control simulation globally.

*... in the ocean*

I adapted the method to prescribe also the forcings by surface wind and poleward oceanic heat transport. There are two differences: First, I output, time-shuffle and input the relevant fields daily instead of 2-hourly. Second, I do not directly input the time-shuffled fields into the model calculations. To prescribe the surface wind field globally, I use flux adjustment to input the time-shuffled fields of 10 m wind velocity, and zonal and meridional wind stress on water and ice. To prescribe the poleward oceanic heat transport, I nudge 3d ocean temperature and salinity fields in a 5°latitude band south of the winter sea-ice edge at 50-55°N in the Atlantic and 40-45°N in the Pacific at a relaxation time of $1e^{-5}$ $s^{-1}$ such that the temperature and salinity fields are replaced within one day.

*Compensation for warming drift*

The experiments in which I decouple the radiative effects and feedbacks of clouds and water vapour show a warming drift in GMST. To compensate for

Table 4.2: Contribution of radiative effects, feedbacks and forcings to the variability of the Arctic sea-ice area annually and seasonally relative to the control experiment (%). Winter (JFM), spring (AMJ), summer (JAS) and fall (OND) are defined according to the seasonal cycle of sea ice, with its annual maximum in March and annual minimum in September.

| Season | Surface albedo | Clouds | Water vapour | Surface wind | Ocean heat transport | All mechanisms |
|---|---|---|---|---|---|---|
| annual | 12.1 | 3.2 | 8.7 | -4.3 | 1.1 | 24.6 |
| JFM | 6.1 | 8.1 | 2.3 | -4.7 | 1.1 | 29.1 |
| AMJ | 13.7 | 5.6 | 5.2 | -10.2 | -2.8 | 27.0 |
| JAS | 20.5 | 0.1 | 19.7 | 2.0 | 4.2 | 23.1 |
| OND | 9.4 | -2.3 | 9.2 | -4.1 | 2.0 | 17.7 |

this warming, which reduces the mean state of Arctic sea-ice area, I reduce the atmospheric $CO_2$ concentration to 205 ppm and 165 ppm, respectively (see Table 4.1). I compensate the warming drift in the experiment in which all mechanisms are non-interactive by reducing the atmospheric $CO_2$ concentration to 165 ppm. By doing so, the GMST in all experiments deviates by less than 0.06°C from the one in the control experiment, and the mean state and seasonal cycle of Arctic sea-ice area are nearly identical in all experiments (Table 4.1). The nearly identical mean state of Arctic sea-ice area is crucial for this study because the sea-ice variability is highly sensitive to the mean sea-ice state. I use the last 200 years of each experiment for all analyses, regarding this first 50 years as spin up.

*Observations and reanalyses*

I use the NSIDC passive microwave sea-ice concentration data set (Fetterer et al. 2017) for sea-ice area, the ERA-Interim reanalysis (Dee et al. 2011) for atmospheric temperature and the ORAS4 reanalysis (Balmaseda et al. 2013) for oceanic temperature. The quadratic trend is removed from all timeseries to account for the exponential evolution of both sea-ice area and atmospheric and oceanic temperatures. Similar correlations between sea-ice area and atmospheric and oceanic temperatures exist when detrending the temperature timeseries linearly (ATM.: r = -0.52, OCEAN: r = -0.21, compare Figure 4.1).

*Testing for significance*

Significance is tested by applying a two-sided Student's t-test at the significance level $\alpha = 0.001$. Usage of different significant levels ($\alpha = 0.01$ or 0.1) leads to qualitatively similar results. The correlation of NSIDC sea-ice concentration with ORAS4 oceanic temperature becomes significant at $\alpha = 0.2$, i.e. a confidence level of 80% (see Figure 4.1).

## 4.4 RADIATIVE EFFECTS, FEEDBACKS AND FORCINGS

To quantify how much the radiative effects and feedbacks related to surface albedo, clouds and water vapour, and the forcings by surface winds and poleward oceanic heat transport contribute to sea-ice variability, I perform and analyse experiments with the global climate model MPI-ESM1.2-LR in which I decouple the aforementioned mechanisms collectively and individu-

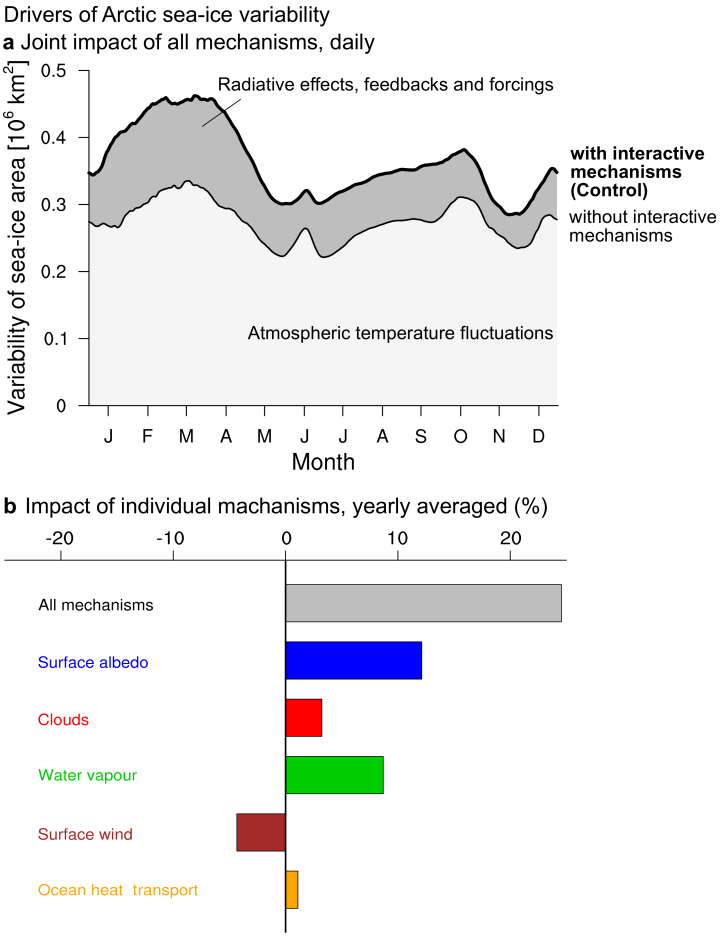



Figure 4.2: Impact of radiative effects, feedbacks and forcings on the variability of the Arctic sea-ice area. **a** Change in multi-year daily standard deviation of Arctic sea-ice area with interactive radiative effects, feedbacks and forcings (Control) and without these mechanisms being interactive throughout the year. **b** Yearly averaged change in multi-year daily standard deviation of the Arctic sea-ice area in percent from the six experiments relative to the control experiment. The joint impact of the radiative effects and feedbacks from surface albedo, clouds and water vapour, and the forcings from surface wind and poleward oceanic heat transport is shown on top.

*Contribution of radiative effects, feedbacks, and forcings to sea-ice variability*

ally (see Section 4.3). I compare the experiments in which one or all mechanisms are decoupled, hence non-interactive, to a fully interactive control experiment (Figure 4.2 and Figure C.1), which allows me to separate how much of Arctic sea-ice variability is driven by atmospheric temperature fluctuations and how much by the other suggested mechanisms. The sum of the annually averaged contributions from the individual mechanisms is similar to their joint impact, but synergies and interactions among the assessed mechanisms exist (Figure 4.2b and Figure C.1). In total, I find that the radiative effects, feedbacks, and forcings explain only 25% of the Arctic sea-ice variability, with the seasonally averaged contributions being larger in winter and spring than in summer and autumn (Table 4.2). Regionally, the assessed mechanisms impact sea-ice variability primarily in regions of the sea-ice edge where the sea-ice concentration varies most in the control experiment (Figure 4.3 and

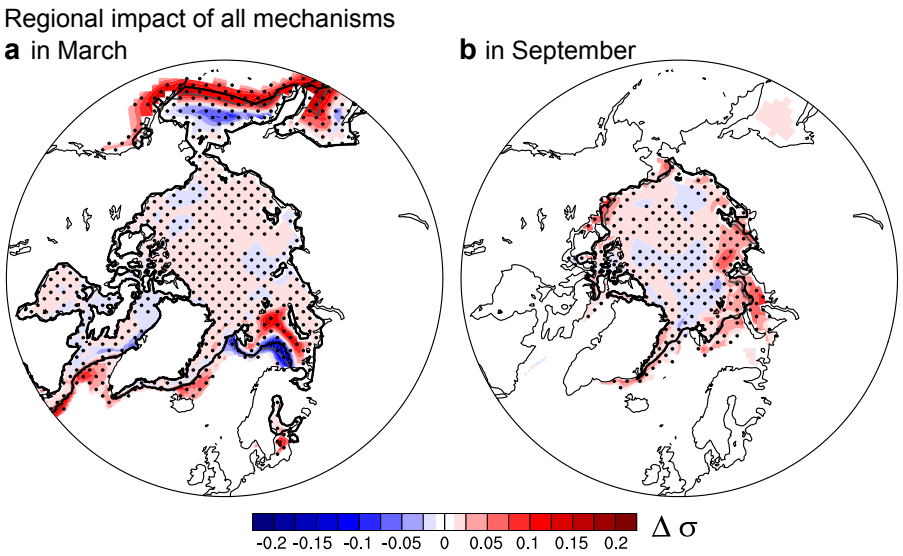



Figure 4.3: Regional impact of radiative effects, feedbacks and forcings on Arctic sea-ice variability. **a** March and, **b** September difference in variability of sea-ice concentration (Control - all mechanisms non-interactive, 99.9% significance where stippled). An increased variability of sea-ice concentration caused by the assessed mechanisms is marked in red, while a decreased variability is shown in blue. Changes in variability smaller than 1% are not shown. The sea-ice edge at 15% concentration displayed as black and grey lines is similar for both experiments.

Figure C.2). In March, the mechanisms lead to an increase of the sea-ice variability mainly in the marginal ice zone south of the sea-ice edge at 15% sea-ice concentration and a decrease further north in the Bering Sea, while the Barents Sea shows the opposite pattern (Figure 4.3a). These changes are mainly caused by the radiative effects of clouds and water vapour, with similar contributions from surface albedo, surface wind, and ocean heat transport in the Barents Sea (Figure C.3). In September, the mechanisms overall increase the sea-ice variability (Figure 4.3b), primarily caused by the surface albedo feedback with the radiative effects from clouds partly opposing (Figure C.3). The overall minor role of radiative effects, feedbacks and forcings suggests that three quarters of Arctic sea-ice variability are directly caused by fluctuations in atmospheric or oceanic temperature.

## 4.5 ATMOSPHERIC AND OCEANIC TEMPERATURE FLUCTUATIONS

*Correlation to atmospheric and oceanic temperature fluctuations*

To quantify the individual contributions from atmospheric and oceanic temperature fluctuations, I investigate their correlations with Arctic sea-ice variability (Figure 4.4). At the surface, the annual mean atmospheric temperature shows a negative correlation with annual mean sea-ice area, because sea ice and atmospheric temperature are tightly coupled within the Arctic boundary layer. The correlation weakens with height in all experiments (Figure 4.4a, black and colored solid lines). Here I only analyse the tropospheric temperature variability above the Arctic boundary layer height (Pavelsky et al. 2011) to avoid a direct influence of sea ice on atmospheric temperature. The tropos-

pheric mean temperature between 850 and 400hPa significantly correlates to sea-ice area at r = -0.49 in the interactive control experiment. By examining the spatial distribution of the correlation, I find that the negative correlation is largest at the sea-ice edge and weakened by positive correlations in the Central Arctic (Figure 4.4b). The presence of radiative effects, feedbacks, and forcings acting collectively strengthens the dominant direct link between mid-tropospheric temperature fluctuations and sea-ice variability (r = -0.49) by about one third compared to the experiment without interactive mechanisms (r = -0.34). The poleward oceanic heat transport strongly contributes to this strengthening, suggesting that oceanic heat released to the atmosphere sustains the strong link between sea-ice variability and atmospheric temper-

*... and its modulation by radiative effects, feedbacks, and forcings*

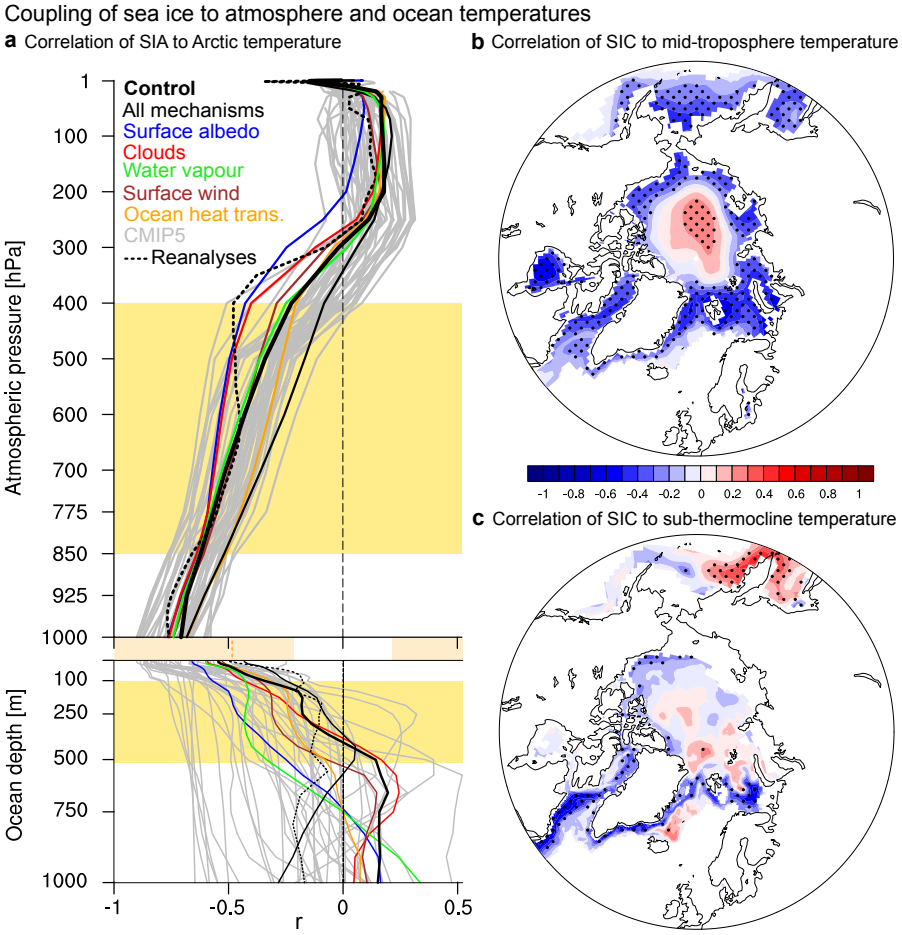



Figure 4.4: Connecting sea-ice variability to atmospheric and oceanic temperature fluctuations. **a** Correlation between annual Arctic sea-ice area (SIA) and annual Arctic atmospheric and oceanic temperature in MPI-ESM1.2-LR with (thick black line) and without (thin lines) interactive radiative effects, feedbacks, and forcings, in 41 CMIP5 models (grey lines, see Table C.1) and reanalyses (dashed line, see Section 4.3). Correlations are significant at 99.9% where marked with the beige bar for all experiments and CMIP5 simulations, and to the left of the beige dotted line for the 38-year long reanalyses. Note that the experiments are named after the individual non-interactive mechanism. Control-experiment correlation of interannual variability of sea-ice concentration (SIC) to **b** atmospheric and **c** oceanic temperature averaged across the mid-troposphere and the sub-thermocline ocean depth as indicated by the yellow boxes in **a**. Correlations significant at 99.9% are stippled.

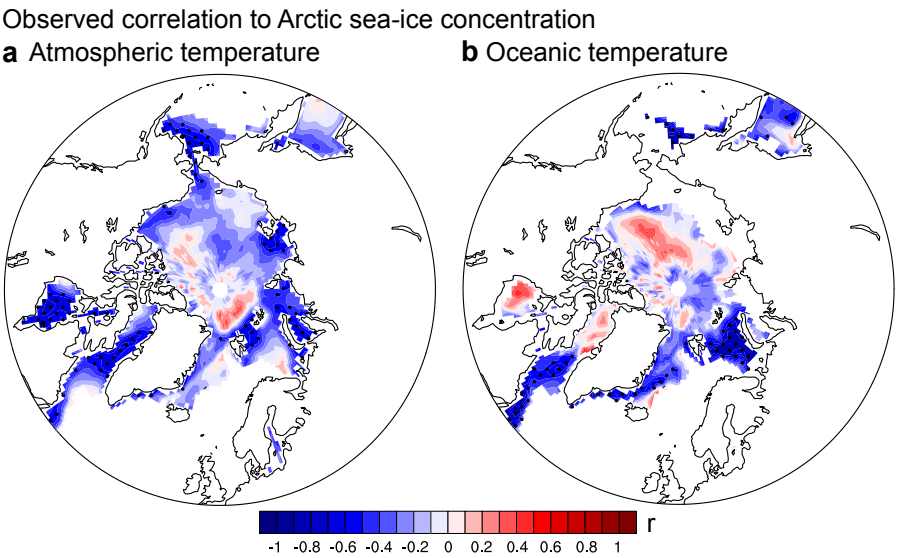



Figure 4.5: Observed correlation of sea-ice concentration to **a** mid-tropospheric temperature and **b** 100-500m depth ocean temperature. The NSIDC passive microwave sea-ice concentration (Fetterer et al. 2017), ERA-Interim reanalysis (Dee et al. 2011) and ORAS4 reanalysis (Balmaseda et al. 2013), which are used here, have been detrended by removing the least squares quadratic trend. Correlations significant at 99.9% are stippled.

ature fluctuations especially in the lower troposphere.

The oceanic temperature near the surface shows a significant negative correlation with annual mean sea-ice area. This significant correlation is unsurprising because sea ice and ocean temperature are tightly coupled above the thermocline. Below the thermocline, however, the oceanic temperature insignificantly correlates to sea-ice area in the control experiment. Regionally, significant negative correlations only occur near the winter sea-ice edge in the Atlantic sector (Figure 4.4c). Whereas an individual non-interactive radiative effect, feedback, or forcing strengthens the negative correlation below the thermocline, the joint impact of all mechanisms on the correlation between sub-thermocline ocean temperature and sea-ice area is small (Figure 4.4a).

*Correlation in MPI-ESM1.2-LR versus in CMIP5*

To test how representative the strength of the correlations is in MPI-ESM1.2-LR, I correlate the Arctic atmospheric and oceanic temperature with Arctic sea-ice area from the preindustrial control simulations for 41 CMIP5 models (Figure 4.4a, grey lines; Table C.1). I find that the CMIP5 models consistently show negative correlations between mid-tropospheric temperature and sea-ice area ranging from r = -0.67 to r = -0.29, and MPI-ESM1.2-LR represents the mean CMIP5 correlation strength (r = -0.48). I further find that the correlation between sub-thermocline ocean temperature and sea-ice area differs between the CMIP5 models (r = -0.75 to r = 0.09). While a few models show a strong negative correlation nearly independent of depth, most models agree with the expected weakening of the correlation strength with depth similar to the behaviour of MPI-ESM1.2-LR. I also find that MPI-ESM1.2-LR matches well the correlations between observations of sea-ice area and reanalyses of temperature especially in the lower and mid troposphere and the upper 400 m of ocean depth (Figure 4.4a, dotted line). Regionally, the correlation pattern

*Comparison to the observed correlation pattern*

of sea-ice concentration to mid-tropospheric temperature resembles the simulated pattern with negative correlations in regions of the sea-ice edge and positive correlations in the Central Arctic (Figure 4.5a and Figure 4.4b). The correlation pattern to sub-thermocline ocean temperature shows stronger negative correlations especially along the sea-ice edge compared to the simulated correlation strength (Figure 4.5b and Figure 4.4c). However, the timeseries for observed sea-ice area and temperature reanalyses are short and rarely show significant correlations. The representativity of both the mean Arctic correlation strengths and the correlation patterns modelled by MPI-ESM1.2-LR to the ones from reanalyses give me confidence that also the radiative effects, feedbacks, and forcings acting collectively in the real world might strengthen and weaken the negative correlations to atmospheric and oceanic temperatures as in MPI-ESM1.2-LR.

## 4.6 ORIGIN OF TROPOSPHERIC TEMPERATURE FLUCTUATIONS

*The poleward atmospheric energy transport*

The dominant role of tropospheric temperature fluctuations in driving sea-ice variability does not reveal the origin of these temperature fluctuations. However, the correlation field between surface pressure and mid-tropospheric temperature in the control experiment is similar to the pattern of the positive phase of the North Atlantic Oscillation, but shifted to the northeast (Figure 4.6). In the Norwegian-Greenland Sea a warm mid-troposphere is associated with a low surface pressure. The anticorrelation indicates baroclinic instability, which suggests that Arctic mid-tropospheric temperature is strongly influenced by poleward atmospheric energy transport in the Atlantic sector (Yang et al. 2010; Zhang 2015). This poleward transport of extratropical air

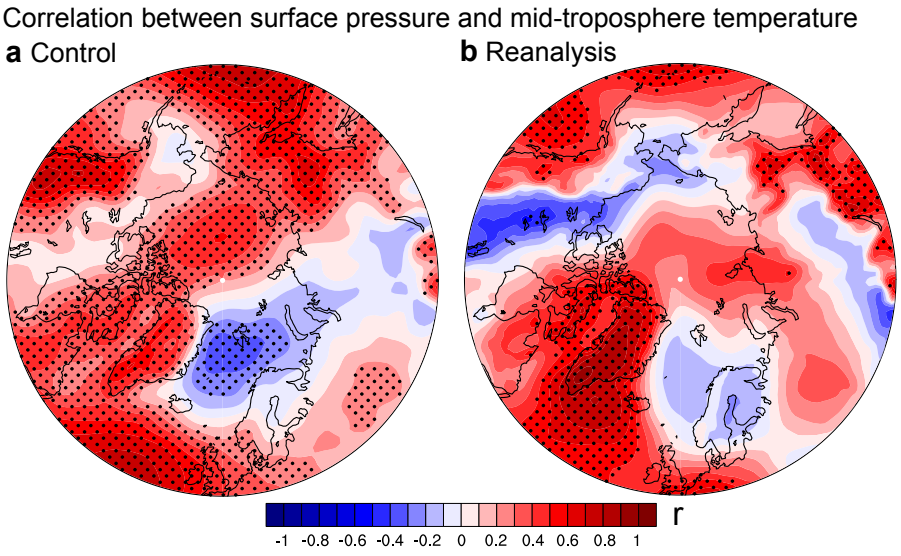



Figure 4.6: Correlation between interannual variability of surface pressure and mid-troposphere temperature. **a** 200-year control experiment. **b** 1979-2016 ERA-Interim reanalysis (Dee et al. 2011) with both surface pressure and mid-troposphere temperature detrended linearly. Correlations significant at 99.9% are stippled.

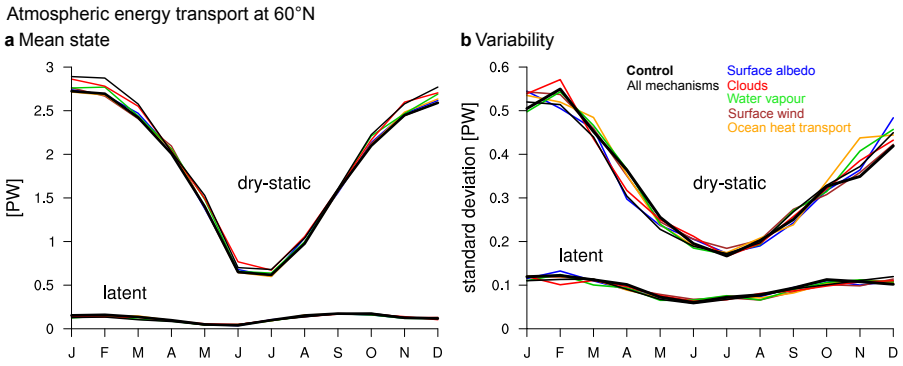



Figure 4.7: Atmospheric energy transport to the Arctic. **a** Mean state and, **b** variability of latent and dry-static atmospheric energy transport at 60°N with (thick black line) and without (thin lines) interactive radiative effects, feedbacks and forcings over the evolution of a year.

*... is insensitive to radiative effects, feedbacks and forcings*

masses is mainly accomplished by extratropical cyclones (Messori et al. 2017; Wernli and Papritz 2018). To examine whether the poleward atmospheric energy transport is affected by the decoupling of individual radiative effects, feedbacks and forcings, I calculate the atmospheric energy transport across 60° N for all experiments. To do so, I calculate the vertically integrated advection of dry-static and latent energy across 60° N as the sum of the contributions from transient eddies, the mean meridional circulation, and stationary eddies (Keith 1995). Surprisingly, I find no considerable difference in neither the mean state nor the variability of dry-static and latent atmospheric energy transport between the experiments (Figure 4.7a,b). I conclude that the preindustrial monthly mean poleward atmospheric energy transport is largely insensitive to the assessed mechanisms, in contrast to the impact of climate feedbacks on poleward atmospheric energy transport in a warming world (Zelinka and Hartmann 2011).

## 4.7 IMPLICATIONS

*... for possible tippings points*

The dominating impact of atmospheric temperature fluctuations on sea ice has fundamental implications for the ongoing scientific debates on possible tipping points in the sea-ice system and on the predictability of interannual sea-ice variability. Previous studies discussed whether radiative effects and feedbacks are responsible for possible tipping points of the Arctic sea-ice cover (e.g., Abbot et al. 2011; Bathiany et al. 2016; Notz 2009; Winton 2006). A tipping point might occur when a small change in forcing triggers a strongly nonlinear response in the internal dynamics of the sea-ice system, which requires processes and feedbacks that could generate such nonlinear dynamics (Lenton 2011). The minor quantitative relevance of the radiative effects and feedbacks in driving Arctic sea-ice variability described here thus counters the physical reasoning for possible feedback-driven tipping points in the sea-ice system and also explains why these mechanisms do not cause the abrupt Arctic sea-ice loss in winter (Bathiany et al. 2016). The strong direct thermal link to the atmosphere thus provides the mechanism for a reversible Arctic

sea-ice cover (e.g., Tietsche et al. 2011; Wagner and Eisenman 2015).

*... for interannual predictability*

The tight link between tropospheric temperature variability and sea-ice variability reasons the lack of predictive skill of Arctic sea-ice area on yearly or longer timescales (e.g., Guemas et al. 2016). On these timescales, atmospheric temperature fluctuations are not predictable due to the nonlinear and chaotic nature of the climate system (Kirtman et al. 2013). The control of sea-ice variability by atmospheric temperature might represent a predictability limit on Arctic sea-ice area (Serreze and Stroeve 2015) that should be further investigated.

The holistic quantifications of the dominant atmospheric temperature fluctuations and the minor radiative effects, feedbacks and forcings provide fundamental understanding of the origin of Arctic sea-ice variability. This understanding allows us to contextualise sea-ice fluctuations in the observed record and to interpret observed record-lows of Arctic sea-ice area as a response to an unusually warm atmosphere.

# 5 CONCLUSIONS

Internal climate variability (ICV) has long been seen as a troubling but unavoidable uncertainty for climate science. Inspired by the increased awareness of the importance of ICV in the climate system, I struck off to quantify and understand the chaotic nature of the Earth's climate. I limited my studies to a few characteristic variables, primarily sea-ice area and near-surface air temperature (SAT), but ensured a broad applicability of my methods. Faced with chaos to understand, I put emphasis on the robustness of my results and the road to get there. I structure my findings according to the eight guiding research questions of this dissertation. I consider these findings as first milestones on a long way we still have to go to understand the ICV on our changing planet. In the final section, I suggest pathways to further unravel processes and to reduce uncertainties in past, present and future estimates of ICV with the tools we now have at hand.

## 5.1 ESTIMATING INTERNAL CLIMATE VARIABILITY

In Chapter 2, I introduced a method that fills the gap to consistently estimate ICV for different forcing periods. Applying this method enables me to answer the following three questions:

**a. How can ICV be consistently estimated for a changing forcing?**

- To infer consistent estimates of ICV that are variable across different forcing periods, I developed a method that relates and thereby complements the existing two approaches based on preindustrial control simulations and single-model ensembles in a multi-model ensemble. The method exploits the quasi-ergodic assumption and hence can deal with a small number of ensemble simulations as provided for most CMIP5 models.
- Applying the method to a wide range of variables allows for valuable insights into how the ICV of the Earth is projected to change under climate change.

**b. How does internal variability of SAT, sea-ice area and sea-ice volume change under climate change?**

- The internal variability of annual globally-averaged SAT remains largely unchanged for historical simulations and might decrease for future simulations with large $CO_2$ forcing. Regionally, the projected changes reveal likely increases in temperature variability in the tropics, subtropics, and polar regions, and extremely likely decreases in mid latitudes.

- The internal variability of sea-ice volume and area remains largely unchanged for historical simulations and decreases likely or extremely likely and proportionally to their mean state under large $CO_2$ forcing. The relationship does not hold for future Arctic sea-ice area, which shows no consistent change across models.

**c. How can climate-model simulations be robustly evaluated?**

- A robust evaluation of climate-model simulations requires to account for at least two uncertainties, namely the simulated ICV of a model and the observational uncertainty. Model-specific estimates of ICV are crucial to consistently evaluate the plausibility of climate-model simulations.
- For sea-ice simulations, the robust evaluation confirms that internal variability can explain most of the models' deviation from observed trends but often not the models' deviation from the observed mean states.

## 5.2 COPING WITH INTERNAL CLIMATE VARIABILITY

In Chapter 3, I quantified the contributions to the future probability of climate extremes from changes in the mean and from changes in higher-order moments. Applying my method from Chapter 2 and an empirical threshold approach to CMIP5 simulations of daily SAT and daily total precipitation allows me to answer the following questions:

**a. How are the background conditions for climate extremes, i.e. the mean and the internal variability of daily SAT and daily total precipitation projected to change?**

- For large $CO_2$ forcing, the future mean state of daily globally-averaged SAT increases extremely likely. This increase is accompanied by an extremely likely decrease in its future internal variability. The decrease primarily originates from the high latitudes that are projected to warm strongest.
- The future mean amount of daily globally-averaged total precipitation increases extremely likely. This increase is accompanied by an extremely likely increase in its future internal variability. The projected increases primarily originate from the tropics and mid latitudes that have large and highly variable precipitation rates, and are opposed by decreases in many subtropical regions, primarily the eastern ocean basins, that have low and little variable precipitation rates.

**b. What is the future probability of climate extremes under strong global warming?**

- Hot extremes will become much more likely everywhere, but particularly in the tropics that have a large signal-to-noise ratio. In some tropical regions, all daily temperatures exceed in the future the temperature that defined a hot extreme in the past, representing a 20-fold increase in probability. The increased probability in mid and high latitudes is substantial, but less strong than in the tropics.
- Cold extremes are projected to vanish in most parts of the world, but remain possible in some mid- and high-latitude regions.
- Heavy precipitation extremes are projected to become up to twice as likely primarily in some tropical and high-latitude regions and much less likely in the subtropics, in particular in the eastern ocean basins, Central America and the Mediterranean region.
- Changes in the future probability of the 1% most extreme temperature and precipitation events are even stronger than for the 5% most extreme events.

**c. How much of the future probability of climate extremes can be attributed to a shift in the mean, and how much to changes in higher-order moments?**

- The future increased probability of hot extremes and the future decreased probability of cold extremes is primarily driven by the strong shift towards higher temperatures. The projected changes in higher-order moments are secondary, but counteract the increased future probability of hot extremes in high latitudes, strengthen the increased future probability of hot extremes in many mid latitudes, and reinforce the decreased future probability of cold extremes in mid- to high latitudes.
- In contrast to temperature extremes, the projected changed probability of heavy precipitation extremes can be primarily attributed to changes in higher-order moments. Slight shifts in the mean towards higher precipitation amounts in some tropical and high-latitude regions and lower precipitation amounts in most subtropical and mid-latitude regions are secondary, but mostly reinforce the changes caused by higher-order moments.

## 5.3 CAUSES OF INTERNAL CLIMATE VARIABILITY

In Chapter 4, I examined possible drivers of Arctic sea-ice variability. Enabled by substantial computational power, a sophisticated approach, and an Arctic-wide perspective, I challenged previous findings and filled the following knowledge gaps:

**a. Which mechanisms primarily drive the variability in Arctic sea-ice area?**

- Most of the substantial year-to-year variability of Arctic sea-ice area is directly driven by atmospheric temperature fluctuations. This finding opposes the wide-spread belief that intricate atmospheric or oceanic effects and feedbacks are important drivers of variability. I find that they explain only 25% of the Arctic sea-ice variability.
- Atmospheric temperature fluctuations drive sea-ice variability across the whole Arctic Ocean but especially at the sea-ice edge, whereas the impact of oceanic temperature fluctuations on sea ice is mainly limited to the Atlantic sea-ice edge. The CMIP5 models and reanalyses agree on the correlation of sea-ice variability to tropospheric temperature fluctuations, while they disagree on the correlation to fluctuations of sub-termocline ocean temperature.

**b. What does the relative role of driving mechanisms imply for the functioning of the Arctic climate system?**

- The major impact of atmospheric temperature fluctuations and the minor role of radiative effects and feedbacks in driving Arctic sea-ice variability counters the physical reasoning for possible feedback-driven tipping points in the sea-ice system. The tight link of sea ice to atmospheric temperatures thus provides the mechanism for a reversible Arctic sea-ice cover.
- The tight link between tropospheric temperature variability and sea-ice variability reasons the lack of predictive skill of Arctic sea-ice area on yearly and longer timescales. On these timescales, atmospheric temperature fluctuations are not predictable and thus set a natural limit to seasonal sea-ice predictability.

Overall, I showed in this dissertation that ensemble simulations within CMIP5 are a powerful tool to study changes and impacts of ICV over time. I hence conclude that for applications like the ones presented here, a few ensemble simulations from multiple models in higher temporal and/or spatial resolution can be more valuable than many ensemble simulations from a single model in lower resolution, because the results are less affected by model biases and therefore more robust.

I further conclude that obtaining a mechanistic understanding of the chaotic nature of the climate system is largely facilitated, and is in fact made possible by the growing resources provided by high-performance computing. Climate modelling is hence an indispensable prerequisite for a proper quantification and understanding of ICV.

Finally, I conclude that the quantification of ICV as a major component of climate variability is key for understanding the relative role and the respective interactions between climate variability and anthropogenic climate change. I

consider further basic research on these complex interactions fundamental for understanding the profound changes that we impose on the Earth's climate system.

## 5.4 THE THREE GOALS OF THIS DISSERTATION

I would have loved to study many more facets of ICV within the Earth system because plenty of questions remain. But science would not be science if we could answer all questions at once. Instead, science is like doing a jigsaw: the full picture grows by adding puzzle pieces.

For me, this dissertation has three goals: The first goal is to add one puzzle piece to our understanding of ICV mainly for the polar regions, but also worldwide, and to thereby become a scientist who scrutinizes findings, develops ideas and who is guided by questions and curiosity.

The second goal is to provide others with new approaches, new knowledge and new ideas that enable them to add further pieces to our picture of the Earth's climate system. Puzzle pieces that I have in mind for which my applied methodologies might be fruitful are the ICV of the Southern Ocean, a region of substantial uncertainty within the climate system, and paleoclimate applications that might open pathways to find robust numbers on the Earth system's variability. Other pieces might add to our understanding of the ICV of the Atlantic Meridional Overturning Circulation or of major variability modes like the El Niño Southern Oscillation and the North Atlantic Oscillation. Complementary to applying the methods used here to other climate variables and phenomena, I expect the joint exploitation of the existing single-model large ensembles to provide further new insights on the changing ICV of our changing planet. There are still plenty of mysteries to disclose.

The third goal is to raise our awareness of the profound changes we impose on the Earth system already now but especially in future if we continue to emit as many greenhouse gases as we do today. Although not all questions in climate science are answered, we have enough knowledge about the far-reaching consequences of this large-scale real-world experiment to act. Beyond the target in global climate policy to limit global warming to below 2°C, the findings of this dissertation caution not to underestimate the future consequences induced by changes in extreme events and climate variability.

# A

# APPENDIX TO CHAPTER 2

## A.1 TRADE-OFF BETWEEN ENSEMBLE SIZE AND TIME-AVERAGING LENGTH

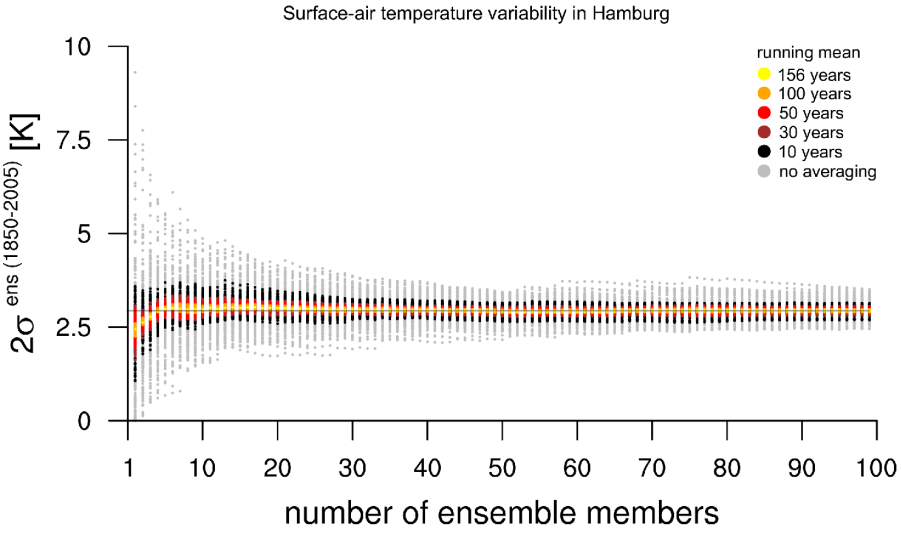



Figure A.1: Ensemble standard deviation of SAT for the grid cell Hamburg calculated for varying ensemble size and time-averaging length from all possible consecutive combinations of the 100 historical simulations from MPI-ESM1.1-LE. The horizontal line marks the best estimate of simulated variability inferred from all 100 ensemble simulations and the full time-averaging length. The spread of estimates on the right tail is conservative because the larger the ensemble sizes, the more the combinations suffer from resampling of ensemble members.

## A.2 REFERENCE DATA AND UNCERTAINTIES

### A.2.1 *Sea-ice volume*

As reference data for the evaluation of Northern Hemisphere sea-ice volume, I use reanalysis data from the Pan-Arctic Ice-Ocean Modelling and Assimilation System (PIOMAS) (Zhang and Rothrock 2003) that cover the period from 1979 to today. PIOMAS is considered as useful for climate-model evaluation (Laxon et al. 2013) as the sea-ice – ocean model a) assimilates sea-ice concentrations from satellite retrievals and is forced by NCEP atmospheric reanalysis data and b) simulates a sea-ice thickness estimate that agrees with past and recent airborne and in-situ point measurements as well as with recent satellite measurements of ICESat (Kwok and Rothrock 2009; Schweiger et al. 2011) and complemented data of CryoSat (Laxon et al. 2013). In a detailed assessment of PIOMAS March sea-ice thickness including additional

satellite, submarine and mooring data, Stroeve et al. (2014) confirmed that PIOMAS is suitable for model evaluation of long-term trends.
Schweiger et al. (2011) discussed the uncertainties in PIOMAS sea-ice volume and provided conservative uncertainty estimates for March and October sea-ice volume and sea-ice volume trends. Based on model sensitivity studies, they stated an uncertainty of the 32-year trend in sea-ice volume of 1.0 $\cdot 10^3$ km$^3$ decade$^{-1}$ and a conservative uncertainty range of 2.25 $\cdot 10^3$ km$^3$ for the mean state of sea-ice volume in March and 1.35 $\cdot 10^3$ km$^3$ in October. I interpolate the uncertainty ranges for the other months by weighting them with the monthly mean sea-ice volume averaged over the period 1979-2010. To reach a smooth curve of monthly uncertainty estimates that is fixed at the March value of 2.25 $\cdot 10^3$ km$^3$ and inspired by the mean seasonal cycle of Arctic sea-ice volume, the July - December values are increased by a factor of 1.12 to 1.25. Adapting the uncertainty estimate in summer rather than in winter is justified by the uncertainty related to melt ponds in the sea-ice concentration products from satellite retrievals that are assimilated to PIOMAS. These monthly uncertainty estimates define the uncertainty ($\delta_{\mathrm{ref}}$) used when evaluating Northern Hemisphere sea-ice volume.

### A.2.2 *Sea-ice area*

As reference data for the evaluation of modeled sea-ice area, I use satellite retrievals of sea-ice concentration. The sea-ice concentration data product used here is the "Climate Data Record of Passive Microwave Sea Ice Concentration" (CDR, Meier 2013). The CDR combines different satellite algorithms that, when applied individually, result in different estimates of sea-ice concentration dependent on the applied transfer function that translates the passive-microwave signature into sea-ice concentration. The reliability of satellite retrievals based on a single algorithm is questioned mainly because of the different treatment of the impact of surface properties (e.g., Lindsay et al. 2014; Titchner and Rayner 2014). The CDR aims to reduce the uncertainty originating from the use of specific algorithms. Therefore, I consider the CDR time series as a best estimate of the 'true' evolution of sea-ice concentration.
To account for the area around the North Pole that is not covered by satellite data, I fill this data hole following the procedure by Olason and Notz (2014). The first satellite observations from 1979 to August 1987 only reached 84.5°N. However for this period, filling the data hole with a sea-ice concentration of 1 is reasonable because the latitudes to the south show a constantly dense sea-ice concentration as well. This assumption does not hold for the period from August 1987 onwards although the observations now reach 87.2°N. The sea-ice concentration starts to become too variable in the central Arctic. I therefore use the mean concentration of the outer rim of the large pre-1987 data hole (i.e., between 84.5°N and 87.2°N) to fill the remaining post-1987 data hole of sea-ice concentration.
As uncertainty estimates ($\delta_{\mathrm{ref}}$) for sea-ice area, I use the standard deviation

estimate provided by the CDR. Note that this standard deviation estimate is only available since August 1987.

## A.3 ABSOLUTE CHANGE IN VARIABILITY OF NEAR-SURFACE AIR TEMPERATURE

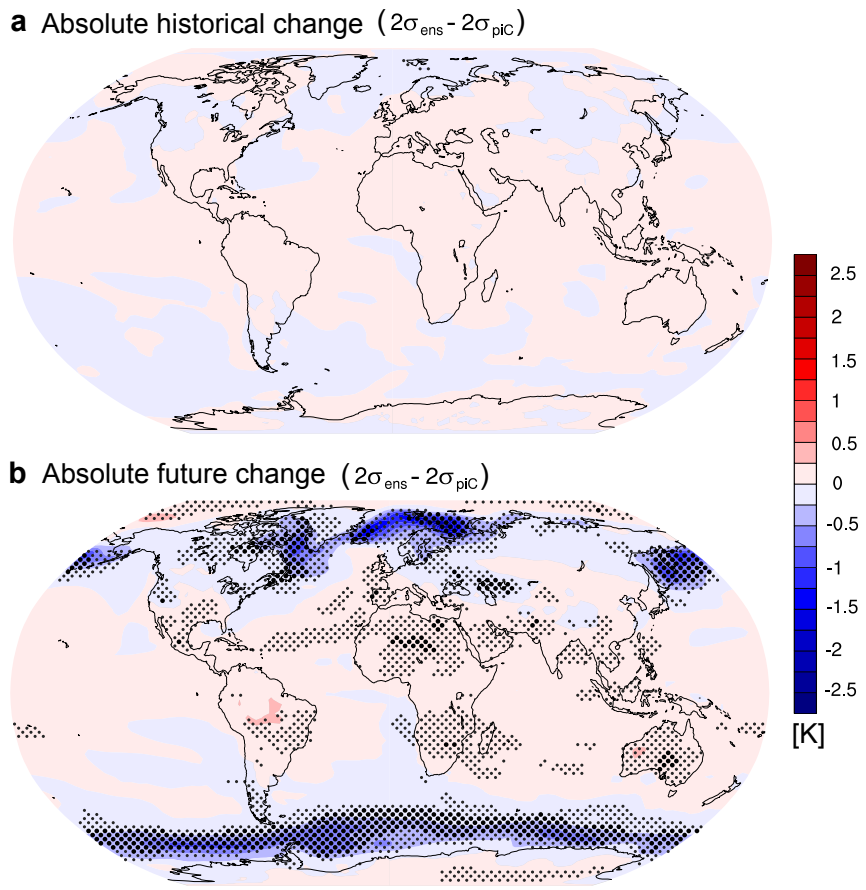



Figure A.2: Regional changes in variability of SAT. **a** The absolute change between the preindustrial SAT variability and the historical SAT variability, and **b** the absolute change between the preindustrial SAT variability and the SAT variability of a future climate forced by the RCP8.5 emission scenario. A possible increase (red shades) and a possible decrease (blue shades) in SAT variability, and likely changes (light stippling) and extremely likely changes (strong stippling) are shown.

# B

# APPENDIX TO CHAPTER 3

## B.1 MEAN STATE AND INTERNAL CLIMATE VARIABILITY IN 1950-2016 AND 2034-2100

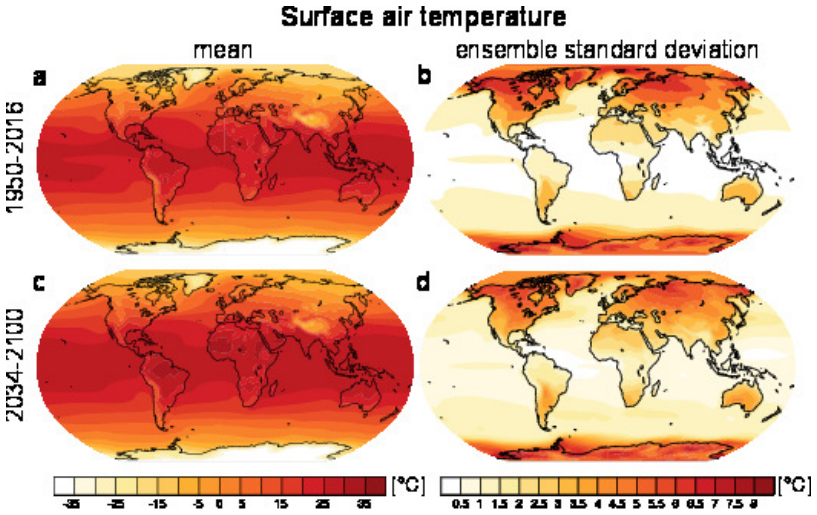



Figure B.1: Multi-model mean of the mean state and the internal variability, as measured by the ensemble standard deviation, of daily SAT for **a** and **b** 1950-2016 and **c** and **d** 2034-2100.

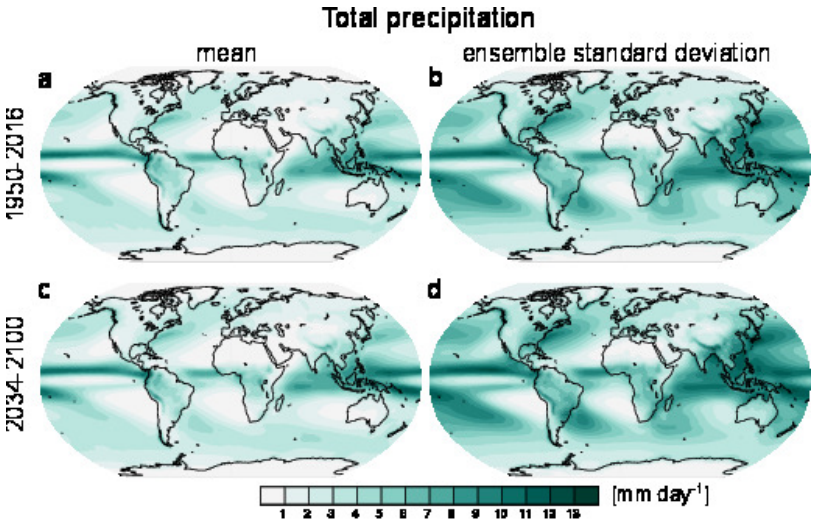



Figure B.2: Multi-model mean of the mean state and the internal variability, as measured by the ensemble standard deviation, of daily total precipitation for **a** and **b** 1950-2016 and **c** and **d** 2034-2100.

## B.2 ALTERNATIVE APPROACH TO QUANTIFY THE CONTRIBUTION FROM HIGHER-ORDER MOMENTS WITHOUT A WARMING SHIFT

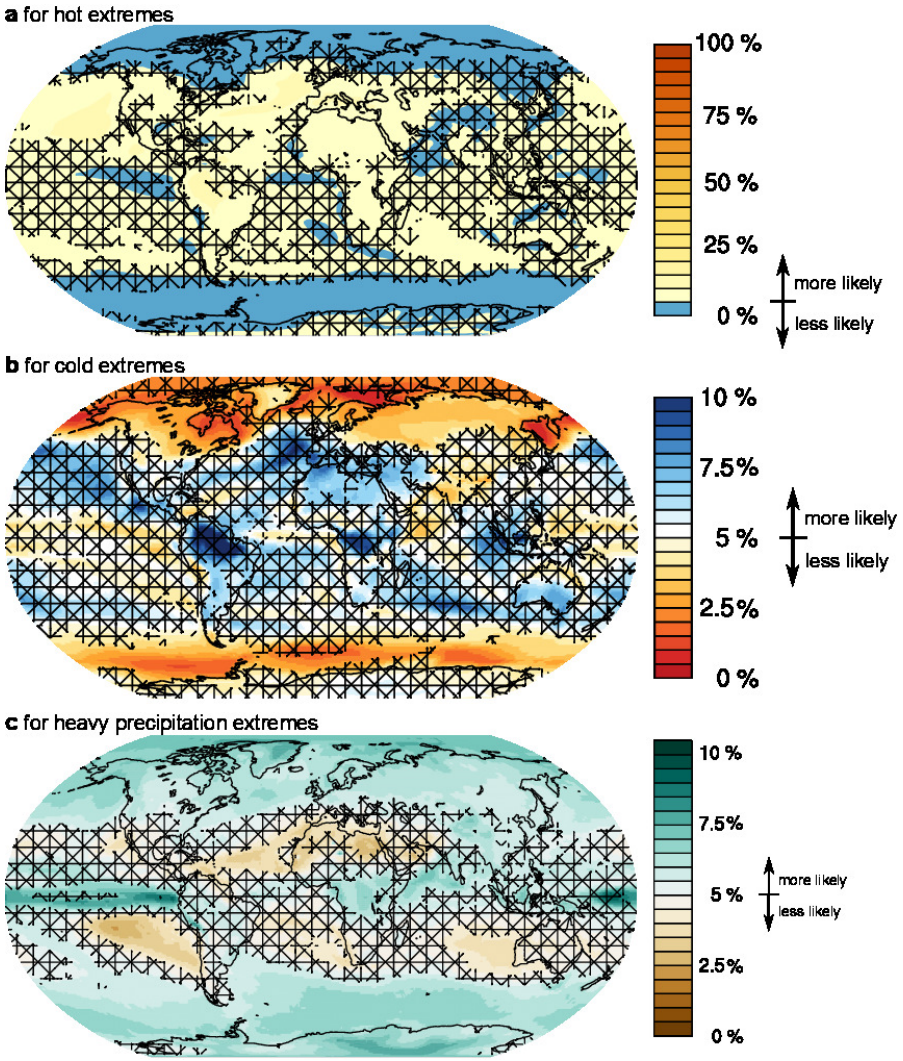



Figure B.3: Alternative approach to quantify the contribution from changes in higher-order moments without a concurrent warming shift to the future probability of **a** hot extremes, **b** cold extremes, and **c** heavy precipitation extremes, averaged across models. Crosshatching marks regions where the projected change is not robust across models.

## B.3 QUANTIFYING AND ATTRIBUTING THE FUTURE PROBABILITY OF THE 1% MOST EXTREME EVENTS

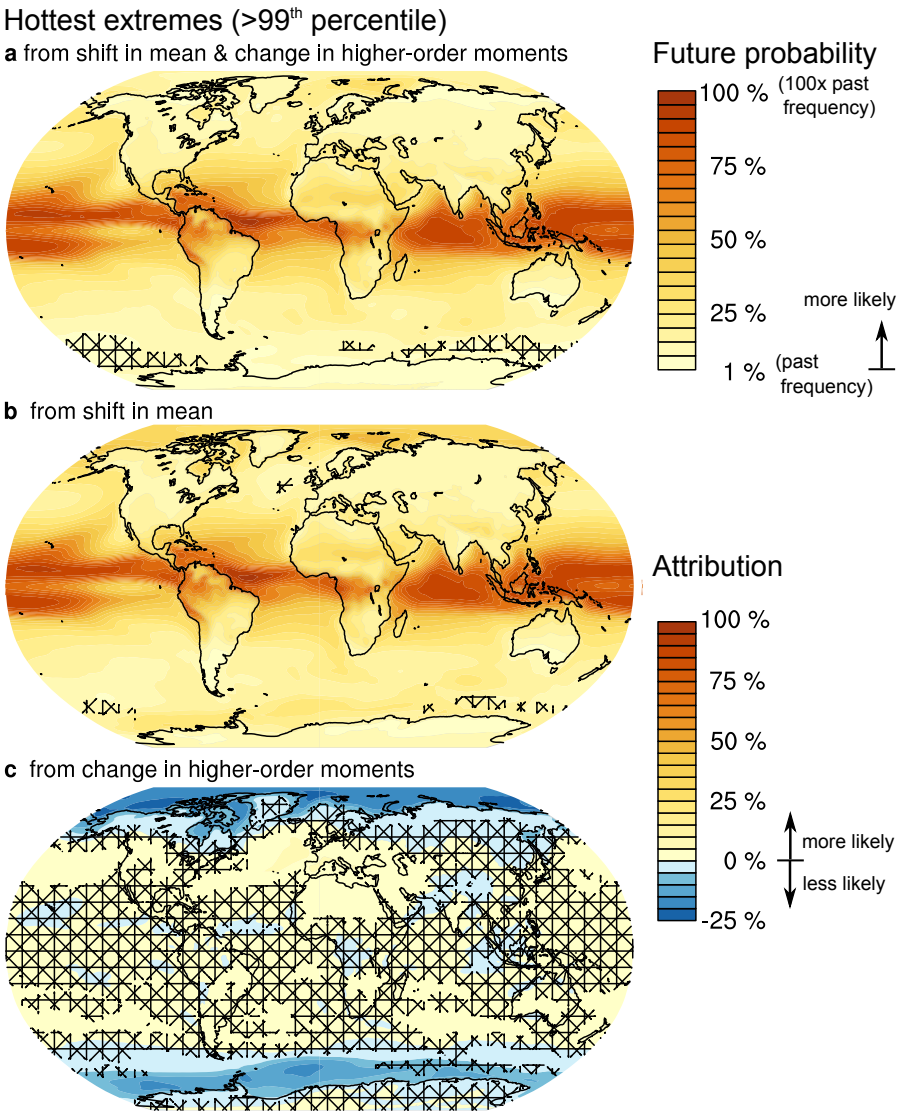



Figure B.4: Attribution of the future probability of the 1% hottest extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of 1% hottest extremes from the full change in the PDF of daily SAT between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

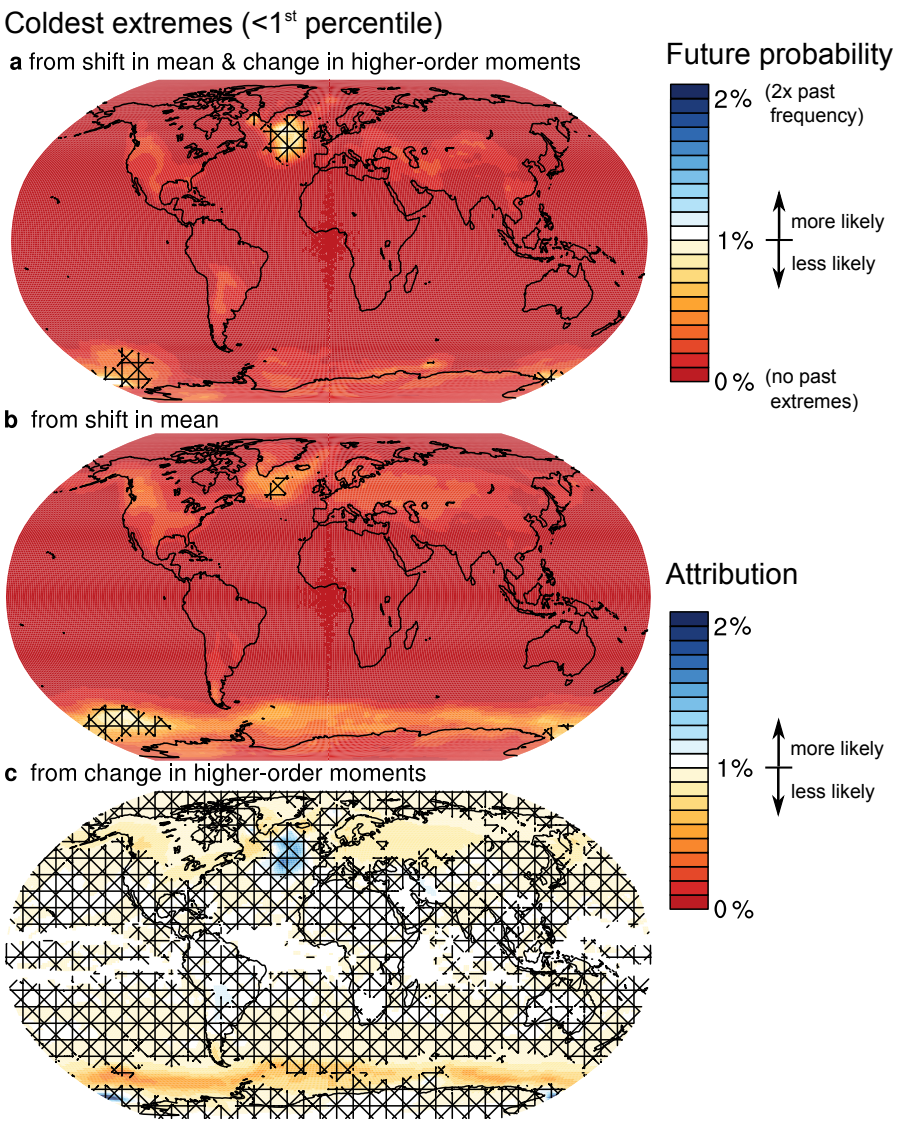



Figure B.5: Attribution of the future probability of the 1% coldest extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of the 1% coldest extremes from the full change in the PDF of daily SAT between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

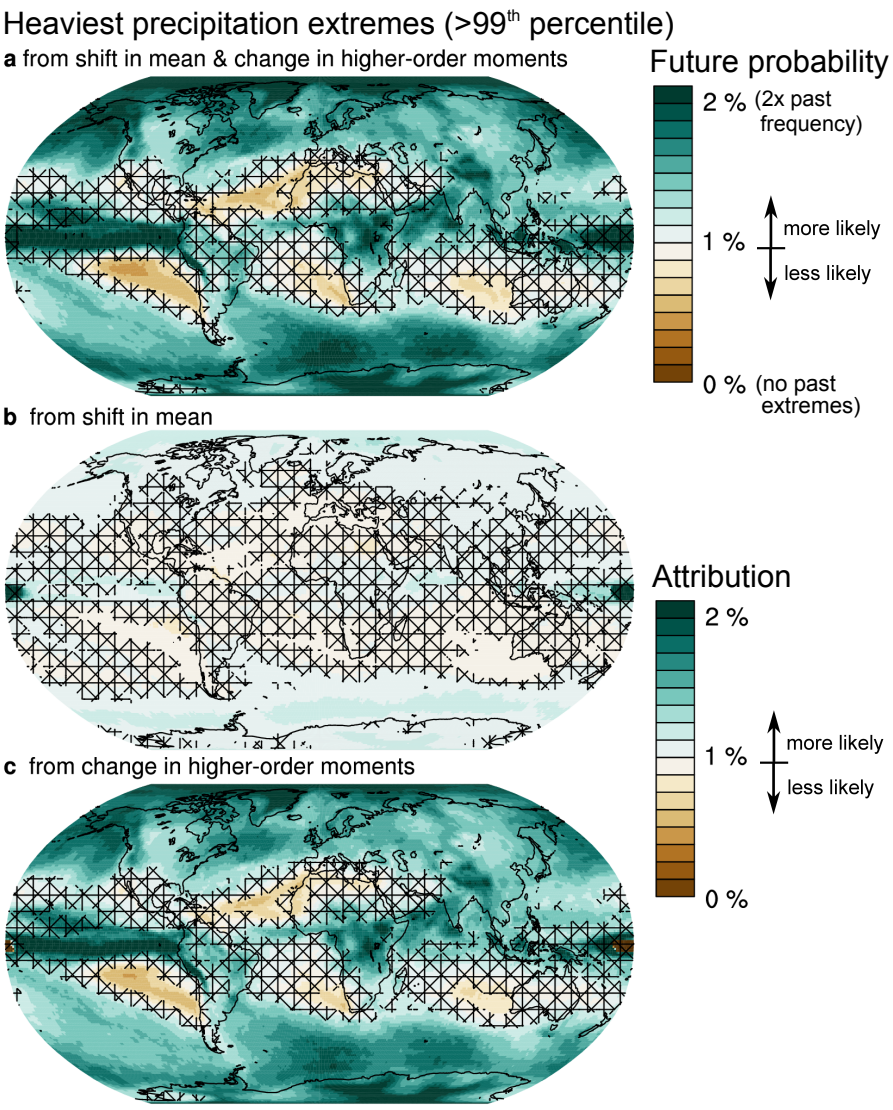



Figure B.6: Attribution of the future probability of the 1% heaviest precipitation extremes to the shift in the mean and to changes in higher-order moments, averaged across models. **a** The future probability of the 1% heaviest precipitation extremes from the full change in the PDF of daily total precipitation between 1950-2016 and 2034-2100 is decomposed into the contribution from **b** the shift in the mean and from **c** the change in higher-order moments of the distribution. Crosshatching marks regions where the projected change is not robust across models.

# C

# APPENDIX TO CHAPTER 4

## C.1 IMPACT OF EACH NON-INTERACTIVE RADIATIVE EFFECT, FEEDBACK OR FORCING

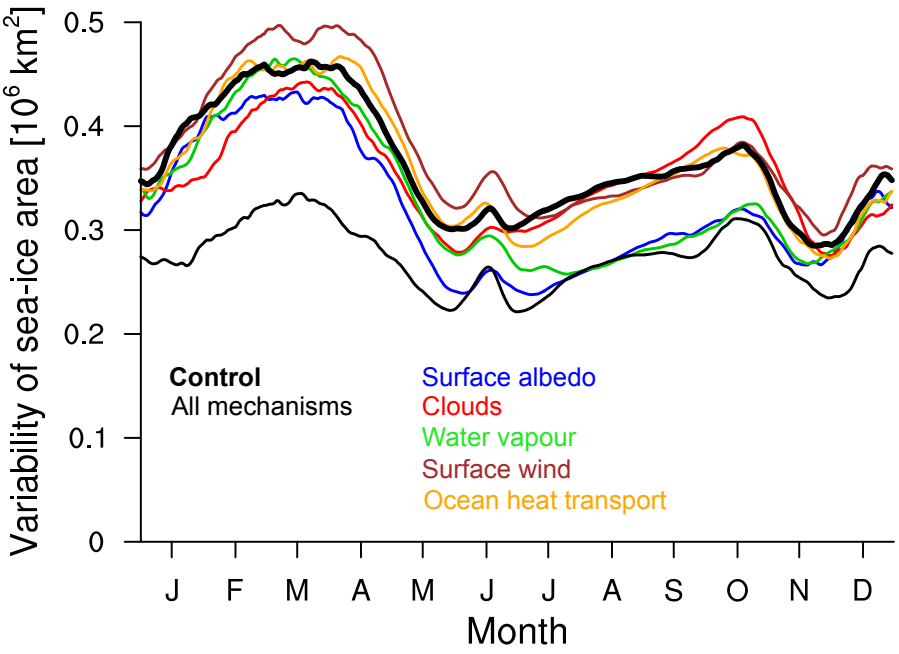



Figure C.1: Impact of each non-interactive radiative effect, feedback or forcing on the multi-year daily standard deviation of Arctic sea-ice area over the course of the year. The effect of a single non-interactive mechanism (thin lines, see legend) is compared to the fully interactive control experiment and the experiment in which all mechanisms are non-interactive as displayed in Figure 4.2a. The yearly averaged impact is shown in Figure 4.2b.

## C.2 ABSOLUTE VARIABILITY OF SEA-ICE CONCENTRATION

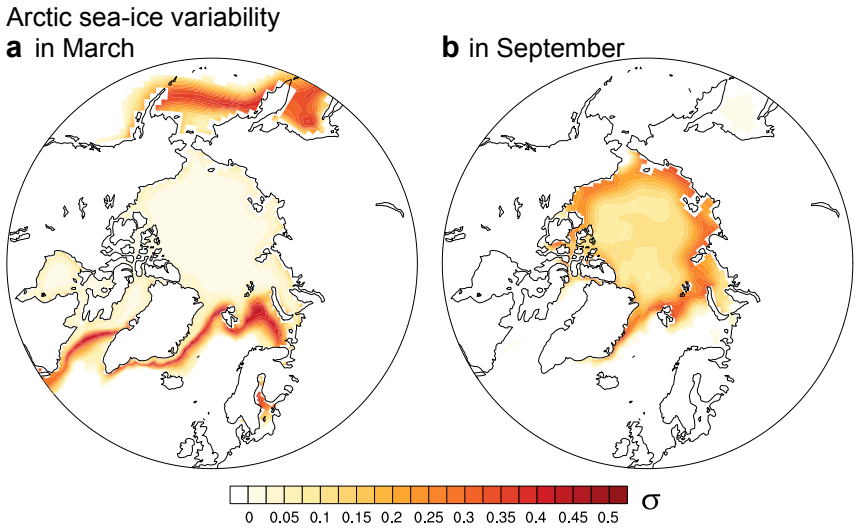



Figure C.2: Variability of sea-ice concentration in the control experiment for **a** March and **b** September.

## C.3 CHANGES IN VARIABILITY OF SEA-ICE CONCENTRATION FOR ALL EXPERIMENTS

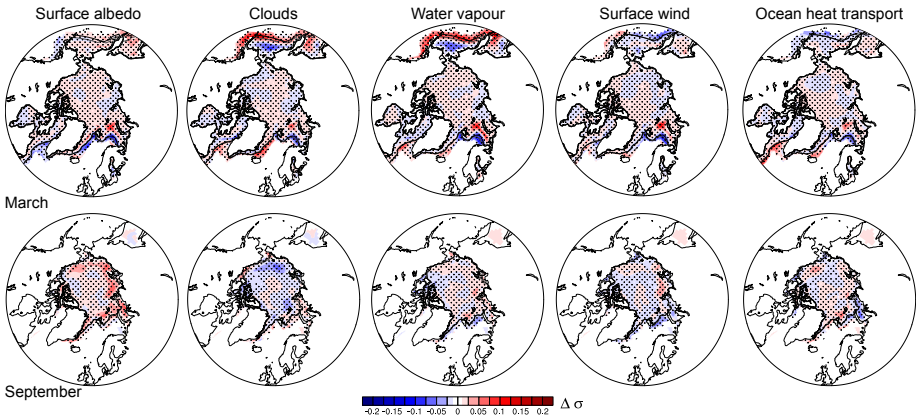



Figure C.3: Regional changes in variability of sea-ice concentration in March (top row) and September (bottom row) for all experiments with a non-interactive radiative effect, feedback or forcing relative to the control experiment (complementary to Figure 4.3). An increased variability of sea-ice concentration caused by the assessed mechanisms is marked in red, while a decreased variability is shown in blue. Changes significant at 99.9% are stippled and changes smaller than 1% are not shown. The sea-ice edges at 15% concentration displayed as black and grey lines for the control experiment and each the named experiment, respectively, are very similar.

Table C.1: CMIP5 models used.

| Model name | Model center | Length of preindustrial control simulation [years] |
|---|---|---|
| ACCESS1.0 | Commonwealth Scientific and Industrial Research Organization, Bureau of Meteorology, Australia | 500 |
| ACCESS1.3 | Commonwealth Scientific and Industrial Research Organization, Bureau of Meteorology, Australia | 500 |
| BCC-CSM1.1 | Beijing Climate Center, China | 500 |
| BCC-CSM1.1(m) | Beijing Climate Center, China | 500 |
| BNU-ESM | College of Global Change and Earth System Science, Beijing Normal University | 559 |
| CanESM2 | Canadian Centre for Climate, Canada | 1096 |
| CCSM4 | National Center for Atmospheric Research, USA | 1051 |
| CESM1(BGC) | National Center for Atmospheric Research, USA | 500 |
| CESM1(CAM5) | National Center for Atmospheric Research, USA | 319 |
| CESM1(FASTCHEM) | National Center for Atmospheric Research, USA | 222 |
| CESM1(WACCM) | National Center for Atmospheric Research, USA | 200 |
| CMCC-CESM | Centro Euro-Mediterraneo per i Cambiamenti, Italy | 277 |
| CMCC-CM | Centro Euro-Mediterraneo per i Cambiamenti, Italy | 330 |
| CMCC-CMS | Centro Euro-Mediterraneo per i Cambiamenti, Italy | 500 |
| CNRM-CM5 | Centre National de Recherches Météorologiques, France | 850 |
| CNRM-CM5.2 | Centre National de Recherches Météorologiques, France | 140 |
| CSIRO Mk3.6.0 | Commonwealth Scientific and Industrial Research Organization, Bureau of Meteorology, Australia | 500 |
| FGOALS-g2.0 | Institute of Atmospheric Physics, CAS, China | 700 |
| FGOALS-s2.0 | Institute of Atmospheric Physics, CAS, China | 501 |
| FIO-ESM | The First Institute of Oceanography, SOA, China | 800 |
| GFDL-CM3 | Geophysical Fluid Dynamics Laboratory, USA | 500 |
| GFDL-ESM2G | Geophysical Fluid Dynamics Laboratory, USA | 500 |
| GFDL-ESM2M | Geophysical Fluid Dynamics Laboratory, USA | 500 |
| GISS-E2-H-CC | Goddard Institute for Space Studies, USA | 251 |
| GISS-E2-R-CC | Goddard Institute for Space Studies, USA | 251 |
| HadCM3 | Met Office Hadley Centre, UK | 1200 |
| HadGEM2-CC | Met Office Hadley Centre, UK | 240 |
| HadGEM2-ES | Met Office Hadley Centre, UK | 175 |
| INM-CM4.0 | Institute for Numerical Mathematics, Russia | 500 |
| IPSL-CM5A-LR | Institut Pierre Simon Laplace, France | 1000 |
| IPSL-CM5A-MR | Institut Pierre Simon Laplace, France | 269 |
| IPSL-CM5B-LR | Institut Pierre Simon Laplace, France | 300 |
| MIROC5 | Model for Interdisciplinary Research On Climate, Japan | 670 |
| MIROC-ESM | Model for Interdisciplinary Research On Climate, Japan | 630 |
| MIROC-ESM-CHEM | Model for Interdisciplinary Research On Climate, Japan | 255 |
| MPI-ESM-LR | Max Planck Institute for Meteorology, Germany | 1000 |
| MPI-ESM-MR | Max Planck Institute for Meteorology, Germany | 1000 |
| MPI-ESM-P | Max Planck Institute for Meteorology, Germany | 1156 |
| MRI-CGCM3 | Meteorological Research Institute, Japan | 500 |
| NorESM1-M | Norwegian Climate Center, Norway | 501 |
| NorESM1-ME | Norwegian Climate Center, Norway | 252 |

# BIBLIOGRAPHY


Abbot, D. S., M. Silber and R. T. Pierrehumbert (2011). 'Bifurcations leading to summer Arctic sea ice loss'. *Journal of Geophysical Research: Atmospheres* 116 (D19), p. D19120. DOI: 10.1029/2011JD015653.

Allan, R. P. and B. J. Soden (2008). 'Atmospheric Warming and the Amplification of Precipitation Extremes'. *Science* 321.5895, pp. 1481–1484. DOI: 10.1126/science.1160787.

Årthun, M. et al. (2012). 'Quantifying the Influence of Atlantic Heat on Barents Sea Ice Variability and Retreat'. *Journal of Climate* 25.13, 4736–4743. DOI: 10.1175/JCLI-D-11-00466.1.

Ballester, J., F. Giorgi and X. Rodó (2010). 'Changes in European temperature extremes can be predicted from changes in PDF central statistics'. *Climatic Change* 98.1, p. 277. DOI: 10.1007/s10584-009-9758-0.

Balmaseda, M. A., K. Mogensen and A. T. Weaver (2013). 'Evaluation of the ECMWF ocean reanalysis system ORAS4'. *Quarterly Journal of the Royal Meteorological Society* 139.674, pp. 1132–1161. DOI: 10.1002/qj.2063.

Banerjee, A., L. M. Polvani and J. C. Fyfe (2017). 'The United States "warming hole": Quantifying the forced aerosol response given large internal variability'. *Geophysical Research Letters* 44.4, 2016GL071567. DOI: 10.1002/2016GL071567.

Barrow, E. M. and M. Hulme (1996). 'Changing probabilities of daily temperature extremes in the UK related to future global warming and changes in climate variability'. *Climate Research* 6.1, pp. 21–31.

Bathiany, S. et al. (2016). 'On the Potential for Abrupt Arctic Winter Sea Ice Loss'. *Journal of Climate* 29.7, pp. 2703–2719. DOI: 10.1175/JCLI-D-15-0466.1.

Bathiany, S. et al. (2016). 'Statistical indicators of Arctic sea-ice stability - prospects and limitations'. *Cryosphere* 10.4, 1631–1645. DOI: 10.5194/tc-10-1631-2016.

Bitz, C. and G. Roe (2004). 'A mechanism for the high rate of sea ice thinning in the Arctic Ocean'. *Journal of Climate* 17.18, 3623–3632. DOI: 10.1175/1520-0442(2004)017<3623:AMFTHR>2.0.CO;2.

Bradley, R. S. and P. D. Jones (1992). *Climate Since A.D. 1500*. Ed. by R. S. Bradley and P. D. Jones. Routledge, London, 606–622.

Briffa, K. R. et al. (1998). 'Influence of volcanic eruptions on Northern Hemisphere summer temperature over the past 600 years'. *Nature* 393.6684, 450–455. DOI: 10.1038/30943.

Brown, P. T. et al. (2017). 'Change in the magnitude and mechanisms of global temperature variability with warming'. *Nature Climate Change* 7.10, pp. 743–748. DOI: 10.1038/nclimate3381.

Bunzel, F. et al. (2016). 'Seasonal climate forecasts significantly affected by observational uncertainty of Arctic sea ice concentration'. *Geophysical Research Letters* 43.2, pp. 852–859. DOI: 10.1002/2015GL066928.

Cavanaugh, N. R. et al. (2015). 'The probability distribution of intense daily precipitation'. *Geophysical Research Letters* 42.5, pp. 1560–1567. DOI: 10.1002/2015GL063238.

Collins, M. et al. (2013). 'Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA'.

Coumou, D. and S. Rahmstorf (2012). 'A decade of weather extremes'. *Nature Climate Change* 2, 491–496. DOI: 10.1038/NCLIMATE1452.

Curry, J. A. et al. (1995). 'Water vapor feedback over the Arctic Ocean'. *Journal of Geophysical Research: Atmospheres* 100 (D7), 14223–14229. DOI: 10.1029/95JD00824.

Dee, D. P. et al. (2011). 'The ERA-Interim reanalysis: configuration and performance of the data assimilation system'. *Quarterly Journal of the Royal Meteorological Society* 137.656, pp. 553–597. DOI: 10.1002/qj.828.

Deser, C, J. Walsh and M. Timlin (2000[a]). 'Arctic sea ice variability in the context of recent atmospheric circulation trends'. *Journal of Climate* 13.3, 617–633. DOI: 10.1175/1520-0442(2000)013<0617:ASIVIT>2.0.CO;2.

Deser, C., J. E. Walsh and M. S. Timlin (2000[b]). 'Arctic Sea Ice Variability in the Context of Recent Atmospheric Circulation Trends'. *Journal of Climate* 13.3, 617–633. DOI: 10.1175/1520-0442.

Deser, C. et al. (2012). 'Uncertainty in climate change projections: the role of internal variability'. *Climate Dynamics* 38.3-4, 527–546. DOI: 10.1007/s00382-010-0977-x.

Deser, C. et al. (2014). 'Projecting North American climate over the next 50 years: Uncertainty due to internal variability'. *Journal of Climate* 27.6, 2271–2296. DOI: 10.1175/JCLI-D-13-00451.1.

Ding, Q. et al. (2017). 'Influence of high-latitude atmospheric circulation changes on summertime Arctic sea ice.' *Nature Climate Change* 7.4, 289–295. DOI: 10.1038/nclimate3241.

Donat, M. G. and L. V. Alexander (2012). 'The shifting probability distribution of global daytime and night-time temperatures'. *Geophysical Research Letters* 39.14, p. L14707. DOI: 10.1029/2012GL052459.

Drijfhout, S., G. J. van Oldenborgh and A. Cimatoribus (2012). 'Is a Decline of AMOC Causing the Warming Hole above the North Atlantic in Observed and Modeled Warming Patterns?' *Journal of Climate* 25.24, pp. 8373–8379. DOI: 10.1175/JCLI-D-12-00490.1.

Eisenman, I. et al. (2011). 'Consistent changes in the sea ice seasonal cycle in response to global warming'. *Journal of Climate* 24.20, 5325–5335. DOI: 10.1175/2011JCLI4051.1.

Esau, I., R. Davy and S. Outten (2012). 'Complementary explanation of temperature response in the lower atmosphere'. *Environmental Research Letters* 7.4. DOI: 10.1088/1748-9326/7/4/044026.

Fang, Z. and J. M. Wallace (1994). 'Arctic Sea Ice Variability on a Timescale of Weeks and Its Relation to Atmospheric Forcing'. *Journal of Climate* 7.12, pp. 1897–1914. DOI: 10.1175/1520-0442.

Fetterer, F. et al. (2017). *Sea Ice Index, Version 3*. Boulder, Colorado USA. NSIDC: National Snow and Ice Data Center. [Accessed: 2017-11-28]. DOI: 10.7265/N5K072F8.

Fischer, E. M. and R. Knutti (2015). 'Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes'. *Nature Climate Change* 5.6, pp. 560–564. DOI: 10.1038/nclimate2617.

Fischer, E. M., U. Beyerle and R. Knutti (2013). 'Robust spatially aggregated projections of climate extremes'. *Nature Climate Change* 3.12, pp. 1033–1038. DOI: 10.1038/nclimate2051.

Flato, G. et al. (2013). 'Evaluation of Climate Models. In Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA'.

Frankcombe, L. M. et al. (2015). 'Separating internal variability from the externally forced climate response'. *Journal of Climate* 28.20, 8184–8202. DOI: 10.1175/JCLI-D-15-0069.1.

Frankignoul, C. and K. Hasselmann (1977). 'Stochastic climate models, Part II Application to sea-surface temperature anomalies and thermocline variability'. *Tellus* 29.4, pp. 289–305. DOI: 10.1111/j.2153-3490.1977.tb00740.x.

Frankignoul, C., G. Gastineau and Y.-O. Kwon (2017). 'Estimation of the SST Response to Anthropogenic and External Forcing and Its Impact on the Atlantic Multidecadal Oscillation and the Pacific Decadal Oscillation'. *Journal of Climate* 30.24, pp. 9871–9895. DOI: 10.1175/JCLI-D-17-0009.1.

Gleckler, P. J., K. E. Taylor and C. Doutriaux (2008). 'Performance metrics for climate models'. *Journal of Geophysical Research* 113.D6. DOI: 10.1029/2007JD008972.

Goosse, H. et al. (2009). 'Increased variability of the Arctic summer ice extent in a warmer climate'. *Geophysical Research Letters* 36. DOI: 10.1029/2009GL040546.

Gregory, J. M. et al. (2002). 'Recent and future changes in Arctic sea ice simulated by the HadCM3 AOGCM'. *Geophysical Research Letters* 29.24. DOI: 10.1029/2001GL014575.

Guemas, V. et al. (2016). 'A review on Arctic sea-ice predictability and prediction on seasonal to decadal time-scales'. *Quarterly Journal of the Royal Meteorological Society* 142.695, 546–561. DOI: 10.1002/qj.2401.

Hall, A (2004). 'The role of surface albedo feedback in climate'. *Journal of Climate* 17.7, 1550–1568.

Harada, Y. et al. (2016). 'The JRA-55 Reanalysis: Representation of Atmospheric Circulation and Climate Variability'. *Journal of the Meteorological Society of Japan. Ser. II* 94.3, pp. 269–302. DOI: 10.2151/jmsj.2016-015.

Hasselmann, K. (1976). 'Stochastic climate models Part I. Theory'. *Tellus* 28.6, pp. 473–485. DOI: 10.1111/j.2153-3490.1976.tb00696.x.

Haumann, F. A., D. Notz and H. Schmidt (2014). 'Anthropogenic influence on recent circulation-driven Antarctic sea ice changes'. *Geophysical Research Letters* 41.23, 8429–8437. DOI: 10.1002/2014GL061659.

Hawkins, E. et al. (2016). 'Irreducible uncertainty in nearterm climate projections'. *Climate Dynamics* 46, 3807–3819. DOI: 10.1007/s00382-015-2806-8.

Hawkins, E. and R. Sutton (2009). 'The potential to narrow uncertainty in regional climate predictions'. *Bulletin of the American Meteorological Society* 90.8, 1095+. DOI: 10.1175/2009BAMS2607.1.

— (2011). 'The potential to narrow uncertainty in projections of regional precipitation change'. *Climate Dynamics* 37.1-2, 407–418. DOI: 10.1007/s00382-010-0810-6.

Hedemann, C. et al. (2017). 'The subtle origins of surface-warming hiatuses'. *Nature Climate Change* 7.5, pp. 336–339. DOI: 10.1038/nclimate3274.

Hingray, B. and M. Said (2014). 'Partitioning internal variability and model uncertainty components in a multimember multimodel ensemble of climate projections'. *Journal of Climate* 27.17, 6779–6798. DOI: 10.1175/JCLI-D-13-00629.1.

Holland, M. M. et al. (2008). 'The role of natural versus forced change in future rapid summer Arctic ice loss'. *Arctic sea ice decline: observations, projections, mechanisms, and implications*. Ed. by DeWeaver, ET and Bitz, CM and Tremblay, LB. Vol. 180. Geophysical Monograph Series, 133–150. DOI: 10.1029/180GM10.

Holmes, C. R. et al. (2016). 'Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection'. *Journal of Climate* 29.6, 2221–2236. DOI: 10.1175/JCLI-D-14-00735.1.

Huang, J. et al. (2017). 'Recently amplified arctic warming has contributed to a continual global warming trend'. *Nature Climate Change* 7.12, p. 875. DOI: 10.1038/s41558-017-0009-5.

Huber, M. and R. Knutti (2014). 'Natural variability, radiative forcing and climate response in the recent hiatus reconciled'. *Nature Geoscience* 7.9, 651–656. DOI: 10.1038/NGEO2228.

Huffman, G. and D. Bolvin (2012). 'Version 1.2 GPCP One-Degree Daily Precipitation Data Set Documentation'. NASA Goddard Space Flight Center.

Huntingford, C. et al. (2013). 'No increase in global temperature variability despite changing regional patterns'. *Nature* 500.7462, 327+. DOI: 10.1038/nature12310.

IPCC (2012). Summary for Policymakers. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. A Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 1-19.

— (2014a). Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, 151 pp.

— (2014b). Annex II: Glossary [Mach, K.J., S. Planton and C. von Stechow (eds.)]. In: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, R.K. Pachauri and L.A. Meyer (eds.)]. IPCC, Geneva, Switzerland, pp. 117-130.

Jakob, D. and D. Walland (2016). 'Variability and long-term change in Australian temperature and precipitation extremes'. *Weather and Climate Extremes* 14, pp. 36–55. DOI: 10.1016/j.wace.2016.11.001.

Kalnay, E. et al. (1996). 'The NCEP/NCAR 40-Year Reanalysis Project'. *Bulletin of the American Meteorological Society* 77.3. NCEP Reanalysis data provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, 437–471. DOI: 10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Känel, L. von, T. L. Frölicher and N. Gruber (2017). 'Hiatus-like decades in the absence of equatorial Pacific cooling and accelerated global ocean heat uptake'. *Geophysical Research Letters* 44.15, 2017GL073578. DOI: 10.1002/2017GL073578.

Kapsch, M.-L., R. G. Graversen and M. Tjernström (2013). 'Springtime atmospheric energy transport and the control of Arctic summer sea-ice extent'. *Nature Climate Change* 3.8, 744–748. DOI: 10.1038/nclimate1884.

Karl, T. R. and R. W. Knight (1998). 'Secular Trends of Precipitation Amount, Frequency, and Intensity in the United States'. *Bulletin of the American Meteorological Society* 79.2, pp. 231–242. DOI: 10.1175/1520-0477(1998)079<0231:STOPAF>2.0.CO;2.

Karl, T. R. et al. (2015). 'Possible artifacts of data biases in the recent global surface warming hiatus'. *Science*, aaa5632. DOI: 10.1126/science.aaa5632.

Kashiwase, H. et al. (2017). 'Evidence for ice-ocean albedo feedback in the Arctic Ocean shifting to a seasonal ice zone'. *Scientific Reports* 7.1. DOI: 10.1038/s41598-017-08467-z.

Katz, R. W. and B. G. Brown (1992). 'Extreme events in a changing climate: Variability is more important than averages'. *Climatic Change* 21.3, pp. 289–302. DOI: 10.1007/BF00139728.

Kay, J. E. et al. (2015). 'The Community Earth System Model (CESM) Large Ensemble Project, A community resource for studying climate change in the presence of internal climate variability'. *Bulletin of the American Meteorological Society* 96.8, 1333–1349. DOI: 10.1175/BAMS-D-13-00255.1.

Kayano, M. T., V. B. Rao and R. V. Andreoli (2005). 'A review of short-term climate variability mechanisms'. *Advances in Space Research*. Fundamentals of Space Environment Science 35.5, pp. 843–851. DOI: 10.1016/j.asr.2004.10.010.

Keith, D. W. (1995). 'Meridional energy transport: uncertainty in zonal means'. *Tellus A: Dynamic Meteorology and Oceanography* 47.1, 30–44. DOI: 10.3402/tellusa.v47i1.11492.

Kharin, V. V. et al. (2013). 'Changes in temperature and precipitation extremes in the CMIP5 ensemble'. *Climatic Change* 119.2, pp. 345–357. DOI: 10.1007/s10584-013-0705-8.

King, A. D. et al. (2015). 'The timing of anthropogenic emergence in simulated climate extremes'. *Environmental Research Letters* 10.9, p. 094015. DOI: 10.1088/1748-9326/10/9/094015.

Kirtman, B. et al. (2013). 'Near-term Climate Change: Projections and Predictability. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change'. [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Knutson, T. R., F. Zeng and A. T. Wittenberg (2013). 'Multimodel Assessment of Regional Surface Temperature Trends: CMIP3 and CMIP5 Twentieth-Century Simulations'. *Journal of Climate* 26.22, 8709–8743. DOI: 10.1175/JCLI-D-12-00567.1.

Knutti, R. and J. Sedláček (2013). 'Robustness and uncertainties in the new CMIP5 climate model projections'. *Nature Climate Change* 3.4, pp. 369–373. DOI: 10.1038/nclimate1716.

Kobayashi, S. et al. (2015). 'The JRA-55 Reanalysis: General Specifications and Basic Characteristics'. *Journal of the Meteorological Society of Japan. Ser. II* 93.1, pp. 5–48. DOI: 10.2151/jmsj.2015-001.

Kwok, R. and D. A. Rothrock (2009). 'Decline in Arctic sea ice thickness from submarine and ICESat records: 1958-2008'. *Geophysical Research Letters* 36. DOI: 10.1029/2009GL039035.

Laxon, S. W. et al. (2013). 'CryoSat-2 estimates of Arctic sea ice thickness and volume'. *Geophysical Research Letters* 40.4, 732–737. DOI: 10.1002/grl.50193.

Lenton, T. M. (2011). 'Early warning of climate tipping points'. *Nature Climate Change* 1.4, p. 201. DOI: 10.1038/nclimate1143.

Letterly, A., J. Key and Y. Liu (2016). 'The influence of winter cloud on summer sea ice in the Arctic, 1983-2013'. *Journal of Geophysical Research: Atmospheres* 121.5, 2178–2187. DOI: 10.1002/2015JD024316.

Lewis, S. C. and A. D. King (2017). 'Evolution of mean, variance and extremes in 21st century temperatures'. *Weather and Climate Extremes* 15, pp. 1–10. DOI: 10.1016/j.wace.2016.11.002.

Lindsay, R. et al. (2014). 'Evaluation of Seven Different Atmospheric Reanalysis Products in the Arctic'. *Journal of Climate* 27.7, 2588–2606. DOI: 10.1175/JCLI-D-13-00014.1.

Lorenz, E. N. (1963). 'Deterministic Nonperiodic Flow'. *Journal of the Atmospheric Sciences* 20.2, pp. 130–141. DOI: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

— (1993). *The Essence of Chaos*. USA: University of Washington Press.

Lu, J., A. Hu and Z. Zeng (2014). 'On the possible interaction between internal climate variability and forced climate change'. *Geophysical Research Letters* 41.8, 2962–2970. DOI: 10.1002/2014GL059908.

Lustenberger, A., R. Knutti and E. M. Fischer (2014). 'Sensitivity of European extreme daily temperature return levels to projected changes in mean and variance'. *Journal of Geophysical Research: Atmospheres* 119.6, pp. 3032–3044. DOI: 10.1002/2012JD019347.

Mahlstein, I. and R. Knutti (2012). 'September Arctic sea ice predicted to disappear near 2 degrees C global warming above present'. *Journal of Geophysical Research* 117. DOI: 10.1029/2011JD016709.

Mahlstein, I., P. R. Gent and S. Solomon (2013). 'Historical Antarctic mean sea ice area, sea ice trends, and winds in CMIP5 simulations'. *Journal of Geophysical Research* 118.11, 5105–5110. DOI: 10.1002/jgrd.50443.

Marotzke, J. and P. M. Forster (2015). 'Forcing, feedback and internal variability in global temperature trends'. *Nature* 517.7536, 565–U291. DOI: 10.1038/nature14117.

Massonnet, F. et al. (2012). 'Constraining projections of summer Arctic sea ice'. *Cryosphere* 6.6, 1383–1394. DOI: 10.5194/tc-6-1383-2012.

Mauritsen, T. et al. (2013). 'Climate feedback efficiency and synergy'. *Climate Dynamics* 41.9, 2539–2554. DOI: 10.1007/s00382-013-1808-7.

McKinnon, K. A. et al. (2016). 'The changing shape of Northern Hemisphere summer temperature distributions'. *Journal of Geophysical Research: Atmospheres* 121.15, 2016JD025292. DOI: 10.1002/2016JD025292.

Mearns, L. O., R. W. Katz and S. H. Schneider (1984). 'Extreme High-Temperature Events: Changes in their probabilities with Changes in Mean Temper-

ature'. *Journal of Climate and Applied Meteorology* 23.12, pp. 1601–1613. DOI: 10.1175/1520-0450(1984)023<1601:EHTECI>2.0.CO;2.

Meehl, G. A. et al. (2000). 'An Introduction to Trends in Extreme Weather and Climate Events: Observations, Socioeconomic Impacts, Terrestrial Ecological Impacts, and Model Projections'. *Bulletin of the American Meteorological Society* 81.3, pp. 413–416. DOI: 10.1175/1520-0477(2000)081<0413:AITTIE>2.3.CO;2.

Meier, W. (2013). 'NOAA/NSIDC Climate Data Record of Passive Microwave Sea Ice Concentration, Version 2. Boulder, Colorado USA: National Snow and Ice Data Center. '

Menary, M. B. and R. A. Wood (2018). 'An anatomy of the projected North Atlantic warming hole in CMIP5 models'. *Climate Dynamics* 50.7, pp. 3063–3080. DOI: 10.1007/s00382-017-3793-8.

Messori, G., C. Woods and R. Caballero (2017). 'On the drivers of wintertime temperature extremes in the High Arctic'. *Journal of Climate*. DOI: 10.1175/JCLI-D-17-0386.1.

Miles, M. W. et al. (2014). 'A signal of persistent Atlantic multidecadal variability in Arctic sea ice'. *Geophysical Research Letters* 41.2, 463–469. DOI: 10.1002/2013GL058084.

Mitchell, J. M. (1976). 'An overview of climatic variability and its causal mechanisms'. *Quaternary Research* 6.4, pp. 481–493. DOI: 10.1016/0033-5894(76)90021-1.

Moss, R. H. et al. (2010). 'The next generation of scenarios for climate change research and assessment'. *Nature* 463.7282, 747–756. DOI: 10.1038/nature08823.

Neumann, J von (1932). 'Proof of the quasi-ergodic hypothesis'. *Proceedings of the National Academy of Sciences U.S.A* 18, 70–82.

Notz, D. (2009). 'The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss'. *Proceedings of the National Academy of Sciences U.S.A* 106.49. doi: 10.1073/pnas.0902356106, 20590–20595. DOI: 10.1073/pnas.0902356106.

— (2009). 'The future of ice sheets and sea ice: Between reversible retreat and unstoppable loss'. *Proceedings of the National Academy of Sciences U.S.A.* 106.49, pp. 20590–20595. DOI: 10.1073/pnas.0902356106.

— (2014). 'Sea-ice extent and its trend provide limited metrics of model performance'. *Cryosphere* 8, 229–243. DOI: 10.5194/tc-8-229-2014.

— (2015). 'How well must climate models agree with observations?' *Philosophical Transactions of the Royal Society A* 373.2052. DOI: 10.1098/rsta.2014.0164.

Ogi, M., K. Yamazaki and J. M. Wallace (2010). 'Influence of winter and summer surface wind anomalies on summer Arctic sea ice extent'. *Geophysical Research Letters* 37.7, L07701. DOI: 10.1029/2009GL042356.

Olason, E. and D. Notz (2014). 'Drivers of variability in Arctic sea-ice drift speed'. *Journal of Geophysical Research* 119.9, 5755–5775. DOI: 10.1002/2014JC009897.

Olonscheck, D. and D. Notz (2017). 'Consistently Estimating Internal Climate Variability from Climate Model Simulations'. *Journal of Climate* 30.23, pp. 9555–9573. DOI: 10.1175/JCLI-D-16-0428.1.

Palmer, M. D. and D. J. McNeall (2014). 'Internal variability of Earth's energy budget simulated by CMIP5 climate models'. *Environmental Research Letters* 9.3. DOI: 10.1088/1748-9326/9/3/034016.

Parey, S., T. T. H. Hoang and D. Dacunha-Castelle (2013). 'The importance of mean and variance in predicting changes in temperature extremes'. *Journal of Geophysical Research: Atmospheres* 118.15, pp. 8285–8296. DOI: 10.1002/jgrd.50629.

Park, H.-S. et al. (2015). 'The Impact of Poleward Moisture and Sensible Heat Flux on Arctic Winter Sea Ice Variability'. *Journal of Climate* 28.13, 5030–5040. DOI: 10.1175/JCLI-D-15-0074.1.

Pavelsky, T. M. et al. (2011). 'Atmospheric inversion strength over polar oceans in winter regulated by sea ice'. *Climate Dynamics* 36.5, pp. 945–955. DOI: 10.1007/s00382-010-0756-8.

Pendergrass, A. G. and E. P. Gerber (2016). 'The Rain Is Askew: Two Idealized Models Relating Vertical Velocity and Precipitation Distributions in a Warming World'. *Journal of Climate* 29.18, pp. 6445–6462. DOI: 10.1175/JCLI-D-16-0097.1.

Pendergrass, A. G. et al. (2017). 'Precipitation variability increases in a warmer climate'. *Scientific Reports* 7.1, p. 17966. DOI: 10.1038/s41598-017-17966-y.

Perkins-Kirkpatrick, S. E. et al. (2017). 'The influence of internal climate variability on heatwave frequency trends'. *Environmental Research Letters* 12.4, p. 044005. DOI: 10.1088/1748-9326/aa63fe.

Pithan, F. and T. Mauritsen (2014). 'Arctic amplification dominated by temperature feedbacks in contemporary climate models'. *Nature Geoscience* 7.3. doi: 10.1038/NGEO2071, 181–184. DOI: 10.1038/NGEO2071.

Rädel, G. et al. (2016). 'Amplification of El Niño by cloud longwave coupling to atmospheric circulation'. *Nature Geoscience* 9.2, 106–110. DOI: 10.1038/ngeo2630.

Rahmstorf, S. et al. (2015). 'Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation'. *Nature Climate Change* 5.5, pp. 475–480. DOI: 10.1038/nclimate2554.

Resplandy, L., R. Seferian and L. Bopp (2015). 'Natural variability of CO2 and O-2 fluxes: What can we learn from centuries-long climate models simulations?' *Journal of Geophysical Research* 120.1, 384–404. DOI: 10.1002/2014JC010463.

Rhines, A. and P. Huybers (2013). 'Frequent summer temperature extremes reflect changes in the mean, not the variance'. *Proceedings of the National Academy of Sciences* 110.7, E546–E546. DOI: 10.1073/pnas.1218748110.

Roberts, C. D. et al. (2015). 'Quantifying the likelihood of a continued hiatus in global warming'. *Nature Climate Change* 5.4, 337–342. DOI: 10.1038/nclimate2531.

Rosenblum, E. and I. Eisenman (2016). 'Faster Arctic sea ice retreat in CMIP5 than in CMIP3 due to volcanoes'. *Journal of Climate*. DOI: 10.1175/JCLI-D-16-0391.1.

Räisänen, J. (2002). 'CO2-Induced Changes in Interannual Temperature and Precipitation Variability in 19 CMIP2 Experiments'. *Journal of Climate* 15.17, pp. 2395–2411. DOI: 10.1175/1520-0442(2002)015<2395:CICIIT>2.0.CO;2.

Santer, B. D. et al. (2008). 'Consistency of modelled and observed temperature trends in the tropical troposphere'. *International Journal of Climatology* 28.13, 1703–1722. DOI: 10.1002/joc.1756.

Santer, B. D. et al. (2014). 'Volcanic contribution to decadal changes in tropospheric temperature'. *Nature Geoscience* 7.3, 185–189. DOI: 10.1038/ngeo2098.

Schär, C. et al. (2004). 'The role of increasing temperature variability in European summer heatwaves'. *Nature* 427.6972, pp. 332–336. DOI: 10.1038/nature02300.

Schindler, A. et al. (2015). 'On the internal variability of simulated daily precipitation'. *Journal of Climate* 28.9, 3624–3630. DOI: 10.1175/JCLI-D-14-00745.1.

Schneider, E. and J. Kinter (1994). 'An examination of internally generated variability in long climate simulations'. *Climate Dynamics* 10.4-5, 181–204. DOI: 10.1007/BF00208987.

Schneider, T., T. Bischoff and H. Plotka (2015). 'Physics of changes in synoptic midlatitude temperature variability (vol 28, pg 2312, 2015)'. *Journal of Climate* 29.9, 3471. DOI: 10.1175/JCLI-D-16-0096.1.

Schweiger, A. et al. (2011). 'Uncertainty in modeled Arctic sea ice volume'. *Journal of Geophysical Research* 116. DOI: 10.1029/2011JC007084.

Screen, J. A. (2014). 'Arctic amplification decreases temperature variance in northern mid- to high-latitudes'. *Nature Climate Change* 4.7, 577–582. DOI: 10.1038/NCLIMATE2268.

Seneviratne, S. I. et al. (2014). 'No pause in the increase of hot temperature extremes'. *Nature Climate Change* 4, p. 161. DOI: 10.1038/nclimate2145.

Serreze, M. C. and R. G. Barry (2011). 'Processes and impacts of Arctic amplification: A research synthesis'. *Global and Planetary Change* 77.1, pp. 85–96. DOI: 10.1016/j.gloplacha.2011.03.004.

Serreze, M. C. and J. Stroeve (2015). 'Arctic sea ice trends, variability and implications for seasonal ice forecasting'. *Philosophical Transactions of the Royal Society A* 373.2045, p. 20140159. DOI: 10.1098/rsta.2014.0159.

Sgubin, G. et al. (2017). 'Abrupt cooling over the North Atlantic in modern climate models'. *Nature Communications* 8. DOI: 10.1038/ncomms14375.

Shindell, D. T. et al. (2003). 'Volcanic and Solar Forcing of Climate Change during the Preindustrial Era'. *Journal of Climate* 16.24, pp. 4094–4107. DOI: 10.1175/1520-0442(2003)016<4094:VASFOC>2.0.CO;2.

Shu, Q., Z. Song and F. Qiao (2015). 'Assessment of sea ice simulations in the CMIP5 models'. *Cryosphere* 9.1, 399–409. DOI: 10.5194/tc-9-399-2015.

Simolo, C. et al. (2011). 'Evolution of extreme temperatures in a warming climate'. *Geophysical Research Letters* 38.16. DOI: 10.1029/2011GL048437.

Stroeve, J. et al. (2014). 'Using records from submarine, aircraft and satellites to evaluate climate model simulations of Arctic sea ice thickness'. *Cryosphere* 8.5, 1839–1854. DOI: 10.5194/tc-8-1839-2014.

Stroeve, J. C. et al. (2012). 'Trends in Arctic sea ice extent from CMIP5, CMIP3 and observations'. *Geophysical Research Letters* 39.L16502. DOI: 10.1029/2012GL052676.

Sutton, R., E. Suckling and E. Hawkins (2015). 'What does global mean temperature tell us about local climate?' *Philosophical Transactions of the Royal Society A* 373.2054. DOI: 10.1098/rsta.2014.0426.

Swanson, K. L., G. Sugihara and A. A. Tsonis (2009). 'Long-term natural variability and 20th century climate change'. *Proceedings of the National Academy of Sciences U.S.A* 106.38, 16120–16123. DOI: 10.1073/pnas.0908699106.

Swart, N. C. et al. (2015). 'Influence of internal variability on Arctic sea-ice trends'. *Nature Climate Change* 5, 86–89. DOI: 10.1038/nclimate2483.

Taylor, K. E., R. J. Stouffer and G. A. Meehl (2012). 'An Overview of CMIP5 and the Experiment Design'. *Bulletin of the American Meteorological Society* 93.4, 485–498. DOI: 10.1175/BAMS-D-11-00094.1.

Thompson, D. W. J. et al. (2015). 'Quantifying the Role of Internal Climate Variability in Future Climate Trends'. *Journal of Climate* 28, 6443–6456. DOI: 10.1175/JCLI-D-14-00830.1.

Tietsche, S. et al. (2011). 'Recovery mechanisms of Arctic summer sea ice'. *Geophysical Research Letters* 38.2, p. L02707. DOI: 10.1029/2010GL045698.

Tingley, M. P. and P. Huybers (2013). 'Recent temperature extremes at high northern latitudes unprecedented in the past 600 years'. *Nature* 496.7444, pp. 201–205. DOI: 10.1038/nature11969.

Titchner, H. A. and N. A. Rayner (2014). 'The Met Office Hadley Centre sea ice and sea surface temperature data set, version 2: 1. Sea ice concentrations'. *Journal of Geophysical Research*. Available online. DOI: 10.1002/2013JD020316.

Trenberth, K. E. (2011). 'Attribution of climate variations and trends to human influences and natural variability'. *Wiley Interdisciplinary Reviews: Climate Change* 2.6, 925–930. DOI: 10.1002/wcc.142.

Trenberth, K. E., J. T. Fasullo and T. G. Shepherd (2015). 'Attribution of climate extreme events'. *Nature Climate Change* 5.8, pp. 725–730. DOI: 10.1038/nclimate2657.

Ukita, J. et al. (2007). 'Northern Hemisphere sea ice variability: lag structure and its implications'. *Tellus A* 59.2, 261–272. DOI: 10.1111/j.1600-0870.2006.00223.x.

Wagner, T. J. W. and I. Eisenman (2015). 'How Climate Model Complexity Influences Sea Ice Stability'. *Journal of Climate* 28.10, pp. 3998–4014. DOI: 10.1175/JCLI-D-14-00654.1.

Wallace, J. M., Y. Zhang and J. A. Renwick (1995). 'Dynamic Contribution to Hemispheric Mean Temperature Trends'. *Science* 270.5237, pp. 780–783. DOI: 10.1126/science.270.5237.780.

Walsh, J. E. et al. (2017). 'A database for depicting Arctic sea ice variations back to 1850'. *Geographical Review* 107.1, pp. 89–107. DOI: 10.1111/j.1931-0846.2016.12195.x.

Wentz, F. J. et al. (2007). 'How Much More Rain Will Global Warming Bring?' *Science* 317.5835, pp. 233–235. DOI: 10.1126/science.1140746.

Wernli, H. and L. Papritz (2018). 'Role of polar anticyclones and mid-latitude cyclones for Arctic summertime sea-ice melting'. *Nature Geoscience*, 1. DOI: 10.1038/s41561-017-0041-0.

Wettstein, J. J. and C. Deser (2014). 'Internal Variability in Projections of Twenty-First-Century Arctic Sea Ice Loss: Role of the Large-Scale Atmospheric Circulation'. *Journal of Climate* 27.2, 527–550. DOI: 10.1175/JCLI-D-12-00839.1.

Winton, M. (2006). 'Does the Arctic sea ice have a tipping point?' *Geophysical Research Letters* 33.23, p. L23504. DOI: 10.1029/2006GL028017.

Woods, C. and R. Caballero (2016). 'The Role of Moist Intrusions in Winter Arctic Warming and Sea Ice Decline'. *Journal of Climate* 29.12, pp. 4473–4485. DOI: 10.1175/JCLI-D-15-0773.1.

Yang, X.-Y., J. C. Fyfe and G. M. Flato (2010). 'The role of poleward energy transport in Arctic temperature evolution'. *Geophysical Research Letters* 37.14. DOI: 10.1029/2010GL043934.

Ylhäisi, J. S. and J. Räisänen (2014). 'Twenty-first century changes in daily temperature variability in CMIP3 climate models'. *International Journal of Climatology* 34.5, pp. 1414–1428. DOI: 10.1002/joc.3773.

Zelinka, M. D. and D. L. Hartmann (2011). 'Climate Feedbacks and Their Implications for Poleward Energy Flux Changes in a Warming Climate'. *Journal of Climate* 25.2, pp. 608–624. DOI: 10.1175/JCLI-D-11-00096.1.

Zhang, J. and D. Rothrock (2003). 'Modeling global sea ice with a thickness and enthalpy distribution model in generalized curvilinear coordinates'. *Monthly Weather Review* 131.5, 845–861. DOI: 10.1175/1520-0493(2003)131<0845:MGSIWA>2.0.CO;2.

Zhang, R. (2015). 'Mechanisms for low-frequency variability of summer Arctic sea ice extent'. *Proceedings of the National Academy of Sciences U.S.A.* 112.15, pp. 4570–4575. DOI: 10.1073/pnas.1422296112.

Zhang, X. et al. (2011). 'Indices for monitoring changes in extremes based on daily temperature and precipitation data'. *Wiley Interdisciplinary Reviews: Climate Change* 2.6, pp. 851–870. DOI: 10.1002/wcc.147.

Zhou, T. and R. Yu (2006). 'Twentieth-Century Surface Air Temperature over China and the Globe Simulated by Coupled Climate Models'. *Journal of Climate* 19.22, pp. 5843–5858. DOI: 10.1175/JCLI3952.1.

Zunz, V., H. Goosse and F. Massonnet (2013). 'How does internal variability influence the ability of CMIP5 models to reproduce the recent trend in Southern Ocean sea ice extent?' *Cryosphere* 7.2, 451–468. DOI: 10.5194/tc-7-451-2013.