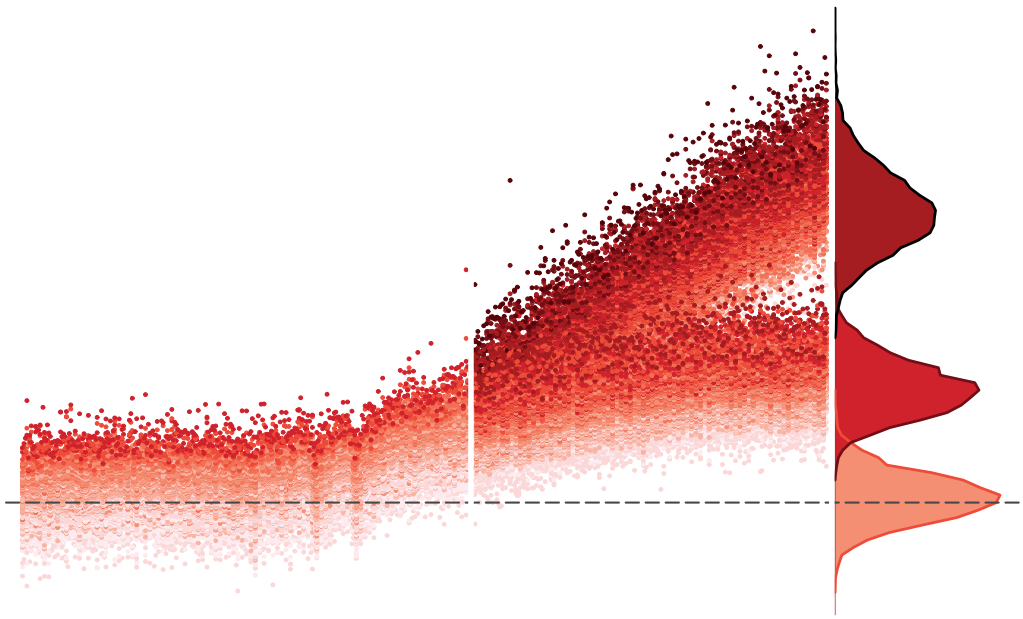
Max-Planck-Institut
für Meteorologie

MAX-PLANCK-GESELLSCHAFT

International Max Planck Research School
on Earth System Modelling

Internal variability in summertime heat extremes under global warming



Laura Suárez Gutiérrez

Hamburg 2019

Berichte zur Erdsystemforschung
Reports on Earth System Science

221
2019

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

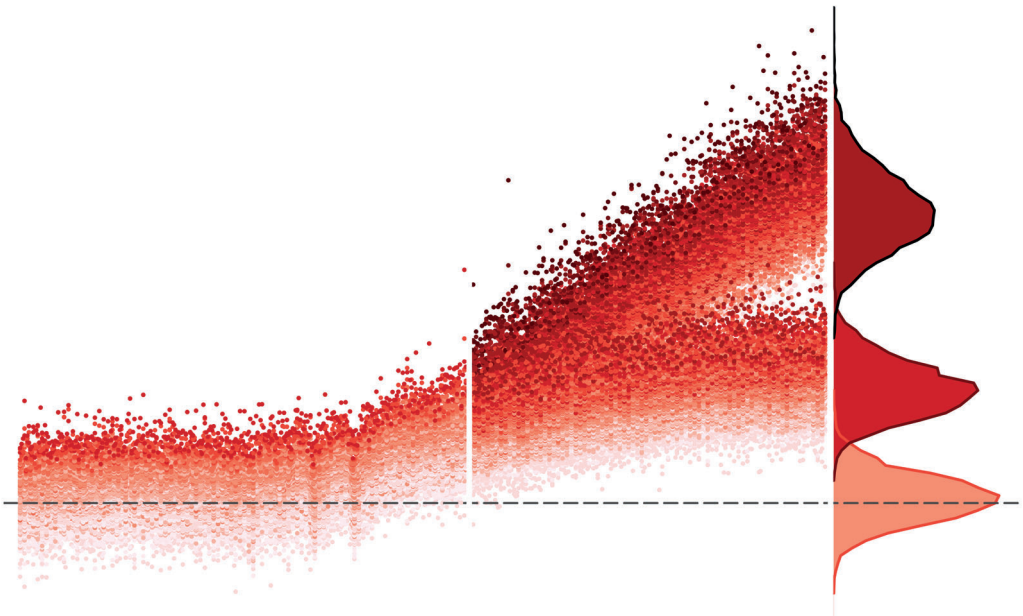
name.surname@mpimet.mpg.de
www.mpimet.mpg.de

Layout

Bettina Diallo and Norbert P. Noreiks
Communication

Copyright

Internal variability in summertime heat extremes under global warming



Laura Suárez Gutiérrez

Hamburg 2019

Laura Suárez Gutiérrez
from León, Spain

Max-Planck-Institut für Meteorologie
The International Max Planck Research School on Earth System Modelling
(IMPRS-ESM)
Bundesstrasse 53
20146 Hamburg

Universität Hamburg
Geowissenschaften
Meteorologisches Institut
Bundesstr. 55
20146 Hamburg

Tag der Disputation:18. Juni 2019

Folgende Gutachter empfehlen die Annahme der Dissertation:
Dr. Wolfgang A. Müller
Prof. Dr. Johanna Baehr

Vorsitzender des Promotionsausschusses:
Prof. Dr. Dirk Gajewski

Dekan der MIN-Fakultät:
Prof. Dr. Heinrich Graener

Berichte zur Erdsystemforschung / Max-Planck-Institut für Meteorologie 221
Reports on Earth System Science / Max Planck Institute for Meteorology 2019

ISSN 1614-1199

ABSTRACT

In this dissertation, I study how internal climate variability shapes the changing characteristics of summertime heat extremes both in Europe and globally as the world warms. A substantial sampling of internal variability is crucial to capture the most extreme events and determine how their magnitude and frequency change in a warming world, and is hence a vital requirement for this evaluation. To achieve this, I use the largest existing ensemble of a comprehensive climate model: The Max Planck Institute Grand Ensemble (MPI-GE). Due to the large ensemble size, MPI-GE is the best tool available to precisely sample the simulated internal variability in a changing climate.

First, I quantify the contribution of different driving mechanisms to extreme summertime heat over Europe, and how changes in these contributions cause an increase of variability of summertime heat in a warmer world. With a multiple regression approach, that simultaneously considers all relevant sources of variability, I identify the large-scale atmospheric dynamics as the main driver of heat extremes over Europe; while the local thermodynamic effect of soil moisture limitation plays a secondary role. Most heat extremes occur under extreme atmospheric conditions, both in current and future climates. However, in the regions where variability increases, heat extremes occur 10-40% less frequently under extreme atmospheric conditions in 21st century, and 40% more frequently under extreme moisture limitation. An increasing number of extremes are driven by moisture limitation under warming, and occur even under a neutral or unfavorable atmospheric state, confirming that the increase in European heat extremes and associated variability increase are dominated by the the thermodynamic effect of moisture limitation.

Second, I evaluate to what extent the increase in extreme European summer heat can be controlled by maintaining global warming below the limits in the UNFCCC Paris Agreement. Due to internal climate variability, only 40% of the summer months over Europe in a 2°C warmer world would exhibit mean temperatures distinguishable from those in a 1.5°C world. This distinguishability is largest over Southern Europe, and decreases to around 10% of the summer months over Eastern Europe. Furthermore, the irreducible uncertainty arising from internal variability narrows the controllability of extreme maximum temperatures to the point that, by limiting global warming to 1.5°C, only the 10% most extreme summer maximum temperatures in a 2°C world could be averted.

Lastly, I investigate where the major risk hotspots emerge under global warming for the main factors defining our vulnerability to extreme heat: maximum temperatures, return periods of extreme temperatures, maximum temperature variability, sustained tropical night temperatures, and extreme wet bulb temperatures. My findings indicate that maintaining global warming below 2°C is vital to minimize the risk of extreme heat and limit the exposure of non-adapted regions to harmful heat levels. However, each metric produces different major risk hotspots — from the highest maximum temperatures over the Arabic Peninsula, to the largest variability increase over India or Central Europe — highlighting the different potential risks and related adaptation measures that need to be considered over different regions.

ZUSAMMENFASSUNG

In dieser Dissertation untersuche ich, wie die interne Klimavariabilität Einfluss auf die sich ändernden Merkmale extremer Hitzeereignisse in Europa und global bei steigenden Treibhausgaskonzentrationen nimmt. Eine umfangreiche Stichprobe der internen Klimavariabilität ist eine unerlässliche Bedingung für die Beschreibung der Extremereignisse und deren sich veränderter Frequenz und Intensität in einer erwärmenden Welt. Um dies zu erreichen, verwende ich das größte existierende Ensemble eines globalen Klimamodells: Das Max-Planck-Institut Grand Ensemble (MPI-GE). Hinsichtlich der Ensemblegröße ist das MPI-GE derzeit das am besten geeignete Werkzeug, um eine repräsentative Stichprobe der simulierten internen Klimavariabilität zu erhalten.

Zuerst quantifiziere ich die Beiträge unterschiedlicher Antriebsmechanismen für extreme Hitzeereignisse über Europa und untersuche wie Änderungen der jeweiligen Beiträge, bedingt durch eine Klimaerwärmung, sich auf die Variabilität der Hitzeereignisse auswirken. Mittels einer multiplen Regression, die gleichzeitig alle relevanten Quellen der Variabilität der Hitzeereignisse berücksichtigt, identifiziere ich die großskalige, atmosphärische Dynamik als den Hauptantrieb für Hitzeereignisse über Europa, während die lokalen thermodynamischen Effekte der Bodenfeuchte von zweiter Ordnung sind. Die meisten Hitzeereignisse werden begünstigt durch extreme atmosphärische Grundzustände, im gegenwärtigen wie auch im zukünftigen Klima. In Regionen, in denen die Variabilität von Hitzeereignissen bedingt durch die Klimaerwärmung ansteigt, reduziert sich allerdings der Beitrag des atmosphärischen Grundzustandes auf extreme Hitzeereignisse um 10-40%, während der Beitrag des lokalen thermodynamischen Effektes durch extreme Bodenfeuchte um 40% ansteigt. Diese, durch Bodenfeuchte verursachten Hitzeereignisse, entstehen sogar dann, wenn die atmosphärischen Grundzustände neutral oder ungeeignet für Extremereignisse sind. Dadurch wird deutlich, dass der Anstieg der extremen Hitzeereignisse über Europa im zukünftigen Klima durch lokale thermodynamische Effekte der Bodenfeuchte dominiert wird.

In einem zweiten Schritt untersuche ich, inwieweit der Anstieg extremer Hitzeereignisse über Europa durch die Einhaltung der Klimaziele des UNFCCC Paris Abkommens kontrollierbar ist. Die interne Klimavariabilität führt dazu, dass sich nur für 40% der Sommermonate über Europa die mittleren Temperaturen bei einer 2°C globalen Erwärmung von jenen einer 1.5°C globalen Erwärmung unterscheiden. Diese Unterscheidbarkeit ist am größten über Südeuropa und reduziert sich auf 10% der Sommermonate für Osteuropa. Zudem beschränkt die nichtreduzierbare Unsicherheit, die durch interne Klimavariabilität entsteht, die Kontrollierbarkeit von Temperaturextremen nur soweit, dass die Reduktion der globalen Klimaerwärmung auf 1.5°C lediglich 10% der extremen Sommermaximaltemperaturen einer 2°C Erwärmung verhindern würde.

In einem letzten Schritt analysiere ich, wo die größten Hotspots bezogen auf die globale Erwärmung und den Hauptfaktoren unserer Anfälligkeit gegenüber ex-

tremen Hitzeereignissen zu finden sind. Hierfür untersuche ich folgende Faktoren: Maximaltemperaturen, Wiederkehrperioden extremer Temperaturen, maximale Temperaturvariabilität, tropische Nachttemperaturen und extreme Feuchtlufttemperaturen. Meine Resultate zeigen, dass die Begrenzung der globalen Erwärmung auf unter 2°C entscheidend für eine Minimierung der Risiken durch extreme Hitzeereignisse ist und die Gefahr von schädlichen Hitzeereignissen in nicht-angepassten Regionen einschränkt. Allerdings produziert jeder Faktor unterschiedliche regionale Hotspots — zum Beispiel die höchste Maximaltemperatur über der arabischen Halbinsel oder die größte Temperaturvariabilität über Indien und Zentraleuropa. Dies unterstreicht die unterschiedlichen potentiellen Risiken und Anpassungsstrategien, die für verschiedene Regionen berücksichtigt werden müssen.

PARTS OF THIS DISSERTATION PRE-PUBLISHED OR INTENDED FOR PUBLICATION

L. Suarez-Gutierrez, C. Li, W. A. Müller and J. Marotzke (2019a). Drivers of internal variability in European summer Heat Extremes. *To be submitted* — **Chapter 1**
Contributions: L.S.G. designed the research, designed and performed the analysis and wrote the manuscript. W.A.M., C.L. and J.M. contributed to the discussion of the results and the manuscript at all stages.

L. Suarez-Gutierrez, C. Li, W. A. Müller and J. Marotzke (2018). Internal variability in European summer temperatures at 1.5°C and 2°C of global warming. *Environ. Res. Lett.* 44, pp. 5709–5719. doi: 10.1002/2017GL073798 — **Chapter 2**
Contributions: The authors designed the research jointly. L.S.G. designed and performed the analysis and wrote the manuscript. W.A.M., C.L. and J.M. contributed to the discussion of the results and the manuscript at all stages.

L. Suarez-Gutierrez, C. Li, W. A. Müller and J. Marotzke (2019b). Heat Hotspots under Global Warming. *In Preparation* — **Chapter 3**
Contributions: L.S.G. designed the research, designed and performed the analysis and wrote the manuscript. W.A.M., C.L. and J.M. contributed to the discussion of the results and the manuscript at all stages.

ACKNOWLEDGEMENTS

I wish to give my most sincere thanks to everyone that has helped me and supported me in the process of writing this dissertation.

First of all to my supervisors, Wolfgang Müller, Jochem Marotzke, and Chao Li, thank you for your invaluable assistance and encouragement during the last three years. Thank you for guiding me through this journey, for your trust, and for always having my back. You have given me the freedom that I need, without letting me get lost in the fog, and for that I am most thankful. Thanks as well to Johanna Baehr, for her grounding and inspiring presence and for her well-thought advise.

I also want to say thanks to all my friends and colleagues at the Max Planck Institute for Meteorology. You are always inspiring and always ready to lend a helping hand, and I am proud to work by your side. This extends to the entire IMPRS group as well. Each of us is on a similar but different adventure, and knowing that and sharing our stories has always given me a sense of belonging and empowerment. Thank you all, and good luck with the challenges lying ahead.

I am also grateful to everyone that directly helped me improve my work. In particular, to Nicola, Nele, Clara and Toni, for reading one or more earlier versions of this dissertation, and giving me valuable suggestions. Thanks as well to the IMPRS office, Antje, Connie and Michi, for caring about us so much and making everything so much easier.

To all my friends, to the ones who are with my everyday and to the ones very far away, thank you for being so supportive and loving. To my Doppelkopf partners, Clara, Tobi Finn, Fabi, Toni, Tobi Becker, Katherine, and many others, thank you for making long work days so much happier.

I want to give a special thank you to all my german friends and adoptive family. When I first came to Germany I never imagined that things would unfold this way. Now, six years after, I have found here my second home. Thank you to Toni, Clara, Sonja, Tobi, Fabi, Antonia, Bernd, Lara and Elena, for welcoming me into your hearts and make me feel at home. Especially to Toni, thank you so much for your love and support. Your help in this process has been invaluable to me. You make my life brighter and I am lucky to have you by my side.

Finally, to my family. Gracias por todo, por quererme y por apoyarme. Siento que cada paso de este viaje me haya alejado de vostros un poco más, y espero algún día poder volver. Mamá, gracias por siempre cuidar de mi y hacerme ver mi valor. Papá, gracias por inspirarme y hacerme la persona que soy hoy. Abuelo Julio, gracias por todo tu apoyo, se que ves en mi muchos de tus sueños hechos realidad y eso me llena de ilusión. Gracias Abuela Mari por hacerme reir, bailar y decir las cosas claras. Gracias Abuela Julia por llevarme de la mano a ver el mundo y hacerme echar a volar. Gracias Maripi por ser mi referente, por tu mano dura y tu cariño al estilo de abuelito. Gracias a Darío, Carla y Pequeña Julita. Quizá estas palabras os inspiren en algún momento de vuestras vidas, y quiero que sepáis que siempre contaréis con mi apoyo. A todos los demás, gracias. Aún en la distancia, no podría haberlo logrado sin todos vosotros.

CONTENTS

LIST OF FIGURES

LIST OF ACRONYMS

AMV	North Atlantic multi-decadal variability
CMIP5	Coupled Model Intercomparison Project phase 5
EuST	European summer monthly mean temperature
EUSTXx	European summer maximum value of daily maximum temperature
EUSTXn	European summer minimum value of daily minimum temperature
EVP	Evapotranspiration
GMST	Global mean surface temperature
IPCC	Intergovernmental Panel on Climate Change
MPI-GE	Max Planck Institute Grand Ensemble
MPI-ESM	Max Planck Institute Earth System Model
RCP	Representative Concentration Pathway
SLP	Sea level pressure
SM	Soil moisture
SRES	Special Report on Emissions Scenarios
TNn	Monthly minimum value of daily minimum temperature
TXx	Monthly maximum value of daily maximum temperature
UNFCCC	United Nations Framework Convention on Climate Change
VIF	Variance Infraction Value
W	Wet bulb temperature
WGBT	Wet Bulb Globe Temperature
Z500	Geopotential height at 500 hPa
Z850	Geopotential height at 850 hPa

INTRODUCTION

INTERNAL VARIABILITY IN SUMMERTIME HEAT EXTREMES

Whatever can happen will happen, if we make trials enough.

— Early version of Murphy's Law,
Augustus De Morgan (1866)

As temperatures continue to rise globally due the accumulation of anthropogenic greenhouse gases in the atmosphere, extreme heat will become more likely, and more extreme (IPCC, 2013; Meehl and Tebaldi, 2004; Russo et al., 2014). Instances of extreme heat lead to increased heat-related mortality and illness, worsening the risk of heat exhaustion, dehydration, and cardio-vascular and kidney diseases (Kjellstrom et al., 2010). Additionally, extreme heat can cause substantial ecological and socio-economical impacts, such as decreased labour productivity, increased risk of wildfires, habitat loss, crop failure, decreased agricultural efficiency, and increased risk of environmental refugees by rendering some regions partially inhabitable (IPCC, 2014; Sherwood and Huber, 2010; Dunne et al., 2013). Already under current global warming levels, the deadly combination of extreme heat and humidity, together with insufficient infrastructure caused the death of thousands in the 2015 heatwaves in India and Pakistan (Wehner et al., 2016). In Europe, the combination of extreme daytime temperatures and lack of nighttime cooling caused more than 70.000 additional deaths over 16 countries during the 2003 summer (Robine et al., 2008; Laaidi et al., 2012); and 55.000 people died due to the 2010 heatwave in Russia alone (Barriopedro et al., 2011). In a world 2°C warmer than preindustrial levels, conditions equivalent to the 2015 heatwave could occur every year over some regions in India and Pakistan (Matthews et al., 2017); and one out of every two summer months are projected to be on average warmer than the 2010 summer over Europe (Suarez-Gutierrez et al., 2018).

Extreme events are, by definition, large deviations from the mean climate state. These quasi-random deviations are implicitly caused by chaotic internal variability, that emerges from spontaneously generated mechanisms and feedbacks occurring

across all components of the climate system. By chance, some of these deviations are so large that they become extreme events. As the mean climate shifts towards a warmer state, these deviations are projected to reach more extreme levels, and events that were extreme in the past are projected to occur at a higher frequency. Given their substantial societal impacts, an extensive evaluation of how global warming exacerbates the strength and frequency of extreme heat events is required. In this dissertation, I go one step further and I evaluate how the irreducible uncertainty emerging from internal variability determines to what extent the risk of extreme heat can be controlled by maintaining global warming below fixed limits. For this evaluation, a substantial sampling of internal variability that captures up to the most extreme of these deviations as the climate changes is crucial. To achieve this, I use a 100-member ensemble of single-model simulations — the Max Planck Institute Grand Ensemble (MPI-GE) — that offers a precise and substantial sampling of the simulated internal variability and allows me to cleanly separate between the changes caused by the forced warming response and the quasi-random fluctuations due to internal variability.

To illustrate the scale of the fluctuations emerging from internal variability in comparison to the response to anthropogenic forcing changes, I use summertime temperatures over Europe (Fig. I.1). European summer temperatures have increased by 1°C on average in the last three decades, illustrating a mean shift towards warmer European summers, that is largely attributed to anthropogenic global warming. Whereas this shift represents the forced effect of anthropogenic global warming, the amplitude of temperature fluctuations on inter-annual to multi-decadal timescales represents the internal variability. Due to this internal variability, the decadal probability distributions of European summer temperature anomalies drawn from the hundred realizations of MPI-GE have a width of more than 4°C. When drawn from observations, which represent the one realization of the real-world climate, the decadal probability distributions are narrower, yet cover generally more than two thirds of the ensemble distributions. Furthermore, observed European summer temperatures not only exhibit large decadal variability, but also large variability on longer and shorter timescales. In some cases, internal variability on multi-decadal timescales drives observed temperature distributions to fluctuate from the upper to the lower tail of simulated distribution from one decade to the next. On inter-annual timescales, internal variability can cause month-to-month differences in observed mean temperatures that are more than twice as large as the 1°C mean increase due to anthropogenic climate change. This is demonstrated in 2010, where an observed mean temperature anomaly at the center of the decadal probability distribution in June, developed into the highest ever recorded European summertime temperature anomaly in July.

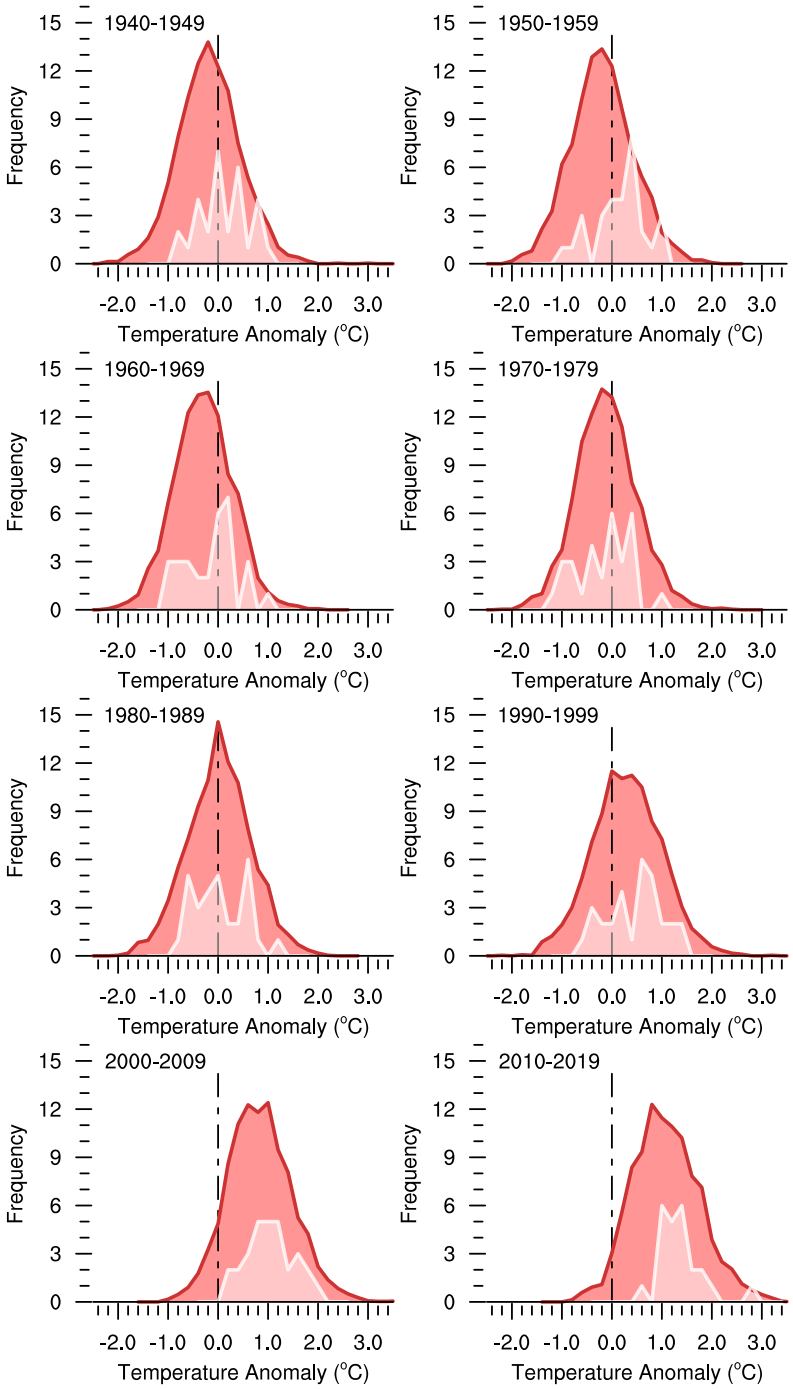


Figure I.1: Decadal variability in European summer temperatures. Probability distribution functions (PDFs) of summer (JJA) monthly mean 2m air temperature anomalies for decades starting from 1940–1949 to 2010–2019 simulated by MPI-GE (red) compared to CRUTEM4 (Jones et al., 2012) observations for the period of 1940–2017 (white). Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2019. Spatially averaged anomalies are calculated over the land points in the region over Europe defined by the [10°W–50°E, 35–68°N] domain. Anomalies are calculated with respect to the climatological baseline defined by the period of 1961–1990. Bin size is 0.2°C. For visualization purposes, the frequency in MPI-GE PDFs is shown as percentage of ensemble months per decade; while the frequency in the observed PDFs is shown in total months per decade.

The core purpose of this dissertation is to understand how these quasi-random fluctuations arising from chaotic internal variability shape the changing characteristics of summertime heat extremes, both in Europe and globally, as the world warms. In Chapter 1 I investigate which driving mechanisms control the variability in extreme European summertime heat, causing some summers to be much warmer than others. For this, I quantify the contributions from both dynamical and thermodynamical sources of variability in driving extreme temperatures simultaneously, separating circumstantial from necessary conditions that lead to extreme heat. Furthermore, I determine how these contributions may change to cause an increase in extreme temperature variability in a warmer world. In Chapter 2, I evaluate how the irreducible uncertainty arising from internal variability defines to what extent an increase in extreme European summertime heat can be averted by limiting global warming to fixed temperature targets. Lastly, in Chapter 3 I investigate where the major risk hotspots for extreme heat occur globally as the world warms. I evaluate how global warming aggravates several relevant elements of extreme summertime heat, and identify the maximum global warming level that allows us to avert dangerous heat levels confidently, once the large irreducible uncertainty introduced by chaotic internal variability is considered.

In the following sections, I provide a more in-depth introduction to the topics discussed in each chapter, and pose the six main research questions that guide this dissertation. But first, I wish to dedicate the next paragraphs of this introduction to the modelling experiment that serves as foundation for this study, the Max Planck Institute Grand Ensemble. New scientific advantages such as large ensemble experiments allow us to not only repeat old methods in new frameworks, but also to expand, and in some cases completely rethink our methodologies. In this spirit, in this dissertation I introduce several novel methodologies that illustrate how the power of large ensemble experiments can be fully exploited beyond custom.

MPI-GE is currently the largest existing single-model ensemble using a comprehensive, fully-coupled Earth System Model, both in terms of forcing scenarios represented and in terms of independent members (Maher et al., 2019). The ensemble consists of sets of 100 independent simulations that evolve under the same forcing conditions and model configuration, but start from different initial climate states. This design generates one hundred potential *Earths that might have been*, that differ from each other only due to the effect of internal variability. Thus, the large ensemble size provides an accurate estimate of the average response to external forcing, that can be cleanly separated from the quasi-random fluctuations caused by chaotic internal variability. Furthermore, the large ensemble size in MPI-GE allows a precise and sub-

stantial sampling of the simulated internal variability. This ensures that very extreme events that occur for example once every hundred years, occur on average every simulated year across the hundred ensemble realizations. This is a vital requirement for this study because it generates large samples of very extreme events that evolve under the changing climate, needed both to empirically evaluate the statistical significance of changes in very rare events, and to robustly characterize the irreducible uncertainty that arises from internal variability.

A more in-depth description of MPI-GE experimental design and how it compares to other large ensemble experiments, as well as some examples demonstrating its power, can be found in Appendix A. In Appendix B, I introduce a new model-evaluation approach to investigate the model's ability to simulate the estimated real-world internal variability. This method is one example of how large-ensemble experiments allow us to expand and rethink our methodologies beyond customary practices. The first advantage of this method, as opposed to other evaluation techniques based on comparisons of mean values or standard deviations, lies in focusing on the evaluation of the higher-order moments of the distribution, offering a more appropriate evaluation of the simulated representation of the magnitude and frequency of extreme events. Second, it allows me to directly identify whether differences between observed and simulated values are due to an incorrect simulation of the mean climate, its response to external forcings, or rather due to an incorrect representation of internal variability. After giving a more detailed explanation of this novel approach, I apply it to evaluate the ability of MPI-GE to capture the estimated real-world variability in mean and maximum surface temperatures in the summer months. In Appendix C, I apply this and other methods to evaluate how MPI-GE simulates the mean and the variability of surface temperatures both globally and, in particular, over Europe.

Due to its unique experimental design, MPI-GE is an unparalleled tool to study extreme events, which mechanisms lead to their development, and how their characteristics change in a changing climate. In the following sections I will formulate the specific scientific questions that motivate each chapter of this dissertation based on these ideas, and how MPI-GE can be utilized to best answer these questions.

Chapter 1:

Drivers of internal variability in European heat extremes

In Europe, as well as in other regions of the world, the frequency and the intensity of extreme summer heat are projected to increase not only as a response to rising global mean temperatures (Stott et al., 2004; Russo et al., 2014; Christidis et al., 2014), but also as a response to increasing summer temperature variability, that results from changes in the driving mechanisms of extreme summer temperatures under warming (Meehl and Tebaldi, 2004; Schär et al., 2004; Fischer et al., 2012; Bathiany et al., 2018). However, it remains unclear which of the two main driving mechanisms of extreme summer temperatures over Europe — either the large-scale dynamical atmospheric state or the local thermodynamic effects of moisture limitation — controls the increase in European summer temperature variability. Several studies attribute this increase to the local thermodynamical effects of limited moisture availability (Seneviratne et al., 2006; Diffenbaugh et al., 2007; Fischer and Schär, 2009; Bathiany et al., 2018). In contrast, other studies attribute it to the large-scale dynamical atmospheric state changing towards more anticyclonic conditions (Meehl and Tebaldi, 2004; D. E. Horton et al., 2015; Holmes et al., 2016; Mann et al., 2018). However, these studies focus on attributing the changing variability in extreme temperatures to either only dynamical or only thermodynamical drivers, and do not quantify the contribution from both types of drivers simultaneously. Thus, it remains unclear how both the dynamical large-scale atmospheric state and local thermodynamical mechanisms simultaneously contribute to heat extremes over Europe, and which of these contributions dominates the increase in extreme temperature variability under warming.

These are the questions that I wish to answer in this chapter. The large ensemble framework of MPI-GE allows me to cleanly separate the changes in heat extremes caused by a shift towards a warmer mean state from those caused by an increase in variability. The latter results in temperature deviations from the mean state that are larger than those in our current climate, caused by changes in the underlaying driving mechanisms of extreme temperatures. To achieve this separation I introduce a novel definition of extreme events, based on a moving threshold with respect to the evolving decadal-mean climate state. This extreme event characterization relies on an accurate characterization of internal variability, and is only well-defined when both the evolving decadal-mean climate and its probability distribution in a transient climate are known accurately, as occurs in large ensemble experiments. I then use these large samples of extreme events under a wide range of background conditions to understand the variability in the response in extreme temperatures to their different driving mechanisms, separating circumstantial from necessary conditions in the development

of heat extremes.

For this purpose I first construct a multiple linear regression model with the dynamical and thermodynamical driving mechanisms as predictors. Using this approach, I can directly quantify how much of the variability in extreme summer temperatures can be explained by each driving mechanism, and by how much the importance of each set of drivers can be overestimated by not considering both sources of variability simultaneously. Second, to identify which driving mechanisms dominate the change in extreme summer temperature variability, I evaluate how often heat extremes develop under extreme atmospheric conditions as opposed to under extreme moisture limitation, and how these frequencies may change under global warming. Based on these two approaches, in chapter 1 of this thesis I answer the two following research questions:

1.1 What are the contributions from large-scale dynamical atmospheric mechanisms and local thermodynamical effects of moisture limitation as drivers of variability in extreme summertime temperatures?

1.2. How do these dynamical and thermodynamical driving contributions change to cause an increase in extreme summer temperature variability in a warmer world?

Chapter 2:

Controllability of European temperatures under warming

Considering the large internal variability in European summer temperatures, and that this variability is projected to increase under further global warming, the next step is to evaluate to what extent the risk of extreme European summer temperatures can be averted by maintaining global warming below fixed limits. I define these limits as 1.5°C and 2°C of global mean surface temperature above pre-industrial levels, as stated in the Paris Agreement at the 21st Conference of the Parties of the United Nations Framework Convention on Climate Change (UNFCCC). For this evaluation I take advantage of both the large number of independent members and the diversity of forcing conditions in MPI-GE, that allows me to use transient climate simulations to construct quasi-stationary samples of the climate conditions at different warming levels. I use a *time-slice* method similar to the methods used in Schleussner et al. (2016) or King and Karoly (2017), with the advantage of being performed on a single-model large ensemble. The large ensemble size of MPI-GE combined with the assumption of

quasi-stationarity allows me to calculate the probability distributions that define the climate conditions at different warming levels empirically, without explicitly parametrizing the tails of the distributions with extreme value statistics. Using these empirical probability distributions based on 3000 simulated years for each climate condition, I can directly calculate maximum temperature levels for extreme events that are well defined for return periods up to 500 years.

To compare European summertime temperatures at 1.5°C of global warming above pre-industrial conditions with those at 2°C of warming, I evaluate changes in summer monthly mean, block maximum, and extreme European summer temperatures, and quantify for the first time changes in extreme events with return periods up to 500 years. However, the most novel aspect of this evaluation emerges from the quantification of the distinguishability and controllability of European summertime temperatures for different levels of global warming. To quantify the distinguishability between two climate states I introduce a novel approach based on the areal overlap of the probability distributions that define each climate. This distinguishability metric allows me to directly quantify how often summer months in a 2°C world could also occur in a 1.5°C world. Furthermore, I examine the controllability of extreme temperatures, by determining to what extent the most extreme European summer temperatures at 2°C of global warming could be avoided in a 1.5°C warmer world. These two concepts are key to determine how the irreducible uncertainty introduced by chaotic internal variability regulates to what extent increases in extreme European summertime heat can be averted by maintaining global warming below fixed limits. In line with this reasoning, Chapter 2 of this thesis concentrates on the following two central questions:

2.1 **To what extent are European summer temperatures at 1.5°C of global warming distinguishable from those at 2°C of warming?**

2.2. **To what extent can the risk an increase in extreme European summer temperatures be controlled by maintaining global warming below fixed global mean temperature limits?**

Chapter 3:

Heat hotspots under global warming

Globally, some of the regions that will suffer the effects of extreme heat the most are highly-populated developing regions, where society is most vulnerable due to scarce infrastructure and relatively low adaptive capacity (Coffel et al., 2018). In the last chapter of this dissertation, I identify where major risk hotspots for extreme heat occur globally under further warming, and what maximum global warming levels allow us to confidently avoid these risks. To investigate how global warming aggravates extreme summertime heat in a global perspective, I consider all of the most relevant elements of our vulnerability to extreme heat. The foremost of these elements are maximum temperatures. However, some of the events with the largest impacts to date, were events that combined the effect of extreme maximum temperatures with other conditions that exacerbate heat stress, such as high humidity or high nighttime temperatures (Laaidi et al., 2012; Wehner et al., 2016). To combine all of these aspects, I evaluate how global warming aggravates five different metrics: maximum reachable temperatures, return periods of very extreme events, maximum temperature variability, sustained tropical night temperatures, and extreme wet bulb temperatures.

Most previous studies which evaluate how some of these heat stress indicators change under global warming are based on smaller multi-model ensembles (e.g., Fischer and Knutti, 2013; Russo et al., 2014; Russo et al., 2017; Matthews et al., 2017; Coffel et al., 2018; Bathiany et al., 2018) or on smaller single-model ensembles (e.g., Sherwood and Huber, 2010). Also, most studies explore changes linked to different forcing scenarios (e.g., Fischer and Knutti, 2013; Russo et al., 2014; Matthews et al., 2017; Coffel et al., 2018; Bathiany et al., 2018), as opposed to changes between different warming levels (e.g., Russo et al., 2017). In contrast, I base my analysis on one very large single-model ensemble, MPI-GE. In addition to the large ensemble size, the diversity of forcing conditions represented by MPI-GE allows me to robustly characterize and compare the climates of five different worlds with global warming levels of 0°C, 1.5°C, 2°C, 3°C and 4°C above pre-industrial conditions, providing a precise sample of the simulated internal variability for each level. A robust sampling of internal variability is key to capturing how the most extreme events reachable under each climate conditions change. Using MPI-GE, I can construct five different heat metrics and five different global warming levels to robustly evaluate the maximum global warming level that allows us to avert dangerous heat levels confidently, considering the irreducible uncertainty introduced by chaotic internal variability. Thus, in Chapter 3, I focus on the following central questions:

3.1 Which regions become summertime heat stress hotspots for these different extreme heat metrics under global warming?

3.2. For which maximum global warming level can the risk of extreme heat conditions under these different metrics be confidently averted?

Chapters 1 and 3 are based on work currently being prepared for publication. Chapter 2 is based on work previously published, jointly with several co-authors (Suarez-Gutierrez et al., 2018), and has been slightly adapted to fit the structure of this dissertation.

CHAPTER 1

DRIVERS OF INTERNAL VARIABILITY IN EUROPEAN SUMMER HEAT EXTREMES

1.1. Summary

We use the 100-member Max Planck Institute Grand Ensemble (MPI-GE) to disentangle the contributions from large-scale dynamic atmospheric conditions and local thermodynamic effects of moisture limitation as drivers of variability in European summer heat extremes. Using a novel extreme event definition we find a 70% increase in heat extremes with respect to the evolving mean decadal climate under a moderate warming scenario, accompanied by a maximum increase in summer temperature variability of 35% during the 21st century. With a multiple regression approach, we find the dynamical mechanisms representing blocking and anticyclonic conditions are the main driver of variability in extreme European summer temperatures, both in past and future climate conditions. By contrast, local thermodynamic drivers play a secondary role in explaining the total variability in extreme temperature. However, considering both sources of variability simultaneously is crucial; assessing the contribution from only one type of driver can explain much less of the extreme temperature variance and leads to an overestimation of its effect on extreme temperatures, particularly when only considering local thermodynamical drivers. Finally, we find that although most European summer heat extremes occur under extreme atmospheric conditions, extremes develop 40% more frequently during the 21st century as a result of extreme moisture limitation, even under neutral or unfavorable atmospheric conditions, confirming that the local thermodynamic effect of moisture limitation dominates the increase in European summer temperature variability and heat extremes with respect to the evolving climate.

1.2. Introduction

The frequency and intensity of extreme summer heat events are projected to increase over Europe as a response to rising global mean temperatures (Stott et al., 2004; Christidis et al., 2014; Suarez-Gutierrez et al., 2018). On the other hand, the frequency and intensity of heat extremes could also be further exacerbated as a response to changes in the driving mechanisms of extreme summer temperatures under global warming, resulting in an increase in summer temperature variability (Schär et al., 2004; Meehl and Tebaldi, 2004; Fischer et al., 2012; Lustenberger et al., 2014; Donat et al., 2017; Bathiany et al., 2018). Although some studies argue that European summer extreme temperatures mostly follow mean summer warming (Ballester et al., 2009; Ballester et al., 2010), the variability in European summer temperatures is large, and there is robust evidence indicating that this variability could increase under global warming (Fischer et al., 2012; Lustenberger et al., 2014; Bathiany et al., 2018). However, it remains unclear which of the two main driving mechanisms of extreme European summer temperatures is responsible for this increase in variability. Several studies attribute the increase in summer temperature variability to the local thermodynamical effects on temperature of moisture limitation (Seneviratne et al., 2006; Diffenbaugh et al., 2007; Fischer and Schär, 2009; Vogel et al., 2017; Donat et al., 2017). In contrast, other studies point to the large-scale dynamical atmospheric state as the main driver of heat extremes, and that changes towards more anticyclonic atmospheric patterns drive the increase in temperature variability (Meehl and Tebaldi, 2004; D. E. Horton et al., 2015; Holmes et al., 2016; Mann et al., 2018). These studies focus on attributing variability changes under global warming to either only dynamical or only thermodynamical drivers, but do not quantify the contribution from both large-scale and local drivers simultaneously. Here we evaluate large samples of extreme summer temperature events simulated by the Max Planck Institute Grand Ensemble (MPI-GE) to robustly quantify the contributions from large-scale dynamical drivers and local thermodynamical drivers to extreme temperature variability, and how these contributions may change in a warmer world.

Considering both large-scale dynamical mechanisms and local thermodynamic mechanisms simultaneously is crucial to understand how heat extremes develop and what drives the variability in extreme European summer temperatures. Both driving mechanisms are not completely independent of each other — but rather capable of intensifying or counteracting one another — and also not completely collinear — but each capable of accounting for part of the variability in extreme temperatures (Della-Marta et al., 2007; Zampieri et al., 2009; R. M. Horton et al., 2016; Sillmann et al., 2017; Bunzel et al., 2017). Studies show that, over the US, both a shift towards more

anticyclonic conditions and drier soils can intensify hot extremes in a warmer world, with the later thermodynamic drivers accounting for more than half of the summer temperature variability (Diffenbaugh and Ashfaq, 2010; Merrifield et al., 2017). In the case of Europe, both dynamical mechanisms connected to atmospheric circulation and thermodynamical effects related to relative humidity have been shown to affect temperature during extremely hot days in both the ERA interim record and historical simulations from different CMIP5 models (Krueger et al., 2015). Case-based studies also indicate that past record-breaking heat waves occurred as a combination of extreme conditions in both dynamic and thermodynamic drivers (Fischer et al., 2007; Miralles et al., 2014). Others find the best predictive skill for temperatures during past observed European heat waves when considering a combination of dynamical and thermodynamical drivers as predictors (Della-Marta et al., 2007). Under future climate conditions, studies also find that both dynamical and thermodynamical drivers contribute to the changes in variability of European summer temperatures on daily to subdaily scales (Cattiaux et al., 2015). On the other hand, studies considering the local downwelling radiation at the surface as a proxy for the local atmospheric state find that downwelling radiation does not influence the change in variability of European summertime mean temperatures substantially (Fischer et al., 2012; Bathiany et al., 2018), or that is not clear whether the relative contributions from downwelling radiation or soil drying dominate the change in variability (Fischer and Schär, 2009). Thus, how both the dynamical large scale atmospheric state and local thermodynamical mechanisms simultaneously contribute to driving heat extremes and which of these contributions dominates the increase in extreme European summer temperature variability under global warming remains unclear.

To understand how these large-scale dynamics and local thermodynamic effects drive extreme temperatures in a warming world, and how relevant internal variability is in the development of heat extremes, we need to analyze large samples of extreme events that develop under a wide range of background conditions, but that are simulated under the same external forcings and model physics. For this we use simulations from the 100-member Max Planck Institute Grand Ensemble (MPI-GE), currently the largest existing ensemble using a fully-coupled Earth System Model (Bittner et al., 2016; Hedemann et al., 2017; Suarez-Gutierrez et al., 2017; Maher et al., 2019). The large size of the ensemble is crucial to robustly sample internal variability and to empirically evaluate the statistical significance of changes in very rare events. An ensemble size of 100 simulations under the same forcing conditions allows 1-in-100-years events to occur on average every simulated year (Suarez-Gutierrez et al., 2018), providing the large samples of extreme events that develop under different driving conditions and global warming levels needed for our study. Another relevant aspect to note is that

most previous studies investigate changes in variability in multi-model ensembles and using standard deviation changes as a proxy (Fischer and Schär, 2009; Fischer et al., 2012; Bathiany et al., 2018). However, this combination does not allow a clear separation between the forced transient warming and the deviations caused by internal variability, and can lead to misleading results. Using a very large single-model ensemble, we are able to instead directly evaluate how temperature deviations from the mean state change under global warming, based on a precise characterization of the simulated internal variability and the forced warming signal that are not confounded by different responses to forcing or model configurations.

We introduce a definition of extreme events with respect to the evolving decadal-mean climate state, 2σ events, that allows us to focus on the extreme events that would pose the biggest challenge to society — even if we manage to adapt to the changing mean climate. We define 2σ events as summer months (JJA) with anomalies of European summer monthly mean temperatures (EuSTs) that deviate from the decadal mean climate state by at least two EuST pre-industrial standard deviations. This moving threshold delimitation of very rare extreme events is only well-defined when both the evolving decadal-mean climate state and the probability distribution of EuSTs in a changing climate are known accurately, as occurs in large ensemble experiments, because it relies on an accurate characterization of the simulated internal variability. This novel characterization allows us to study changes in the number of heat extremes not only caused by the shift in the distribution towards higher temperatures, but by variability changes leading to temperature deviations from the mean state that are larger than those in our current climate, and that are caused by changes in the underlaying driving mechanisms of extreme temperatures.

To represent the main driving mechanisms of summer temperature variability we include large-scale dynamical atmospheric conditions, represented by geopotential height at 500 hPa (Z500) and sea level pressure (SLP), as well as the local thermodynamical effects of moisture availability, represented by soil moisture fraction (SM) and evapotranspiration (EVP). Large Z500 and SLP anomalies are associated with the persistent blocking and anticyclonic conditions that act as a dynamical driver of extreme temperatures (Meehl and Tebaldi, 2004; Della-Marta et al., 2007; Pfahl, 2014; R. M. Horton et al., 2016; Sillmann et al., 2017; Schaller et al., 2018). On the other hand, large negative EVP anomalies indicate that less water is evaporated into the atmosphere. This may occur as a response to lower temperatures and less heat causing the evaporation of less moisture, showcasing a positive EVP-EuST relation. However, lower evapotranspiration can also relate to limited moisture conditions that have a local thermodynamic driving effect on high temperatures. This driving mechanism,

characterized by a negative EVP-EuST relation, is initiated when moisture is limited for evaporation, increasing the fraction of radiative energy that is transformed into sensible — rather than latent — heat and thus resulting in higher temperatures. Similarly, low SM anomalies can occur both as a response to high temperatures that may dry out the soil through evaporation, or as an indicator of the low moisture conditions that initiate the local thermodynamic driving mechanism leading to higher temperatures (Seneviratne et al., 2006; R. M. Horton et al., 2016; Sillmann et al., 2017; Vogel et al., 2017).

In this study, we investigate the internal variability in the response of extreme temperatures to different background states in a warming world, separating circumstantial from necessary conditions in the development of large samples of heat extremes. First, in section 1.3 we introduce the MPI-GE and evaluate its ability to simulate European heat extremes, as well as elaborate on the details of our approach and methods. In section 1.4.1, we analyze extreme temperature events with respect to evolving decadal climate over Europe, and how these extremes change under warming in MPI-GE. In section 1.4.2, we quantify the contributions from both large-scale dynamical drivers and local thermodynamical drivers to the variability in extreme European summer temperatures by constructing a multiple linear regression model with these dynamical and thermodynamical drivers as predictors in both current and future climate conditions. We also construct multiple regression models based on only dynamical and only thermodynamical drivers, to characterize the importance of simultaneously considering both sources of variability. In section 1.4.3, we identify the driving mechanisms that dominate the change in extreme summer temperature variability by evaluating how often heat extremes develop under extreme atmospheric conditions, as opposed to under extreme moisture limitation, and how these frequencies may change under global warming. In section 1.5 we discuss the caveats and limitations of our approach, followed by a summary of our main results and conclusions in section 1.6.

1.3. *Data and Methods*

We use transient climate simulations from the Max Planck Institute Grand Ensemble (MPI-GE) under historical and RCP4.5 forcing conditions (Maher et al., 2019). The ensemble consists of 100 realizations based on the same model physics and parametrizations and driven by the same external forcings, but each starting from a different initial climate state taken from different points of the model's pre-industrial control run. The MPI-GE uses the model version MPI-ESM1.1 in the low resolution (LR) configuration, with resolution T63 and 47 vertical levels in the atmosphere (Giorgetta

et al., 2013) and 1.5°resolution and 40 vertical levels in the ocean (Jungclaus et al., 2013). MPI-ESM1.1 is fairly similar to the the CMIP5 version of MPI-ESM (Taylor et al., 2012), but has a slightly lower equilibrium climate sensitivity of 2.8°C (Giorgetta et al., 2013; Flato et al., 2013), and a new 5-layer soil hydrology scheme (Hagemann and Stacke, 2015) implemented in the land-surface model component of MPI-ESM1.1, JSBACH (Raddatz et al., 2007). Observational data from the CRUTEM4.6 (Jones et al., 2012) dataset are used for comparing the MPI-GE simulations to current climate conditions.

Despite its low resolution, comparable to most models in the CMIP5 ensemble, the MPI-GE captures observed temperature variability adequately (Suarez-Gutierrez et al., 2018). In Appendix C, we find that MPI-GE offers an adequate representation of the observed estimate of internal variability in European summer temperatures, particularly in the upper tail of the temperature distribution. Temporal resolution is also relatively limited in MPI-GE, with only monthly output available. Ideally, we would use hourly to daily values to capture the amplitude of internal variability more precisely and to separate between specific conditions leading to extreme temperatures from those that occur as a response to them. However, both dynamical and thermodynamical mechanisms leading to temperature extremes are based on the persistence of either anticyclonic or dry conditions, making their signal still clearly identifiable from monthly mean values. Lastly, MPI-ESM1.1 uses a simple although relatively improved 5-layer soil hydrology scheme. Compared to the previous soil hydrology scheme in MPI-ESM, a one layer bucket scheme (Roeckner et al., 2003) that tends to overestimate evapotranspiration leading to excessively dry conditions, the new 5-layer scheme offers a better representation of soil moisture memory (Hagemann and Stacke, 2015), and more realistic simulations of the large-scale atmospheric patterns in 500 hPa geopotential height that lead to an improved representation of extreme temperature events (Bunzel et al., 2017). Despite its caveats, MPI-ESM has been shown in model evaluation studies to adequately simulate the relevant dynamic and thermodynamic mechanisms contributing to the development of past observed extreme European summer temperatures (Krueger et al., 2015).

European summer temperatures (EuSTs) are defined here as monthly mean near-surface 2m air temperature anomalies for the summer months (JJA) over land-only grid cells in the region defined by the [10°W–50°E, 35–68°N] latitude-longitude domain. We then define temperature extremes as anomalies of EuST that deviate from the decadal mean climate state by at least two EuST pre-industrial standard deviations (2σ events). Since we expect standard deviation to change with time under warming, we use the pre-industrial standard deviation for the period 1850–1899 averaged across all ensemble members as a deviation threshold for extremes; while the mean decadal

climate state is defined as the centered decadal running ensemble mean. Analogously, the variables representing the main driving mechanisms are defined as monthly mean anomalies of Z500, SLP, SM and EVP over land-only grid cells in the same domain. The effect of the thermal expansion of the lower troposphere under global warming is removed by detrending Z500 anomalies at each grid cell, subtracting the ensemble mean Z500 averaged over the domain for each time step. Since the subtracted field is spatially constant, the spatial patterns of Z500 that define large-scale dynamics are not affected by this procedure (Cattiaux et al., 2013). SM is defined as the fraction between water accumulated in the soil versus the maximum water storage capacity for each grid cell. Lastly, EVP is defined as the flux of water going from the soil and vegetation into the atmosphere; positive EVP values indicate net gain of water in the atmosphere and net loss in the soil. All anomalies are calculated with respect to the climatological period of 1961–1990.

We use these variables, as well as global mean 2m surface air temperature (GMST), as predictors in a stepwise multiple regression model to account for the variability in large samples of extreme 2σ EuST events simulated by MPI-GE. Our multiple regression model consists of several steps, starting with a forward selection of variables as predictors ranked by their individual correlation to extreme EuSTs, followed by a backward elimination of redundant predictors, accounting for multicollinearity and non-significance (Storch and Zwiers, 1999). In the first step, we rank all variables in decreasing order of correlation to extreme EuSTs for each grid cell, and select the variable with the highest correlation as the first predictor in the regression model. In the next step, the variable with the next highest correlation is evaluated for multicollinearity, and only if the multiple correlation coefficient between the considered variable and any of the predictors already introduced in the model is below 0.95, corresponding to an variance inflation factor (VIF) of 10 (O'brien, 2007), the variable is then selected for the regression model. Once each new variable is added to the regression model, we evaluate if the addition improves the model significantly. If the p-value of the newly added predictor is above 0.05 or the fraction of unexplained variance is not reduced compared to the step before, the variable is again eliminated from the model. If, on the contrary, these conditions hold, then the remaining predictors in the model are tested for significance, and removed from the model if their p-values are above the 0.05 threshold. This forward-selection backward-elimination procedure is repeated until no other variables can be added to the regression model.

1.4. Results

1.4.1. Changes in European heat extremes

Under the moderate warming scenario RCP4.5, corresponding in MPI-GE to a global warming level of 2.25°C above preindustrial conditions by the end of the 21st century (Suarez-Gutierrez et al., 2018), the MPI-GE simulates an average increase in European summer temperatures (EuST) of almost 3°C compared to 20th century conditions (Fig. 1.1a). MPI-GE projects that, by the end of the 21st century, the average European summer month will be comparable to the warmest month observed in this region, during the 2010 European heatwave. Furthermore, temperature anomalies during the most extreme summer months could reach values twice as large as those recorded in 2010. We find that the MPI-GE offers an adequate representation of the observed estimate of internal variability and of the frequency and amplitude of extreme European summer temperature events. The large size of MPI-GE makes it well capable of simulating events comparable to the most extreme European summer temperatures on record, unlike other large ensemble experiments (Schaller et al., 2018). Some of the simulated events exhibit even substantially larger EuST deviations from the decadal mean state that those observed. However, these simulated events showcasing large EuST deviations have return periods of over hundreds of years, and the observational record may just be too short to determine whether or not the ensemble overestimates the amplitude of very extreme events.

We use a novel moving-threshold definition that evolves with the changing decadal climate to characterize extreme events in the ensemble simulations. These extreme events with respect to the evolving decadal mean climate, 2σ events, correspond to months with spatially averaged EuST anomalies that deviate from the decadal mean climate state by more than two EuST pre-industrial standard deviations. Such events occur in the hundred MPI-GE simulations a total of 860 times during the 20th century, and increase by 70% to 1483 extremes during the 21st century (Fig. 1.1a). This calculation is influenced by the baseline period used to calculate anomalies. Using the climatological average defined by the period of 1961–1990 to calculate anomalies we find approximately 25% less extremes during the second half of the 20th century compared to the two previous 50-year periods (Table 1.1). In contrast, using the pre-industrial climatological average defined by the period of 1851–1880 leads to an homogeneous number of extremes during the 19th and 20th century, and to a slightly lower relative increase in extremes during the 21st century of 55% (Table 1.1). Although the choice of 1961–1990 as climatology period can inflate the relative increase in extreme events, we maintain this period as climatological baseline to facilitate the comparison to observations. Observed EuST anomalies, also with respect to the 1961-

1990 baseline, occur 11 times above the 2σ threshold during the second half of the 19th century, 8 times during the 20th century, and twice in the 21st century so far.

2σ Extremes	Anomalies wrt. 1961–1990	Anomalies wrt. 1851–1880
1850–1899	491	466
1900–1949	489	481
1950–1999	371	447
Total 20th century	860	928
Total 21st century	1483	1431

Table 1.1: Number of 2σ extremes for different climatological periods. Number of summer months with extreme anomalies with respect to the evolving decadal mean climate simulated by the 100-member MPI-GE during different periods, for anomalies calculated with respect to the 1961–1990 climatological average versus for anomalies calculated with respect to the 1851-1880 average.

The increase in extreme events with respect to the evolving decadal mean climate does not occur homogeneously over Europe. We find a maximum increase in the number of 2σ extreme events of more than 100% over Central Eastern Europe, accompanied by a lower increase of 50-100% over most of the central part of the continent. On the other hand, some Southern and Northern European regions show a moderate decrease in extreme events with respect to mean climate conditions of less than 25% (Fig. 1.1b). But not only do extreme events with EuSTs above the 2σ threshold occur more often over most of Europe in the 21st century; these events also exhibit larger deviations from the mean decadal climate than those under 20th century conditions, due to an increase in EuST variability. The change in variability, illustrated by changes in the width of the EuST distribution, exhibits a similar pattern to the change in the number of extreme events (Fig. 1.1c). By the end of the 21st century, we find a maximum increase in summer temperature variability of 35% over Central Eastern Europe; while other Southern and Northern regions present no substantial change to a slight decrease compared with early 20th century conditions.

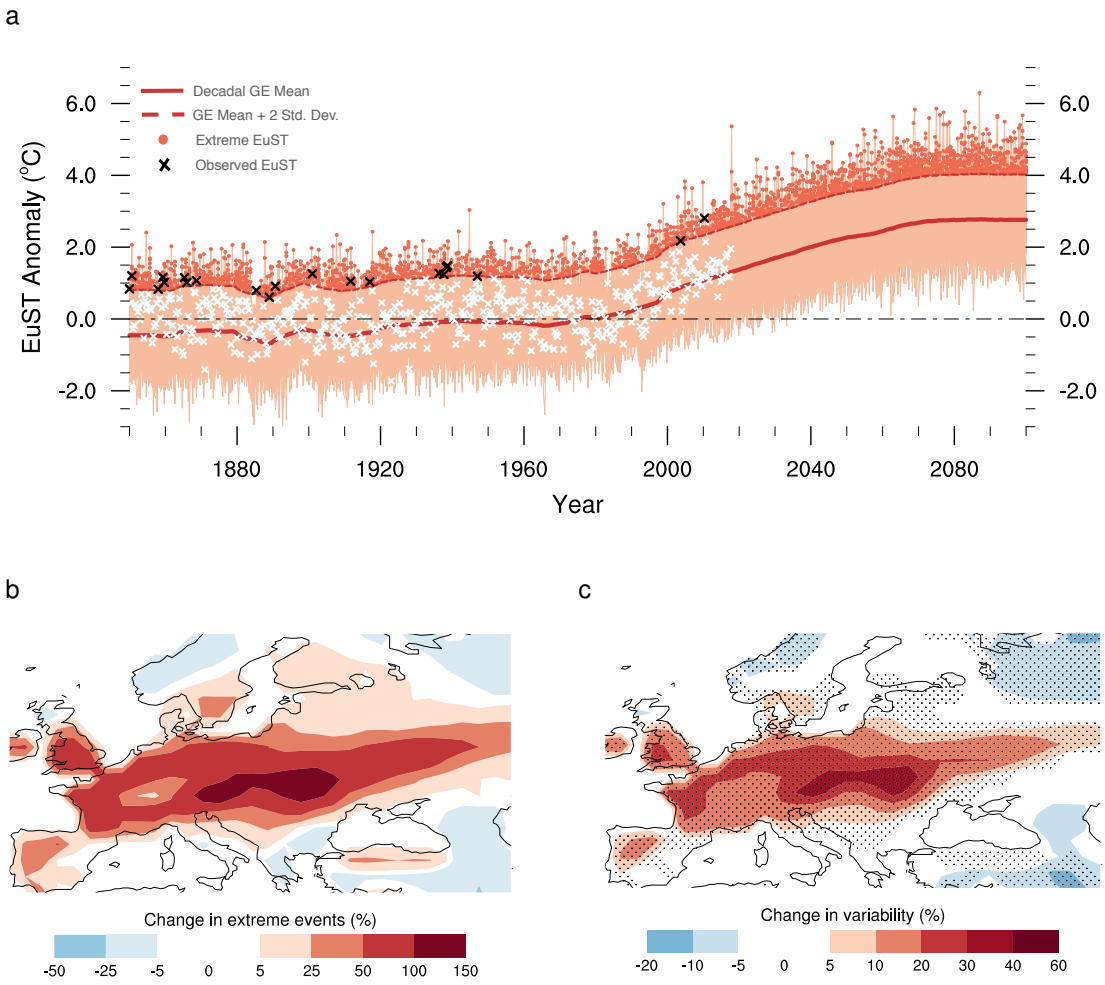


Figure 1.1: Change in European heat extremes and variability. **(a)** Time series of simulated EuST anomalies for the period of 1850–2099. MPI-GE simulations (orange) are compared to CRUTEM4 observations of EuST anomalies (black and white crosses) for the period 1850–2017.Extreme EuST 2σ events (orange dots) are defined as simulated EuST anomalies that deviate from the decadal mean climate state by more than two EuST pre-industrial standard deviations (moving threshold; dashed red line). The decadal mean climate state is defined by the 10-year running ensemble mean (thick red line). The CRUTEM4 EuST anomalies that exceed the 2σ moving threshold are highlighted in black. **(b)** Relative change in number of local 2σ extreme EuST events during the 21st century (2000–2099) relative to the 20th century (1900–1999). **(c)** Relative change in variability based on change in EuST probability density distribution width (2.5th–97.5th percentiles) for early 20th century (1900–1929) compared to late 21st century (2070–2099) for each grid cell. Stippling shows significance for late 21st century PDF widths larger (or smaller) than all the possible 30-year PDFs in the 20th century. Simulations are historical runs for the period 1850–2005 and RCP4.5 runs for the period 2006–2099 from the MPI-GE. All anomalies are calculated with respect to the period of 1961–1990. Spatially averaged temperatures are calculated for the land points in the [10°W–50°E, 35–68°N] domain.

Using MPI-GE we find variability changes that are comparatively smaller and constricted to smaller regions than those found from standard deviation change assessments in multi-model ensembles (Schär et al., 2004; Fischer and Schär, 2009; Fischer et al., 2012). Note that these previous studies were based on higher forcing scenarios more comparable to RCP8.5 than to RCP4.5 (namely SRES A2 and A1B), and that they use the same climatology period to calculate temperature anomalies, 1961–1990, as period of reference to calculate changes in variability, which leads to an underestimation of the variability in the reference period (Supplementary Information; SI Fig. S.1). To avoid this artificial increase in variability we use the period of 1900–1929 as reference for variability changes. As long as the reference period does not overlap with the period used to define the climatology baseline, choosing a different 30-year reference period in the 20th century does not substantially alter our results (SI Fig. S.1).

1.4.2. *Drivers of variability in European heat extremes*

To understand the background conditions that lead to heat extremes and which mechanisms dominate the variability of extreme summer temperatures over Europe, we develop a multiple regression model based on the variables describing the large-scale dynamic atmospheric state (Z500, SLP) and local thermodynamical effects of moisture availability (SM, EVP) as predictors of extreme EuSTs, as well as the annual global mean surface temperature (GMST). Using this novel approach, we can account for up to 90% of the extreme EuST variability. Fig. 1.2 shows the results of our multiple regression analysis, with the point-to-point standardized regression coefficients for each of these predictors. The standardized regression coefficient refers the power of each predictor to affect EuSTs, and illustrates the change in EuST in standard deviations (σ) driven by a change of one standard deviation in the predictor when all other predictors are held constant.

We find that Z500 is the multiple regression predictor that presents the strongest relation to extreme summer temperatures, with 1σ deviations in Z500 driving above 1.2σ deviations in temperature over Northern Europe and slightly less southward. The decrease over Southern Europe may arise from the competing effects of thermodynamical drivers over this region; however it may also arise from the fact that the blocking centers over southern Europe are not necessarily colocated over the center of maximum temperatures, but are shifted westwards due to advection of warm air (Pfahl, 2014), decreasing the Z500-EuST point-to-point regression coefficient in this region. The high Z500-EuST regression coefficient, combined with the fact that geopotential height at the 500 hPa level remains mostly independent from surface temperature conditions, indicates that large Z500 anomalies and their associated persistent blocking events are

the main individual driver of extreme summer temperature variability over Europe in MPI-GE.

In contrast, the other atmospheric variable considered, SLP, presents a regression coefficient to temperature that is slightly lower than for Z500, and negative. Although the notion that temperatures are higher where SLP is lower may appear counterintuitive, it relates to the fact that the peak temperatures are generally not reached where the centre of the high pressure system is located, but rather westwards from this point, where advection of warm air is strongest. In this way, although high-pressure systems are the driver of the anticyclonic conditions that cause high summer temperatures, the highest temperature anomalies would tend not to occur over the points with the highest SLP anomalies, causing this negative relation. Additionally, SLP does not remain as unaffected by the surface conditions as Z500, and part of this negative relation may also appear as a response to the high surface temperatures causing a local low-pressure area. This may occur over regions where intense surface temperatures cause the heating of the air in the lower atmosphere above, resulting in air that is less dense than the air in the neighboring regions and tends to rise, generating a local low-pressure area or thermal low. Our monthly-mean based analysis cannot completely disentangle the driver and response relations between SLP and high temperatures. However, the fraction of explained EuST variance drops substantially when SLP is excluded from the multiple regression model, demonstrating that SLP plays a relevant role in characterizing the background conditions leading to heat extremes, and in explaining part of the EuST variability. Our results indicate that the large-scale dynamical mechanisms combining the effects of Z500 and SLP are the dominating driver of extreme European summer temperature variability.

On the other hand, the local thermodynamical effects of moisture limitation also play a significant, albeit smaller, role explaining extreme summer temperature variability. SM exhibits a negative regression coefficient that is significant mainly only over Northern-Central Europe, where a 1σ decrease in SM corresponds to a maximum increase in temperature of 0.5σ. However, comparable to the case of SLP, it is also not directly clear whether the SM-EuST relation indicates limited moisture availability as a driver of extreme temperatures, or whether soil moisture becomes limited as a response to the high temperatures. Limited moisture availability can therefore be considered a driver of high temperatures only when evapotranspiration remains low, and considered a response when otherwise, although it remains challenging to disentangle these two processes on monthly timescales.

The distinction between driver and response to high temperatures is more straightforward in the case of EVP. The negative EVP-EuST relation is slightly larger than for SM, and significant only over Southern Europe, where limited moisture conditions are commonplace. We find that a 1σ negative anomaly in EVP leads to an increase in temperature from 0.25σ to 1σ. This negative EVP-EuST relation indicates that less water is evaporated into the atmosphere when high temperatures occur. This process can only be triggered by limited moisture availability that limits the amount of evapotranspiration. Under this limited moisture conditions less radiative energy is then transformed into latent heat, increasing the amount of energy available for surface heating and increasing temperature, indicating EVP as a driver of extreme EuSTs over this region. By contrast, in the case of the smaller positive EVP-EuST relation over Northern Europe, the typically large moisture reservoirs in the soil and vegetation over this region result in larger positive EVP anomalies that occur as a response to high temperatures.

Lastly, we include GMST as predictor accounting for the global warming trend and to avoid a spurious increase in the regression coefficients of predictors that also exhibit a similar trend. GMST presents a slightly positive relation to extreme EuSTs, which increases around twofold during the 21st century. This increase derives from EuSTs closely following and contributing to the increasing trend in global temperatures, but we have found no indication that anomalously warm years globally result in higher European summer temperatures (Suarez-Gutierrez et al., 2018).

In contrast to the 90% explained EuST variability and fraction of unexplained variance mostly under 0.1 in the full multiple regression model, we find that performing the same multiple regression analysis but considering only dynamical drivers, Z500 and SLP, as predictors leads to an overestimation of the relations of these predictors to extreme temperatures, while explaining only around 60% of the EuST variability (Fig. 1.3, left column). On the other hand, a multiple regression model with only SM and EVP as drivers of extreme EuSTs leads to a more than twofold overestimation of their relation to extreme temperatures, as well as the area of significance of this relation, while only explaining about 25% of the EuST variability (Fig. 1.3, right column). These results highlight the importance of considering both dynamical and thermodynamical sources of variability simultaneously in order to understand which driving mechanisms dominate the variability in EuSTs and to account for the effect of multicollinearity between the different drivers.

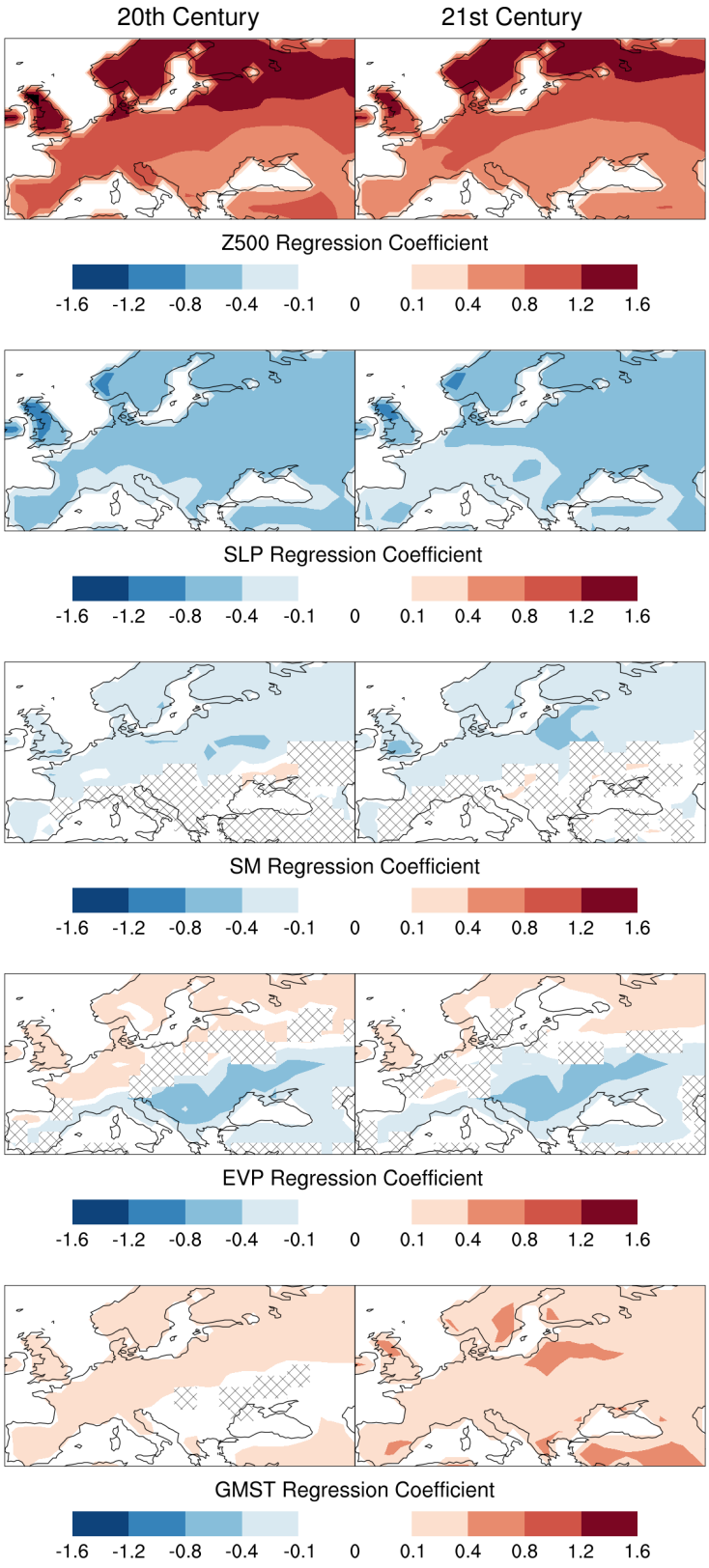


Figure 1.2: Standardized regression coefficients from multiple regression analysis. Point-to-point standardized regression coefficients between 2σ extreme EuST and different drivers from multiple regression analysis for the 20th century (left column) compared to for the 21st century (right column). Hatching represents regions where the variable is excluded from the regression model either because its contribution is not significant or because it exhibits too high multicollinearity with the remaining predictors in the model.

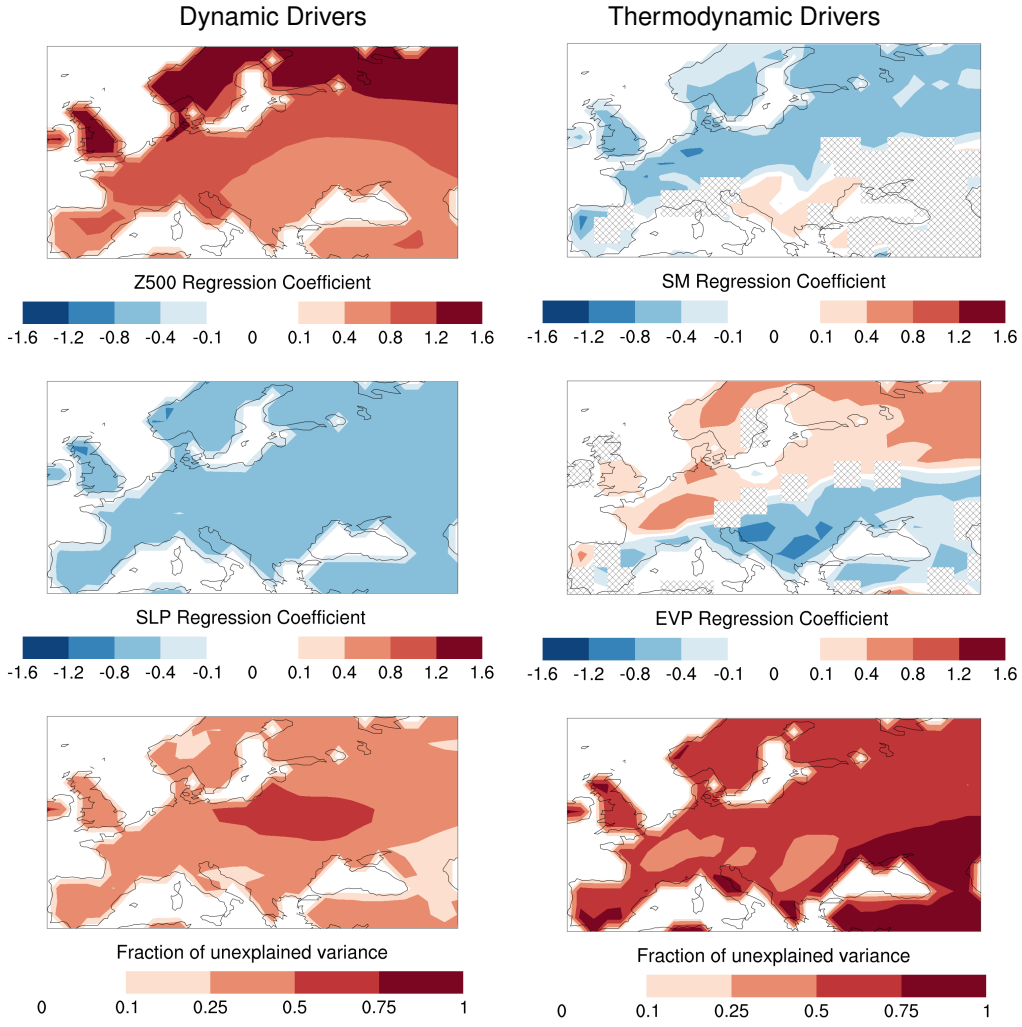


Figure 1.3: Regression coefficients from multiple regression analysis for each set of drivers. Point-to-point standardized regression coefficients for the 21st century from multiple regression analysis between 2σ extreme EuST and large-scale dynamical drivers only (left column) and between 2σ extreme EuST and local thermodynamical drivers only (right column). Hatching represents regions where the variable is excluded from the regression model either because its contribution is not significant or because it exhibits too high multicollinearity with the remaining predictors in the model. The last row represents the fraction of unexplained variance in EuST remaining for each multiple regression model.

Our results indicate that the large-scale dynamical atmospheric conditions, combining the effects of geopotential height at 500 hPa and sea level pressure, are the dominating factor driving extreme European summer temperature variability. By contrast, the local thermodynamical effects caused by limited moisture availability, particularly relevant over Southern and Central Europe, play a secondary role in accounting for this variability. Regarding how these relations may change with warming, we find only minor changes in either the strength or the pattern of the relations of the driving mechanisms with extreme temperatures with warming, with the exception of GMST. These changes include a slight decrease in the regression coefficient to the

dynamical drivers over Southern Europe, and a slight northward expansion of the area where evapotranspiration acts as a significant driver.

1.4.3. *Drivers of variability change in European heat extremes*

The multiple regression analysis in the previous section is a useful tool to quantify the contribution from each driver to the total variability in large samples of EuST extremes. However, it does not allow us to robustly identify the dominating driver of the EuST variability change, because both the strength and pattern of the relation between extreme temperatures and their main driving mechanisms under warming exhibit only minor changes. To explain which mechanisms drive the increase in extreme temperature variability and the 70% increase in extreme temperature events with respect to the changing mean climate during the 21st century, we evaluate how often extreme temperature events occur under extreme atmospheric conditions as opposed as to under extreme moisture limitation, in past compared to future climate conditions.

Extreme 2σ EuST events occur most often, up to 75% of the cases, accompanied by extreme Z500 anomalies, particularly over Northern Europe (Fig. 1.4, top row). During the 21st century this frequency decreases to under 25% in most of Central Europe, while increasing in some Mediterranean regions. On the other hand, we find that during the 20th century fewer than 25% of the extreme EuST events occur under limited moisture conditions of low SM and EVP; while increasing to up to 50% by the end of the 21st century (Fig. 1.4, bottom row). The area of this maximum increase corresponds with the area where we find the largest increase in temperature extremes and variability, and also with the high moisture variability transition zone between dry climate conditions to the south — where commonplace moisture limitation constrains evapotranspiration — and wetter climate conditions to the north — with large moisture reservoirs and where evapotranspiration is limited by radiation (Koster et al., 2009; Fischer et al., 2012). We find that although most extreme EuST events develop under extreme Z500 anomalies, an increasing number of heat extremes develop under neutral atmospheric conditions in the 21st century; while up to 40% more of the extremes develop under extreme moisture limitation.

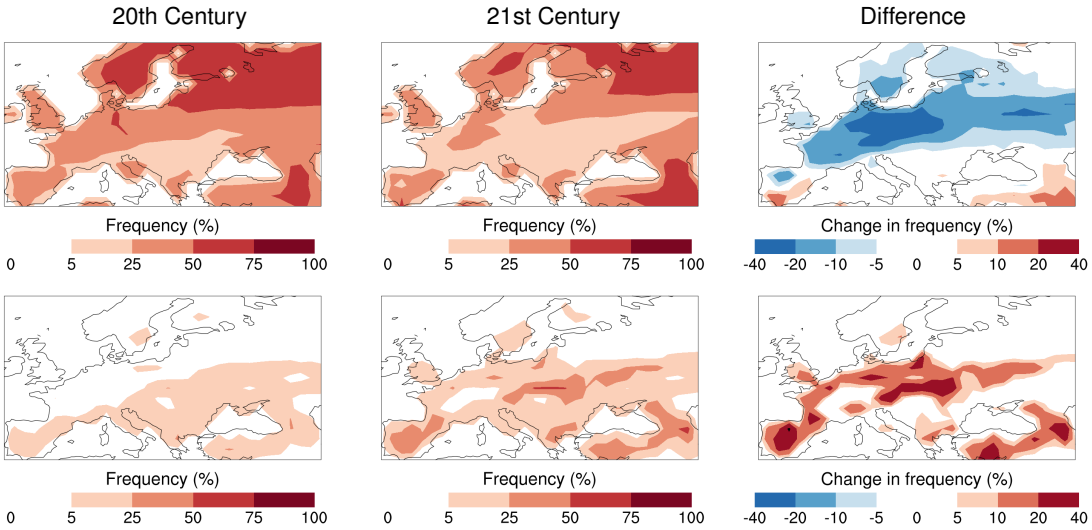


Figure 1.4: Extreme temperatures under extreme atmospheric or moisture conditions. Percentage of extreme 2σ EuST events that exhibit simultaneous extreme atmospheric conditions represented by 2σ extreme Z500 positive anomalies (top row) during the 20th century (1900–1999; left column), during the 21st century (2000–2099; middle column) and difference in this frequency in the 21st century minus in the 20th century (right column). Percentage of extreme 2σ EuST events that exhibit simultaneous extreme moisture conditions represented by extreme 2σ extreme negative EVP and SM anomalies (bottom row) during the 20th century (1900–1999; left column), during the 21st century (2000–2099; middle column) and difference in this frequency in the 21st century minus in the 20th century (right column). Z500, SM and EVP extremes are defined as anomalies that deviate by more than two standard deviations from the average preindustrial conditions in each driver, defined for the period 1850–1899.

We analyze the specific atmospheric and moisture conditions during extreme EuST events at grid cell level over the regions with large increases in EuST extremes and find that the most extreme summer temperatures develop when both persistent anticyclonic conditions and dryness occur, illustrated by high Z500 and low EVP values respectively (Fig. 1.5). These results indicate that extreme temperature events with respect to the mean climate state become more intense during the 21st century under both extreme atmospheric conditions and limited moisture. However, by the end of the 21st century we find that heat extremes develop more frequently as a result of moisture limitation, even under neutral or unfavorable atmospheric conditions, confirming the local thermodynamic effect of moisture limitation as the main driver of the increasing 2σ summer temperature extremes over Europe.

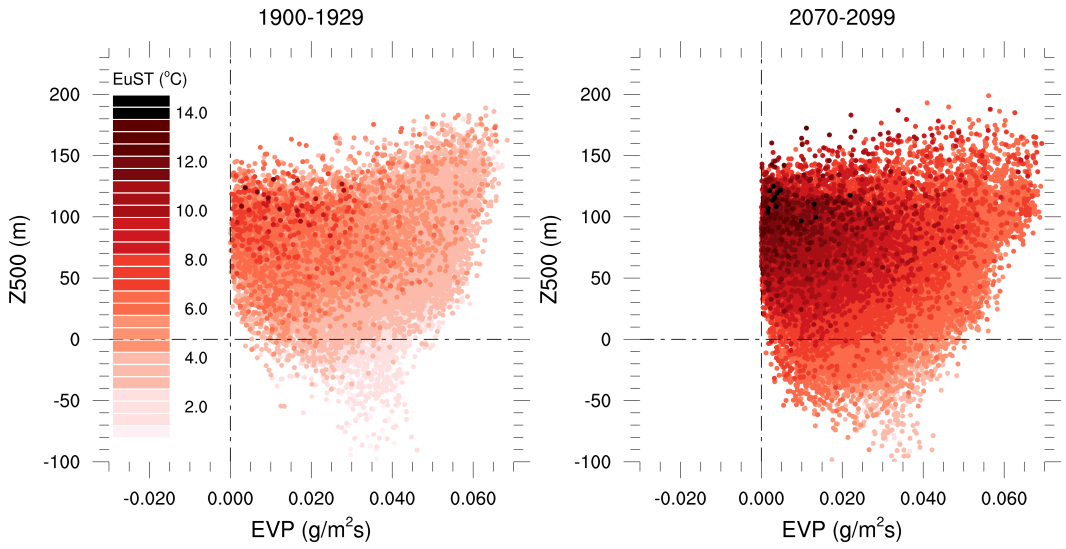


Figure 1.5: Atmospheric and moisture conditions during extreme temperature events. Z500 against EVP absolute values during EuST extreme events for the early 20th century, period 1900–1929 **(a)**, and for the late 21st century, period 2070–2099 **(b)** for each grid cell with an increase in extreme EuST events during the 21st century of 50% or larger. Each point represents one extreme event at grid cell level, defined for EuST anomalies larger than the 2σ threshold for the given grid cell. The color gradient represents the EuST anomaly for each event. Z500 and EVP represent absolute values, with the thermal expansion effect removed from Z500.

1.5. Discussion

Our analysis is based on monthly mean values due to the temporal resolution limitations in MPI-GE. The fact that key characteristic of the anticyclonic or dry conditions that act as driving mechanisms of heat extremes is their persistency, makes their signal still clearly identifiable on monthly values. However, we would ideally use hourly to daily values to capture the amplitude and duration of extremes more precisely, and to directly separate between the specific conditions causing heat extremes from those occurring as a result to them. This is particularly relevant for mechanisms regarding sea level pressure and soil moisture, that can both act as a driver of extremes and result as a response to high temperatures. Although our results agree with the theoretical understanding of how heat extremes develop, it remains challenging to disentangle cause and effect processes on monthly timescales.

The distinction between cause and effect can be partly achieved by applying a multiple regression model that includes all potential sources of variability and that is based on large enough samples of heat extremes. This method quantifies the internal variability in the response of temperature to different background conditions,

and can to some extent separate conditions that are necessary to the development of extremes from those that may be circumstantial. Using a multiple regression analysis also allows us to account for the multicollinearity effect of different drivers on each other. This multicollinearity refers to one or more of the drivers, additionally to having an effect on temperature, having also an intensifying or counteracting effect on the other drivers. Multicollinearity is inherit to the highly complex climate system and cannot be simply removed; but its effect on our analysis can be reduced. To avoid an inflation of the explained variability due to high multicollinearity we exclude from our multiple regression model variables with a variance inflation factor of more than 10. More conservative multicorrelation thresholds may eliminate relevant variables from the model and lead to spurious or less robust results, but do not substantially influence our conclusions (SI Fig. S2)

To achieve the best combination of predictors for our multiple regression model, we evaluated several other variables representing the main driving mechanisms. For representing the large-scale dynamic drivers we additionally evaluated 850 hPa geopotential height (Z850) and North Atlantic jet stream position. First, Z850 presents a strong relation with EuST in MPI-GE, however it also exhibits high correlation with Z500 and SLP. This high multicollinearity with the other atmospheric variables leads to the exclusion of Z850 from the multiple regression model in the majority of the domain when Z500 and SLP are included as predictors. On the other hand, excluding SLP and, especially, Z500 to include Z850 instead results in a substantial decrease in explained EuST variance. Second, the North Atlantic jet stream position, represented as the latitude where the monthly-average zonal wind maxima occur, exhibits only a minor relation to extreme EuSTs in MPI-GE. Including jet stream position as predictor in the model has only a minor effect on the explained EuST variance. For representing the effects of moisture limitation we also considered soil moisture fraction in the month preceding the extremes as predictor. However, similarly to the case of Z850, this lagged predictor exhibits too high collinearity to SM in the extreme month, while explaining less of the EuSTs variance.

Another potential caveat of our approach is the limitation to point-to-point relations in the multiple regression analysis, that may lead to an underestimation of non-local effects, particularly relevant for of large-scale drivers. This is to some extent counteracted by the use of monthly averages of the dynamic atmospheric conditions as predictors. However, it is possible that the contribution from large-scale dynamical drivers, that we find dominates extreme European summer temperature variability, is still underestimated in our study due to an underrepresentation of their non-local

contribution to extreme temperatures.

Our results highlight the relevant role that internal variability plays in the development of heat extremes in a warming world. To quantify the role of the main mechanisms driving extreme summer temperatures, we need to understand to what extent internal variability influences the development of heat extremes under different background conditions. Large ensemble experiments like MPI-GE are a great tool for this purpose, because they provide large samples of extreme events that develop under different background conditions, but under the same external forcings and model physics. However, basing our analysis on large ensembles of simulations with the same climate models implies that the results and conclusions drawn from this analysis may be subject to the flaws and uncertainties inherent to any single-model study. Our conclusions on how different driving mechanisms contribute to heat extremes and changes in temperature variability are characterized by how the different driving mechanisms perform and affect temperatures in MPI-ESM. This may occur differently in different climate models, and there is valuable knowledge to be gained from repeating this analysis on other large ensemble experiments.

In particular, the thermodynamic effect of moisture availability as a driving mechanism of simulated heat extremes is affected by the soil hydrology scheme included in the model. Although the 5-layer scheme in MPI-GE represents an improvement with respect to previous versions of MPI-ESM, biases in soil moisture memory may remain. Also the location of the transition zone between wet and dry climates may vary in different models, causing differences on where the maximum increase in 2σ extremes occurs as a response to high moisture variability leading to extremely dry conditions. However, our conclusions are supported by robust evidence of a trend towards dryer summer conditions over Europe (Briffa et al., 2009; Ruosteenoja et al., 2018; Spinoni et al., 2018), and of a poleward shift in the transition zone, where moisture is highly-variable, that can affect temperature variability under global warming (Seneviratne et al., 2006; Fischer et al., 2012).

On the other hand, there is no clear evidence of whether the blocking and anticyclonic conditions that act as dynamical large-scale driving mechanism of heat extremes will occur with more or less frequency in the future (Woollings et al., 2018). While some studies find significant positive trends in the frequency of anticyclonic circulations (D. E. Horton et al., 2015); most climate models show a decline in blocking conditions with relatively good agreement (Matsueda and Endo, 2017; Woollings et al., 2018). However, the atmospheric circulation over Europe is not realistically captured by most climate models. Climate models tend to consistently underestimate the occurrence and

persistency of blocking events compared to observed estimates (Davini and D'Andrea, 2016), a tendency that is also shown by MPI-ESM (Müller et al., 2018). Despite these potential biases, the relationship between blocking and heat extremes is often well captured by models, particularly in large ensemble experiments (Schaller et al., 2018). This indicates that, although the effect of blocking as a driver of heat extremes may be well captured in MPI-GE, a biased representation of blocking conditions or its future occurrence may cause biases in the frequency or the intensity of heat extremes in MPI-GE. If this were the case, more frequent blocking conditions could lead to a larger increase of extreme events under future warming, additionally to the increasing number of extremes that occur under unfavorable atmospheric conditions due to the effect of moisture limitation.

Regardless of the potential limitations of our study, we demonstrate that to obtain robust quantifications of the contributions from different drivers of heat extremes, we need to account for the complex multicollinearity between these driving mechanisms, and this can only be achieved by considering all sources of variability simultaneously. We also demonstrate that the main driver of variability in European summer temperatures may be different from the dominant driver of variability change, and that large samples of extreme events like those provided by large ensemble experiments are crucial to detect and understand these changes.

1.6. Summary and Conclusions

We use the 100-member MPI Grand Ensemble to disentangle the contributions from large-scale dynamic atmospheric conditions and local thermodynamic effects of moisture limitation as drivers of variability in large samples of European summer heat extremes. Using a novel extreme event definition we find a 70% increase in heat extremes with respect to the evolving decadal climate during the 21st century. This extreme event characterization allows us to study changes in the number of extremes not only caused by the shift in the distribution towards higher temperatures, but by changes in temperature variability that result from changes in the underlaying driving mechanisms of heat extremes. The increase in extremes reaches maximum values of more than 100% in central Eastern Europe, where summer temperature variability increases by 35% in the MPI-GE. In agreement with previous studies (Fischer and Schär, 2009; Fischer et al., 2012; Bathiany et al., 2018), our results indicate that summer temperature variability is increasing under global warming over some parts of Europe. This increase in variability results on temperature extremes that present larger deviations from the mean average conditions, thus contradicting the assumption

that European summer extreme temperatures mostly follow mean summer warming (Ballester et al., 2009; Ballester et al., 2010).

With a multiple regression approach we find that the large-scale dynamical mechanism representing blocking and anticyclonic conditions is the main driver of variability in extreme European summer temperatures in MPI-GE, both under past and future climate conditions. In particular, we find that 500 hPa geopotential height is the dominating individual driver of extreme temperature variability, especially over Northern Europe. On the other hand, local thermodynamic drivers play a secondary role in explaining total extreme temperature variability. Furthermore, we find that considering both sources of variability simultaneously is crucial to understand extreme temperature variability. Assessing the contribution from only one type of drivers can explain much less of the extreme temperature variability and leads to an overestimation of the effect of the drivers on extreme temperatures, particularly when considering only the local thermodynamical drivers.

Our results also show that the dynamical atmospheric mechanisms that act as dominant driver of extreme summer temperature variability are not the dominant driver of variability change. We find that most European summer heat extremes occur under extreme 500 hPa geopotential height conditions, and that the most extreme summer temperatures develop when both persistent anticyclonic conditions and dryness occur. However, we find that heat extremes develop 40% more frequently as a result of moisture limitation during the 21st century, even under neutral or unfavorable atmospheric conditions, confirming that the increase in European summer heat extremes with respect to the evolving decadal climate is driven by the the local thermodynamic effect of moisture limitation.

CHAPTER 2

INTERNAL VARIABILITY OF EUROPEAN SUMMER TEMPERATURES AT 1.5°C AND 2°C OF GLOBAL WARMING

2.1. Summary

We use the 100-member Grand Ensemble with the climate model MPI-ESM to evaluate the controllability of mean and extreme European summer temperatures with the global mean temperature targets in the Paris Agreement. We find that European summer temperatures at 2°C of global warming are on average 1°C higher than at 1.5°C of global warming with respect to pre-industrial levels. In a 2°C warmer world, one out of every two European summer months. would be warmer than ever observed in our current climate. Daily maximum temperature anomalies for extreme events with return periods of up to 500 years reach return levels of 7°C at 2°C of global warming and 5.5°C at 1.5°C of global warming. The largest differences in return levels for shorter return periods of 20 years are over southern Europe, where we find the highest mean temperature increase. In contrast, for events with return periods of over 100 years these differences are largest over central Europe, where we find the largest changes in temperature variability. However, due to the large effect of internal variability, only four out of every ten summer months in a 2°C warmer world present mean temperatures that could be distinguishable from those in a 1.5°C world. The distinguishability between the two climates is largest over Southern Europe, while decreasing to around 10% distinguishable months over Eastern Europe. Furthermore, we find that 10% of the most extreme and severe summer maximum temperatures in a 2°C world could be avoided by limiting global warming to 1.5°C.

2.2. *Introduction*

Recent decades have been marked by an increasing number of extremely warm summers over the European continent (Schär et al., 2004; Christidis et al., 2014), and this rising tendency, largely attributed to anthropogenic climate change, is expected to be accentuated under further global warming (Stott et al., 2004; Meehl and Tebaldi, 2004; Christidis et al., 2014; Vautard et al., 2014). In the framework of the Paris Agreement, it is crucial to evaluate which of the risks and impacts of climate change would be reduced by limiting global warming to 1.5°C (hereafter 1.5°C target) in contrast to by limiting warming to 2°C (hereafter 2°C target). Here we examine to what extent the most extreme European summer temperatures at 2°C of global warming could be avoided in a 1.5°C warmer world. In other words, we examine to what extent extreme European summer temperatures could be controlled by limiting global warming to the global mean temperature limits of the Paris Agreement. To evaluate the controllability of European summer temperatures with global mean temperature limits, it is necessary to robustly characterize the irreducible uncertainty that arises from chaotic internal variability (Sriver et al., 2015; Hawkins et al., 2016). For this purpose, we use a state-of-the-art tool to sample internal variability: the 100-member Max Planck Institute Grand Ensemble (MPI-GE) (Bittner et al., 2016; Hedemann et al., 2017; Suarez-Gutierrez et al., 2017; Maher et al., 2019).

Summer monthly mean and daily maximum temperatures at 2°C of global warming are projected to become around 1°C higher over Europe than at 1.5°C of warming (Schleussner et al., 2016; Perkins-Kirkpatrick and Gibson, 2017; King and Karoly, 2017; Sanderson et al., 2017). Sanderson et al. (2017) and Wehner et al. (2017) also find differences of around 1°C between 20-year return levels of maximum temperatures at 1.5°C versus at 2°C of global warming. These studies are based on climate modelling experiments of different nature: the Coupled Model Intercomparison Project phase 5 (CMIP5) multi-model ensemble (Schleussner et al., 2016; Perkins-Kirkpatrick and Gibson, 2017; King and Karoly, 2017), and ensemble-experiments such as the Half a Degree Additional warming, Prognosis and Projected Impacts project (HAPPI; Mitchell et al., 2017) atmosphere-only runs (Wehner et al., 2017), and the Community Climate 10-member CESM1 ensemble (Sanderson et al., 2017).

A key factor in evaluating the differences between the climates for the two targets is to consider the magnitude of the response in the Earth's climate to half a degree more of warming relative to the signal of internal variability. For this purpose, large ensembles of simulations based on the same coupled climate models (like the experiments described in Deser et al., 2012; Kay et al., 2015; Rodgers et al., 2015; Fyfe et al.,

2017; Sanderson et al., 2017 and Suarez-Gutierrez et al., 2017) are the best available tools, because they provide unambiguous characterizations of the simulated internal variability in a changing climate without being confounded by different model configurations. MPI-GE has 100 independent realizations, which start from different times of a pre-industrial control run but are driven by the same external forcings, and is currently the largest existing ensemble from a fully-coupled Earth System Model. The large size of the ensemble is a crucial requirement to robustly sample internal variability and to empirically evaluate the statistical significance of changes. An ensemble size of 100 simulations under the same forcing conditions allows 1-in-100-years events to occur on average every simulated year, which provides the large samples of extreme events under different warming conditions that are necessary for the purpose of our study.

We use MPI-GE simulations not only to evaluate average changes in summer monthly mean, block maximum, and extreme temperatures, but also to quantify the irreducible uncertainty in European summer temperatures that arises through internal variability. We evaluate the controllability of European summer temperatures with global temperature targets by quantifying the distinguishability of European summer months at 2°C of global warming with respect to those in a 1.5°C warmer world. We also quantify for the first time changes in the magnitude of extreme summer temperature events with return periods of up to 500 years.

2.3. *Data and Methods*

We use transient climate simulations from the 100-member MPI-GE under historical, RCP2.6 and RCP4.5 forcing conditions. The Grand Ensemble uses the model version MPI-ESM1.1 in the low resolution (LR) configuration, with resolution T63 and 47 vertical levels in the atmosphere (Giorgetta et al., 2013) and 1.5°resolution and 40 vertical levels in the ocean (Jungclaus et al., 2013). MPI-ESM has an equilibrium climate sensitivity of 2.8°C and a transient climate response of 1.57°C, values that are slightly below the values for the CMIP5 version of MPI-ESM (Flato et al., 2013). Each of the 100 realizations in the ensemble is based on the same model physics and parametrizations, and is driven by the same external forcings. The realizations differ only in their initial climate state, taken from different points of the model's pre-industrial control run.

MPI-ESM-LR has a relatively low resolution, comparable to most of the models in the CMIP5 experiment, which can influence the model's ability of simulating small-scale processes and affect the reliability of our projections. However, we find

that MPI-GE offers an adequate representation of the observed estimate of internal variability in global mean temperatures and European summer temperatures. The evaluation of MPI-GE in Appendix C indicates a slight overestimation of the frequency of colder than average summer months in the ensemble simulations that may be partially caused by biases in precipitation variability (Appendix Figures A.4, A.5 and A.6). Observational data from the HadCRUT4.5 (Morice et al., 2012; Osborn and Jones, 2014) and the CRUTEM4.6 (Jones et al., 2012) datasets are used for comparisons to current climate conditions.

Global mean surface temperature (GMST) is defined as the annually averaged, global mean, near-surface 2m air temperature anomaly. European monthly mean summer temperature (EuST) is defined as the monthly averaged 2m air temperature anomaly for the summer months (JJA), averaged over the land-only grid cells in the region defined by the [10°W–50°E, 35–60°N] latitude-longitude domain. Ideally, we would use daily temperatures to capture the amplitude of internal variability more precisely. However, monthly frequency is the highest output frequency in the Grand Ensemble simulations. We also use the summer maximum value of daily maximum temperature (EuSTXx) as the block maximum temperature anomaly reached each summer at each grid cell averaged over the land-only grid cells in the same domain. All anomalies are calculated with respect to the pre-industrial conditions defined by the period 1851–1880. Observed temperature anomalies are transformed to anomalies with respect to pre-industrial levels following the estimates in Hawkins et al. (2017).

We construct representative samples of the quasi-stationary climate conditions at 1.5°C and 2°C of global warming from transient climate simulations with a *time-slice* method that is similar to the methods used in Schleussner et al. (2016) or King and Karoly (2017). We select years of 0°C, 1.5°C, and 2°C of global warming with respect to pre-industrial levels from all historical, RCP2.6, and RCP4.5 100-member simulations from MPI-GE. Global mean temperatures deviate from the long-term mean state on year-to-year timescales due to the effect of internal variability. Therefore, we calculate centered decadal-averaged GMST to robustly define global warming levels. We define years of 0°C of global warming as those years in which the centered decadal-averaged GMST is in the range of 0°C plus-minus one standard deviation of GMST, 0.13°C. Similarly, for 1.5°C and 2°C of global warming above pre-industrial levels we select years in which the centered decadal-averaged GMST is in the range of 1.5±0.13°C and 2±0.13°C, respectively.

Based on this assumption of quasi-stationarity, we are able to calculate well-defined return levels of maximum daily temperature anomalies for extreme events

with return periods of up to 500 years based on the probability distributions of spatially averaged EuSTXx. The high ensemble resolution provides large enough samples of extreme events, with 3000 simulated years for each climate conditions, that allow us to empirically calculate probability distributions and eliminate the need to parametrize the tails of the distributions with extreme value statistics. We calculate the return period of any given return level as the inverse of the probability of exceedance of this return level per year. We obtain this probability of exceedance directly as one minus the cumulative probability of the given return level.

2.4. Results

We evaluate three sets of 100 transient climate simulations under historical, RCP2.6 and RCP4.5 forcings from MPI-GE. Figure 2.1 illustrates the characterization of the global warming levels analyzed here. Within the 21st century, around half of the ensemble members show GMST projections that remain below the 2°C warming target for the RCP4.5 scenario, while most of the GMST projections following the RCP2.6 scenario are in agreement with the 1.5°C warming target. To achieve a sample size of around 3000 simulated years for each climate, we choose only simulated years within the periods marked by the black dashed lines in Figure 2.1.

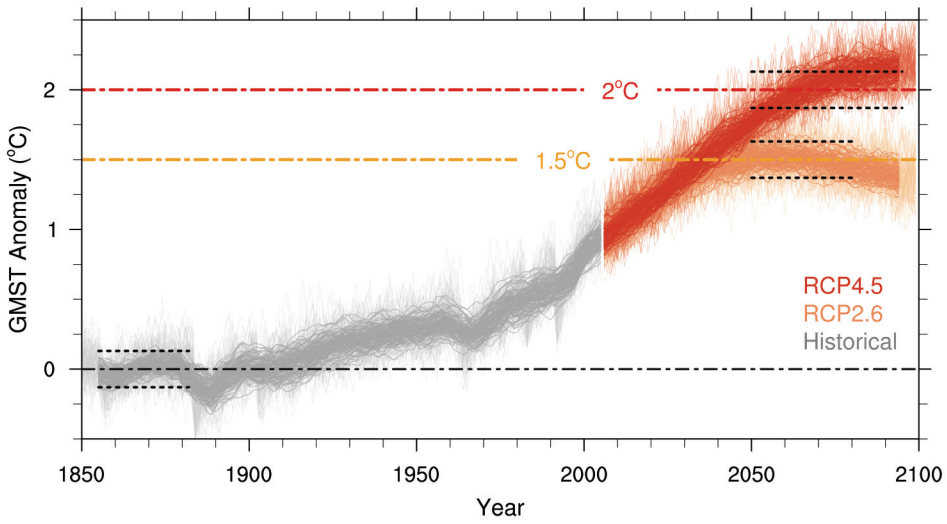


Figure 2.1: Global mean surface temperature (GMST) in MPI-GE. Time series of annually averaged GMST anomalies (colored thin lines) and centered decadal-averaged GMST anomalies (colored thick lines) for the period 1850–2099, simulated by MPI-GE. Simulations are historical runs for the period 1850–2005 (gray lines), and RCP2.6 (orange lines) and RCP4.5 (red lines) for the period 2006–2099. The black dashed lines show the periods of sampling for each warming level and the standard deviation of GMST from the long-term warming levels for pre-industrial and warming targets of 1.5°C and 2°C.

2.4.1. *European summer monthly mean temperatures*

The empirical probability distributions of GMST for the years selected from the Grand Ensemble simulations show that the three samples describing pre-industrial climate conditions and the climates for the two targets are significantly distinguishable from each other (Figure 2.2a). The distribution of GMST for pre-industrial conditions, with a width of around 0.9°C, presents no overlap with either the 1.5°C or the 2°C distributions. The two target distributions overlap over 5% of their area and have a width of around 1°C. Using a narrower range in decadal-averaged GMST for defining each climate reduces the sample size of the distributions, but does not substantially influence our results.

Whereas the climates for the two warming targets are distinguishable at the global level, European summer temperatures have substantially larger internal variability than global mean temperatures. The probability distributions of summer monthly mean temperature anomalies for the selected years describing the two target climates present a width more than four times larger than the GMST distributions (Figure 2.2b). In contrast to the GMST distributions, and due to the large influence of internal variability, all three distributions of European summer temperatures for different warming levels present some fraction of areal overlap. This overlap is largest when comparing the two target climates, with a 60% areal overlap between the EuST probability distributions at 1.5°C and at 2°C.

With respect to the relative GMST distributions, the EuST distributions for both warming targets are also shifted towards higher mean temperature anomalies. The probability distribution of EuST is centered around 2°C anomalies at 1.5°C of global warming, and for 2°C of global warming is centered around anomalies of 3°C, indicating a pattern of regional amplification of global warming over Europe that is in line with previous expectations (IPCC, 2013).

The two highest observed European summer monthly mean temperatures in July 2010 and August 2017 are marked in Figure 2.2b for comparison, as well as the value for August 2003. The estimates for 2003 and 2017 are comparable to the average European summer month in a 1.5°C warmer world, and also comparable to summer months in the upper tail of the pre-industrial distribution. This result is consistent with previous findings that project the 2003 summer temperatures to become commonplace around the 2040s (Stott et al., 2004). On the other hand, the 2010 value is comparable to the average summer month in a 2°C warmer world. This result indicates that, under 2°C of global warming, every other European summer month would be warmer than the

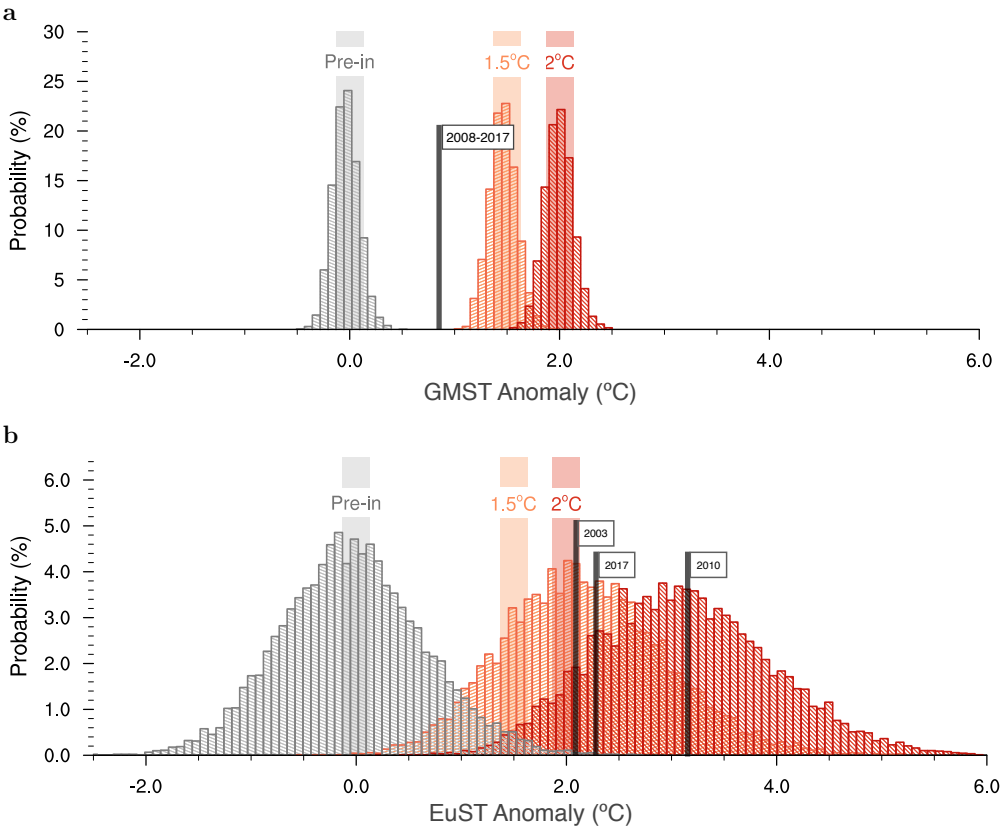


Figure 2.2: Probability distributions at different global warming levels. (a) Probability distribution of global mean surface temperature (GMST) anomalies for pre-industrial conditions (gray), and for global warming levels of 1.5°C (orange) and 2°C (red) above pre-industrial levels, simulated by MPI-GE. The shaded areas indicate the range of ± one standard deviation of GMST around the mean state of 0°C (gray), 1.5°C (orange) and 2°C (red). Each distribution has a sample size of around 3000 simulated years. The black reference line marks the observed decadal-average GMST anomaly for 2008-2017 from HadCRUT4 data. **(b)** Probability distribution of European monthly mean summer temperature (EuST) anomalies as in (a). Each distribution has a sample size of around 9000 summer months. The black reference lines mark the observed EuST monthly mean anomaly for August 2003, August 2017 and July 2010 from CRUTEM4 data. Bin size is 0.075°C; frequencies are normalized to unity and translated to percentage.

warmest summer month on record in current climate conditions; while the other half of European summer months in a 2°C world would be more similar to current climate conditions. However, the 60% of areal overlap between the two target distributions indicates that less than half of the European summer months in a 2°C world, four months out of every ten, would be distinguishable from those in a 1.5 °C world.

Figure 2.3 illustrates these results locally, presenting differences in mean temperatures and in temperature variability as well as distinguishability between the two warming limits per grid cell. Mean temperature differences are around 1°C, consistent with the results shown by King and Karoly (2017) and Sanderson et al. (2017),

and largest over the Mediterranean region (Figure 2.3e). However, the temperature variability, measured as the width of the local probability distributions between the 97.5th and 2.5th percentiles, is very large over these regions (Figure 2.3b, d). This irreducible spread caused by internal variability of up to 10°C is much larger than the average temperature changes. However, the change in variability in European summer temperatures is relatively small in comparison with the mean temperature changes and is localized over central-northern Europe (Figure 2.3f). This pattern of change in variability, shown here as the difference in spread, is also present when variability changes are portrayed as ratios (Supporting Figure S.3 in Appendix D).

Figures 2.3g and 2.3h show how often European summer months at 2°C of global warming could be distinguishable from those at 1.5°C of warming for each grid cell. We define the distinguishability between the two target climates as the percentage of summer months in a 2°C world that could not be part of the temperature distribution of the 1.5°C world. For Figure 2.3g, we base this estimate of distinguishability between the two climates on the areal overlap between the two temperature distributions at grid-cell level. This distinguishability is largest over Southern Europe, with a maximum of around 35% distinguishable summer months, and decreases to around 10% for Eastern Europe. We also include a second, more conservative measure of distinguishability in Figure 2.3h, based on the percentage of summer months in a 2°C world that present EuSTs larger than the 95th percentile in the EuST distribution at 1.5°C of warming. This measure of distinguishability yields values of around 5% to 20% months with distinguishable mean temperatures. For both measures, we find the minimum distinguishability between the two climates over Eastern Europe, where summer temperature variability is largest.

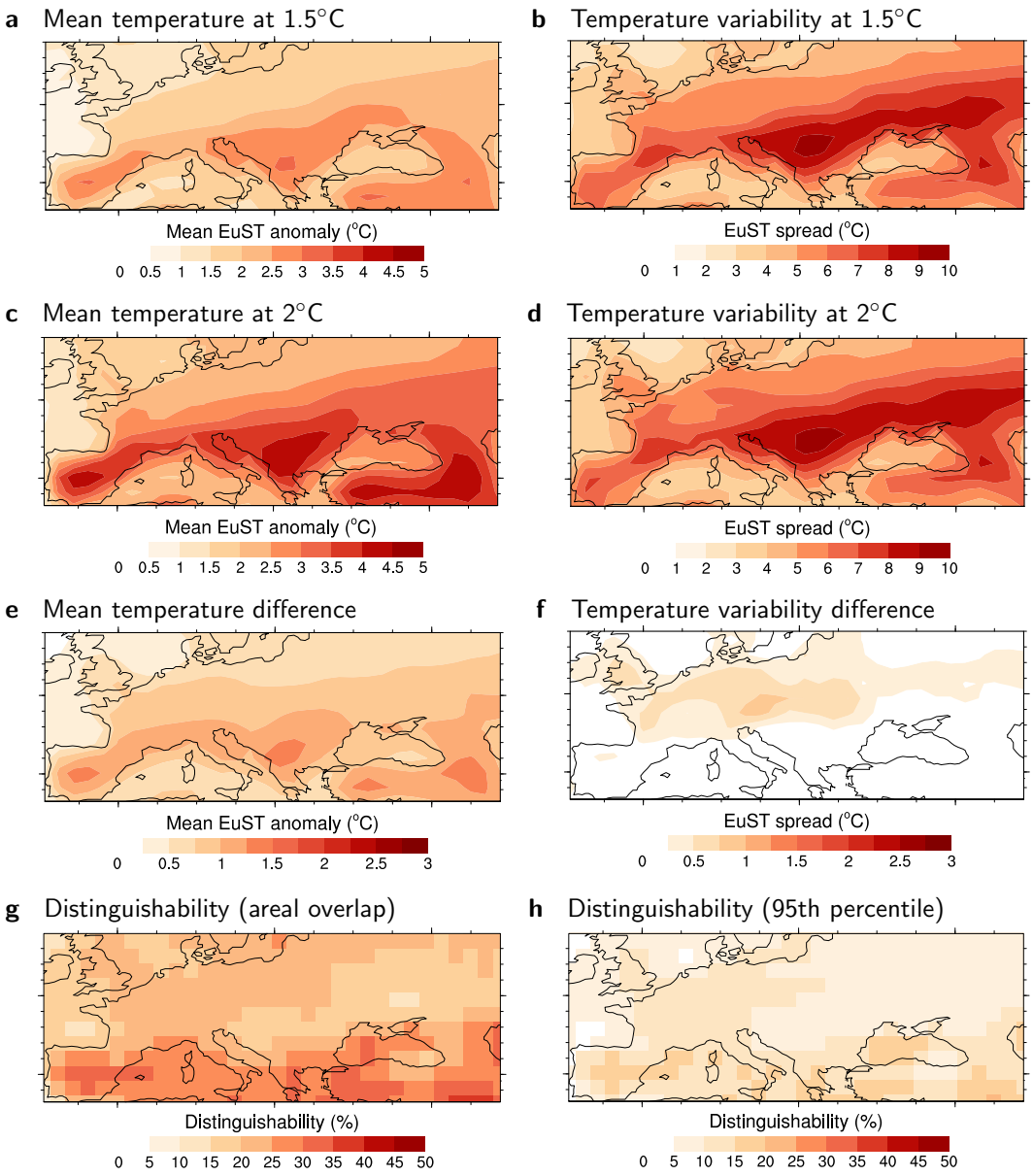


Figure 2.3: Mean temperatures and temperature variability at different global warming levels. (a) Average monthly mean temperature anomaly at 1.5°C of global warming. **(b)** Spread in monthly mean temperature anomalies at 1.5°C of global warming, measured as the difference between the 97.5th and 2.5th percentiles. **(c)** Average monthly mean temperature anomaly at 2°C of global warming. **(d)** Spread in monthly mean temperature anomalies at 2°C of global warming, as in (b). **(e)** Difference in average monthly mean temperature anomaly at 2°C of global warming minus at 1.5°C of global warming. **(f)** Difference in the spread in monthly mean temperature anomalies at 2°C of global warming minus at 1.5°C of global warming, as in (b). **(g)** Summer months at 2°C of global warming with temperatures that could be distinguishable from those at 1.5°C based on the areal overlap between the two distributions. **(h)** Summer months at 2°C of global warming with temperatures that could be distinguishable from those at 1.5°C based on the percentage of months at 2°C of warming above the 95th percentile in the 1.5°C distribution.

2.4.2. *Return levels in European summer maximum temperatures*

In this section we evaluate how the return levels of extreme European summer temperatures change under global warming. We base our analysis on block maximum values of daily maximum temperature (EuSTXx) simulated by MPI-GE, shown in Figure 2.4a against simultaneous global mean temperatures for the three sampled climate conditions. EuSTXx anomalies exhibit large variability, with a spread of more than 4°C. EuSTXx anomalies are centered around 3°C in a 1.5°C warmer world, and around 4°C in a 2°C warmer world. The EuSTXx distributions for 1.5°C and 2°C of warming present an areal overlap of 70% (not shown), indicating that only three out of every ten of the summer maximum temperatures in a 2°C warmer world could be distinguishable from those in a 1.5°C world. Due to this large overlap and the large variability in block maximum daily maximum temperatures, only 10% of the most extreme EuSTXx values at 2°C of warming would be out of reach in a 1.5°C warmer world (with 99% confidence).

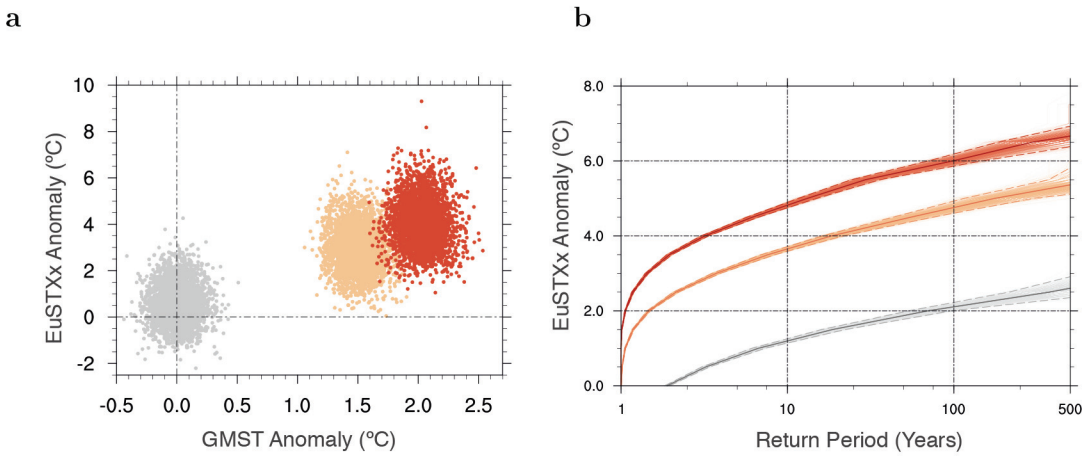


Figure 2.4: European summer maximum value of daily maximum temperature (EuSTXx) at different global warming levels. (a) EuSTXx anomalies against GMST anomalies for pre-industrial conditions (gray), and for global warming levels of 1.5°C (orange) and 2°C (red) above pre-industrial conditions, simulated by MPI-GE. **(b)** Return levels of EuTXx summer block maxima against their return period, represented by the thick solid lines in gray for pre-industrial conditions, in orange for global warming levels of 1.5°C and in red for 2°C above pre-industrial conditions. Uncertainty in these return levels is estimated by bootstrap-resampling with replacement. The colored thin lines represent 1000 individual bootstrap estimates; the colored dashed lines represent the 95% confidence intervals.

The return levels of summer maximum daily maximum temperature anomalies for events with return periods from 1 to 500 years are presented in Figure 2.4b for the three sampled climate conditions. 2°C events, which have return periods of around 100 years under pre-industrial conditions, are projected to occur every one to two years

in both target climates. 10-year return period events present values of around 3.5°C in a 1.5°C warmer world and of around 5°C in a 2°C warmer world. This difference of 1°C to 1.5°C in return levels of extreme summer temperatures between the two target climates is roughly maintained for increasingly longer return periods. For 500-year return periods, we find return levels that reach values of almost 7°C at 2°C of global warming and of around 5.5°C at 1.5°C of global warming. We reach similar results by basing this analysis on the summer minimum value of daily minimum temperatures (EuSTXn; Supporting Figure S.4 in Appendix D).

The differences between return levels for different return periods are not homogeneously distributed over Europe. 20-year return levels of EuTXx anomalies present a maximum over south-eastern Europe and are generally largest in southern Europe for both target climates (Figure 2.5a, b). The difference in 20-year return levels between the two target climates is also largest in southern Europe, with values around 1.5°C (Figure 2.5c), overlapping with the region of largest mean temperature increase in Figure 2.3e. Similarly, 100-year return levels also present a maximum over south-eastern Europe (Figure 2.5d, e). In contrast to events with shorter return periods, extreme events with return periods of 100 to 500 years (not shown) present differences between the two target climates that are largest in central Europe, over the regions where we find the largest increase in temperature variability in Figure 2.3f (Figure 2.5f).

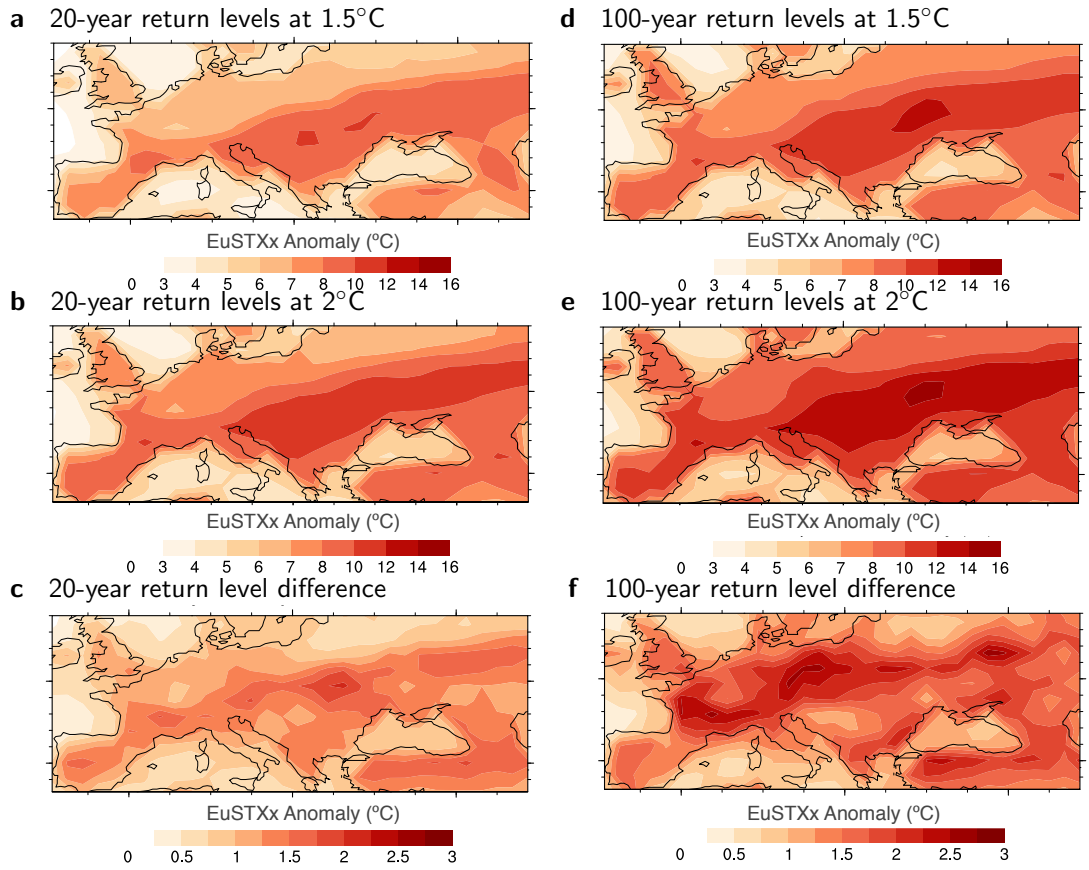


Figure 2.5: Return levels of summer block maximum daily temperatures at different global warming levels. **(a)** Daily maximum temperature anomalies for events with 20-year return periods at 1.5°C of global warming. **(b)** Daily maximum temperature anomalies for events with 20-year return periods at 2°C of global warming. **(c)** Daily maximum temperature anomalies for events with 20-year return periods at 2°C of global warming minus at 1.5°C of global warming. **(d)** Daily maximum temperature anomalies for events with 100-year return periods at 1.5°C of global warming. **(e)** Daily maximum temperature anomalies for events with 100-year return periods at 2°C of global warming. **(f)** Daily maximum temperature anomalies for events with 100-year return periods at 2°C of global warming minus at 1.5°C of global warming.

2.5. *Summary and Conclusions*

We use the state-of-the-art MPI-GE to evaluate the controllability of monthly mean, block maximum, and extreme European summer temperatures under the global warming limits of the Paris Agreement. We find that at 2°C of global warming, one out of every two European summer months is projected to be warmer than ever observed in our current climate. We find European summer monthly mean temperature differences of around 1°C between the 2°C and the 1.5°C warmer worlds, in line with previous results by Schleussner et al. (2016), Perkins-Kirkpatrick and Gibson (2017), King and Karoly (2017) and Sanderson et al. (2017). We also find differences of around 1°C in maximum daily temperature anomalies for extreme events with return periods of up to 500 years, which reach values of almost 7°C at 2°C of global warming and of around 5.5°C at 1.5°C of global warming. For 20-year return period events these differences are consistent with the differences in 20-year return levels shown by Sanderson et al. (2017) and Wehner et al. (2017), reaching values of around 1.5°C. These differences in 20-year return levels are largest in southern Europe, over the regions where we find the largest mean temperature increase. For events with return periods of 100 to 500 years these differences reach values of more than 2.5°C and are largest in central Europe, over regions where we find the largest temperature variability increase.

Our results indicate that due to the irreducible uncertainty in European summer temperatures caused by internal variability, only 40% of the European summer monthly mean temperatures in a 2°C warmer world would be distinguishable from those in a 1.5°C warmer world. This distinguishability between the two climates is largest over Southern Europe, and decreases to around 10% over Eastern Europe. Furthermore, we find that only 10% of the most extreme summer maximum temperatures in a 2°C world would be avoided at 1.5°C of global warming. However, although only 10% of the most extreme temperatures could be avoided, these events would correspond to the most extreme and severe heat waves, the ones with the most critical consequences.

Although these results may be subject to uncertainties inherent to any single-model study as well as to the relatively low resolution of MPI-ESM-LR, we believe that the concepts and methods at the core of our analysis can serve as blueprint for future studies with focus on other regions and phenomena. Our findings highlight the limited controllability of the amplitude of extreme temperature events at regional levels by establishing global mean temperature limits, and emphasize the importance of considering the irreducible uncertainty introduced by chaotic internal variability to evaluate the impacts of climate change.

CHAPTER 3

GLOBAL HEAT HOTSPOTS UNDER GLOBAL WARMING

3.1. Summary

We use the Max Planck Institute Grand Ensemble (MPI-GE) to determine the maximum global warming level that allows us to avert the risk of extreme heat, once the irreducible uncertainty introduced by internal variability is considered. We find that limiting global warming to 2°C above pre-industrial levels can substantially reduce the risk of extreme heat hotspots in all five extreme heat metrics considered — maximum reachable temperatures, return periods of extreme temperatures, maximum temperature variability, sustained tropical night temperatures, and extreme wet bulb temperatures. At 2°C of warming, MPI-GE projects maximum reachable summer temperatures below 50°C almost all over the world; whilst at 4°C, temperatures higher than 50°C could occur in all continents. Very extreme events that occur once every hundred years in pre-industrial conditions could occur every 10 to 25 years at 1.5°C. At 4°C of warming, these 1-in-100-years events could happen every year almost all over the world. Maximum temperature variability increases relative to pre-industrial levels up to 50% in some regions at 2°C. At 4°C, the large increases in the maximum temperature variability in North America or Central Europe could imply year-to-year variations of maximum temperatures up to 14°C, and above 18°C in India; whilst maximum temperature variability decreases by 10 to 35% in high latitudes and regions such as in Southern Europe or North America. We also find that, for warming beyond 2°C, conditions that aggravate heat stress spread substantially over non-adapted regions. At 4°C of global warming, tropical night hotspots expand polewards globally and prevail for at least 95% of the summer months; while extreme wet bulb temperatures that only rarely occur in current climates could occur in all continents. Our results indicate that limiting global warming to 2°C is vital to limit the exposure of non-adapted regions to potentially dangerous heat levels.

3.2. Introduction

Extreme heat will become more likely and more extreme under global warming (Meehl and Tebaldi, 2004; Russo et al., 2014). As temperatures continue to rise due to the accumulation of anthropogenic greenhouse gases in the atmosphere, the risk of dangerous heat levels will increase in many regions of the world (IPCC, 2013; Kjellstrom et al., 2010; Sherwood and Huber, 2010; Matthews et al., 2017; Coffel et al., 2018). Extreme heat leads to increased heat-related mortality and illness, worsening the risk of heat exhaustion, dehydration, and cardio-vascular and kidney diseases (Kjellstrom et al., 2010; Hanna et al., 2011; De Blois et al., 2015). Additionally, it can also lead to ecological and socio-economical impacts, such as decreased labour productivity, increased risk of wildfires, habitat loss, crop failure, decreased agricultural efficiency, and increase the risk of environmental refugees by rendering some regions partially inhabitable (IPCC, 2014; Kjellstrom et al., 2010; Sherwood and Huber, 2010; Dunne et al., 2013; Gourdji et al., 2013; Bowman et al., 2017). In this study, we investigate how global warming aggravates extreme summertime heat, and evaluate for which maximum global warming level the risk of extreme heat conditions under five different metrics can be confidently averted.

Some of the regions that will suffer the effects of extreme heat the most are highly-populated developing regions, where society is most vulnerable due to scarce infrastructure and relatively low adaptive capacity (Mishra et al., 2017; Coffel et al., 2018; Newth and Gunasekera, 2018). Already under current global warming levels, the deadly combination of extreme heat and humidity together with insufficient infrastructure caused the death of thousands in the 2015 heatwaves in India and Pakistan (Wehner et al., 2016). But also in wealthy and developed countries the impacts of extreme heat are already palpable. The combination of extreme daytime temperatures and lack of nighttime cooling caused more than 70.000 additional deaths over 16 European countries during the 2003 summer (Robine et al., 2008; Laaidi et al., 2012); and 55.000 people died due to the 2010 heatwave in Russia alone (Barriopedro et al., 2011). As global warming continues to rise, such extreme events will occur more frequently. By 2°C of global warming above preindustrial levels, one out of every two summer months are projected to be on average warmer than the 2010 summer over Europe (Suarez-Gutierrez et al., 2018); and over some regions in India and Pakistan, conditions equivalent to the 2015 heatwave could occur every year (Matthews et al., 2017).

Extreme events are large deviations from the mean climate state that occur due to quasi-random chaotic internal variability, emerging from spontaneously generated

mechanisms and feedbacks across all components of the climate system. These extreme deviations are, by definition, rare, and occur intrinsically by chance. In particular, the internal variability in maximum summertime temperatures is so large that can make their probability distributions at clearly differentiable global warming levels hard to distinguish at the regional level (Suarez-Gutierrez et al., 2018). Thus, to evaluate how the strength and frequency of extreme heat events change in a warming climate, it is vital to sufficiently sample internal variability to adequately capture the very extreme events at the tails of the distribution. To achieve this, we use the largest ensemble of single-model simulations from a comprehensive, fully-coupled model: the 100-member Max Planck Institute Grand Ensemble (MPI-GE; Maher et al., 2019)

Furthermore, to evaluate how global warming aggravates extreme heat events it is also crucial to consider all the most relevant elements that define our vulnerability to extreme heat. The foremost element of extreme heat events are maximum temperatures. However, some of the events with the largest impacts to date were events that combined the effect of extreme temperatures with other conditions that exacerbate heat stress, such as high humidity or nighttime temperatures (Laaidi et al., 2012; Wehner et al., 2016). To combine all these aspects, we evaluate the risk of extreme summertime heat with five different metrics — maximum reachable temperatures, return periods of very extreme temperatures, maximum temperature variability, sustained tropical night temperatures, and extreme wet bulb temperatures. Here, we use these five metrics to identify which regions become major extreme heat risk hotspots under global warming. But going one step further, we evaluate what maximum global warming level allows us to avert the risk of extreme heat, once we take into account the irreducible uncertainty that arises from chaotic internal variability.

We investigate the maximum reachable temperatures under different global warming levels to identify where summertime maximum temperatures become most extreme. We also investigate how the return periods of maximum temperatures that are very extreme under preindustrial conditions change under different global warming levels. Another key aspect of our vulnerability to extreme heat that we investigate is maximum temperature variability. In climates where temperature variability is small, such as the tropics, a shift towards a warmer mean state can imply that society and ecosystems need to adapt to new average conditions totally outside the range they are used to, with relatively small temperature fluctuations resulting in large impacts. In climates where temperature variability is large, such as mid-latitude continental interiors, society and ecosystems need to prepare for a broad range of conditions, and drastic fluctuations can happen in periods too short to allow for adaptation.

Moreover, regardless of the absolute amplitude of maximum temperature variability under pre-industrial conditions, heat stress can worsen if variability changes under warming. An increase in variability, as projected over certain regions (Fischer et al., 2012; Donat et al., 2017; Bathiany et al., 2018; Suarez-Gutierrez et al., 2018), leads to heat extremes with increased amplitude and frequency, and to overall larger temperature deviations from the mean state that impose bigger adaptational challenges. On the other hand, a decrease in variability implies that the effects of global warming are less likely to be temporarily counteracted by internal variability on any given summer, resulting on year-to-year summer temperatures that are consistently warmer. Most previous studies use standard deviation changes as a proxy for variability changes under warming, and are based on multi-model ensembles (Fischer et al., 2012; Donat et al., 2017; Bathiany et al., 2018). However, this combination does not allow a clear separation between the forced transient warming and the deviations caused by internal variability, and can lead to misleading results. Using a very large single model ensemble, we are able to instead directly evaluate how the range of maximum temperatures reachable every year changes under different levels of global warming, based on an precise characterization of the simulated internal variability that is not confounded by different responses to forcing or model configurations.

Alongside maximum temperature reached during the day, nighttime minimum temperature is another aspect that exacerbates the impacts of extreme heat. Minimum temperatures above 20°C to 25°C, known as tropical night conditions, result in an absence of nighttime cooling that impedes organisms to recover from extreme heat during daytime. The human body can, with time, adapt to these conditions, that occur currently in many tropical and equatorial regions of the world. However, for unadapted individuals and accompanied by extreme maximum temperatures, this lack of restorative cooling can aggravate heat stress and is directly linked to increased heat-related hospitalizations and mortality, particularly if sustained over several days (Basu and Samet, 2002; Laaidi et al., 2012; Royé, 2017; Murage et al., 2017). Although there is an overall agreement that a shift towards a warmer mean state results in higher minimum temperatures (Russo and Sterl, 2011), it remains unclear to what extent sustained tropical night conditions can be averted by limiting global warming. We investigate the probability of sustained tropical night conditions expressed as daily minimum temperatures that exceed the tropical night threshold of 20°C for an entire month, and quantify the maximum global warming levels that allows us to avoid these conditions.

Lastly, we investigate extreme heat conditions involving simultaneous high temperature and humidity. Even at very high air temperatures, the human body can

efficiently loose heat through evaporative cooling, as long as humidity levels remain low. Under high air temperature and humidity conditions, evaporative cooling looses efficiency, and we become unable to maintain a stable core temperature. Several studies highlight an increase in these hot and humid conditions, particularly in highly-populated, vulnerable areas such as South East Asia, India or West and Central Africa (Dunne et al., 2013; Pal and Eltahir, 2015; Im et al., 2017; Newth and Gunasekera, 2018; Coffel et al., 2018). Under high emission scenarios, the portion of the population exposed to potentially deadly heat stress could increase from the 30% currently exposed to 75% by the end of the century (Mora et al., 2017). Very extreme humid heatwaves that never occur in current conditions and very rarely occur at 2°C are expected to occur every other year at 4°C of global warming (Russo et al., 2017). Furthermore, under very strong warming levels above 10°C, the combination of heat and humidity could reach levels so high that it renders large regions of the globe inhabitable (Sherwood and Huber, 2010).

A variety of indexes are used to measure the combined impact of temperature and humidity on heat stress (Willett and Sherwood, 2012; Buzan et al., 2015). One of the most commonly used of these indexes, and the one that we focus on, is wet bulb temperature (Sherwood and Huber, 2010; Pal and Eltahir, 2015; Im et al., 2017; Coffel et al., 2018). Wet bulb temperature (W) is defined as the temperature that an air parcel would reach through evaporative cooling once fully saturated. As opposed to comfort-based heat indexes (Russo et al., 2017; Matthews et al., 2017; Li et al., 2018) or more complex heat stress measures considering the effect of wind chill and solar irradiation, such as wet bulb globe temperature (WGBT; Dunne et al., 2013; Fischer and Knutti, 2013; Newth and Gunasekera, 2018), W establishes a clear thermodynamic threshold on heat transfer for which health impacts cannot be overcome by adaptation (Sherwood and Huber, 2010). For W values similar to the normal skin temperature of 35°C, evaporative cooling is significantly less effective and the body starts to accumulate heat. For W values above 35°C during periods as short as a few hours, core body temperatures reach values that are lethal, even for acclimated healthy individuals. Under current warming levels, W almost never exceeds values above 31°C (Sherwood and Huber, 2010). However, harmful to deadly levels of heat stress can occur at lower W depending on health conditions, age, and level of physical activity. The moderate risk threshold for vulnerable individuals occurs at W around 28°C; and at W above 32°C physical labor becomes unsafe also for healthy individuals (Dunne et al., 2013; Buzan et al., 2015; Coffel et al., 2018). Here we investigate maximum reachable W under different global warming levels, and determine the maximum global warming for which dangerous levels of W can be avoided.

Most previous studies evaluating how some of these heat stress indicators change under global warming are based on smaller multi-model ensembles subsampling CMIP5 (Dunne et al., 2013; Fischer and Knutti, 2013; Russo et al., 2014; Donat et al., 2017; Mishra et al., 2017; Matthews et al., 2017; Russo et al., 2017; Mora et al., 2017; Newth and Gunasekera, 2018; Coffel et al., 2018; Li et al., 2018; Bathiany et al., 2018), smaller regional model ensembles (Pal and Eltahir, 2015; Im et al., 2017), or smaller single-model ensembles (Sherwood and Huber, 2010; Willett and Sherwood, 2012; Mishra et al., 2017). Also, most studies explore changes linked to different forcing scenarios (e.g., Dunne et al., 2013; Fischer and Knutti, 2013; Russo et al., 2014; Matthews et al., 2017; Coffel et al., 2018; Li et al., 2018; Bathiany et al., 2018), as opposed to changes between different warming levels (e.g., Russo et al., 2017). In contrast, we base our analysis on one very large single-model ensemble, MPI-GE, currently the largest existing initial-condition ensemble using a comprehensive, fully-coupled Earth System Model — largest both in terms of forcing scenarios represented and in terms of independent members (Maher et al., 2019). MPI-GE consists of sets of 100 independent realizations under the same forcing conditions but starting from different initial states, that allow 1-in-100-years events to occur on average every simulated year. The large ensemble size in MPI-GE is crucial to robustly sample internal variability and to empirically evaluate the statistical significance of changes in the characteristic of extreme events. Another crucial ingredient in MPI-GE unique design is its diversity of forcing conditions. This diversity allows us to robustly characterize and compare the climates at global warming levels of 0°C, 1.5°C, 2°C, 3°C and 4°C above pre-industrial conditions, providing a precise sample of the simulated internal variability for each climate. Using MPI-GE we construct five different heat metrics and five different global warming levels to robustly evaluate the irreducible risk of very extreme heat conditions that arises due to chaotic internal variability.

3.3. Data and Methods

We use transient climate simulations from the Max Planck Institute Grand Ensemble (MPI-GE) under historical forcing and three future representative concentration pathways (RCP), namely RCP2.6, RCP4.5, and RCP8.5 (Maher et al., 2019; Bittner et al., 2016; Hedemann et al., 2017; Suarez-Gutierrez et al., 2017). The ensemble consists of sets of 100 realizations based on the same model physics and parametrizations and driven by the same external forcings, but each starting from different initial climate states, taken from different points of the model's pre-industrial control run. The MPI-GE uses the model version MPI-ESM1.1 in the low resolution (LR) configuration, with resolution T63 and 47 vertical levels in the atmosphere (Giorgetta et al., 2013) and

1.5°resolution and 40 vertical levels in the ocean (Jungclaus et al., 2013). MPI-ESM1.1 is fairly similar to the the CMIP5 version of MPI-ESM (Taylor et al., 2012), but has a slightly lower equilibrium climate sensitivity of 2.8°C (Mauritsen et al., 2019; Flato et al., 2013). MPI-ESM-LR has a relatively low resolution, comparable to most of the models in the CMIP5 ensemble, which can influence the model's ability to simulate small-scale processes and affect the reliability of our projections.

The large ensemble size of MPI-GE makes it well capable of simulating very extreme temperature events, such as the 2003 and 2010 European heat waves (Suarez-Gutierrez et al., 2018), unlike other large ensemble experiments of smaller size (Schaller et al., 2018). To evaluate the ability of MPI-GE to simulate observed current climate conditions globally, we compare it to the 1° x 1° gridded data from Berkeley Earth Surface Temperatures (BEST) climatology and monthly maximum temperature anomaly for the period 1850–2018 (Rohde et al., 2012). The average maximum temperatures for current climate conditions defined by the period of 1990–2018 simulated by MPI-GE are within the range of BEST observations in most regions of the world. MPI-GE average maximum temperatures are larger than observations for regions over Australia, West Asia, or North and South America; while the simulated average is smaller than the observations over East Asia and most tropical regions (Supporting Fig. S.5).

Whereas averaged summertime monthly maximum temperatures in MPI-GE are in good agreement with observations; maximum temperature variability appears to be larger in MPI-GE than the observed estimates in most regions (Appendix Fig. A.4). We find that MPI-GE adequately simulates extremes, particularly warm extremes, in several relevant regions. In some regions such as Europe or North America, warm extremes are generally adequately represented, while cold extremes appear to be overestimated. This tendency to overestimate the variability in maximum temperatures may indicate that MPI-GE also overestimates future projections of maximum temperatures. On the other hand, observed values generally occur within the ensemble range, with some exceptions in Central Africa or East Asia. Our evaluation indicates that MPI-GE does not underestimate extreme events in most regions, but may tend to overestimate the magnitude and frequency of extremes particularly in the lower tail of the distribution. However, summer maximum temperatures exhibit generally large internal variability, and thus the observational record may be too short to determine whether the amplitude and frequency of extreme events is adequately captured in MPI-GE.

We define global mean surface temperature (GMST) as the annually averaged, global mean, near-surface 2m air temperature anomaly with respect to pre-industrial

conditions, defined by the period 1851–1880. We focus on summer months defined as JJA for the Northern Hemisphere and DJF for the Southern Hemisphere. We use the summer maximum value of daily maximum temperature (TXx), defined as the block maximum temperature reached each summer at each grid cell. We evaluate the likelihood of experiencing sustained tropical night conditions during an entire month, for summer months with block minimum values of daily minimum temperatures (TNn) above 20°C. We construct monthly wet-bulb temperature estimates (W) using summer monthly averages of near-surface 2m air temperatures and relative humidity based on the method described in Stull (2011). Ideally, to obtain the most accurate results W should be calculated instantaneously at the model time step. However, this is not possible in MPI-GE, with only monthly mean relative humidity output available. Calculating monthly W using monthly mean temperature and humidity, as opposed to calculating monthly W averages based on instantaneous data, can lead to a maximum overestimation of up to 1.5°C (Buzan et al., 2015). Although this overestimation varies with temperature, its 90% confidence range remains below 0.5°C, and its median is in the 0.005-0.2°C range for all temperatures considered (Buzan et al., 2015). To counteract this potential bias, we subtract the maximum median overestimation of 0.2°C from the monthly W estimates in this study. Although this correction does not alter our conclusions, results for uncorrected W values are shown in the Supporting Figures S.6 and S.7.

We construct representative samples of the climate conditions at 0°C, 1.5°C, 2°C, 3°C and 4°C of mean global warming with respect to pre-industrial levels using MPI-GE transient climate simulations. GMST deviates from the long-term mean state on year-to-year timescales due to the effect of internal variability. Therefore, we calculate centered decadal-averaged GMST to robustly define each global warming level (Suarez-Gutierrez et al., 2018). We define years of 0°C of global warming as those years in which the centered decadal-averaged GMST is in the range of 0 ± 0.25°C in the historical MPI-GE simulations. Analogously, for the remaining global warming levels we select years in which the centered decadal-averaged GMST is in the range of 1.5±0.25°C from RCP2.6 simulations, 2±0.25°C from RCP4.5 simulations, and 3±0.25°C and 4±0.25°C from RCP8.5 simulations. This *time-slice* method to define global warming levels from transient simulations is similar to the methods used in Schleussner et al. (2016) and King and Karoly (2017) or Suarez-Gutierrez et al. (2018), but we defined each level based on a slightly larger range of decadal averaged GMST, to reach an adequate and homogeneous sample size of around 1000 simulated years for each warming level.

The climates at 0°C, 1.5°C and 2°C of global warming are defined from simulations in a near-equilibrium state. On the other hand, due to the lack of near-equilibrium sim-

ulations for higher warming levels, the climates at 3°C and 4°C of global warming are defined from highly transient simulations. Although similar definitions of fixed global warming levels from highly transient simulations have been used before (Schleussner et al., 2016; King and Karoly, 2017), the climate conditions sampled from transient runs may differ from the near-equilibrium conditions at said warming level, such as in different warming patterns or different ocean heat content distributions (Gregory et al., 2015). The use of highly transient runs also implies a higher probability that climates with slightly higher or lower levels of warming may be oversampled. Additionally, part of the differences between each warming level sampled from different RCPs may arise from differences beyond CO_2 atmospheric concentrations, such as different land use changes or aerosol forcings.

3.4. *Results and Discussion*

We use MPI-GE simulations under historical, RCP2.6, RCP4.5, and RCP8.5 forcings to construct samples of the climate conditions at five different global warming levels (Fig. 3.1). To achieve homogeneous sample sizes of around 1000 simulated years for each climate conditions, we restrict our selection to the periods marked by the dashed black lines in Figure 3.1a. The empirical probability distributions of GMST for the simulated years selected show the effect of sampling near-equilibrium conditions, as for GMST levels of 0°C, 1.5°C, or 2°C, in comparison to sampling highly transient conditions, as for 3°C and 4°C of GMST (Fig. 3.1b). The latter exhibit GMST values that are more variable, resulting in wider probability distributions. However, the distributions are correctly centered around the representative GMST levels and present no substantial overlap, indicating that each sample distribution is distinguishable from the others and offers an adequate representation of the climate conditions of each warming level.

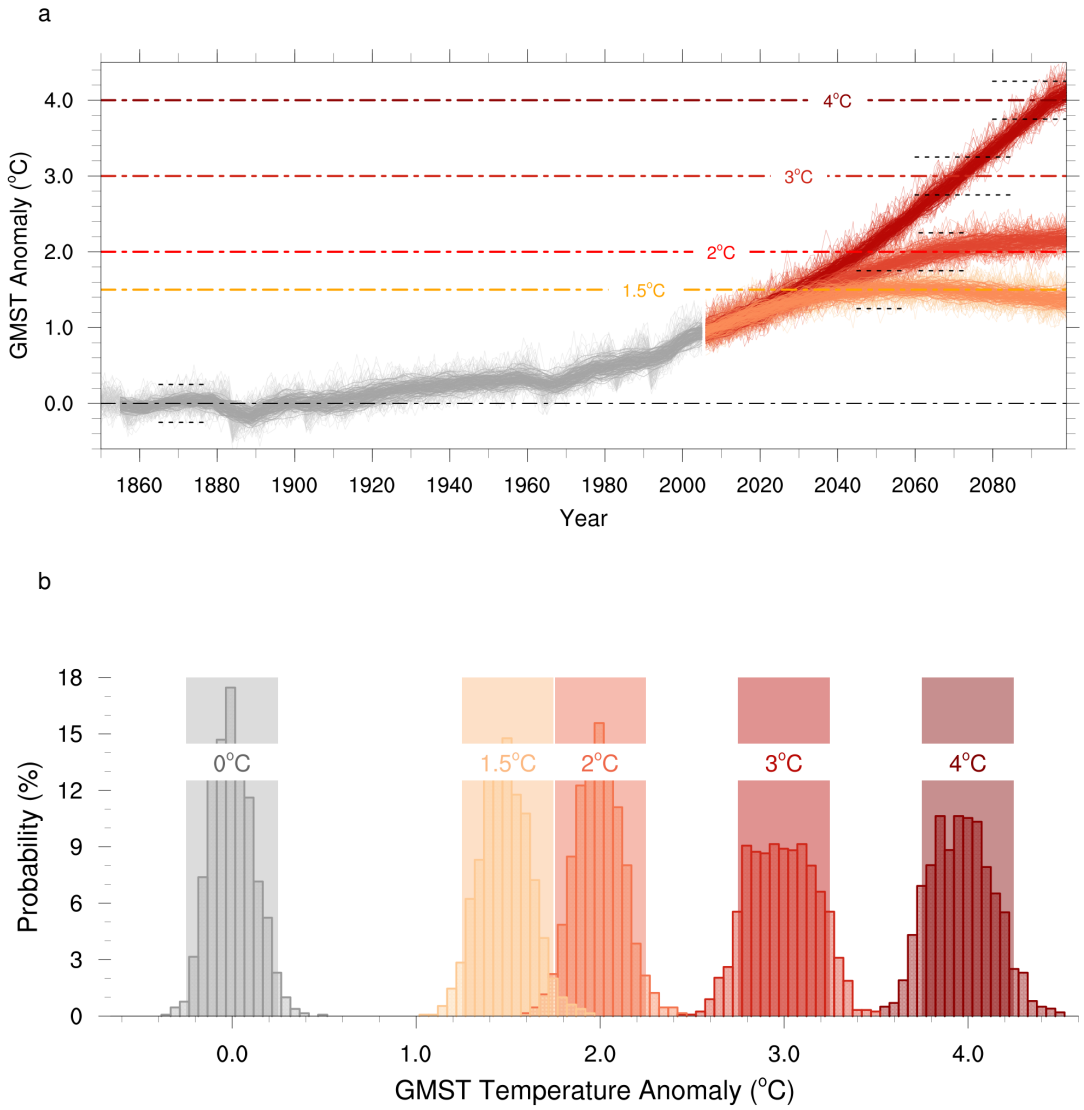


Figure 3.1: Global mean surface temperature (GMST) in MPI-GE. (a) Time series of annually averaged GMST anomalies (colored thin lines) and centered decadal-averaged GMST anomalies (colored thick lines) for the period 1850–2099, simulated by the MPI-ESM Grand Ensemble. Simulations are historical runs for the period 1850–2005 (gray lines), and RCP2.6 (orange lines), RCP4.5 (red lines) and RCP8.5 (dark red lines) for the period 2006–2099. The black dashed lines show the periods of sampling for each warming level. **(b)** Probability distribution of GMST anomalies for pre-industrial conditions at 0°C of warming (gray; sample size n=1300), and for future global warming levels of 1.5°C (yellow; n=1300), 2°C (orange; n=1297), 3°C (red; n=1225) and 4°C (dark red; n= 997) above pre-industrial conditions. The shaded bars represent the range of ±0.25°C around each GMST mean state. Bin size is 0.05°C; frequencies are normalized to unity and translated to percentage.

3.4.1. *Maximum Reachable Temperatures*

We evaluate the maximum summer values of maximum monthly temperatures (TXx) reached under each climate conditions in MPI-GE and compare them to maximum temperatures observed at our current warming level conditions of around 1°C above pre-industrial levels (Hawkins et al., 2017). The observed estimates represent the maximum value of the spatial average of maximum temperatures in each grid cell. Thus, localized record high temperatures may be smoothed within each grid-cell, leading to maximum temperatures per grid cell that are slightly lower than the record high temperatures for specific locations. With this consideration in mind, we find that although the observed temperature patterns are well represented in MPI-GE, the maximum temperatures reachable under pre-industrial conditions in the MPI-GE simulations are similar or higher than those observed under current global warming levels (Fig. 3.2, top row).

In some regions, such as North America, Argentina, Western Asia or Australia, this may occur because MPI-GE simulates maximum temperatures on average warmer than those observed (Supporting Fig. S.5). This can also occur due to a potential overestimation of extreme temperatures over these regions in MPI-GE, that may result from an overestimation of temperature variability over some continental areas (Appendix Fig. A.4). However, the large ensemble size in MPI-GE allows for simulated extreme events with return periods over hundreds of years, and the observational record may just be too short to determine whether the ensemble overestimates very extreme temperatures or whether the Time of Emergence (Hawkins and Sutton, 2012) has not yet been reached for TXx over these regions.

For higher warming levels of 1.5°C and 2°C, we find that the areas where maximum TXx values reach 45°C to 50°C increase in comparison to pre-industrial conditions, particularly over North and West Africa (Fig. 3.2, middle row). For 2°C of warming and below, we find TXx values mostly lower than 50°C all over the world, with some exceptions in the Arabic Peninsula or Northern India and Pakistan. For 3°C and 4°C of global warming maximum temperatures could exceed 50°C across all continents (Fig. 3.2, bottom row). The MPI-GE projects the highest summer maximum temperatures over Western Asia, reaching maximum values above 60°C in countries such as Pakistan, Iraq or Saudi Arabia.

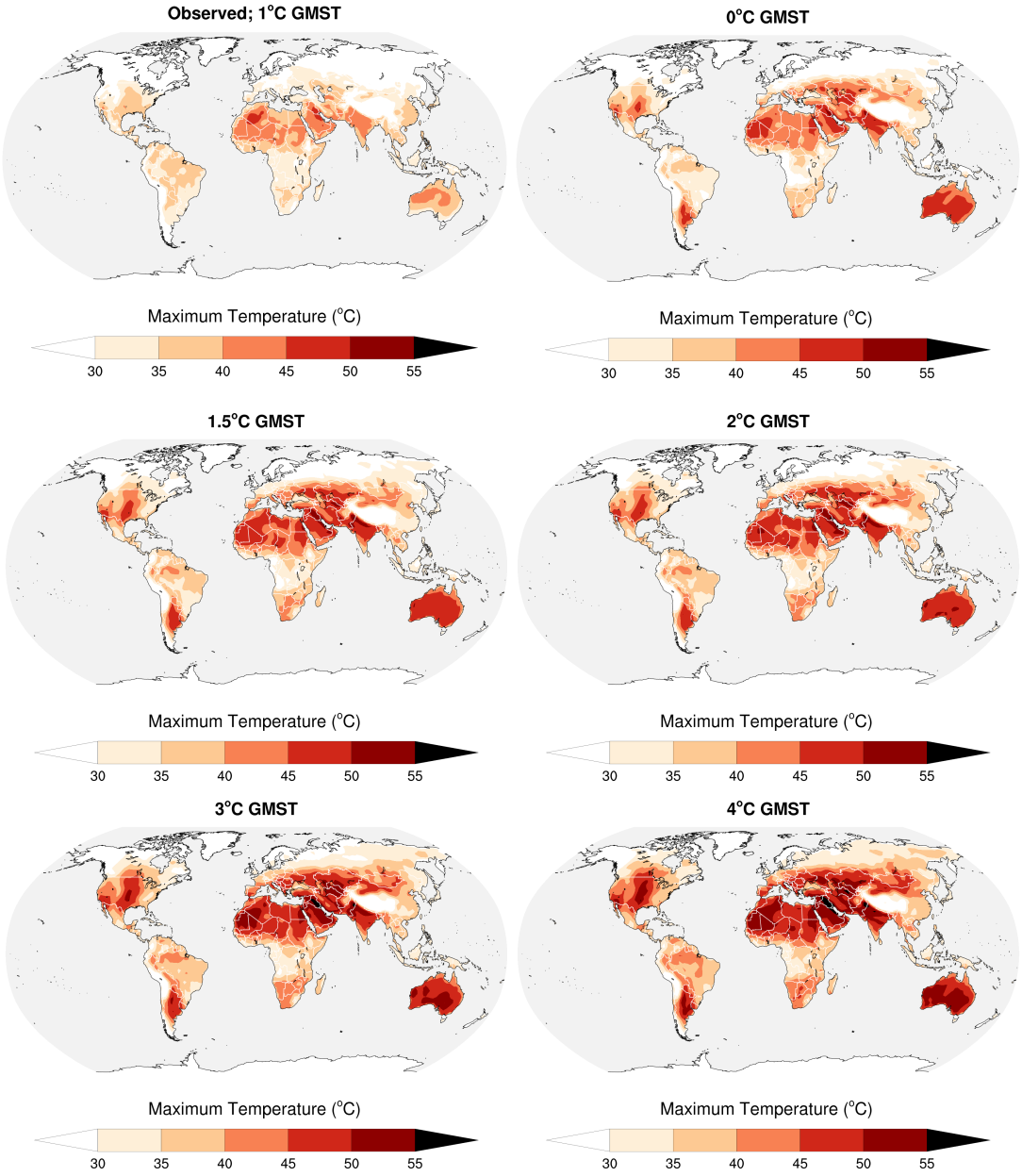


Figure 3.2: Maximum reachable summer maximum temperatures at different global warming levels. Absolute maximum summer maximum value of monthly maximum temperature (TXx) reached under different global warming levels simulated by MPI-GE compared to observed maximum temperatures in the BEST dataset. The observed estimates represent the maximum value of the spatial average of maximum temperature anomaly plus the climatology for the respective month in each grid cell for the period 1850–2018. The simulated maximum temperatures represent the 99.5th percentile value for each distribution at each grid cell.

3.4.2. *Return Periods of Very Extreme Temperatures*

Under rising levels of global warming we expect heat extremes not only to exhibit higher maximum temperatures, but also to occur more frequently. In this section we investigate how the frequency of events that are extreme under pre-industrial conditions changes with global warming. As reference, we choose extreme maximum summertime temperatures that occur on average once every hundred years under pre-industrial conditions (1-in-100-years events; Fig. 3.3, top). Temperatures during simulated 1-in-100-years events under pre-industrial climate conditions in MPI-GE are similar to the highest observed summer maximum temperatures in current climate conditions (Fig. 3.2, top left). As global warming levels increase, the temperature levels characteristic of these events occur more frequently than once every hundred years, thus becoming 1-in-x-years events as the world warms.

At 1.5°C of global warming, these 1-in-100-years events could occur every 10 to 25 years in most regions of the globe, and up to every 1 to 2 years in North Africa. At 2°C of warming, these very extreme events are projected to occur more often than every 10 years over most of the world (Fig. 3.3, middle row). At 3°C of global warming, these pre-industrial 1-in-100-years heat extremes could occur generally every 2 to 5 years in most regions, and every year in East Asia, North Africa and North America. By the point when global warming reaches 4°C above pre-industrial levels these very extreme events could occur more often than every two years almost all over the world (Fig. 3.3, bottom row).

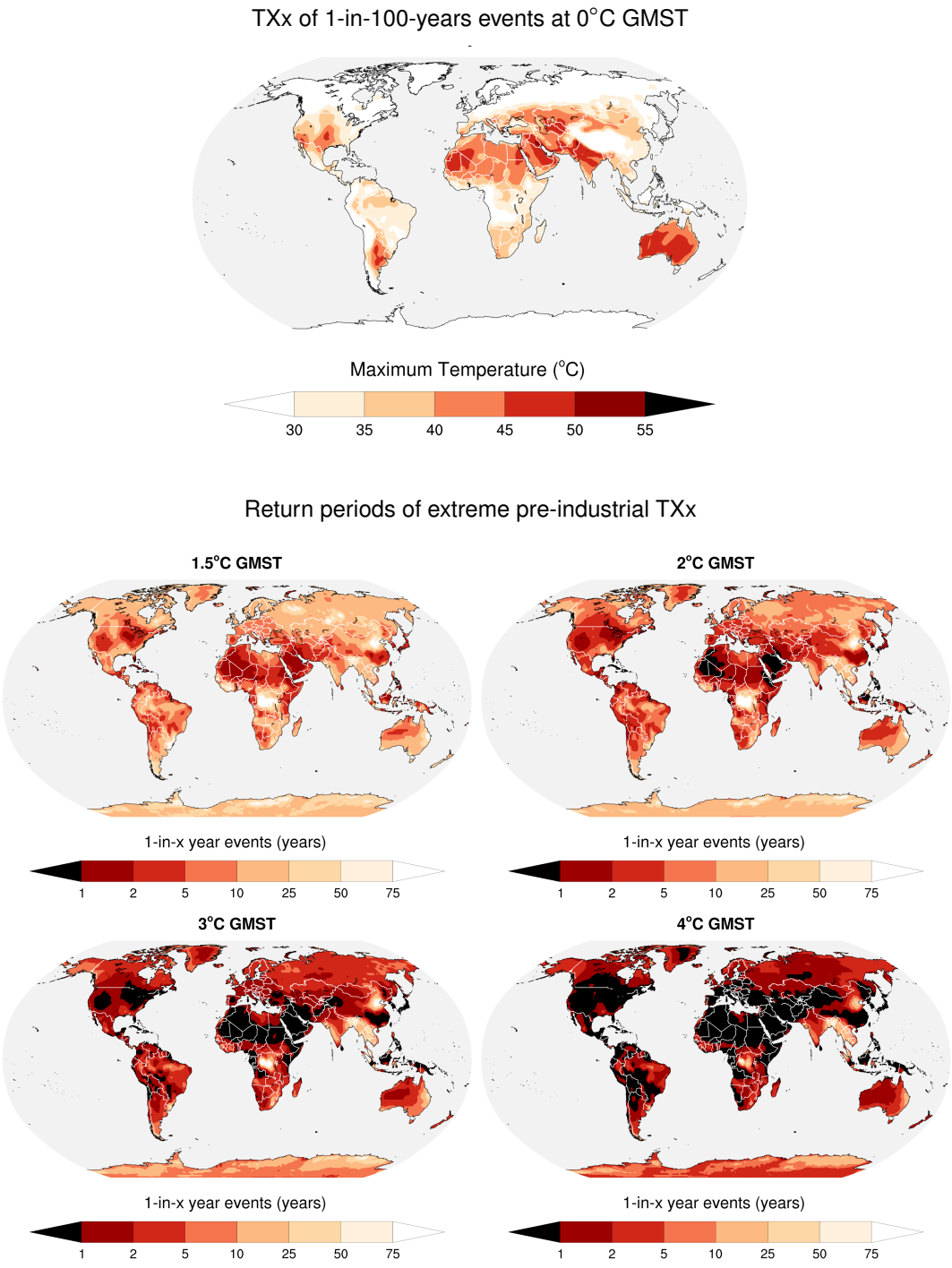


Figure 3.3: Return periods of very extreme summer maximum temperatures at different global warming levels. Return levels of TXx for events with return periods of 100 years under pre-industrial conditions defined at 0°C GMST (top row). Return periods of TXx levels of pre-industrial 1-in-100 years events under different levels of global warming (middle and bottom rows).

3.4.3. *Maximum Temperature Variability*

In this section we evaluate how the year-to-year variability in summer maximum temperatures changes under global warming. Under pre-industrial conditions, in agreement with what we could expect, we find that absolute TXx variability simulated by MPI-GE is larger on mid and high latitudes, particularly in the Northern Hemisphere, and smaller in tropical regions (Fig. 3.4, top row). TXx variability reaches values well above 10°C under pre-industrial conditions in regions such as Eastern India, central Eurasia and other mid-latitude continental interiors. Most equatorial and tropical regions exhibit lower maximum temperature variability under pre-industrial climates, ranging from below 4°C to around 8°C.

For global warming of 1.5°C and 2°C, the change in TXx variability is dominated by a relative increase. This TXx variability increase is similar for both warming levels and remains mostly under 50%, reaching its maximum values in Central South America, North America and India (Fig. 3.4, middle row). Above 2°C of warming, TXx variability increase in these regions reaches values above 50%. Apart from the exceptions of Australia and some parts of Africa and East Asia, where TXx variability does not change substantially, at 4°C of global warming we find a large increase in the variability of summer maximum temperatures in large continental areas all across the globe, with maximum relative increase well above 100% (Fig. 3.4, bottom row). For regions that exhibit low TXx variability under pre-industrial conditions, such as Central South America, this doubling of variability results on absolute TXx variability mostly below 8°C at 4°C of global warming. However, in regions of larger pre-industrial TXx variability, the doubling in TXx variability translates in maximum year-to-year variations of summer maximum temperatures of up to 14°C in North America or Central Europe, and up to 18°C in India.

We also find a decrease in TXx variability ranging from 10 to 35% at 4°C, most prominent in regions such as Greenland, Southern Europe or North America (Fig. 3.4, bottom row). The variability decrease over high latitude regions is most likely dominated by ice melt. Over middle latitude regions, TXx variability may decrease as a result of overall dryer conditions and reduced moisture variability, that restrict temperature variability in the lower tail of the distribution by limiting evaporative cooling (Fischer et al., 2012). These results stand in contrast to results from previous studies evaluating standard deviation changes in multi-model ensembles, that find a consistent increase in summertime monthly mean temperature variability over land under global warming, particularly in the Northern Hemisphere (Bathiany et al., 2018).

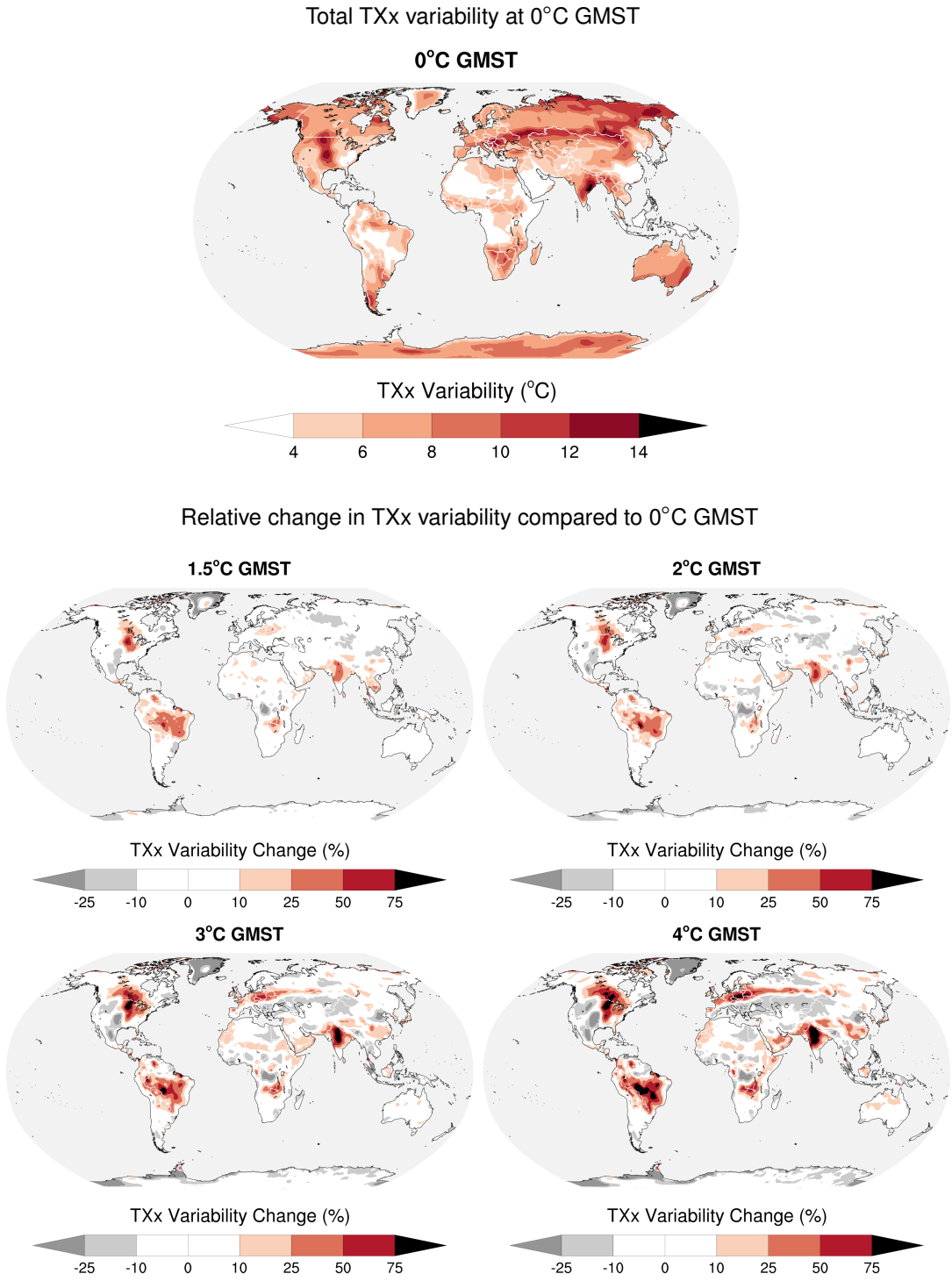


Figure 3.4: Variability in summer maximum temperatures at different global warming levels. Variability in TXx under pre-industrial conditions measured as the difference between the 97.5th and the 2.5th percentiles in the TXx distribution at 0°C GMST (top row). Relative change in variability based on change in TXx probability distribution width (2.5th–97.5th percentiles) at different global warming levels relative to pre-industrial conditions at 0°C GMST (middle and bottom rows).

3.4.4. *Sustained Tropical Night Temperatures*

In this section we evaluate the risk of sustained tropical night temperatures, expressed as minimum temperatures that exceed the tropical night temperature threshold of 20°C for the entirety of the month. These sustained tropical night conditions occur normally over most tropical regions in the pre-industrial climate with at least a 10% probability, but most often with probabilities larger than 95% (Fig. 3.5, top row). For 1.5°C and 2°C of global warming the risk of sustained tropical night conditions exhibits similar values and expands globally polewards, with the most marked increase over the Amazon region and North Africa. In contrast, for warming levels of 3°and 4°C sustained tropical night conditions expand substantially, particularly over mid-latitude regions such as North America, South Africa and Central Eurasia (Fig. 3.5, middle and bottom rows).

Figure 3.6 illustrates the maximum global warming level that allows us to avert the risk of sustained summertime tropical night conditions (90% confidence). Our results indicate that this conditions occur normally in the MPI-GE pre-industrial climate over large low-latitude regions all over the world, where the population is likely acclimated to the lack of nighttime cooling. Under global warming levels below 2°C, the risk of sustained tropical night conditions presents only minor changes. However, for warming levels beyond 2°C there is risk of sustained tropical conditions across all continents. At 4°C of global warming, sustained tropical night conditions could occur over most regions of the globe, with the exceptions of high-latitude regions particularly in the Northern Hemisphere. Additionally to expanding polewards across all continents, sustained tropical night conditions at 4°C of warming prevail consistently, for at least 95% of the summer months. Our results indicate that to limit the exposure of non-adapted regions to sustained tropical night temperatures that significantly aggravate heat stress, limiting global warming to 2°C under pre-industrial conditions is crucial.

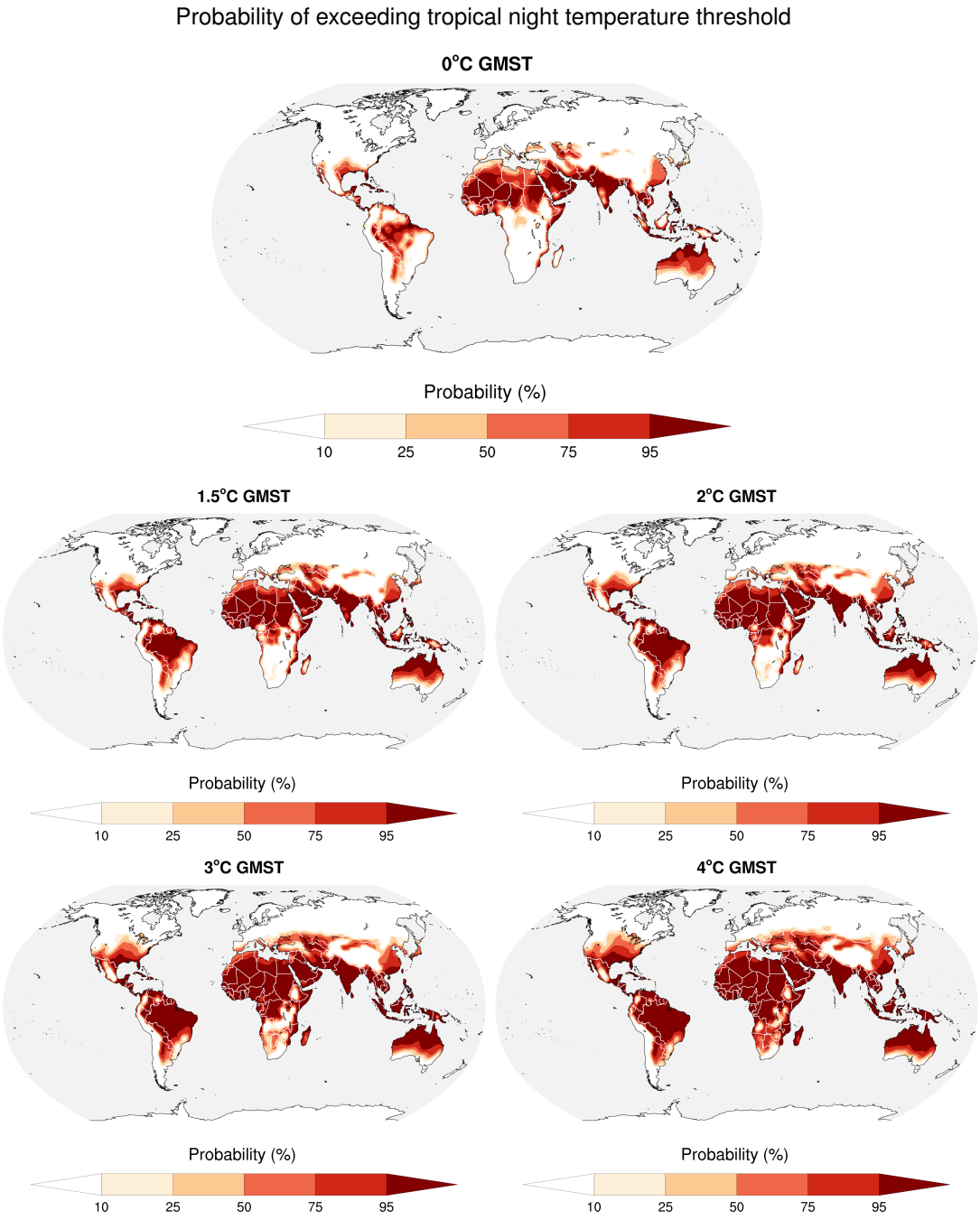


Figure 3.5: Risk of sustained tropical night temperatures at different global warming levels. Probability of sustained exceedance of the tropical night threshold for monthly minimum temperatures (TNn > 20°C) at different global warming levels.

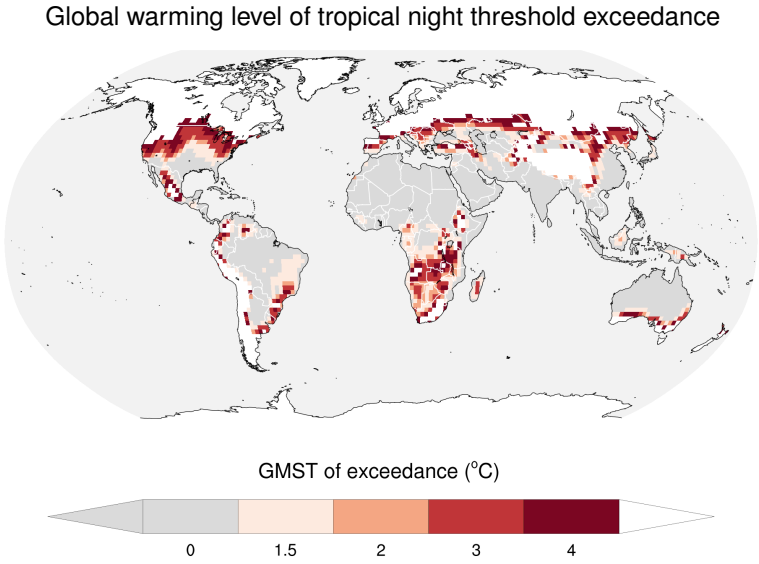


Figure 3.6: Global warming level of sustained tropical night temperatures. Global warming level measured as GMST that exhibits sustained exceedance of the tropical night threshold (minimum monthly temperature TNn > 20°C) with probability of 10% or higher.

3.4.5. *Extreme Wet Bulb Temperatures*

In this section we evaluate how the combination of high temperatures and simultaneous high humidity conditions measured by the wet bulb temperature index (W) changes under different levels of warming. Under pre-industrial climate conditions in MPI-GE, monthly W reaches its maximum values of around 27°C over Northern India and Pakistan; while remaining generally below 24°on the majority of the world (Fig. 3.7, top). Beyond 1.5°C of global warming, projections of maximum reachable monthly W above 26°C spread over Northern India and East Asia; while W above 24°C are projected to occur across all continents (Fig. 3.7, middle row). At 4°C of global warming, monthly W levels above 26°C could occur over large land fractions across all continents; while projections surpass the 28°C threshold over parts of East China, the Arabic Peninsula, Pakistan and Northern India. (Fig. 3.7, bottom row).

These values are comparable to harmful heat and humidity levels for vulnerable individuals at W larger than 28°C, but remain comparatively smaller than the thresholds for moderate and fatal risks at 32°C and 35°C respectively. However, instantaneous W values could exceed monthly estimates by several degrees. Monthly W estimates from MPI-GE under pre-industrial and 1.5°C climate conditions are indeed several degrees lower than current instantaneous W estimations based on reanalysis

data (Sherwood and Huber, 2010; Im et al., 2017). These differences are particularly large over several extreme W hotspots such as Central South America, West Africa or South Asia (Sherwood and Huber, 2010; Im et al., 2017). Whereas the maximum monthly W values simulated by MPI-GE remain generally below 30°C even at 4°C of global warming, daily maximum W values exceed the 29°C threshold under current climate conditions in large parts of India, Pakistan and East China (Im et al., 2017). Over these regions, daily W estimates under RCP8.5 forcing are projected to exceed the fatal 35°C threshold by the end of the century (Im et al., 2017).

Although part of the differences between monthly W estimates in MPI-GE and daily W estimates from previous studies may arise from remaining biases and model differences, this comparison seems to indicate that our projections using monthly W estimates may be somewhat conservative. Additionally, it is also important to note that the maximum daily W values from previous studies are obtained from ensembles much smaller than MPI-GE, with a smaller probability of capturing very extreme events. Therefore comparisons based on larger ensembles may yield even larger differences between maximum daily and monthly W estimates. On the other hand, we find good agreement between the regions of largest W increase under warming in MPI-GE and in previous studies (Sherwood and Huber, 2010; Im et al., 2017). This indicates that, although our monthly W estimates may underestimate the risk of reaching harmful instantaneous W levels within a month, the good agreement on the regions of largest W increase supports our conclusions regarding which regions become major heat-stress hotspots due to the combination of extreme temperature and humidity in a warmer world.

Figure 3.8 illustrates the maximum global warming level that allows us to avert the risk of extreme hot and humid conditions characterized by monthly W above 26°C. These extreme W conditions, that occur rarely in the pre-industrial climate of MPI-GE, can be averted over most regions of the world by limiting global warming to levels below 2°C (99% confidence), with some exceptions over Northern India or East China. For warming levels beyond 2°C, the risk of extreme W conditions spreads over large land fractions across almost all continents, from West Africa to central United States. At 4°C of global warming, extreme W conditions spread not only over large low-latitude and tropical regions, but also over mid-latitude regions in North America and East Asia. Our results show that limiting global warming to 2°C under pre-industrial conditions is vital to avoid the exposure of large non-adapted regions to the combination of extreme temperature and humidity, one of the factors that exacerbates heat stress the most.

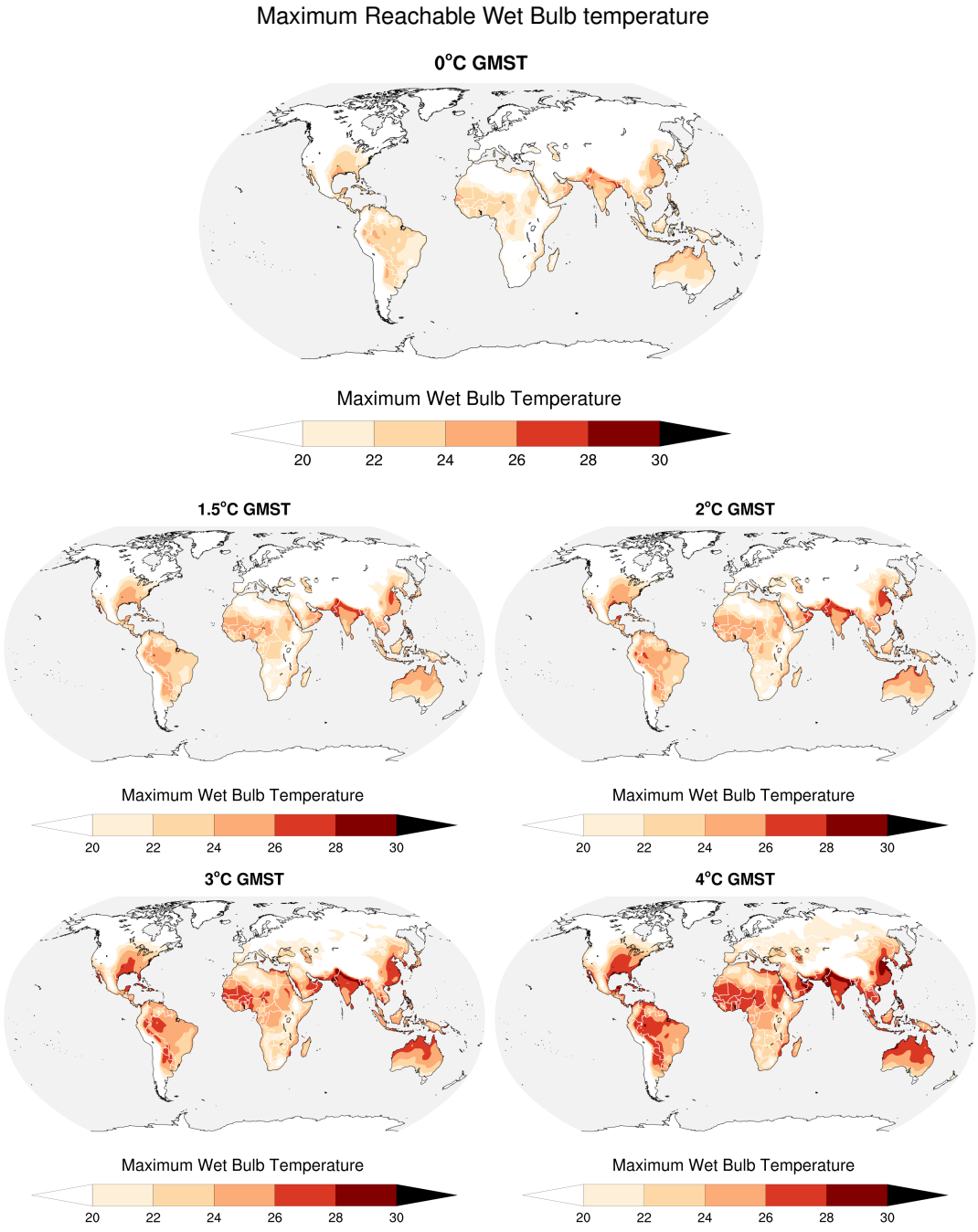


Figure 3.7: Maximum reachable extreme Wet Bulb temperatures at different global warming levels. Maximum reachable monthly Wet Bulb temperatures at different global warming levels. The simulated maximum represents the 99.5th percentile value for each distribution at each grid cell.

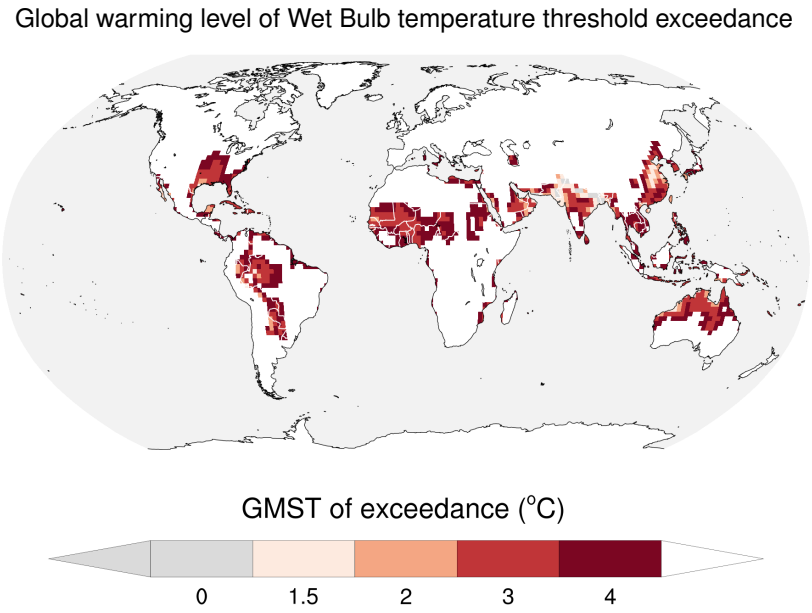


Figure 3.8: Global warming level of exceedance of extreme Wet Bulb temperature threshold. Global warming level measured as GMST that exhibits exceedance of the wet bulb temperature threshold (W > 26°C) with probability of 1% or higher.

3.5. *Summary and Conclusions*

Using the 100-member MPI-GE, we identify global summertime heat stress hotspots for five different extreme heat metrics under five different warming levels, and determine the maximum global warming level for which the risk of extreme heat conditions can be confidently averted. We find that MPI-GE adequately simulates the pattern of observed maximum reachable temperatures. However, absolute reachable maximum temperatures under preindustrial conditions are over some regions larger than those observed at the current warming level of 1°C. This may indicate that MPI-GE overestimates either average maximum temperatures or maximum temperature variability over these regions. However, due to the relatively short length of the observational record, this can also indicate that extreme temperatures with return periods as long as those simulated by MPI-GE have not yet been recorded over these regions.

MPI-GE projects that, for global warming levels below 2°C, maximum reachable summer temperatures stay below 50°C generally all over the world, with some exceptions in the Arabic Peninsula, Northern India and Pakistan. However, for warming levels above 2°C, this threshold could be overshot in all continents, with temperatures reaching values above 60°C over Pakistan, Iraq or Saudi Arabia. We find that very

extreme events that occur under pre-industrial conditions on average once every 100 years could occur every 10 to 25 years at 1.5°C of warming, and more often than once every ten years at 2°C. At 4°C of warming, these 1-in-100-years events could happen every year almost all over the world.

Our results also indicate that maximum summer temperature variability changes substantially under warming in large regions of the globe. Summer maximum temperature variability increases relative to pre-industrial levels up to 50% under 2°C of global warming, mostly in Central South America and North America, Central Europe and India. At 4°C of global warming we find a large increase in the maximum temperature variability in large continental areas, with maximum relative increase well above 100%. This 100% increase translates into maximum year-to-year variations of summer maximum temperatures of up to 14°C in North America or Central Europe, and up to 18°C in India. For regions such as Australia or large parts of Africa and East Asia, maximum temperature variability does not change substantially under warming. For other high latitude regions and parts of Southern Europe or North America, we find that maximum temperature variability decreases by 10 to 35% at 4°C of global warming. These results stand in contrast to previous results, that indicate a consistent and substantial increase in summertime monthly mean temperature variability especially in the Northern Hemisphere (Bathiany et al., 2018).

We find that for warming levels beyond 2°C above pre-industrial conditions, heat stress could be substantially aggravated by high minimum temperatures under sustained tropical night conditions over non-adapted regions. At 4°C of global warming, tropical night hotspots spread polewards over large regions across all continents, particularly in central North America, South Africa, or Central Eurasia, and prevail consistently for at least 95% of the summer months. Similarly, hot and humid conditions under extreme wet bulb temperatures that occur very rarely under pre-industrial conditions are projected to occur in almost all continents at 1.5°C of warming, and could spread over most tropical regions and some regions at mid latitudes in a 4°C warmer world.

Our results indicate that maintaining global warming levels below 2°C above pre-industrial conditions is vital in order to limit the risk of extreme heat conditions in all metrics considered. For each of these heat metrics, we identify major hotspots over different regions — from the highest maximum summertime temperatures projected to occur over the Arabic Peninsula, to the largest maximum temperature variability increase over India or Central Europe. These different heat-stress hotspots highlight the different potential risks and related adaptation measures that are necessary over

different regions. With time and within limits, our society, economy, ecosystems, and even our bodies, are able to adapt to a warmer mean climate state. However, deviations from these mean climate conditions in the form of extreme events that arise due to internal variability challenge our range of adaptability, potentially to its limits.

CONCLUSIONS

To conclude, in this final section I wish to summarize the main findings and conclusions drawn from this dissertation, as well as the overarching implications that emerge from my research.

1. *Drivers of internal variability in European heat extremes*

In Chapter 1, I quantify the contributions from the large-scale dynamic atmospheric mechanisms and the local thermodynamic effects of moisture limitation to the development of extreme heat over Europe, and how these contributions change in a warmer world to cause an increase in summer temperature variability. For this purpose, I introduce a novel definition of extreme events based on a moving threshold with respect to the evolving mean decadal climate. This extreme event definition relies on an accurate characterization of both the simulated internal variability and forced signal only readily available in large ensemble experiments. I then use these large samples of heat extremes simulated by MPI-GE to distinguish between conditions that are necessary for the development of extreme heat from those that may be circumstantial. In doing so, I find the following answers to my research questions.

1.1 What are the contributions from large-scale dynamical atmospheric mechanisms and local thermodynamical effects of moisture limitation as drivers of variability in extreme summertime temperatures?

Using a multiple regression approach that simultaneously considers both of the main driving mechanisms of extreme heat, I identify the dynamical mechanism as the dominating driver of variability in extreme European summer temperatures. On the other hand, the local thermodynamic effect of limited moisture plays a secondary role in driving total extreme temperature variability. Most heat extremes occur under favorable atmospheric conditions, both in current and future climate conditions, and the highest temperatures arise when both persistent anticyclonic conditions and dryness occur.

1.2. How do these dynamical and thermodynamical driving contributions change to cause an increase in extreme summer temperature variability in a warmer world?

In the regions where variability increases, heat extremes occur 10-40% less frequently under extreme atmospheric conditions during the 21st century, and 40% more frequently under extreme moisture limitation. In a warmer world, an increasing number of heat extremes are dominantly driven by the effect of moisture limitation, and occur even under a neutral or unfavorable atmospheric state. My results indicate that the dynamical atmospheric mechanism that acts as dominant driver of total extreme summer temperature variability is not the dominant driver of variability change under warming. Instead, the increase in European summer heat extremes relative to the evolving decadal climate and associated variability increase are dominated by the the local thermodynamic effect of moisture limitation.

This study combines two crucial novel aspects to account for the complex multi-collinearity between each set of drivers: considering all relevant sources of variability simultaneously, and evaluating large samples of extreme events developing under a wide range of background conditions. This approach helps to clarify the long-debated discussion of whether the increased variability in extreme European summer temperatures emerges from the soil or rather descends from the atmosphere. However, in the broader framework of understanding the causality relations between and beyond the main driving mechanisms of extreme summertime heat, our knowledge remains limited. Future research should aim at understanding the complex interdependencies between both mechanisms, such as to what extent soil-moisture variability is actually controlled by atmosphere-borne variability, or how the effects of local moisture evaporation, or lack thereof, influence the large-scale atmospheric dynamics.

Furthermore, beyond the cause and effect relationships between mechanisms, the question of whether an overarching pattern of large-scale variability that controls the variability of extreme European temperatures by controlling the variability of its drivers can be found arises. On decadal to multi-decadal timescales, the North Atlantic multi-decadal variability (AMV) is a strong contender for this overarching role. The ocean heat transport variability and recurring circulation patterns associated with the AMV have been shown to influence mean European summer temperatures (Sutton and Hodson, 2005; Ghosh et al., 2017). However, the contribution from these multi-decadal

variations in the North Atlantic to European extreme heat characteristics and driving mechanisms has not yet been quantified. This is the topic that I plan to address in the next phase of my scientific career. During the course of the last year, I have collaborated with my advisors in the development of a postdoctoral research proposal based on this line of research. Our proposal received positive evaluations during the first phase of the selection process, and we await its resolution in the coming months.

2. *Controllability of European temperatures under warming*

In Chapter 2, I evaluate how the irreducible uncertainty arising from internal variability defines to what extent an increase in extreme European summertime temperatures can be averted by maintaining global warming below fixed limits. I define these limits as the 1.5°C and 2°C global mean surface temperature targets established in the UNFCCC Paris Agreement. I use MPI-GE transient climate simulations to construct quasi-stationary samples for each climate condition, that I then use to robustly compare the different warming levels. The large ensemble size of MPI-GE combined with the assumption of quasi-stationarity allows me to base this comparison on empirically calculated probability distributions and extreme events that are well defined for return periods up to 500 years. This comparison highlights stark differences between the two warming levels on absolute terms. For example, extreme summers that occur on average once every 500 years, exhibit maximum temperature anomalies below 3°C under pre-industrial conditions. These 1-in-500-years anomalies reach 5.5°C at 1.5 °C of global warming, and could rise to 7°C at 2°C of global warming. However, the true innovative aspect of this comparison relies on understanding how the irreducible uncertainty introduced by internal variability could blur these stark differences between the two warming levels. To do so, I answer the following questions.

2.1 To what extent are European summer temperatures at 1.5°C of global warming distinguishable from those at 2°C of warming?

I introduce a novel distinguishability metric which identifies the summer months in a 2°C warmer world that could not be part of the 1.5°C world distribution. In a 2°C warmer world, 40% of the summer months over Europe are projected to exhibit mean temperatures distinguishable from those in a 1.5°C warmer world. This distinguishability is largest over Southern Europe, and decreases to around 10% of the summer months over Eastern Europe.

2.2. To what extent can the risk of extreme European summer temperatures be controlled by maintaining global warming below fixed global mean temperature limits?

Internal variability narrows the controllability of extreme maximum temperatures to the point that, by limiting global warming to 1.5°C, only the 10% most extreme summer maximum temperatures in a 2°C world could be avoided.

These findings highlight the limited controllability of the amplitude of extreme temperature events at regional levels by maintaining global warming below fixed limits, and emphasize the importance of considering the irreducible uncertainty introduced by chaotic internal variability in evaluating the impacts of climate change.

3. *Heat hotspots under global warming*

In Chapter 3 I investigate where the major heat-stress hotspots occur globally as the world warms for five summertime heat metrics — maximum reachable temperatures, return periods of extreme temperatures, maximum temperature variability, sustained tropical night temperatures, and extreme wet bulb temperatures. Based on these metrics, I then identify the maximum global warming that allows us to avert harmful heat levels. I use MPI-GE to construct well-defined samples for five different global warming levels, ensuring the robust sampling of internal variability that is vital to evaluate changes in the most extreme events reachable under different warming levels. By doing so, I find the following answers to my research questions.

3.1 Which regions become summertime heat stress hotspots for these different extreme heat metrics under global warming?

For maximum temperatures, MPI-GE projections stay below 50°C generally all over the world for global warming levels below 2°C. However, for warming levels above 2°C, this threshold could be overshot over large regions in all continents, in particular North Africa and Central Eurasia; while in some regions such as Pakistan, Iraq or Saudi Arabia maximum temperatures could surpass the 60°C threshold.

Similar hotspots emerge for return periods of extreme temperatures, Summer maximum temperatures reached on average once every 100 years under pre-industrial conditions could occur every 10 to 25 years at 1.5°C of warming, and as often as every two years in North Africa. At 4°C of warming, these 1-in-100-years events could happen more often than every two years almost all over the world.

Summer maximum temperature variability increases relative to pre-industrial levels up to 50% under 2°C of warming, in regions such as South America or Central Europe. At 4°C of warming, variability increases over large continental areas, with a maximum relative increase well above 100% that translates into maximum year-to-year variations of summer maximum temperatures of up to 14°C in North America or Central Europe, and up to 18°C in India. For some high latitude regions and parts of Southern Europe or North America, maximum temperature variability decreases by 10-35% at 4°C of warming.

Heat stress could be substantially aggravated by high minimum temperatures, with sustained tropical night conditions over large non-adapted regions under global warming beyond 2°C. At 4°C of global warming, tropical night hotspots spread polewards over large regions across all continents, particularly in central North America, South Africa, or Central Eurasia. In addition to this substantial spread, sustained tropical night conditions also prevail consistently for at least 95% of the summer months.

Hot and humid conditions characterized by extreme wet bulb temperatures that occur rarely in the pre-industrial climate are projected to expand to almost all continents at 1.5°C of warming, and spread to the point of covering most tropical regions and some regions at mid latitudes in a 4°C world.

3.2. For which maximum global warming level can the risk of extreme heat conditions under these different metrics be confidently averted?

Different metrics exhibit different levels of risk under warming. For example, maintaining global warming below 2°C ensures that unprecedented levels of extreme hot and humid conditions can be confidently averted in most regions of the world. On the other hand, at 2°C of warming very extreme events could occur with frequencies already more than 20 times higher than under pre-industrial conditions.

My results indicate that completely averting some of the heat-related risks that arise under further global warming may not be possible. However, I find that maintaining global warming levels below 2°C above pre-industrial conditions is vital in order to minimize these risks and limit the exposure of non-adapted regions to potentially dangerous heat levels.

4. *Concluding remarks*

In this dissertation, I use the largest existing ensemble from a fully coupled Earth system model to investigate how chaotic internal variability shapes the changing characteristics of summertime heat extremes in a warming world. I investigate which mechanisms drive extreme heat variability, and find that the mechanisms causing some summers to be much warmer than others may change under global warming. I also evaluate how internal variability narrows the distinguishability and controllability of extreme summertime heat regionally, even if we succeed to maintain global warming below fixed limits. Lastly, I identify where the major heat hotspots emerge under further global warming — from the highest maximum summertime temperatures projected over the Arabic Peninsula, to the largest maximum temperature variability increase over India or Central Europe.

My findings emphasize the importance of considering the irreducible uncertainty introduced by internal variability to evaluate the impacts of anthropogenic climate change. Due to this irreducible uncertainty, completely averting the risks of extreme heat that arise due to global warming may not be possible. Our best chances for doing so rest on putting a stop to the global mean temperature increase. Maintaining global warming below 2°C above pre-industrial conditions is vital to minimize the risk of several manifestations of extreme heat, and limit the exposure of non-adapted regions to harmful heat levels. Given time and within limits, our society, economy, ecosystems, and even our bodies, are able to adapt to a warmer mean climate state. However, the deviations from these mean climate conditions in the form of extreme events caused by internal variability challenge our range of adaptability, potentially to its limits.

APPENDIX A

THE MAX PLANCK INSTITUTE GRAND ENSEMBLE

The Max Planck Institute Grand Ensemble (MPI-GE) is currently the largest existing single-model ensemble using a comprehensive, fully-coupled Earth System Model — largest both in terms of forcing scenarios represented and in terms of independent members (Maher et al., 2019). MPI-GE consists of sets of 100 simulations that evolve under the same model physics and parametrizations and are driven by the same external forcings, but that start from different initial climate states, sampled from different points of the model's pre-industrial control run. This design implies that each simulation differs from the rest only due to the effect of internal variability, thus allowing a clear separation between the changes caused by the external forcing and the quasi-random fluctuations caused by the chaotic internal variability of the climate system. Furthermore, the large ensemble size in MPI-GE offers a substantial sampling of internal variability in a transient climate. This substantial sampling of internal variability allows us to empirically calculate well-defined probability distributions, and provides the large samples of extreme events under different warming conditions that are vital for this study. In this Appendix, I present the experimental design of MPI-GE, how it relates to other large ensemble experiments, and some examples on how to utilize its power.

The diversity of forcing pathways used in MPI-GE is one of the key ingredients in its unique experimental design. The 100-member sets of MPI-GE simulations are available for historical forcing conditions as well as for three future scenarios and one scenario featuring a 1% CO_2 increase per year. The future scenarios are based on three Representative Concentration Pathways (RCP2.6, RCP4.5 and RCP8.5) from the CMIP5 framework, that describe alternative trajectories for CO_2 emissions and the resulting atmospheric concentration of CO_2 based on different assumptions about population, economic growth, energy consumption and sources, and land use (van Vuuren et al., 2011). RCP2.6 can be interpreted as an early mitigation scenario, where CO_2 emissions are drastically reduced after 2020 and the atmospheric CO_2 concentration reaches its maximum at around 440 ppm. The low-emission scenario RCP2.6 corresponds in MPI-GE to a global warming level by the end of the century below 1.5°C above pre-industrial conditions, well in compliance with the Paris Agreement (Fig. A.1). RCP4.5 represents a moderate warming scenario where emissions continue at current rates until the mid century and then decline, resulting in a 2.2°C warming in MPI-GE

by the end of the century. Lastly, RCP8.5 can be interpreted as a *business as usual* scenario, where CO_2 concentrations reach 950 ppm by the end of the century and continue increasing for another 100 years (van Vuuren et al., 2011; IPCC, 2013). Under this high-emission scenario MPI-GE projects global warming levels of more than 4°C above pre-industrial conditions by the end of the 21st century (Fig. A.1). MPI-GE allows, for the first time, a comparison of the climates under these three possible future trajectories in a comprehensible, large-ensemble setting.

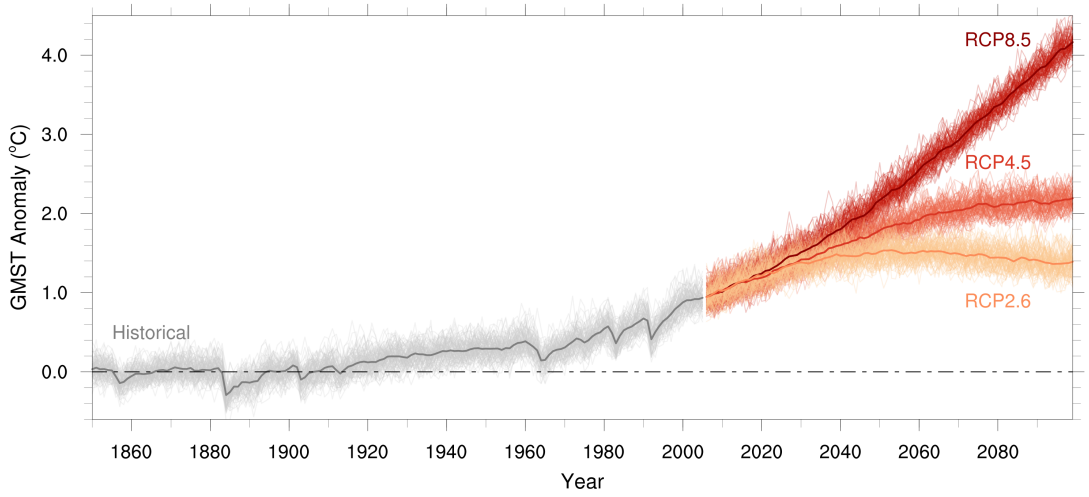


Figure A.1: Global mean surface temperature (GMST) in MPI-GE. Time series of annually averaged GMST anomalies (colored thin lines) and centered ensemble means (colored thick lines) for each forcing conditions for the period 1850–2099, simulated by the MPI-GE. Simulations are historical runs for the period 1850–2005 (gray lines), and RCP2.6 (orange lines), RCP4.5 (red lines) and RCP8.5 (dark red lines) for the period 2006–2099. Anomalies are calculated with respect to pre-industrial levels defined by the period of 1851–1880.

The second key ingredient of MPI-GE is its large ensemble size. An ensemble size of 100 simulations under the same forcing conditions allows 1-in-100-years events to occur on average every simulated year. Thus, it enables the simulation and characterization of large samples of extreme events with return periods of over hundreds of years that develop under different background conditions and different levels of global warming. This precise sampling of the simulated internal variability allows us to empirically calculate probability distributions, without the explicit need to parametrize the tails of the distributions with extreme value statistics. This is a crucial requirement to empirically evaluate the statistical significance of changes in very rare events. Ultimately, MPI-GE offers a precise characterization of the irreducible uncertainty that arises from the simulated chaotic internal variability in a transient climate

that is not confounded by different model configurations or forcing responses.

Several studies use MPI-GE to evaluate this irreducible uncertainty that arises due to internal variability and disentangle it from the response to external forcings in different quantities (Bittner et al., 2016; Hedemann et al., 2017; Suarez-Gutierrez et al., 2017; Suarez-Gutierrez et al., 2018; Maher et al., 2018; Marotzke, 2019). For example, Hedemann et al. (2017) determine that surface warming hiatuses occur 3% of the time in MPI-GE due to internal variability, and that these warming hiatuses can have both oceanic and atmospheric origins. Suarez-Gutierrez et al. (2017) investigate the long-debated topic of whether the real-world upper tropospheric warming in the tropics as described by observations is well captured in model simulations. They show that differences between observed and simulated tropical tropospheric warming, rather than being caused by incorrect model performance, are dominated by observational uncertainty and internal variability. Marotzke (2019) quantifies the degree of irreducible uncertainty around whether emissions reduction will cause the desired climate response over a given timescale, and finds a probability of about one-third that the warming rate in the period of 2021–2035 will increase in RCP2.6 despite the reduction in emissions.

Currently, there are several other large ensemble experiments with different comprehensive climate models available, with sizes ranging from a few members to up to more than 50 (Deser et al., 2012; Kay et al., 2015; Rodgers et al., 2015; Kirchmeier-Young et al., 2017; Frankignoul et al., 2017; Sanderson et al., 2017; Maher et al., 2019). As large ensembles from fully-coupled climate models are gaining traction as the preferred method to robustly estimate the internal variability and forced response in a transient climate, it becomes more relevant to understand how simulated internal variability behaves in different models. However, in-depth multi-model evaluation and comparison of the simulated internal variability of different quantities in different single-model large ensembles, similar to that shown by Maher et al. (2018) and Schaller et al. (2018), remains limited.

Another emerging approach to investigate internal variability in a transient climate are atmosphere-only large-ensemble experiments with prescribed SSTs, such as the Half a Degree Additional warming, Prognosis and Projected Impacts project (HAPPI; Mitchell et al., 2017). This approach has the advantage of requiring fewer computational resources than their fully-coupled counterparts, thus making it much more rapidly accessible and deployable to a large number of modeling groups. However, it also has known limitations. A finite set of prescribed SST patterns offers a limited range of climate states that does not completely sample ocean-driven vari-

ability (Hawkins et al., 2016; Sanderson et al., 2017; Fischer et al., 2018). In contrast, large ensembles from fully-coupled climate models offer a better estimation of the real-world variability, because they sample a wider range of ocean states and include the influence of the ocean-borne variability (Hawkins et al., 2016). Furthermore, fully-coupled large ensembles also offer a more realistic representation of heat extremes over land than atmosphere-only large-ensembles, even if the later offer a larger number of independent simulations (Fischer et al., 2018).

MPI-GE Model Description

MPI-GE uses the Max Planck Institute Earth System Model (MPI-ESM; Giorgetta et al., 2013). The version used, MPI-ESM1.1, is run in the low resolution (LR) configuration, with resolution T63 and 47 vertical levels in the atmosphere component (ECHAM6.3; Giorgetta et al., 2013) and 1.5°resolution and 40 vertical levels in the ocean component (MPIOM1.6; Jungclaus et al., 2013). It also includes the land-surface model component JSBACH-3 (Raddatz et al., 2007) and the biogeochemistry model HAMOCC (Ilyina et al., 2013). MPI-ESM1.1 has some similarities to the the CMIP5 version of MPI-ESM (Taylor et al., 2012; Maher et al., 2019), but has a slightly lower equilibrium climate sensitivity of 2.8°C (Mauritsen et al., 2019; Flato et al., 2013), and a new 5-layer soil hydrology scheme (Hagemann and Stacke, 2015) implemented in JSBACH. MPI-GE has a relatively low spatial resolution, comparable to most of the models in the CMIP5 experiment, which can influence the model's ability to simulate small-scale processes and affect the reliability of its projections. Temporal resolution is also relatively limited in MPI-GE, with only monthly output available, which limits the analysis of processes in shorter time-scales.

APPENDIX B

NEW APPROACH FOR EVALUATING INTERNAL VARIABILITY

In this section I introduce a novel method to investigate how the simulated internal variability in MPI-GE compares to real-world estimates. This method is based on a simple approach: evaluating whether the observed estimates occur across the whole range of the ensemble simulations, and whether they always stay within this range. This model-evaluation method, first applied in Suarez-Gutierrez et al. (2018) and further demonstrated in Maher et al. (2019), presents two main advantages. First, as opposed to other evaluation techniques based on comparisons of mean values or standard deviations, our approach focuses on the model's ability to simulate higher moments of the distribution, offering a more appropriate evaluation of the simulated representation of the strength and frequency of extreme events. Second, this method allows us to directly identify whether differences between observed and simulated values are due to an incorrect simulation of the mean climate or its response to external forcings, or rather due to an incorrect representation of internal variability. Next I will explain this method in more detail and apply it to evaluate the ability of MPI-GE to simulate observed mean and maximum surface temperatures in the summer months.

In the ideal case that real-world variability is perfectly simulated and the observational record is sufficiently long, the observed values would occur across the whole ensemble spread with similar frequency. To illustrate this, I choose the central 75^{th} percentile range of the ensemble. I determine that MPI-GE adequately simulates the observed estimated variability for the regions where observations fall within this 75^{th} percentile range around 75% of the time. On the other hand, I determine that MPI-GE overestimates the observed variability in the regions where more than 75% of the observations crowd within these bounds. This is illustrated by the gray shading in Figure A.2 for simulated summertime monthly mean surface temperatures compared to HadCRUT4 observations (Morice et al., 2012). The variability in monthly mean temperatures is not overestimated for regions such as Europe or North America; whilst is somewhat overestimated in regions such as Central Africa or Australia. The time series for specific grid points where variability is adequately captured (Fig. A.2a) reveal that the observed values indeed occupy the whole ensemble spread, including the ensemble limits, and occur only occasionally outside the ensemble. For the regions

where variability is slightly overestimated, the time series show that observed values fall mostly in the central region of the ensemble (Fig. A.2b). However, in the example case over Australia in Fig. A.2b, extreme observed values occur on the ensemble upper limit in several months, indicating that the ensemble adequately captures the strength and occurrence of warm events, while overestimating cold events.

The second aspect of our evaluation is whether observed estimates always occur within the ensemble range, indicating that MPI-GE simulated variability is sufficient to always capture observed anomalies. The red and blue shading in Figure A.2 indicate the percentage of observed values of summertime monthly mean temperatures that fall above or below the ensemble limits, respectively. In red regions, I expect an underestimation of the magnitude and frequency of extreme warm events that can lead to underestimated projections for future climates; while the same is true for extreme cold events in the blue regions. I find that observations lie outside the ensemble around 10% of the months in most tropical regions, as well as regions over Asia, Europe or North America. Observations may lie both above and below the ensemble limits because MPI-GE underestimates internal variability in certain regions. However, other factors may also play a relevant role. For regions where observations occur systematically either below or above the ensemble limits, the cause may be that the average climate state is not well represented either due to an incorrect response to external forcings, or to an incorrect representation of relevant processes, that can also lead to a correct representation of one tail of the distribution but not the other. Thus, this approach highlights regions of interest that require a careful case-to-case investigation.

As an example, Fig. A.2c illustrates the time series for a grid point in the Arabian Peninsula, where observations lie both above and below the ensemble limits, indicating a potential underestimation of internal variability in MPI-GE. Comparing MPI-GE to the observational record over this region, that starts around 1950, I identify two behaviors. The first half of the record indicates a possible underestimation of variability, with observed values often close to the ensemble limits or below the ensemble minimum. However, after 1990 the warming signal in the observed record is substantially larger than in MPI-GE, resulting in observed values lying more than 1°C above the ensemble maximum. In the case of Fig. A.2d, representing a grid point over China, it appears that the differences between observed and simulated values are not caused by an incorrect representation of the response to external forcing as in the previous case, but rather by an underestimation of internal variability, particularly on decadal to multidecadal time scales.

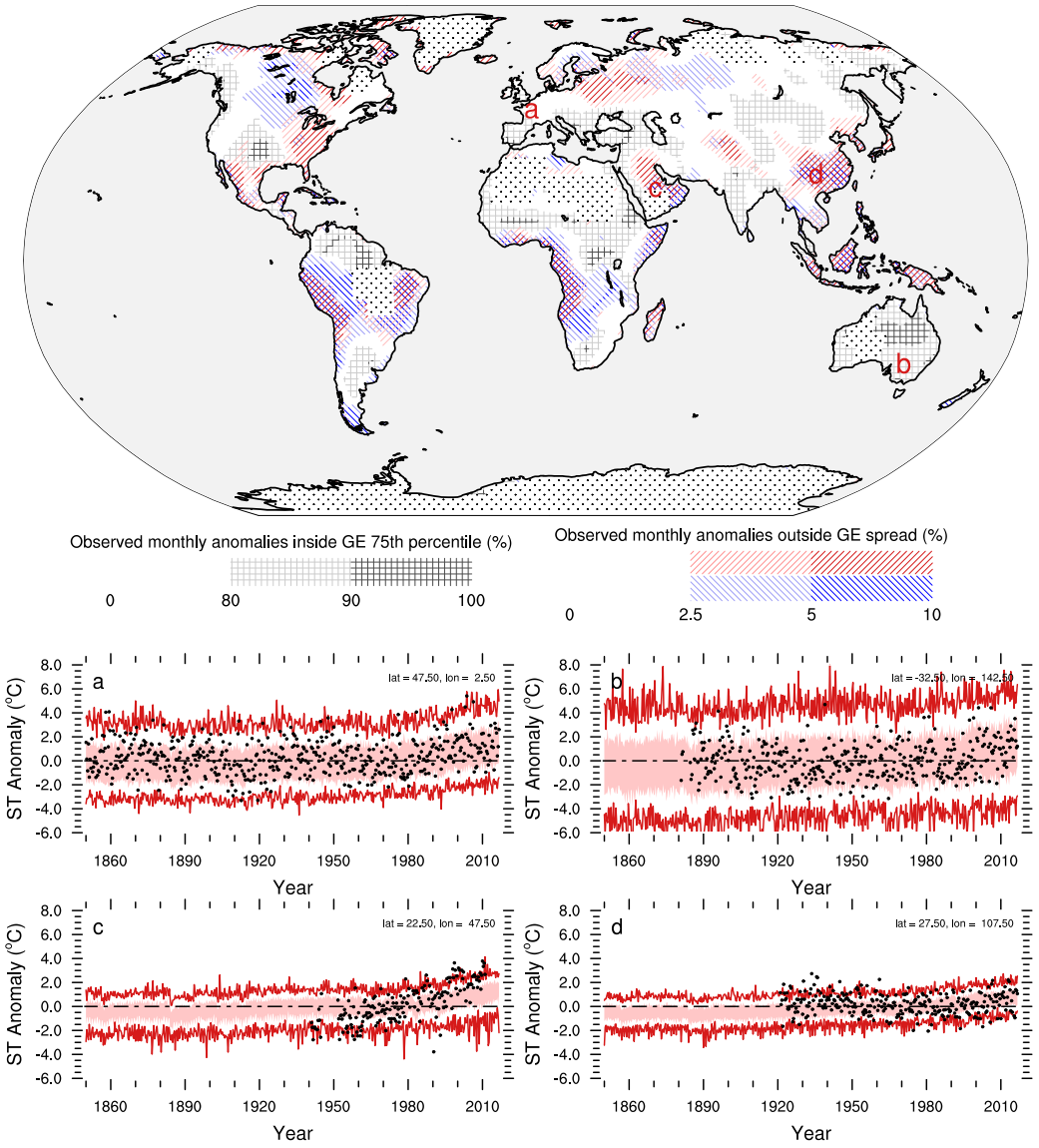


Figure A.2: Summer mean temperature variability in MPI-GE vs. observations. Global map representing the variability of summer monthly mean 2m air surface temperature anomalies from HadCRUT4 data (Morice et al., 2012) compared to the variability in MPI-GE for the period 1850–2018. Red shading represents regions where observed anomalies are larger than the ensemble maximum. Blue shading represents where observed anomalies are smaller than the ensemble minimum. Gray hatching indicates regions where MPI-GE simulated variability is larger than the observed estimate; with the percentage of observed monthly anomalies occurring inside the 75th percentile of the MPI-GE distribution. Black dots represent regions where no observations are available or are available only for less than ten months. Time series for specific grid points (a-d) show ensemble maximum and minimum (red lines) and 75th percentile bounds (red shading) compared to observed anomalies (black points). Anomalies are calculated with respect to climatological levels defined by the period of 1961–1990. Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2018 and are adapted to the HadCRUT4 grid. Summer months are defined as JJA for the Northern Hemisphere and DJF for the Southern Hemisphere.

Whereas variability in mean temperatures is mostly not overestimated in MPI-GE, when investigating maximum temperature I find that observations occur within the 75th percentile bounds of MPI-GE for more than 85% of the months over most regions (Fig. A.3). This comparison reveals that MPI-GE tends to overestimate the spread in summertime monthly maximum surface temperatures as estimated from the BEST data (Rohde et al., 2012). However, a more detailed investigation shows that MPI-GE performs adequately in simulating extremes in several relevant regions, particularly in the upper tail of the distribution. In the case of Europe (Fig. A.3a), warm extremes are generally adequately represented, while cold extremes appear to be overestimated in MPI-GE. This overestimation of summertime cold extremes occurs in other regions, such as North America (Fig. A.3b), while other relevant regions present a similar behavior for both types of extremes (Fig. A.3c and A.3d). This tendency to overestimate the variability in maximum temperatures may indicate that MPI-GE also overestimates future projections of maximum temperatures. On the other hand, observed values generally occur within the ensemble range, with some exceptions in Central Africa or East Asia, indicating that MPI-GE does not underestimate the magnitude and frequency of extreme events in most regions. However, summer maximum temperatures generally exhibit larger variability than summer monthly mean temperatures (Suarez-Gutierrez et al., 2018), and thus the observational record may be too short to determine whether the amplitude and frequency of extreme events is adequately captured in MPI-GE.

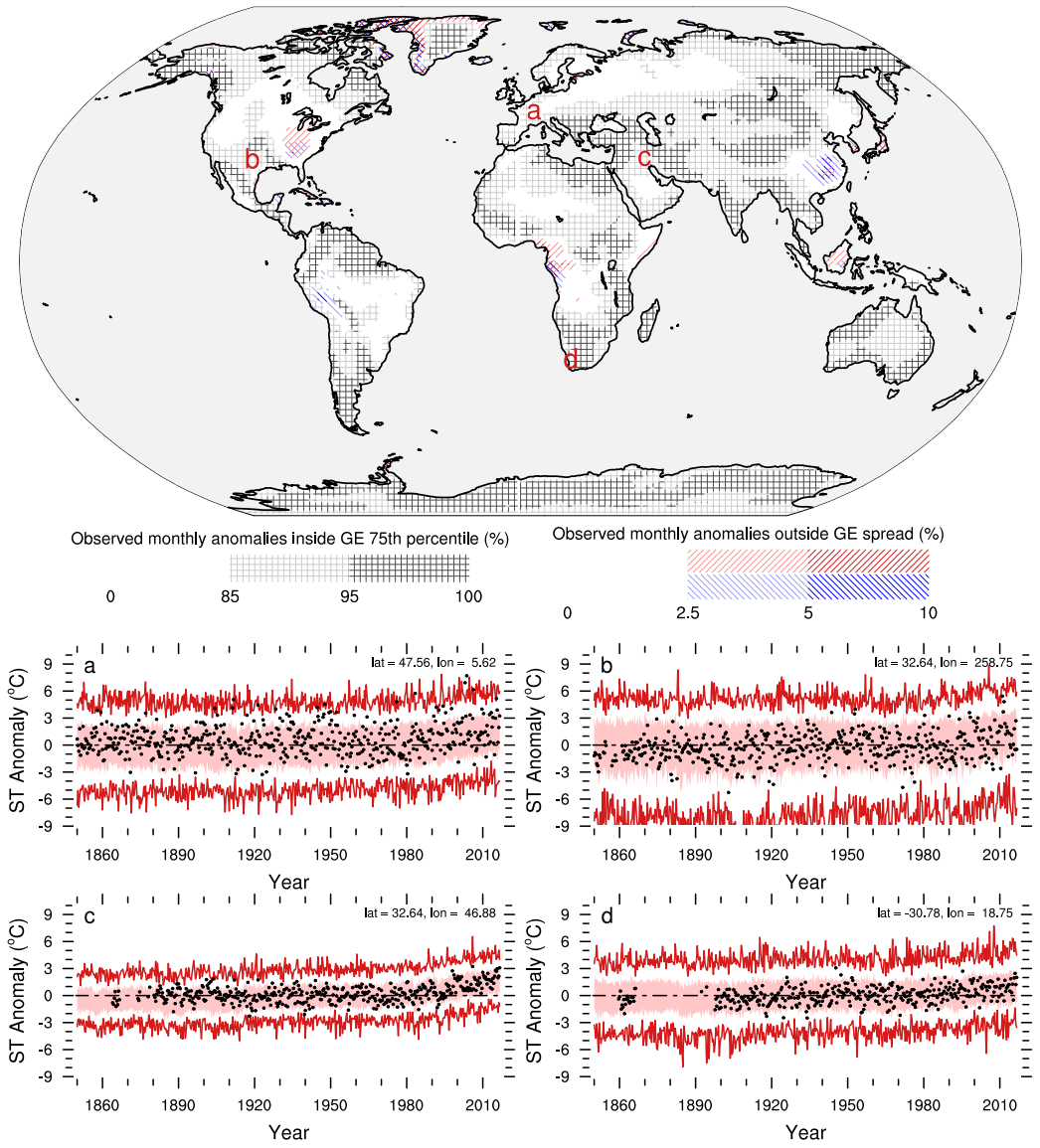


Figure A.3: Summer maximum temperature variability in MPI-GE vs. observations. Global map representing the variability of summer monthly maximum 2m air surface temperature anomalies from BEST data (Rohde et al., 2012) compared to the variability in MPI-GE for the period 1850–2018. Red shading represents regions where observed anomalies are larger than the ensemble maximum. Blue shading represents where observed anomalies are smaller than the ensemble minimum. Gray hatching indicates regions where MPI-GE simulated variability is larger than the observed estimate; with the percentage of observed monthly anomalies occurring inside the 75th percentile of the MPI-GE distribution. Black dots represent regions where no observations are available or are available only for less than ten months. Time series for specific grid points (a-d) show ensemble maximum and minimum (red lines) and 75th percentile bounds (red shading) compared to observed anomalies (black points). Time series for specific grid points (a-d) show ensemble maximum and minimum (red lines) and 75th percentile bounds (red shading) compared to observed anomalies (black points). Anomalies are calculated with to the climatology baseline, defined in BEST data by the period of 1951–1980. Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2018. BEST data are adapted to the coarser resolution MPI-GE grid. Summer months are defined as JJA for the Northern Hemisphere and DJF for the Southern Hemisphere.

APPENDIX C

EVALUATION OF THE MAX PLANCK INSTITUTE GRAND ENSEMBLE

In this section I investigate the ability of the MPI-GE to simulate observed surface temperatures globally and in particular over Europe. Figure A.4 illustrates how MPI-GE captures the observed variability in global mean surface temperature (GMST) and European summer monthly mean temperature (EuST) anomalies. GMST observations occur within the ensemble spread for the majority of the record and generally across the whole ensemble width with no preferred pattern (Fig. A.4a). To illustrate this more clearly, I calculate the place, or rank, that the observed estimate would take in a list of the ensemble members ordered by ascending GMST anomalies for each year. The rank of observed anomalies that are smaller than those from all ensemble members is 1; while the rank of observed anomalies larger than the anomalies of all 100 members is 101. The histogram of these ranks for observed GMST data presents a pattern that is reasonably flat and continuous, indicating that observations occupy all ranks with no preferred frequency (Fig. A.4b). Similarly, the probability distribution function of simulated GMSTs for the reference period of 1981–2010 shows good agreement the with the observed estimates in both the shape and the amplitude of the distribution.

In the case of observed temperatures averaged over Europe, EuST anomalies occur rarely outside the ensemble limits, but tend to cluster in the upper half of the ensemble (Fig. A.4c). The rank histogram for EuST shows a skewness toward higher ranks, indicating that observations fall towards the ensemble maximum with higher frequency as towards the ensemble minimum (Fig. A.4d). However, this skewness exhibits some decadal variability, and is less prominent in the period of 1981–2010, for which the observed and simulated distributions present good agreement in both shape and amplitude (Fig. A.4f). This tendency may indicate an overestimation of the frequency and amplitude of low European summer temperatures in MPI-GE during some periods. However, the observational record may not be sufficiently long to allow for a complete characterization of the large multi-decadal variability in European summer temperatures.

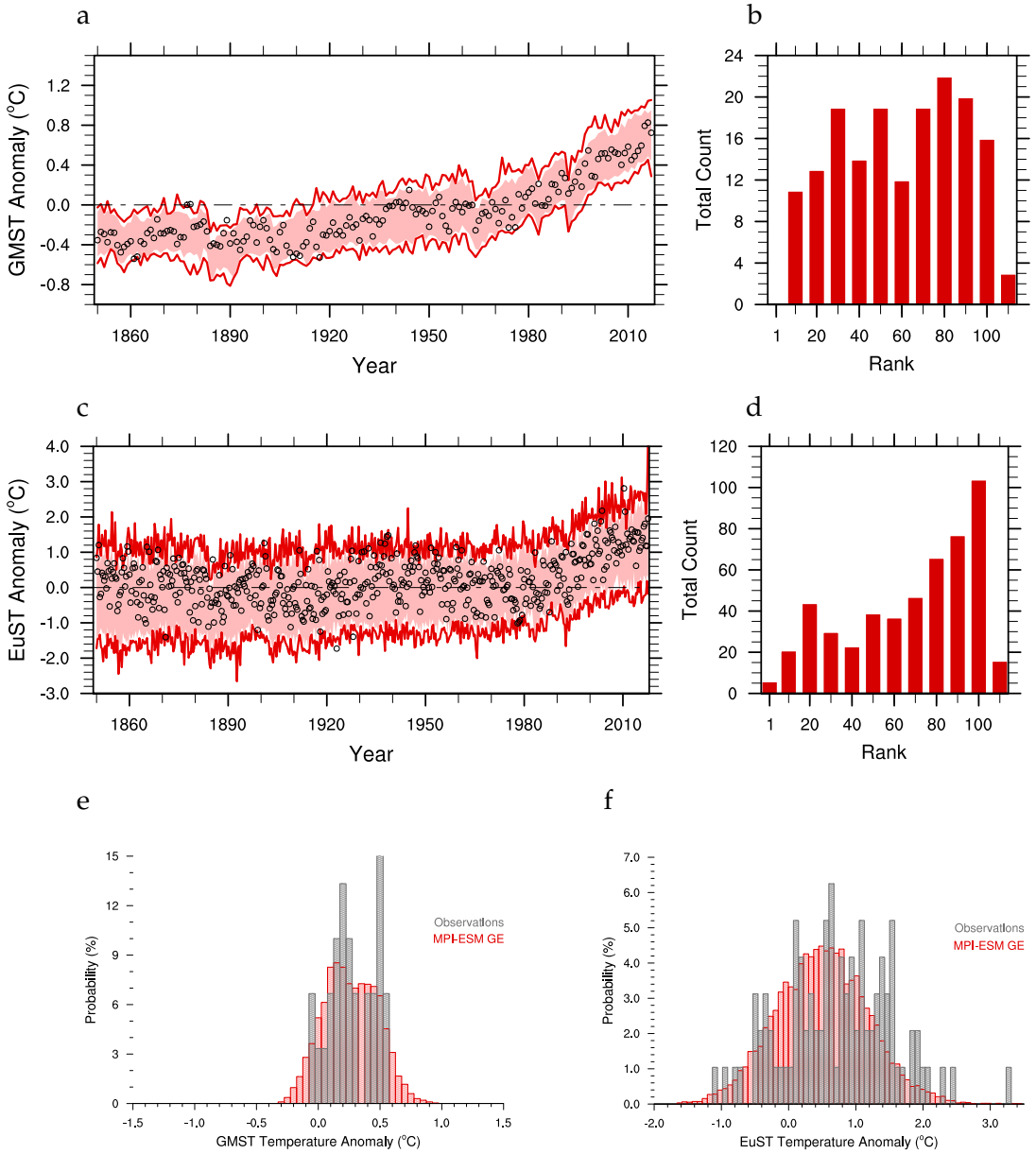


Figure A.4: Global and European temperatures in MPI-GE vs. observations. (a) Time series of GMST anomalies simulated by the MPI-GE (red) compared to HadCRUT4 observations (black). **(b)** Rank histogram for the HadCRUT4 GMST observations as a member of the Grand Ensemble for the period of 1850–2016. **(c)** Time series of EuST anomalies simulated by the MPI-GE (red) compared to CRUTEM4 observations (black), as in (a). **(d)** Rank histogram for the CRUTEM4 EuST observations as a member of MPI-GE for the period of 1850–2017. **(e)** Probability distribution of GMST anomalies simulated by the MPI-GE (red) compared to HadCRUT4 observations (gray) for the period of 1981–2010. **(f)** Probability distribution of EuST anomalies simulated by the MPI-GE (red) compared to CRUTEM4 observations (gray) for the period of 1981–2010. Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2017. Anomalies are calculated with respect to climatology baseline defined by the the period of 1961–1990. Simulated data are subsampled to grid cells where observations are available.

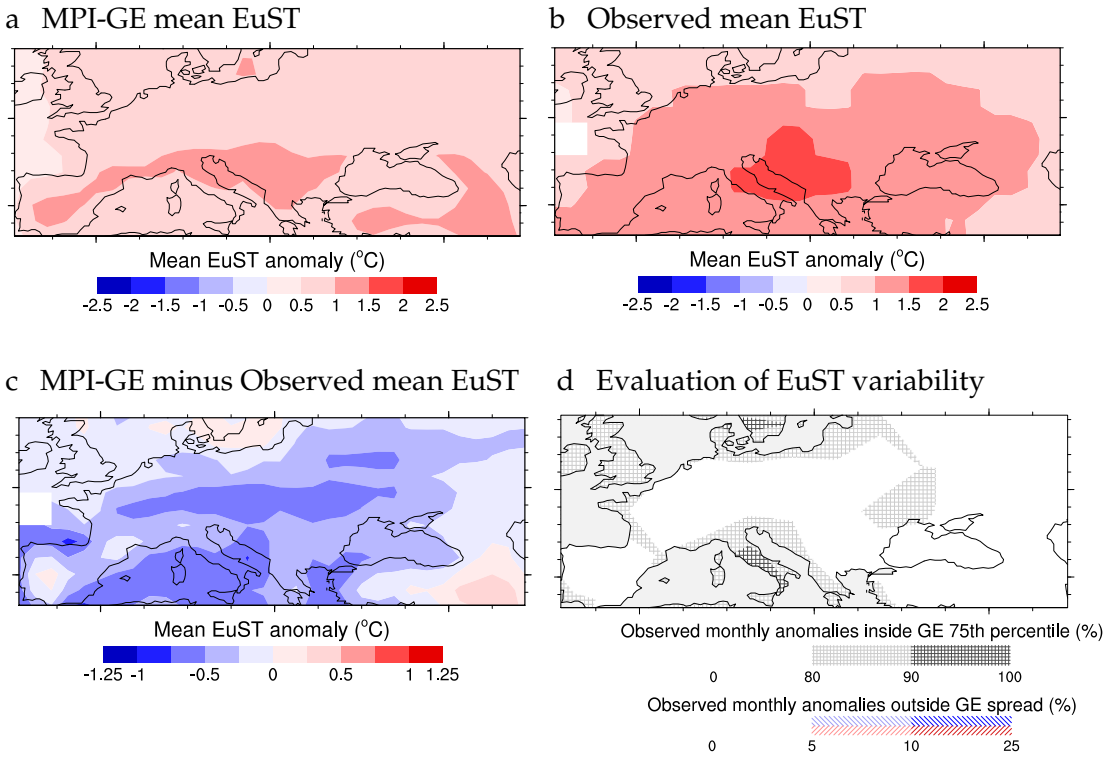


Figure A.5: European summer temperatures in MPI-GE vs. observations. (a) MPI-GE mean EuST anomaly over the period 1990–2017. **(b)** CRUTEM4 mean EuST anomaly over the period 1990–2017. **(c)** Difference between mean EuST anomaly in MPI-GE minus CRUTEM4 over the period 1990–2017. **(d)** EuST variability in MPI-GE compared to CRUTEM4 observed variability for the period 1850–2017. Gray hatching shows where the estimated observed variability is smaller than the simulated variability, as the percentage of observations occurring within the 75th ensemble percentile. Red and blue shading represents regions where the observed estimated variability is larger than the simulated variability. Red represents the percentage of observations larger than the ensemble maximum anomaly at that time step; blue represents the percentage of observations smaller than the ensemble minimum. Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2017. Anomalies are calculated with respect to the climatological level defined by the the period of 1961–1990.

Next, I evaluate how the ability of the MPI-GE to simulate the mean observed summer monthly mean temperatures and their variability varies over Europe (Fig. A.5). For average temperatures in current climate conditions, represented by mean EuSTs over the period of 1990–2017, the ensemble mean temperatures are slightly lower than the observed EuSTs, with the largest differences around 0.5°C (Fig. A.5c). Although the ensemble presents a slight cold bias in the last decades, it captures the estimated variability in the observed record adequately (Fig. A.5d). Observed anomalies rarely occur outside the ensemble limits, and fall mostly across the whole ensemble spread. In some regions like Italy, observations occur within the 75th percentile bounds of the ensemble for more than 90% of the summer months. This indicates that simulated temperature variability is slightly larger than the observed estimate in these areas. In

case this analysis is performed for the frequency of observational estimates within the ensemble's 50th percentile, we find good agreement between observed and simulated variability estimates, with around 50% frequency for the whole domain.

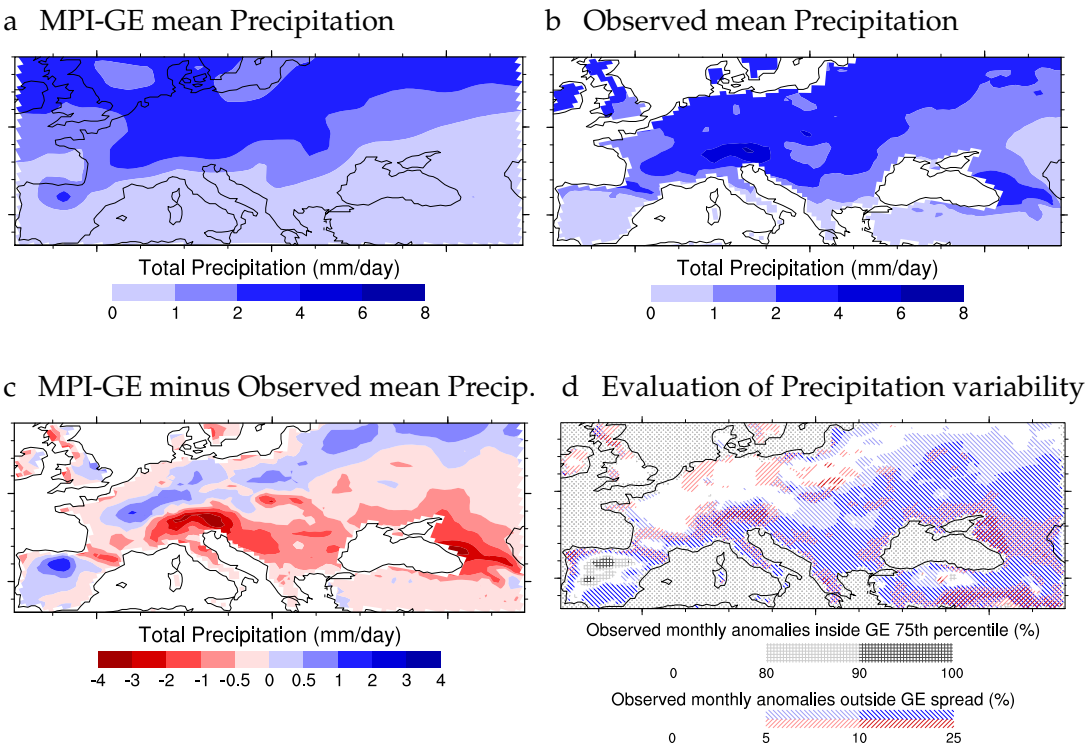


Figure A.6: Precipitation in MPI-GE vs. observations. (a) MPI-GE mean precipitation anomaly over the period 1990–2017. **(b)** E-OBS mean precipitation anomaly over the period 1990–2017. **(c)** Difference between mean precipitation anomaly in MPI-GE minus E-OBS in the period 1990–2017. **(d)** Precipitation variability in MPI-GE compared to E-OBS observed variability for the period 1950–2017. Gray hatching shows where the estimated observed variability is smaller than the simulated variability, as the percentage of observations occurring within the 75th ensemble percentile. Red and blue shading represents regions where the observed estimated variability is larger than the simulated variability. Red represents the percentage of observations larger than the ensemble maximum anomaly at that time step; blue represents the percentage of observations smaller than the ensemble minimum. Simulations are historical runs for the period 1850–2005 and RCP4.5 for the period 2006–2017. Anomalies are calculated with respect to the climatological level defined by the the period of 1961–1990.

One of the reasons that may cause an overestimation in MPI-GE of the amplitude and frequency of colder European summers is an incorrect representation of precipitation. Performing an analogous analysis for total precipitation over Europe I find that the mean precipitation as well as its variability can be misrepresented in MPI-GE (Fig. A.6). The ensemble mean precipitation for the period 1990–2017 is around 1 mm/day larger than the observed average over the northern part of the domain and part of the Iberian Peninsula, while being around 1 to 4 mm/day lower than the observed average

in Southern Europe, particularly over alpine regions (Fig. A.6c). The simulated precipitation variability in the ensemble is smaller that the estimated observed variability over most of Europe (Fig. A.6d). Observed precipitation anomalies occur outside the ensemble limits up to more than 25% of the summer months, indicating that the frequency and amplitude of extremely wet and, particularly, extremely dry events is underestimated in MPI-GE. These results agree with our findings in Fig. A.4d and point to an overestimation of the frequency of colder than average summer months in the ensemble simulations that may be partially caused by biases in precipitation variability. Albeit these potential biases in the simulation of cold extremes, I find that MPI-GE captures the frequency and amplitude of extremes in the upper tail of the observed distribution adequately, and is a suitable tool for the investigation of extremely warm summers.

APPENDIX D

SUPPORTING INFORMATION

Supporting Figures to Chapter 1

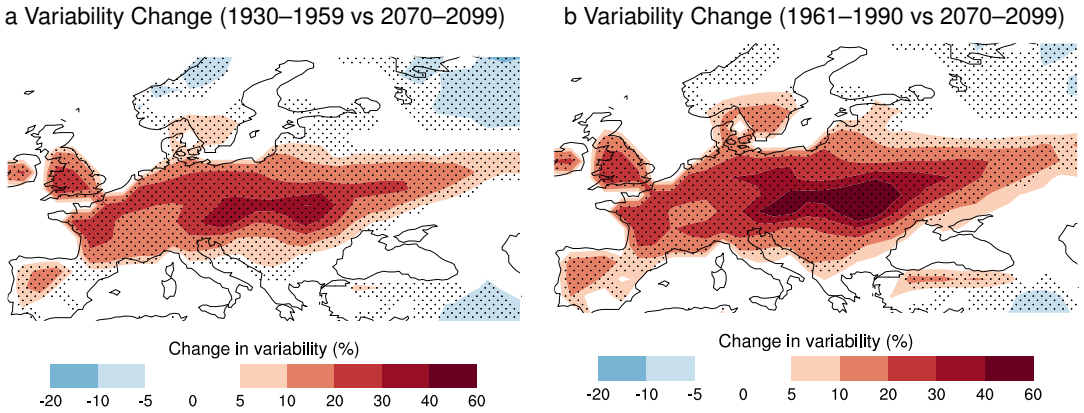


Figure S.1: Change in European summer temperature variability relative to different reference periods. (a) Relative change in variability based on change in EuST probability density distribution width (2.5th–97.5th percentiles) for mid 20th century (1930–1959) compared to late 21st century (2070–2099) for each grid cell. **(b)** Relative change in variability based on change in EuST probability density distribution width (2.5th–97.5th percentiles) for the 20th century climatological period (1961–1990) compared to late 21st century (2070–2099) for each grid cell. Stippling shows significance for late 21st century PDF widths larger (or smaller) than all the possible 30-year PDFs in the 20th century. Simulations are historical runs for the period 1850–2005 and RCP4.5 runs for the period 2006–2099 from the MPI-GE. All anomalies are calculated with respect to the climatological period of 1961–1990.

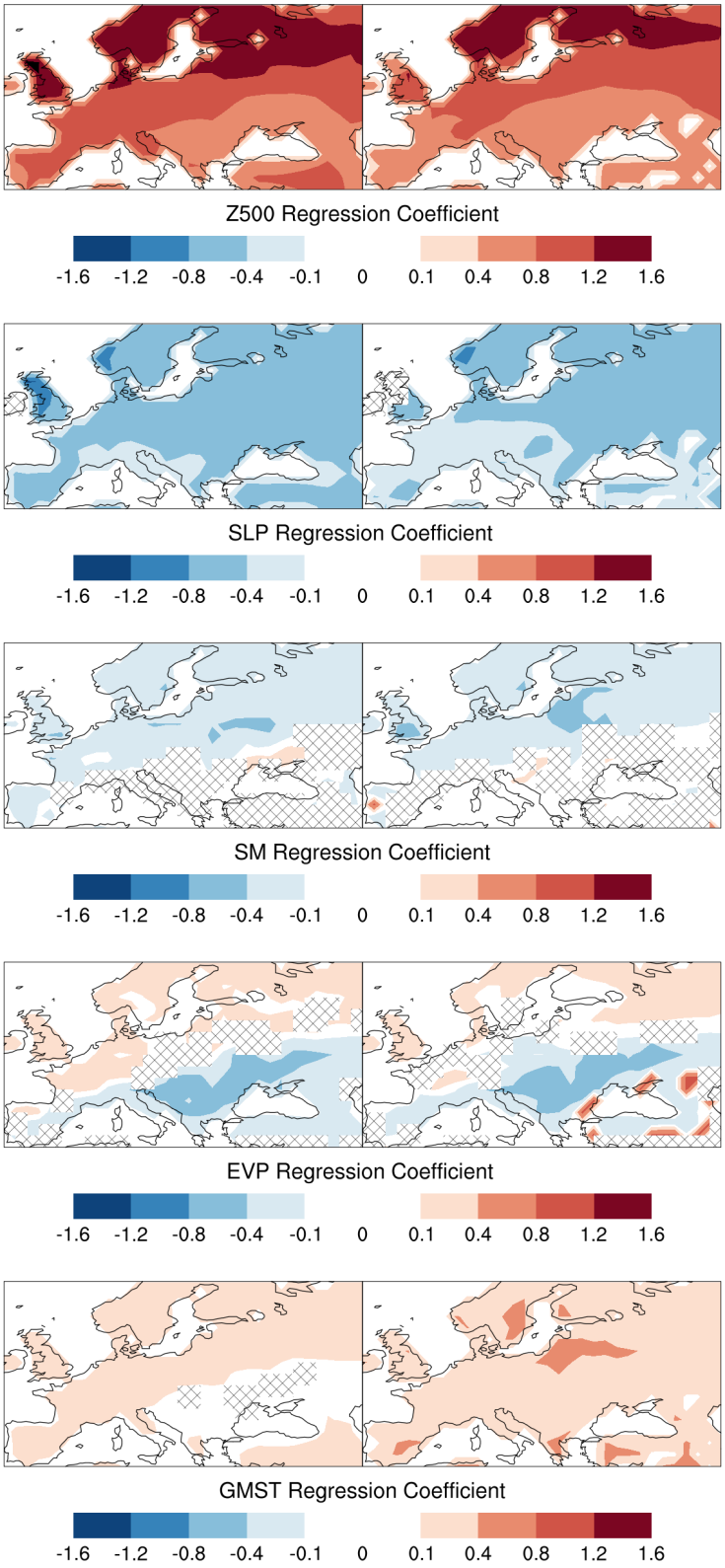


Figure S.2: Standardized regression coefficients from multiple regression with higher multicorrelation threshold. Point-to-point standardized regression coefficients between 2σ extreme EuST and different drivers from multiple regression analysis for the 20th century (left column) compared to for the 21st century (right column) for a more conservative multicorrelation threshold of 0.9, corresponding to a VIF of 5. Hatching represents regions where the variable is excluded from the regression model either because its contribution is not significant or because it exhibits too high multicollinearity with the remaining predictors in the model.

Supporting Figures to Chapter 2

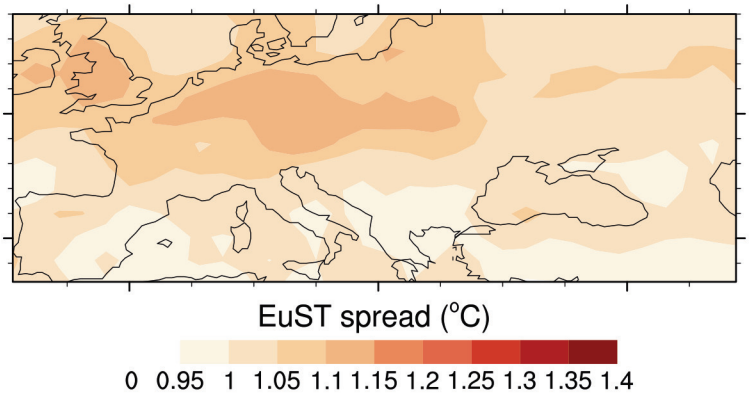


Figure S.3: Variability change in European summer monthly mean temperatures. Ratio of the spread of EuST anomalies at 2°C of global warming over EuST spread at 1.5°C of global warming above pre-industrial conditions, simulated by the MPI-ESM Grand ensemble.

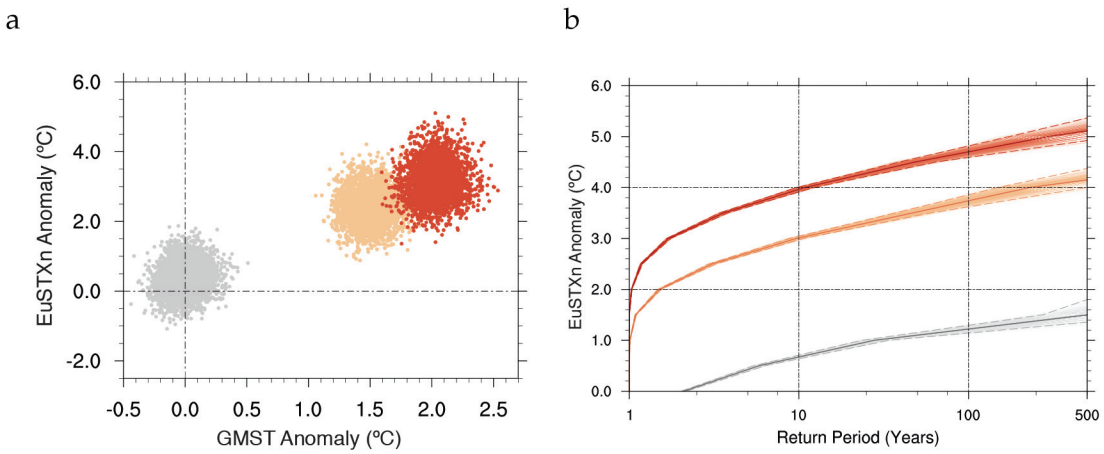


Figure S.4: European summer minimum daily minimum temperature (EuSTXn) at different warming levels. **(a)** EuSTXn anomalies against GMST anomalies for pre-industrial conditions (gray), and for global warming levels of 1.5°C (orange) and 2°C (red) above pre-industrial conditions, simulated by the MPI-ESM Grand ensemble. **(b)** Return levels of EuSTXn summer block minima against their return period, represented by the thick solid lines in gray for pre-industrial conditions, in orange for global warming levels of 1.5°C and in red for 2°C above pre-industrial conditions. Uncertainty in these return levels is estimated by bootstrap-resampling with replacement. The colored thin lines represent 1000 individual bootstrap estimates; the colored dashed lines represent the 95% confidence intervals.

Supporting Figures to Chapter 3

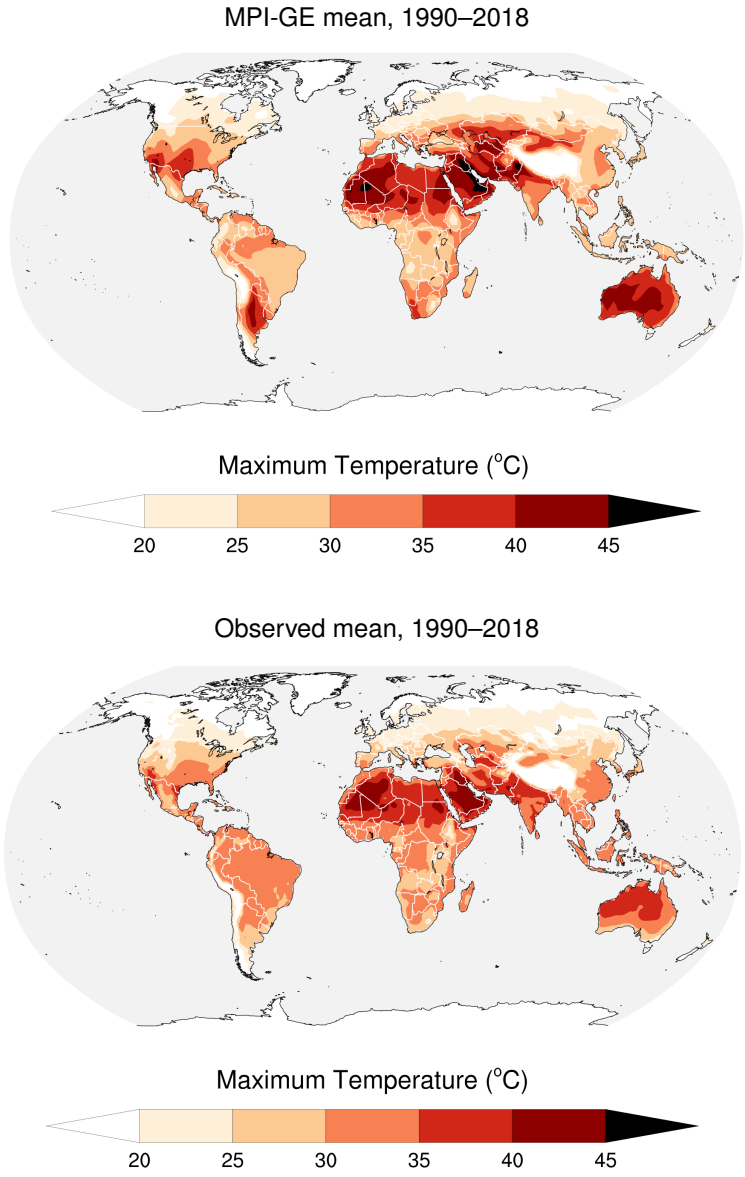


Figure S.5: Maximum temperatures in MPI-GE vs. observations. Absolute summertime monthly maximum temperatures averaged for the period 1990-2018 for MPI-GE simulations compared to observed maximum temperatures in the BEST dataset (Rohde et al., 2012). The observed estimates represent the maximum value of the spatial average of maximum temperature anomaly plus the climatology for the respective month in each grid cell for the period 1951–1980. MPI-GE simulations are historical runs for the period 1990-2005 and RCP4.5 for the period 2005-2018.

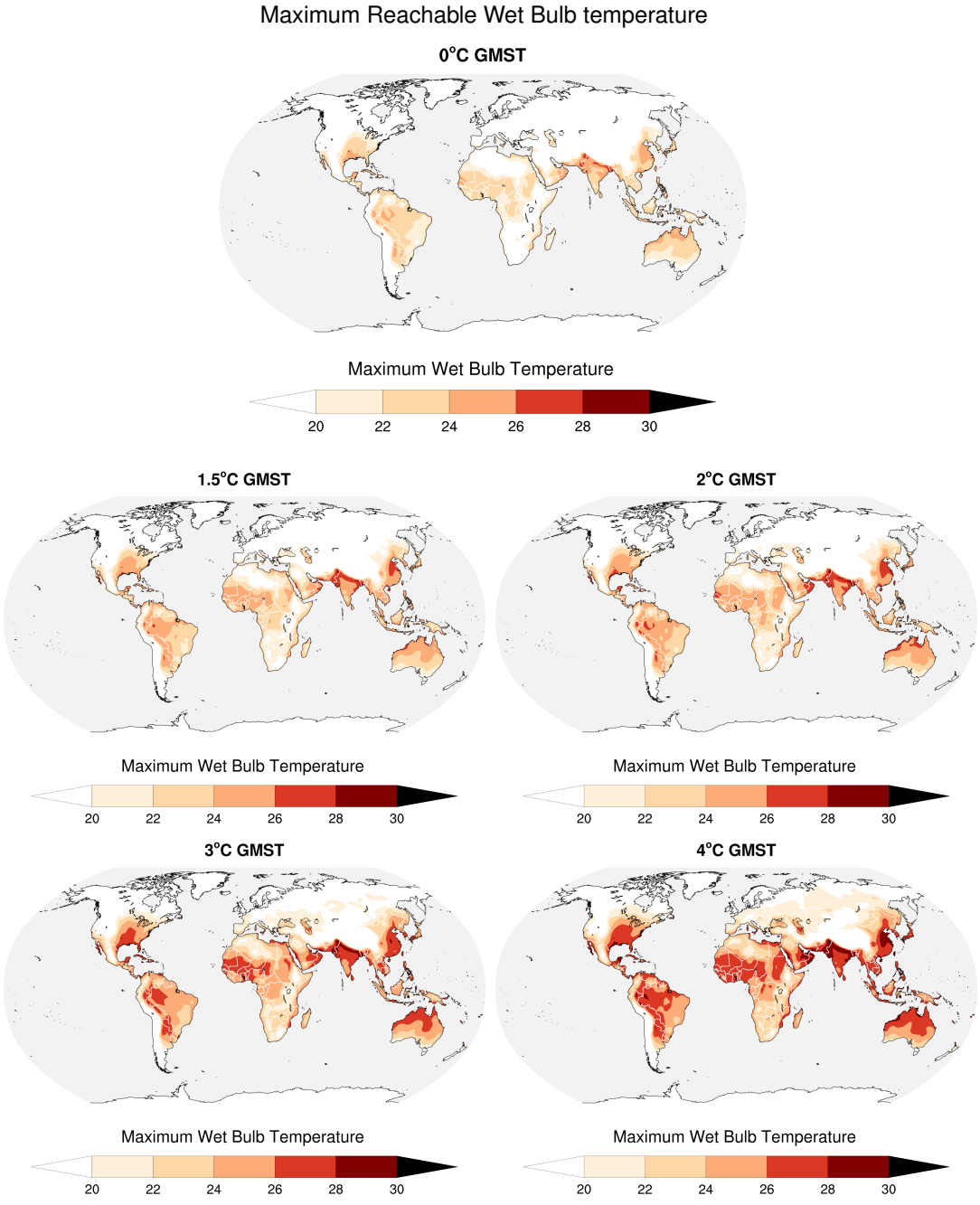


Figure S.6: Uncorrected maximum reachable extreme Wet Bulb temperatures at different global warming levels. Maximum uncorrected Wet Bulb temperatures at different global warming levels,

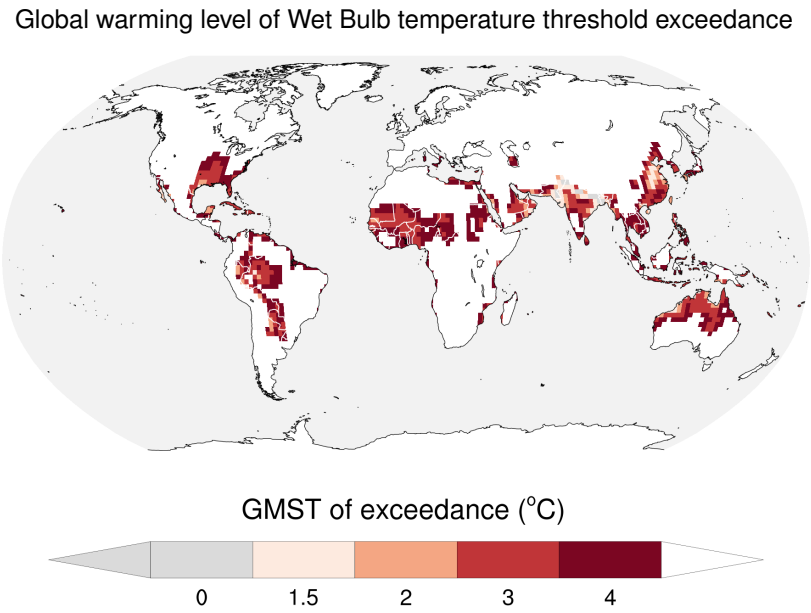


Figure S.7: Global warming level of exceedance of extreme uncorrected Wet Bulb temperature threshold. Global warming level measured as GMST that exhibits exceedance of the uncorrected wet bulb temperature threshold (W > 26°C) with probability of 1% or higher.

BIBLIOGRAPHY

Ballester, J., Giorgi, F. and Rodó, X. (2009). Changes in European temperature extremes can be predicted from changes in PDF central statistics. *Climatic Change* 98.1, p. 277. issn: 1573-1480. doi: 10.1007/s10584-009-9758-0.

Ballester, J., Rodó, X. and Giorgi, F. (2010). Future changes in Central Europe heat waves expected to mostly follow summer mean warming. *Climate Dynamics* 35, pp. 1191–1205. doi: 10.1007/s00382-009-0641-5.

Barriopedro, D., Fischer, E. M., Luterbacher, J., Trigo, R. M. and García-Herrera, R. (2011). The Hot Summer of 2010: Redrawing the Temperature Record Map of Europe. *Science* 332.6026, pp. 220–224. issn: 0036-8075. doi: 10.1126/science.1201224.

Basu, R. and Samet, J. M. (2002). Relation between Elevated Ambient Temperature and Mortality: A Review of the Epidemiologic Evidence. *Epidemiologic Reviews* 24.2, pp. 190–202. doi: 10.1093/epirev/mxf007.

Bathiany, S., Dakos, V., Scheffer, M. and Lenton, T. M. (2018). Climate models predict increasing temperature variability in poor countries. *Science Advances* 4.5. doi: 10.1126/sciadv.aar5809.

Bittner, M., Schmidt, H., Timmreck, C. and Sienz, F. (2016). Using a large ensemble of simulations to assess the Northern Hemisphere stratospheric dynamical response to tropical volcanic eruptions and its uncertainty. *Geophys. Res. Lett.* 43, pp. 9324–9332. doi: 10.1002/2016GL070587.

Bowman, D. M. J. S., Williamson, G. J., Abatzoglou, J. T., Kolden, C. A., Cochrane, M. A. and Smith, A. M. S. (2017). Human exposure and sensitivity to globally extreme wildfire events. *Nature Ecology &Amp; Evolution* 1. doi: 10.1038/s41559-016-0058.

Briffa, K. R., Schrier, G. van der and Jones, P. D. (2009). Wet and dry summers in Europe since 1750: evidence of increasing drought. *International Journal of Climatology* 29.13, pp. 1894–1905. doi: 10.1002/joc.1836.

Bunzel, F., Müller, W. A., Dobrynin, M., Fröhlich, K., Hagemann, S., Pohlmann, H., Stacke, T. and Baehr, J. (2017). Improved seasonal prediction of European summer temperatures with new five-layer soil-hydrology scheme. *Geophys. Res. Lett* 45, pp. 346–353. doi: 10.1002/2017GL076204.

Buzan, J. R., Oleson, K. and Huber, M. (2015). Implementation and comparison of a suite of heat stress metrics within the Community Land Model version 4.5. *Geoscientific Model Development* 8.2, pp. 151–170. doi: 10.5194/gmd-8-151-2015.

Cattiaux, J., Douville, H. and Peings, Y. (2013). European temperatures in CMIP5: origins of present-day biases and future uncertainties. *Climate Dynamics* 41.11, pp. 2889–2907. doi: 10.1007/s00382-013-1731-y.

Cattiaux, J., Douville, H., Schoetter, R., Parey, S. and Yiou, P. (2015). Projected increase in diurnal and interdiurnal variations of European summer temperatures. *Geophys. Res. Lett.* 42.3, pp. 899–907. doi: 10.1002/2014GL062531.

Christidis, N., Jones, G. S. and Stott, P. A. (2014). Dramatically increasing chance of extremely hot summers since the 2003 European heatwave. *Nature Clim. Ch.* 5, pp. 46–50. doi: 10.1038/NCLIMATE2468.

Coffel, E. D., Horton, R. M. and Sherbinin, A. de (2018). Temperature and humidity based projections of a rapid rise in global heat stress exposure during the 21st century. *Environmental Research Letters* 13.1, p. 014001. doi: 10.1088/1748-9326/aaa00e.

Davini, P. and D'Andrea, F. (2016). Northern Hemisphere Atmospheric Blocking Representation in Global Climate Models: Twenty Years of Improvements? *Journal of Climate* 29.24, pp. 8823–8840. doi: 10.1175/JCLI-D-16-0242.1.

De Blois, J., Kjellstrom, T., Agewall, S., Ezekowitz, J. A., Armstrong, P. W. and Atar, D. (2015). The Effects of Climate Change on Cardiac Health. *Cardiology* 131.4, pp. 209–217. doi: 10.1159/000398787.

Della-Marta, P. M., Luterbacher, J., Weissenfluh, H. von, Xoplaki, E., Brunet, M. and Wanner, H. (2007). Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability. *Climate Dynamics* 29.2, pp. 251–275. issn: 1432-0894. doi: 10.1007/s00382-007-0233-1.

Deser, C., Phillips, A., Bourdette, V. and Teng, H. (2012). Uncertainty in climate change projections: the role of internal variability. *Clim. Dyn.* 38, pp. 527–546. doi: 10.1007/s00382-010-0977-x.

Diffenbaugh, N. S. and Ashfaq, M. (2010). Intensification of hot extremes in the United States. *Geophysical Research Letters* 37.15. doi: 10.1029/2010GL043888.

Diffenbaugh, N. S., Pal, J. S., Giorgi, F. and Gao, X. (2007). Heat stress intensification in the Mediterranean climate change hotspot. *Geophys. Res. Lett.* 34.11. doi: 10.1029/2007GL030000.

Donat, M. G., Pitman, A. J. and Seneviratne, S. I. (2017). Regional warming of hot extremes accelerated by surface energy fluxes. *Geophysical Research Letters* 44.13, pp. 7011–7019. doi: 10.1002/2017GL073733.

Dunne, J. P., Stouffer, R. J. and John, J. G. (2013). Reductions in labour capacity from heat stress under climate warming. *Nature Climate Change* 3. doi: 10.1038/nclimate1827.

Fischer, E. M., Beyerle, U., Schleussner, C. F., King, A. D. and Knutti, R. (2018). Biased Estimates of Changes in Climate Extremes From Prescribed SST Simulations. *Geophysical Research Letters* 45.16, pp. 8500–8509. doi: 10.1029/2018GL079176.

Fischer, E. M. and Knutti, R. (2013). Robust projections of combined humidity and temperature extremes. *Nature Climate Change* 3. doi: 10.1038/nclimate1682.

Fischer, E. M., Rajczak, J. and Schär, C. (2012). Changes in European summer temperature variability revisited. *Geophysical Research Letters* 39.19. doi: 10.1029/2012GL052730. url: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012GL052730.

Fischer, E. M. and Schär, C. (2009). Future changes in daily summer temperature variability: driving processes and role for temperature extremes. *Climate Dynamics* 33.7, p. 917. doi: 10.1007/s00382-008-0473-8.

Fischer, E. M., Seneviratne, S. I., Vidale, P. L., Lüthi, D. and Schär, C. (2007). Soil Moisture–Atmosphere Interactions during the 2003 European Summer Heat Wave. *J. Clim.* 20.20, pp. 5081–5099. doi: 10.1175/JCLI4288.1.

Flato, G. J., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S. C., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V. and coauthors (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*, pp. 741–866. doi: 10.1017/CBO9781107415324.020.

Frankignoul, C., Gastineau, G. and Kwon, Y.-O. (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.

Fyfe, J. C., Derksen, C., Mudryk, L., Flato, G. M., Santer, B. D., Swart, N. C., Molotch, N. P., Wan, H., Arore, V. K., Scinocca, J. and Jiao, Y. (2017). Large near-term projected snowpack loss over the western United States. *Nature Comms.* 8. 14996. doi: 10.1038/ncomms14996.

Ghosh, R., Müller, W. A., Baehr, J. and Bader, J. (2017). Impact of observed North Atlantic multidecadal variations to European summer climate: A linear baroclinic response to surface heating. *Climate Dynamics* 48, pp. 3547–3563. doi: 10.1007/s00382-016-3283-4.

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Bottinger, M., Brovkin, V., Crueger, T., Esch, M., Fieg, K. and coauthors (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *JAMES* 5, pp. 572–597. doi: 10.1002/jame.20038.

Gourdji, S. M., Sibley, A. M. and Lobell, D. B. (2013). Global crop exposure to critical high temperatures in the reproductive period: historical trends and future projections. *Environmental Research Letters* 8.2, p. 024041. doi: 10.1088/1748-9326/8/2/024041.

Gregory, J. M., Andrews, T. and Good, P. (2015). The inconstancy of the transient climate response parameter under increasing CO₂. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 373.2054, p. 20140417. doi: 10.1098/rsta.2014.0417.

Hagemann, S. and Stacke, T. (2015). Impact of the soil hydrology scheme on simulated soil moisture memory. *Climate Dynamics* 44.7, pp. 1731–1750. issn: 1432-0894. doi: 10.1007/s00382-014-2221-6.

Hanna, E. G., Kjellstrom, T., Bennett, C. and Dear, K. (2011). Climate Change and Rising Heat: Population Health Implications for Working People in Australia. *Asia Pacific Journal of Public Health* 23.2_suppl, 14S–26S. doi: 10.1177/1010539510391457.

Hawkins, E., Ortega, P., Suckling, E., Schurer, A., Hegerl, G., Jones, P., Joshi, M., Osborn, T., Masson-Delmotte, V., Mignot, J., Thorne, P. and Oldenborgh, G. van (2017). Estimating Changes in Global Temperature since the Preindustrial Period. *BAMS* 98.9, pp. 1841–1856. doi: 10.1175/BAMS-D-16-0007.1.

Hawkins, E., Smith, R. S., Gregory, J. M. and Stainforth, D. A. (2016). Irreducible uncertainty in near-term climate projections. *Clim. Dyn.* 46, pp. 3807–3819. doi: 10.1007/s00382-015-2806-8.

Hawkins, E. and Sutton, R. T. (2012). Time of emergence of climate signals. *Geophysical Research Letters* 39.1. doi: 10.1029/2011GL050087.

Hedemann, C., Mauritsen, T., Jungclaus, J. and Marotzke, J. (2017). The subtle origins of surface-warming hiatuses. *Nature Clim. Ch.* 7, pp. 336–339. doi: 10.1038/nclimate3274.

Holmes, C. R., Woolings, T., Hawkins, E. and Vries, H. de (2016). Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *J. Clim.* doi: 10.1175/JCLI-D-14-00735.1.

Horton, D. E., Johnson, N. C., Singh, D., Swain, D. L., Rajaratnam, B. and Diffenbaugh, N. S. (2015). Contribution of changes in atmospheric circulation patterns to extreme temperature trends. *Nature* 522, pp. 465–469. doi: 10.1038/nature14550.

Horton, R. M., Mankin, J. S., Lesk, C., Coffel, E. and Raymond, C. (2016). A Review of Recent Advances in Research on Extreme Heat Events. *Current Clim. Ch. Rep.* 2.4, pp. 242–259. issn: 2198-6061. doi: 10.1007/s40641-016-0042-x.

Ilyina, T., Six, K. D., Segschneider, J., Maier-Reimer, E., Li, H. and Núñez-Riboni, I. (2013). Global ocean biogeochemistry model HAMOCC: Model architecture and performance as component of the MPI-Earth system model in different CMIP5 experimental realizations. *JAMES* 5.2, pp. 287–315. doi: 10.1029/2012MS000178.

Im, E.-S., Pal, J. S. and Eltahir, E. A. B. (2017). Deadly heat waves projected in the densely populated agricultural regions of South Asia. *Science Advances* 3.8. doi: 10.1126/sciadv.1603322.

IPCC (2013). Summary for Policymakers Climate Change 2013: The Physical Science Basis. *(IPCC, Cambridge Univ. Press)* (eds Stocker, T. F. et al.) Pp. 571–657.

IPCC (2014). Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. [Field, C.B. and Barros, V.R. and Dokken, D.J. and Mach, K.J. and Mastrandrea, M.D. and Bilir, T.E. and Chatterjee, M. and Ebi, K.L. and Estrada, Y.O. and Genova, R.C. and Girma, B. and Kissel, E.S. and Levy, A.N. and MacCracken, S. and Mastrandrea, P.R. and White, L.L. (eds.)] *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA*, 1132 pp.

Jones, P. D., Lister, D. H., Osborn, T. J., Harpham, C., Salmon, M. and Morice, C. P. (2012). Hemispheric and large-scale land-surface air temperature variations: An extensive revision and an update to 2010. *J. Geophys. Res.: Atmos.* 117. D05127. doi: 10.1029/2011JD017139.

Jungclaus, J. H., Fischer, N., Haak, H., Lohmann, K., Marotzke, J., Matei, D., Mikolajewicz, U., Notz, D. and Storch, J. S. von (2013). Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *JAMES* 5.2, pp. 422–446. issn: 1942-2466. doi: 10.1002/jame.20023.

Kay, J. E., Deser, C., Phillips, A., Mai, A., Hannay, C., Strand, G., Arblaster, J. M., Bates, S. C., Danabasoglu, G., Edwards, J., Holland, M., Kushner, P., Lamarque, J.-F., Lawrence, D., Lindsay, K., Middleton, A., Munoz, E., Neale, R., Oleson, K., Polvani, L. and Vertenstein, M. (2015). The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *BAMS* 96.8, pp. 1333–1349. doi: 10.1175/BAMS-D-13-00255.1.

King, A. D. and Karoly, D. (2017). Climate extremes in Europe at 1.5 and 2 degrees of global warming. *Env. Res. Lett.* 114031. doi: 10.1088/1748-9326/aa8e2c.

Kirchmeier-Young, M. C., Zwiers, F. W. and Gillett, N. P. (2017). Attribution of Extreme Events in Arctic Sea Ice Extent. *Journal of Climate* 30.2, pp. 553–571. doi: 10.1175/JCLI-D-16-0412.1.

Kjellstrom, T., Butler, A. J., Lucas, R. M. and Bonita, R. (2010). Public health impact of global heating due to climate change: potential effects on chronic non-communicable diseases.

International Journal of Public Health 55.2, pp. 97–103. doi: 10.1007/s00038-009-0090-2.

Koster, R. D., Guo, Z., Yang, R., Dirmeyer, P. A., Mitchell, K. and Puma, M. J. (2009). On the Nature of Soil Moisture in Land Surface Models. *J. Clim.* 22.16, pp. 4322–4335. doi: 10.1175/2009JCLI2832.1.

Krueger, O., Hegerl, G. C. and Tett, S. F. B. (2015). Evaluation of mechanisms of hot and cold days in climate models over Central Europe. *Environ. Res. Lett.* 10.1. doi: 10.1088/1748-9326/10/1/014002.

Laaidi, K., Zeghnoun, A., Dousset, B., Bretin, P., Vandentorren, S., Giraudet, E. and Beaudeau, P. (2012). The Impact of Heat Islands on Mortality in Paris during the August 2003 Heat Wave. *Environmental Health Perspectives* 120.2, pp. 254–259. doi: 10.1289/ehp.1103532.

Li, J., Chen, Y. D., Gan, T. Y. and Lau, N.-C. (2018). Elevated increases in human-perceived temperature under climate warming. *Nature Climate Change* 8.1, pp. 43–47. doi: 10.1038/s41558-017-0036-2. url: https://doi.org/10.1038/s41558-017-0036-2.

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

Maher, N., Matei, D., Milinski, S. and Marotzke, J. (2018). ENSO Change in Climate Projections: Forced Response or Internal Variability? *Geophysical Research Letters* 45.20, pp. 11, 390–11, 398. doi: 10.1029/2018GL079764.

Maher, N., Milinski, S., Suarez-Gutierrez, L., Botzet, M., Dobrynin, M., Kornblueh, L., Kröger, J., Takano, Y., Ghosh, R., Hedemann, C., Li, C., Li, H., Manzini, E., Notz, D., Putrasahan, D., Boysen, L., Claussen, M., Ilyina, T., Olonscheck, D., Raddatz, T., Stevens, B. and Marotzke, J. (2019). The Max Planck Institute Grand Ensemble: Enabling the Exploration of Climate System Variability. *JAMES* 11.7, pp. 2050–2069. doi: 10.1029/2019MS001639.

Mann, M. E., Rahmstorf, S., Kornhuber, K., Steinman, B. A., Miller, S. K., Petri, S. and Coumou, D. (2018). Projected changes in persistent extreme summer weather events: The role of quasi-resonant amplification. *Science Advances* 4.10. doi: 10.1126/sciadv.aat3272.

Marotzke, J. (2019). Quantifying the irreducible uncertainty in near-term climate projections. *Wiley Interdisciplinary Reviews: Climate Change* 10.1, e563. doi: 10.1002/wcc.563.

Matsueda, M. and Endo, H. (2017). The robustness of future changes in Northern Hemisphere blocking: A large ensemble projection with multiple sea surface temperature patterns. *Geophysical Research Letters* 44.10, pp. 5158–5166. doi: 10.1002/2017GL073336.

Matthews, T. K. R., Wilby, R. L. and Murphy, C. (2017). Communicating the deadly consequences of global warming for human heat stress. *Proceedings of the National Academy of Sciences*. issn: 0027-8424. doi: 10.1073/pnas.1617526114.

Mauritsen, T., Bader, J., Becker, T., Behrens, J., Bittner, M., Brokopf, R., Brovkin, V., Claussen, M., Crueger, T., Esch, M., Fast, I., Fiedler, S., Fläschner, D., Gayler, V., Giorgetta, M., Goll, D. S., Haak, H., Hagemann, S., Hedemann, C., Hohenegger, C., Ilyina, T., Jahns, T., Jimenez de la Cuesta Otero, D., Jungclaus, J., Kleinen, T., Kloster, S., Kracher, D., Kinne, S., Kleberg, D., Lasslop, G., Kornblueh, L., Marotzke, J., Matei, D., Meraner, K., Mikolajewicz, U., Modali, K., Möbis, B., Müller, W. A., Nabel, J. E. M. S., Nam, C. C. W., Notz, D., Nyawira, S.-S., Paulsen, H., Peters, K., Pincus, R., Pohlmann, H., Pongratz, J., Popp, M., Raddatz, T., Rast, S., Redler, R., Reick, C. H., Rohrschneider, T., Schemann, V., Schmidt, H., Schnur, R., Schulzweida, U., Six, K. D., Stein, L., Stemmler, I., Stevens, B., Storch, J.-S. von, Tian, F., Voigt, A., Vrese, P. de, Wieners, K.-H., Wilkenskjeld, S., Winkler, A. and Roeckner, E. (2019). Developments in the MPI-M Earth System Model version 1.2 (MPI-ESM 1.2) and its response to increasing CO2. *JAMES*. doi: 10.1029/2018MS001400.

Meehl, G. A. and Tebaldi, C. (2004). More Intense, More Frequent, and Longer Lasting Heat Waves in the 21st Century. *Science* 305.5686, pp. 994–997. doi: 10.1126/science.1098704.

Merrifield, A., Lehner, F., Xie, S.-.-.-P. and Deser, C. (2017). Removing Circulation Effects to Assess Central U.S. Land-Atmosphere Interactions in the CESM Large Ensemble. *Geophys. Res. Lett.* 44.19. doi: 10.1002/2017GL074831.

Miralles, D. G., Teuling, A. J., Heerwaarden, C. C. van and Vila-Guerau de Arellano, J. (2014). Mega-heatwave temperatures due to combined soil desiccation and atmospheric heat accumulation. *Nature geosci.* 7, pp. 345–349. doi: https://doi.org/10.1038/ngeo2141.

Mishra, V., Mukherjee, S., Kumar, R. and Stone, D. A. (2017). Heat wave exposure in India in current, 1.5°C, and 2.0°C worlds. *Environmental Research Letters* 12.12, p. 124012. doi: 10.1088/1748-9326/aa9388.

Mitchell, D., AchutaRao, K., Allen, M., Bethke, I., Ciavarella, A., Forster, P., Fuglestvedt, J., Gillett, N., Haustein, K., Ingram, W., Iversen, T., Kharin, V., Klingaman, V., Massey, N., Fischer, E., Schleussner, C.-F., Scinocca, J., Seland, Ø., Shiogama, H., Shuckburgh, E., Sparrow, S., Stone, D., Uhe, P., Urs, B., Wallom, D., Wehner, M. and Zaaboul, R. (2017). Half a degree additional warming, prognosis and projected impacts (HAPPI): background and experimental design. *Geosc. Model Dev.* 10.2, pp. 571–583. doi: 10.5194/gmd-10-571-2017.

Mora, C., Dousset, B., Caldwell, I. R., Powell, F. E., Geronimo, R. C., Bielecki, C. R., Counsell, C. W. W., Dietrich, B. S., Johnston, E. T., Louis, L. V., Lucas, M. P., McKenzie, M. M., Shea, A. G., Tseng, H., Giambelluca, T. W., Leon, L. R., Hawkins, E. and Trauernicht, C. (June 2017). Global risk of deadly heat. *Nature Climate Change* 7, 501 EP. doi: 10.1038/nclimate3322.

Morice, C. P., Kennedy, J. J., Rayner, N. A. and Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The hadCRUT4 data set. *J. Geophys. Res.: Atmos.* 117. D08101. doi: 10.1029/2011JD017187.

Müller, W. A., Jungclaus, J. H., Mauritsen, T., Baehr, J., Bittner, M., Budich, R., Bunzel, F., Esch, M., Ghosh, R., Haak, H., Ilyina, T., Kleine, T., Kornblueh, L., Li, H., Modali, K., Notz, D., Pohlmann, H., Roeckner, E., Stemmler, I., Tian, F. and Marotzke, J. (2018). A Higher-resolution Version of the Max Planck Institute Earth System Model (MPI-ESM1.2-HR). *Journal of Advances in Modeling Earth Systems* 10.7, pp. 1383–1413. doi: 10.1029/2017MS001217.

Murage, P., Hajat, S. and Kovats, R. S. (2017). Effect of night-time temperatures on cause and age-specific mortality in London. *Environmental Epidemiology* 1.2. doi: 10.1097/EE9.0000000000000005.

Newth, D. and Gunasekera, D. (2018). Projected Changes in Wet-Bulb Globe Temperature under Alternative Climate Scenarios. *Atmosphere* 9.5, p. 187. issn: 2073-4433. doi: 10.3390/atmos9050187.

O'brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity* 41.5, pp. 673–690. issn: 1573-7845. doi: 10.1007/s11135-006-9018-6.

Osborn, T. J. and Jones, P. D. (2014). The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth Sys. Sci. Data* 6, pp. 61–68. doi: 10.5194/essd-6-61-2014.

Pal, J. S. and Eltahir, E. A. B. (2015). Future temperature in southwest Asia projected to exceed a threshold for human adaptability. *Nature Climate Change* 6. doi: 10.1038/nclimate2833.

Perkins-Kirkpatrick, S. E. and Gibson, P. B. (2017). Changes in regional heatwave characteristics as a function of increasing global temperature. *Scientific Reports* 7. 12256. doi: doi:10.1038/s41598-017-12520-2.

Pfahl, S. (2014). Characterising the relationship between weather extremes in Europe and synoptic circulation features. *Natural Hazards and Earth System Sciences* 14.6, pp. 1461–1475. doi: 10.5194/nhess-14-1461-2014.

Raddatz, T., Reick, C., Knorr, W., Kattge, J., Roeckner, E., Schnur, R., Schnitzler, K.-G., Wetzel, P. and Jungclaus, J. (2007). Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century? *Clim. Dyn.* 29. doi: 10.1007/s00382-007-0247-8.

Robine, J.-M., Cheung, S. L. K., Roy, S. L., Oyen, H. V., Griffiths, C., Michel, J.-P. and Herrmann, F. R. (2008). Death toll exceeded 70,000 in Europe during the summer of 2003. *Comptes Rendus Biologies* 331.2, pp. 171–178. issn: 1631-0691. doi: 10.1016/j.crvi.2007.12.001.

Rodgers, K. B., Lin, J. and Frölicher, T. L. (2015). Emergence of multiple ocean ecosystem drivers in a large ensemble suite with an Earth system model. *Biogeosc.* 12.11, pp. 3301–3320. doi: 10.5194/bg-12-3301-2015.

Roeckner, E., Bäum, G., Bonaventura, L., Brokopf, R., Esch, M., Giorgetta, M., Hagemann, S., Kornblueh, L., Schlese, U., Schulzweida, U., Kirchner, I., Manzini, E., Rhodin, A. and Tompkins, A. (2003). The atmospheric general circulation model ECHAM 5 PART I: Model description. *Report/ MPI für Meteorologie* 349.

Rohde, R., Muller, R., Jacobsen, R., Muller, E., Perlmutter, S., Rosenfeld, A., Wurtele, J., Groom, D. and Wickham, C. (2012). A New Estimate of the Average Earth Surface Land Temperature Spanning 1753 to 2011. *Geoinfor Geostat: An Overview 1:1*. doi: 10.4172/2327-4581.1000101.

Royé, D. (2017). The effects of hot nights on mortality in Barcelona, Spain. *International Journal of Biometeorology* 61.12, pp. 2127–2140. issn: 1432-1254. doi: 10.1007/s00484-017-1416-z.

Ruosteenoja, K., Markkanen, T., Venäläinen, A., Räisänen, P. and Peltola, H. (2018). Seasonal soil moisture and drought occurrence in Europe in CMIP5 projections for the 21st century. *Clim. Dyn.* 50.3, pp. 1177–1192. doi: 10.1007/s00382-017-3671-4.

Russo, S., Dosio, A., Graversen, R. G., Sillmann, J., Carrao, H., Dunbar, M. B., Singleton, A., Montagna, P., Barbola, P. and Vogt, J. V. (2014). Magnitude of extreme heat waves in present climate and their projection in a warming world. *Journal of Geophysical Research: Atmospheres* 119.22, pp. 12, 500–12, 512. doi: 10.1002/2014JD022098.

Russo, S., Sillmann, J. and Sterl, A. (2017). Humid heat waves at different warming levels. *Scientific Reports* 7.1, p. 7477. doi: 10.1038/s41598-017-07536-7.

Russo, S. and Sterl, A. (2011). Global changes in indices describing moderate temperature extremes from the daily output of a climate model. *Journal of Geophysical Research: Atmospheres* 116. doi: 10.1029/2010JD014727.

Sanderson, B. M., Xu, Y., Tebaldi, C., Wehner, M., O'Neil, B., Jahn, A., Pendergrass, A. G., Lehner, F., Strand, W. G., Lin, L., Knutti, R. and Lamarque, J. F. (2017). Community Climate Simulations to assess avoided impacts in 1.5°C and 2°C futures. *Earth Syst. Dynam. Discuss.* Pp. 827–847. doi: 10.5194/esd-2017-42.

Schaller, N., Sillmann, J., Anstey, J., Fischer, E. M., Grams, C. M. and Russo, S. (2018). Influence of blocking on Northern European and Western Russian heatwaves in large climate model

ensembles. *Environmental Research Letters* 13.5, p. 054015. doi: 10.1088/1748-9326/aaba55.

Schär, C., Virale, P. L., Lüthi, D., Frei, C., Häberli, C., Liniger, M. A. and Appenzeller, C. (2004). The role of increasing temperature variability in European summer heatwaves. *Nature* 427, pp. 332–336. doi: 10.1038/nature02300.

Schleussner, C.-F., Lissner, T. K., Fischer, E. M., Wohland, J., Perrette, M., Golly, A., Rogelj, J., Childers, K., Schewe, J., Frieler, K., Mengel, M., Hare, W. and Schaeffer, M. (2016). Differential climate impacts for policy-relevant limits to global warming: the case of 1.5°C and 2°C. *Earth Syst. Dyn.* 7.2, pp. 327–351. doi: 10.5194/esd-7-327-2016.

Seneviratne, S. I., Lüthi, D., Litschi, M. and Schär, C. (2006). Land–atmosphere coupling and climate change in Europe. *Nature* 443.7108, pp. 205–209. doi: 10.1038/nature05095.

Sherwood, S. C. and Huber, M. (2010). An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences* 107.21, pp. 9552–9555. doi: 10.1073/pnas.0913352107.

Sillmann, J., Thorarinsdottir, T., Keenlyside, N., Schaller, N., Alexander, L. V., Hegerl, G., Seneviratne, S., Vautard, R., Zhang, X. and Zwiers, F. W. (2017). Understanding, modeling and predicting weather and climate extremes: Challenges and opportunities. *Weather and Climate Extremes* 18, pp. 65–74. issn: 2212–0947. doi: https://doi.org/10.1016/j.wace.2017.10.003.

Spinoni, J., Vogt, J. V., Naumann, G., Barbosa, P. and Dosio, A. (2018). Will drought events become more frequent and severe in Europe? *International Journal of Climatology* 38.4, pp. 1718–1736. doi: 10.1002/joc.5291.

Sriver, R. L., Forest, C. E. and Keller, K. (2015). Effects of initial conditions uncertainty on regional climate variability: An analysis using a low-resolution CESM ensemble. *Geophys. Res. Lett.* 42, pp. 5468–5476. doi: 10.1002/2015GL064546.

Storch, H. v. and Zwiers, F. W. (1999). *Statistical Analysis in Climate Research*. Cambridge University Press. doi: 10.1017/CBO9780511612336.

Stott, P. A., Stone, D. A. and Allen, R. M. (2004). Human contribution to the European heatwave of 2003. *Nature* 432, pp. 610–614. doi: 10.1038/nature03089.

Stull, R. (2011). Wet-Bulb Temperature from Relative Humidity and Air Temperature. *Journal of Applied Meteorology and Climatology* 50.11, pp. 2267–2269. doi: 10.1175/JAMC-D-11-0143.1.

Suarez-Gutierrez, L., Li, C., Müller, W. A. and Marotzke, J. (2018). Internal variability in European summer temperatures at 1.5°C and 2°C of global warming. *Environ. Res. Lett.* 44, pp. 5709–5719. doi: 10.1002/2017GL073798.

Suarez-Gutierrez, L., Li, C., Müller, W. A. and Marotzke, J. (2019a). Drivers of internal variability in European summer Heat Extremes. *To be submitted*.

Suarez-Gutierrez, L., Li, C., Müller, W. A. and Marotzke, J. (2019b). Heat Hotspots under Global Warming. *In Preparation*.

Suarez-Gutierrez, L., Li, C., Thorne, P. W. and Marotzke, J. (2017). Internal variability in simulated and observed tropical tropospheric temperature trends. *Geophys. Res. Lett.* 44, pp. 5709–5719. doi: 10.1002/2017GL073798.

Sutton, R. T. and Hodson, D. L. R. (2005). Atlantic Ocean Forcing of North American and European Summer Climate. *Science* 309, pp. 115–118. doi: 10.1126/science.1109496.

Taylor, K. E., Stouffer, R. J. and Meehl, G. A. (2012). An Overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.* 93, pp. 485–498. doi: 10.1175/BAMS-D-11-00094.1.

van Vuuren, D. P., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G. C., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S. J. and Rose, S. K. (2011). The representative concentration pathways: an overview. *Climatic Change* 109.1, p. 5. issn: 1573-1480. doi: 10.1007/s10584-011-0148-z.

Vautard, R., Gobiet, A., Sobolowski, S., Kjellström, E., Stegehuis, A., Watkiss, P., Mendlik, T., Landgren, O., Nikulin, G., Teichmann, C. and Jacob, D. (2014). The European climate under a 2°C global warming. *Env. Res. Lett.* 9.034006, 11pp. doi: 10.1088/1748-9326/9/3/034006.

Vogel, M. M., Orth, R., Cheruy, F., Hagemann, S., Lorenz, R., Hurk, B. J. J. M. van den and Seneviratne, S. I. (2017). Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophysical Research Letters* 44.3, pp. 1511–1519. issn: 1944-8007. doi: 10.1002/2016GL071235.

Wehner, M., Stone, D., Mitchell, D., Shiogama, H., Fischer, E., Graff, L. S., Kharin, V. V., Lierhammer, L., Sanderson, B. and Krishnan, H. (2017). Changes in extremely hot days under stabilized 1.5°C and 2.0°C global warming scenarios as simulated by the HAPPI multi-model ensemble. *Earth Syst. Dynam. Discuss.* Pp. 187–195. doi: 10.5194/esd-2017-89.

Wehner, M., Stone, D., Krishnan, H., AchutaRao, K. and Castillo, F. (2016). The Deadly Combination of Heat and Humidity in India and Pakistan in Summer 2015. *Bulletin of the American Meteorological Society* 97.12, S81–S86. doi: 10.1175/BAMS-D-16-0145.1.

Willett, K. M. and Sherwood, S. C. (2012). Exceedance of heat index thresholds for 15 regions under a warming climate using the wet-bulb globe temperature. *International Journal of Climatology* 32.2, pp. 161–177. doi: 10.1002/joc.2257.

Woollings, T., Barriopedro, D., Methven, J., Son, S.-W., Martius, O., Harvey, B., Sillmann, J., Lupo, A. R. and Seneviratne, S. (2018). Blocking and its Response to Climate Change. *Current Clim. Ch. Rep.* 4.3, pp. 287–300. issn: 2198-6061. doi: 10.1007/s40641-018-0108-z.

Zampieri, M., D'Andrea, F., Vautard, R., Ciais, P., Noblet-Ducoudré, N. de and Yiou, P. (2009). Hot European Summers and the Role of Soil Moisture in the Propagation of Mediterranean Drought. *Journal of Climate* 22.18, pp. 4747–4758. doi: 10.1175/2009JCLI2568.1.

Versicherung an Eides statt

Declaration of oath

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel: „Internal variability in summertime heat extremes under global warming“ selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, den 26.04.2019

_____

Laura Suárez Gutiérrez

Hinweis / Reference

Die gesamten Veröffentlichungen in der Publikationsreihe des MPI-M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199

sind über die Internetseiten des Max-Planck-Instituts für Meteorologie erhältlich:
http://www.mpimet.mpg.de/wissenschaft/publikationen.html

All the publications in the series of the MPI -M
„Berichte zur Erdsystemforschung / Reports on Earth System Science“,
ISSN 1614-1199

are available on the website of the Max Planck Institute for Meteorology:
http://www.mpimet.mpg.de/wissenschaft/publikationen.html