

The disproportionate rates of change between extreme and mean temperatures over land

Author:

Gross, Mia

Publication Date:

2019

DOI:

<https://doi.org/10.26190/unsworks/21291>

License:

<https://creativecommons.org/licenses/by-nc-nd/3.0/au/>

Link to license to see what you are allowed to do with this resource.

Downloaded from <http://hdl.handle.net/1959.4/62700> in <https://unsworks.unsw.edu.au> on 2024-05-01

The disproportionate rates of change between extreme and mean temperatures over land

Mia H. Gross

A thesis in fulfilment of the requirements for the degree of
Doctor of Philosophy



Climate Change Research Centre
ARC Centre of Excellence for Climate System Science
Faculty of Science
The University of New South Wales, Australia

June 2019



FOR OFFICE USE ONLY Date of completion of requirements for Award: _____

ORIGINALITY STATEMENT

'I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.'

Signed

Date

COPYRIGHT STATEMENT

'I hereby grant the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all proprietary rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstract International (this is applicable to doctoral theses only).

I have either used no substantial portions of copyright material in my thesis or I have obtained permission to use copyright material; where permission has not been granted I have applied/will apply for a partial restriction of the digital copy of my thesis or dissertation.'

Signed

Date

AUTHENTICITY STATEMENT

'I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis. No emendation of content has occurred and if there are any minor variations in formatting, they are the result of the conversion to digital format.'

Signed

Date

INCLUSION OF PUBLICATIONS STATEMENT

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in their thesis in lieu of a Chapter if:

- The student contributed greater than 50% of the content in the publication and is the “primary author”, ie. the student was responsible primarily for the planning, execution and preparation of the work for publication
- The student has approval to include the publication in their thesis in lieu of a Chapter from their supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis

Please indicate whether this thesis contains published material or not.

☐

This thesis contains no publications, either published or submitted for publication

☐

Some of the work described in this thesis has been published and it has been documented in the relevant Chapters with acknowledgement

☒

This thesis has publications (either published or submitted for publication) incorporated into it in lieu of a chapter and the details are presented below

CANDIDATE'S DECLARATION

I declare that:

- I have complied with the Thesis Examination Procedure
- where I have used a publication in lieu of a Chapter, the listed publication(s) below meet(s) the requirements to be included in the thesis.

Name	Signature	Date (dd/mm/yy)
Mia Gross		

Postgraduate Coordinator's Declaration

I declare that:

- the information below is accurate
- where listed publication(s) have been used in lieu of Chapter(s), their use complies with the Thesis Examination Procedure
- the minimum requirements for the format of the thesis have been met.

PGC's Name	PGC's Signature	Date (dd/mm/yy)

For each publication incorporated into the thesis in lieu of a Chapter, provide all of the requested details and signatures required

Details of publication #1:					
<i>Full title:</i> The sensitivity of daily temperature extremes and variability to dataset choice					
<i>Authors:</i> Gross, M. H., Donat, M. G., Alexander, L. V., Sisson, S. S.					
<i>Journal or book name:</i> Journal of Climate					
<i>Volume/page numbers:</i> 31/1337-1359					
<i>Date accepted/ published:</i> 15 February, 2018					
Status	<i>Published</i>	<input checked="" type="checkbox"/>	<i>Accepted and In press</i>	<input type="checkbox"/>	<i>In progress (submitted)</i>
The Candidate's Contribution to the Work					
I was responsible for almost all the writing and analysis during the preparation and the review process. All authors are accredited with the study concept and manuscript revisions, both before submission and during the review process.					
Location of the work in the thesis and/or how the work is incorporated in the thesis:					
Chapter 2					
Primary Supervisor's Declaration					
I declare that:					
<ul style="list-style-type: none"> the information above is accurate this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form. 					
<i>Supervisor's name</i>		<i>Supervisor's signature</i>		<i>Date (dd/mm/yy)</i>	
Lisa Alexander					

Details of publication #2:					
<i>Full title:</i> Changes in daily temperature extremes relative to the mean in Coupled Model Intercomparison Project phase 5 models and observations					
<i>Authors:</i> Gross, M. H., Donat, M. G., Alexander, L. V.					
<i>Journal or book name:</i> International Journal of Climatology					
<i>Volume/page numbers:</i> N/A					
<i>Date accepted/ published:</i> 10 June 2019					
Status	<i>Published</i>	<input type="checkbox"/>	<i>Accepted and In press</i>	<input checked="" type="checkbox"/>	<i>In progress (submitted)</i>
The Candidate's Contribution to the Work					
I was responsible for all the writing and analysis during the preparation and the review process. All authors are accredited with the study concept and manuscript revisions, both before submission and during the review process.					
Location of the work in the thesis and/or how the work is incorporated in the thesis:					
Chapter 3					
Primary Supervisor's Declaration					
I declare that:					
<ul style="list-style-type: none"> the information above is accurate this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form. 					
<i>Supervisor's name</i>		<i>Supervisor's signature</i>		<i>Date (dd/mm/yy)</i>	
Lisa Alexander					

Details of publication #3:

Full title: Enhanced warming of seasonal cold extremes relative to the mean in the Northern Hemisphere extratropics

Authors: Gross, M. H., Donat, M. G., Alexander, L. V., Sherwood, S.

Journal or book name: Climate Dynamics

Volume/page numbers: N/A

Date accepted/ published: 7 February 2019 (submitted)

Status	<i>Published</i>		<i>Accepted and In press</i>		<i>In progress (submitted)</i>	x
---------------	------------------	--	------------------------------	--	--------------------------------	----------

The Candidate's Contribution to the Work

I was responsible for almost all the writing and analysis during the preparation and submission process. All authors are accredited with the study concept and manuscript.

Location of the work in the thesis and/or how the work is incorporated in the thesis:

Chapter 4

Primary Supervisor's Declaration

I declare that:

- the information above is accurate
- this has been discussed with the PGC and it is agreed that this publication can be included in this thesis in lieu of a Chapter
- All of the co-authors of the publication have reviewed the above information and have agreed to its veracity by signing a 'Co-Author Authorisation' form.

<i>Supervisor's name</i> Lisa Alexander	<i>Supervisor's signature</i>	<i>Date (dd/mm/yy)</i>
--	-------------------------------	------------------------

Abstract

The world is warming, with both extreme and mean temperatures getting warmer. But how extremes are changing and will change relative to the mean remains less clear. Disproportionate rates of change between extreme and mean temperatures need to be better understood because this changes the shape of the distribution, affecting the probability of extreme events and thus their impacts. Therefore, the goal of this thesis is to understand if, when and where disproportionate rates of change occur in the past and future and it does this by using observations, reanalyses and climate models.

I find that for past decades, cold extremes have been warming faster than the mean for much of the Northern Hemisphere extratropics, while warm extremes have been warming faster than the mean in some subtropical regions. Future changes are systematic and robust across a range of climate model simulations. The most striking disproportionate changes are in the Northern Hemisphere mid- to high-latitudes, where cold extremes are projected to warm substantially faster than mean temperatures in all seasons except boreal summer.

Exploring conditions on or leading into the day of the projected cold extremes reveals that the disproportionate warming is driven by different mechanisms in different seasons. In boreal winter, reduced cold air advection is the dominant driver, circulating anomalously warm temperatures from the Arctic to lower

latitudes. But during spring and autumn, it is mostly due to feedbacks related to decreases in snow cover.

Analyses of temperature extremes have inherent uncertainties. I evaluate several commonly used reanalyses with a gridded *in situ*-based daily temperature dataset to assess sensitivities related to dataset choice. Trends in extremes and statistical moments, other than the mean, exhibit sensitivity. However, the conclusions drawn in this thesis remain robust irrespective of dataset choice, and regardless of methodological choice including choice of base period and how extremes are defined.

Ultimately, this thesis provides a comprehensive understanding of changes in daily temperature extremes relative to the mean and their drivers. In turn, this provides essential information for decision-makers who can act to reduce the negative impacts stemming from extreme temperatures.

Preface

The leading author of each of the below manuscripts (MHG) was responsible for all of the analysis, and almost all of the writing during the preparation and the review process (where applicable). All authors are accredited with the study concept and manuscript, both before submission and during the review process.

Gross, M. H., M. G. Donat, L. V. Alexander, and S. A. Sisson, 2018: The sensitivity of daily temperature variability and extremes to dataset choice. *Journal of Climate*, **31**, 1337-1359, doi: 10.1175/JCLI-D-17-0243.1.

Gross, M. H., M. G. Donat and L. V. Alexander, 2019: Changes in global daily temperature extremes relative to the mean in CMIP5 models and observations, *International Journal of Climatology*, accepted.

Gross, M. H., L. V. Alexander, M. G. Donat, and S. Sherwood, 2019: Enhanced warming of cold extremes relative to the mean in the Northern Hemisphere extratropics, *Climate Dynamics*, submitted.

This PhD was supported by the Australian Research Council grant CE110001028, the Australian Government Research Training Program (RTP) scholarship and the Climate Change Research Centre top-up scholarship.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisors, Associate Professor Lisa Alexander, Dr Markus Donat and Professor Steven Sherwood. Their guidance has been invaluable. In particular, Lisa, thank you for your support and encouragement during these past few years. You have truly been a source of inspiration and motivation throughout my PhD and Honours year. Markus, I'd like to thank you for always having a spare minute, or 60, to discuss any problems that came up throughout my PhD, and for your continued support from abroad. Your advice has never gone underappreciated.

I am extremely grateful to the Australian Research Council Centre of Excellence for Climate System Science (ARCCSS), ARC Centre of Excellence for Climate Extremes (CLEX) and Climate Change Research Centre (CCRC) for their support during my PhD. Thank you to ARCCSS for providing me with funding to attend various conferences and workshops throughout my PhD, where I was able to network and meet like-minded scientists. The annual workshops were especially motivating and inspiring. I would like to thank the CCRC for providing me with a scholarship throughout my PhD, and for providing an exceptionally supportive workplace to do research.

Many thanks to Professor Scott Sisson at the School of Mathematics and Statistics (UNSW) for your help and input towards *Chapter 2*. Thank you to the administrative team at the CCRC for helping me with any logistical issues, and for

the Computational Modelling Services (CMS) team for their technical help. In particular, thanks to Paola Petrelli for all your help with finding and downloading CMIP5 data. I'd also like to acknowledge the World Climate Research Programme's Working Group on Climate Modelling, which is responsible for CMIP, and the modelling groups for producing and making their model output available.

To all my friends and colleagues at the CCRC, thank you for the many conversations, coffee chats, social events and good times. Many of you, particularly fellow PhD students (past and present), have helped me get through the toughest of times, even by just having been there for a coffee break. Thank you for your continued support and encouragement.

Finally, I would like to thank my family and friends. To my partner, Andrew, thank you for your love and patience, and for always managing to take my mind off things. Thank you also to the rest of the Murray's for your support. Huge thanks to my friends and my bandmates, having a laugh and playing music has been integral in getting me through these past few years. Last but not least, I would like to thank my family for their love and support, and for truly believing that I could do it.

Contents

List of Figures	xvii
List of Tables	xix
1 Introduction	1
1.1 Overview.....	1
1.2 Review of the literature.....	3
1.3 Structure of the thesis	11
2 The sensitivity of daily temperature variability and extremes to dataset choice	15
2.1 Introduction.....	16
2.2 Data and methods	19
2.2.1 Observational data.....	19
2.2.2 Reanalysis data	19
2.2.3 Statistical methods.....	20
2.3 Results	25
2.3.1 Global and regional PDFs.....	25
2.3.2 Time series of the statistical moments.....	28
2.3.3 Time series of decadal extreme value fits.....	34
2.3.4 Changes in extremes relative to changes in the mean.....	42
2.4 Discussion	46
2.5 Conclusions	50

3 Changes in daily temperature extremes relative to the mean in	
CMIP5 models and observations	53
3.1 Introduction	54
3.2 Data and methods	58
3.2.1 Observational and CMIP5 datasets	58
3.2.2 Methods	59
3.3 Results	62
3.3.1 Observed and simulated excess changes in recent decades	62
3.3.2 Projected excess changes between future and past decades	70
3.3.3 Spatial correlations of recent and future excess changes.....	74
3.4 Discussion and conclusions	79
4 Enhanced warming of seasonal cold extremes relative to the mean	
in the Northern Hemisphere extratropics	87
4.1 Introduction	88
4.2 Data and methods	92
4.2.1 Observational and CMIP5 datasets	92
4.2.2 Methods	93
4.3 Results	96
4.3.1 Recent excess changes in cold extremes relative to the mean.....	96
4.3.2 Projected excess changes in cold extremes	100
4.3.3 Projected changes in cold air temperature advection.....	105
4.3.4 Projected changes in snow cover and surface albedo	107
4.3.5 Projected changes in the timing of anomalously cold days	111
4.4 Discussion and conclusions.....	113
5 Methodological sensitivities related to the analysis of daily	
temperature extremes	121
5.1 Introduction	122

5.2 Parametric uncertainties	126
5.3 Structural uncertainties	133
5.3.1 Absolute temperatures	133
5.3.2 Methods for defining extremes	134
5.4 Uncertainty due to different levels of radiative forcing.....	136
5.5 Discussion and conclusions.....	137
6 Conclusions and future work	141
6.1 Summary of findings	141
6.2 Future work	148
6.3 Concluding remarks	150
 References	 153
 S1 Supplementary material for Chapter 2	 171
 S2 Supplementary material for Chapter 3	 199
 S3 Supplementary material for Chapter 4	 235

List of Figures

1.1	The effect on temperature extremes from a changing climate.....	5
1.2	Extreme and mean temperature changes associated with a 2°C global warming target.....	7
2.1	The 26 regions specified in the IPCC SREX.....	21
2.2	Empirical PDFs for all SREX regions for daily maximum temperature anomalies.....	26
2.3	As Figure 2.2, but for daily minimum temperature anomalies.....	27
2.4	Time series of statistical moments for the globe.....	29
2.5	As Figure 2.4, but for SREX region 3 (west North America)	31
2.6	As Figure 2.4, but for SREX region 10 (south-east South America)	31
2.7	As Figure 2.4, but for SREX region 13 (Mediterranean)	32
2.8	As Figure 2.4, but for SREX region 18 (North Asia/Russia)	32
2.9	Global time series of the extreme value parameters.....	35
2.10	As Figure 2.9, but for SREX region 3 (west North America)	37
2.11	As Figure 2.9, but for SREX region 10 (south-east South America)	39
2.12	As Figure 2.9, but for SREX region 13 (Mediterranean)	40
2.13	As Figure 2.9, but for SREX region 18 (North Asia/Russia)	41
2.14	Excess trends for the high tails of temperature anomalies.....	43
2.15	As Figure 2.14, but for the low tails.....	45
3.1	Recent observed excess changes in HadGHCND for the warm tails.....	63
3.2	As Figure 3.1, but for the CMIP5 multi-model mean.....	65
3.3	As Figure 3.1, but for the cold tails.....	67

3.4	As Figure 3.2, but for the cold tails.....	69
3.5	Future excess changes in the warm tails in CMIP5.....	71
3.6	As Figure 3.5, but for the cold tails.....	73
3.7	Globally-averaged pattern correlations for December – February.....	75
3.8	As Figure 3.7, but for March – May.....	76
3.9	As Figure 3.7, but for June – August.....	78
3.10	As Figure 3.7, but for September – November.....	79
4.1	Recent excess changes in cold extremes for December – February.....	98
4.2	As Figure 4.1, but for March – May.....	99
4.3	As Figure 4.1, but for September – November.....	100
4.4	Future excess changes in cold extremes for December – February.....	102
4.5	As Figure 4.4, but for March – May.....	103
4.6	As Figure 4.4, but for September – November.....	104
4.7	Future changes in actual and excess cold air temperature advection..	106
4.8	Scatter plots of annual values of excess temperature in cold extremes against annual values of snow cover.....	109
4.9	As Figure 4.8, but for surface albedo.....	110
4.10	Changes in the timing of the anomalously coldest days.....	112
5.1	Differences in future excess changes depending on the climatological reference period.....	128
5.2	Differences in projected changes in snow cover and the timing of cold extremes depending on climatological reference period.....	132
5.3	Differences in future excess change depending on choice of method....	134
5.4	Comparison of RCP8.5 and RCP4.5 emissions scenarios.....	136

List of Tables

2.1	Number of HadGHCND stations within each of the 26 SREX regions, and number of grid cells with observational data per region.....	23
3.1	CMIP5 models used in this study and their modelling group.....	60
4.1	List of CMIP5 models used in this study and their institution.....	93

CHAPTER 1

Introduction

1.1 Overview

While global mean temperature is a valuable indicator of global climate change, statistical analysis shows changes in the variability of daily temperature have the greatest impact on the probability of extreme events (Katz and Brown 1992). Many studies point to global-scale changes in extremes being driven by a shift in the mean alone (e.g. Simolo et al. 2010; Donat and Alexander 2012; Rhines and Huybers 2013; Tingley and Huybers 2013), however, there is some evidence for regional-scale changes in variability (e.g. Schär et al. 2004; Huntingford et al. 2013; McKinnon et al. 2016; Rhines et al. 2017).

To date, a number of studies have focused on observed and projected changes in regional temperature variability and the disproportionate increases of hot extremes in summer, mainly over Europe and North America (e.g. Schär et al. 2004; Fischer and Schär 2009; Simolo et al. 2011; Cattiaux et al. 2015; Holmes et al. 2016). Others tend to focus on individual aspects of temperature change, such as global mean temperature (e.g. Sutton et al. 2015; Hawkins and Sutton 2016; Rahmstorf et al. 2017), or global or regional changes in the annual hottest or coldest days (e.g. Alexander et al. 2006; Coumou and Robinson 2013; Christidis et al. 2011). However,

addressing changes in means or extremes alone does not allow the extent of changes in several other aspects of the temperature distribution to be fully understood. Considering changes in extremes relative to mean changes is a simple approach to better characterise distributional changes and to understand how the most impacts-relevant parts of the distribution behave relative to average conditions. These analyses rely on the accuracy of observational datasets which are then used to evaluate climate model simulations. However, the quality and robustness of the observational datasets used for analyses, that are often spatially or temporally incomplete, remain a source of uncertainty, and therefore, these products are not necessarily reliable for evaluating climate models.

Due to the uncertainties surrounding commonly used observations-based datasets, the first aim of this thesis is to understand whether dataset choice has an impact on analyses of changes in temperature variability and extremes. Differences in the datasets when assessing trends in extremes relative to mean temperatures are also explored. The same *in situ*-based observational product is then used to assess the disproportionate rates of change in seasonal extremes relative to seasonal mean temperatures alongside a suite of climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5) archive (Taylor et al. 2012), which are then used to project future changes. This is systematically investigated for all seasons and all global land areas, from the mid-20th century until the end of the 21st century. Where models agree on the most robust changes, I delve into the physical mechanisms behind those changes by examining conditions on or prior to the day the extreme occurs. Various sensitivities related to the analyses of temperature extremes are also explored in order to reduce uncertainties related to methodological choices. The methods used in this thesis will be shown to be robust,

providing an easily reproducible approach for investigating the disproportionate rates of change in extremes relative to mean temperatures.

While some of the findings presented here complement previous research, the individual chapters within the thesis combine to provide a more holistic and complete sense of how temperature extremes are changing and might change in the future. This is explored specifically in relation to changes in the mean of the distribution of daily temperatures, thereby providing information on several aspects of changes in the temperature distribution. This chapter serves to summarise and identify gaps in the relevant scientific literature and outline the individual aims and structure of this thesis.

1.2 Review of the literature

Due to their potential for adverse impacts on society and ecosystems, temperature extremes have been a prominent focus in the scientific literature in recent years. The majority of studies tend to focus on changes in extremes and mean temperatures separately, making it difficult to infer how the distribution of temperature is changing and whether variability is changing on top of shifts in the mean. Climate model simulations project consistent increases in annual mean temperature for land regions across the globe into the 21st century (e.g. Kirtman et al. 2013; Seneviratne et al. 2016; Lewis and King 2017). This is coupled with projected increases in the frequency and intensity of warm extremes and decreases in the frequency of cold extremes (e.g. Coumou and Robinson 2013; Kharin et al. 2013; Kirtman et al. 2013; Sillmann et al. 2013a). Overall, this is consistent with a shift in the distribution towards warmer temperatures.

Changes in the symmetry of the temperature distribution have also been shown using observational data and climate model simulations (e.g. Kodra and Ganguly 2014; Matiu et al. 2016; Lewis and King 2017), with the distribution becoming skewed towards hotter temperatures, especially for parts of Europe (e.g. Matiu et al. 2016), but also for North America, Asia and Australia (Lewis and King 2017). However, some have argued that the methods used to calculate temperature anomalies in some of these studies are causing inflated changes in extremes (Rhines and Huybers 2013; Sippel et al. 2015). Moreover, discrepancies with observations as well as among the climate models themselves are apparent, particularly for cold extremes (Kharin et al. 2007, 2013; Lewis and King 2017). Despite the uncertainties, future simulated warming of both extreme and mean temperatures is significant and robust for much of the globe (Kirtman et al. 2013; Collins et al. 2013; Sillmann et al. 2013a), and will likely exacerbate the impacts stemming from extremes if strict measures to manage the risks are not implemented (Intergovernmental Panel on Climate Change (IPCC) 2012).

Importantly, it is unlikely that the rates of change between different aspects of temperature are uniform. In fact, observations and climate model simulations indicate that cold extremes are warming faster than warm extremes for much of the globe (e.g. Kharin and Zwiers 2000, 2005; Alexander et al. 2006; Donat and Alexander 2012; Hartmann et al. 2013; Kharin et al. 2007, 2013, Donat et al. 2013a, 2016). This is significant because disproportionate rates of change between different aspects of temperature affect the overall shape of the distribution, which affects the probability and frequency of extreme events and thus their impacts (Mearns et al. 1984).

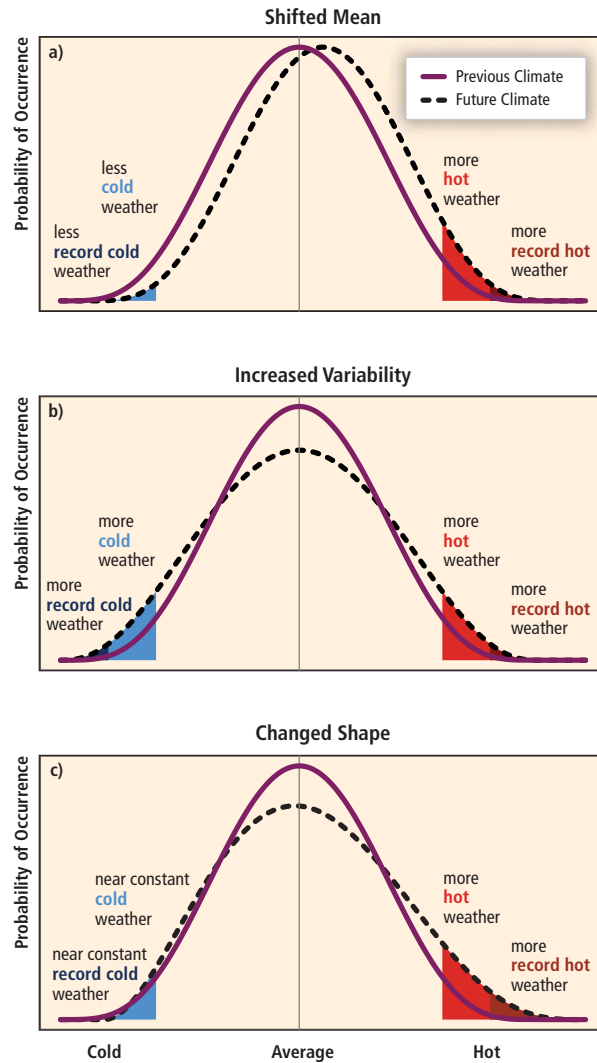


Figure 1.1 The effect on temperature extremes from a changing climate, showing (a) a shift in the distribution towards warmer temperatures, (b) an increase in temperature variability, and (c) change in the shape of the distribution (source: Lavell et al. 2012, Figure 1.2, p. 41).

To demonstrate how temperature extremes might be affected in a changing climate, the IPCC use a schematic to describe changes in the distribution of temperature (Figure 1.1). An increase in the mean alone would shift the entire distribution to warmer temperatures, with hot extremes getting hotter and more frequent, and cold extremes becoming warmer and less frequent (Figure 1.1a). Changes in the variance alone would lead to an increased probability of both cold and warm extremes (Figure 1.1b), while a change in the shape of the distribution can skew temperatures towards the hotter or cooler part of the distribution (Figure 1.1c).

Theoretically, such a schematic is useful to demonstrate the possible effects that overall changes in temperature distributions can have on extremes, or vice versa, emphasising the need to consider changes in the tails relative to the entire distribution (Katz et al. 2013; Sardeshmukh et al. 2015). Quantifying past changes and predicting future changes in extremes relative to mean temperatures would provide crucial information that can be used by decision-makers and planners to reduce the risks of negative impacts from extremes.

While some studies have looked at changes in temperature extremes relative to the mean, there are key differences in the methodologies and measures used that make it difficult to get a complete sense of how these rates of change differ across all land regions for the globe. For instance, most of these studies compare changes in extremes relative to global mean temperatures, with some focusing primarily on globally averaged extremes (e.g. Kharin and Zwiers 2005; Kharin et al. 2007), and others comparing global mean temperature with changes in local extremes (e.g. Seneviratne et al. 2016; Vogel et al. 2017). Despite being more useful to identify and explore the physical processes driving regional distributional changes, only limited studies have looked at changes in local extremes relative to changes in local mean or median temperatures (e.g. Brown et al. 2008; Orłowsky and Seneviratne 2012; Donat et al. 2017).

Figure 1.2, from Seneviratne et al. 2016, shows changes in extreme and mean temperatures associated with a 2°C target of global mean temperature increase. Both hot and cold extremes, and mean temperatures, are projected to warm across the globe relative to 2°C of global warming. From the figure, it is clear that the different variables warm at different magnitudes, and regional differences can be substantial, especially for the coldest annual temperatures (Figure 1.2b). While

assessing local changes in extremes relative to global mean temperature can be useful for regional decision-making and setting emissions targets, global mean temperature is a rather abstract measure that does not directly impact society (Sutton et al. 2015; Seneviratne et al. 2016). Rather, it is changes in the local mean temperature that people and the environment actually experience. Assessing changes in local extremes relative to corresponding local mean temperatures is also vital to understand the physical processes and relationships that contribute to disproportionate rates of warming.

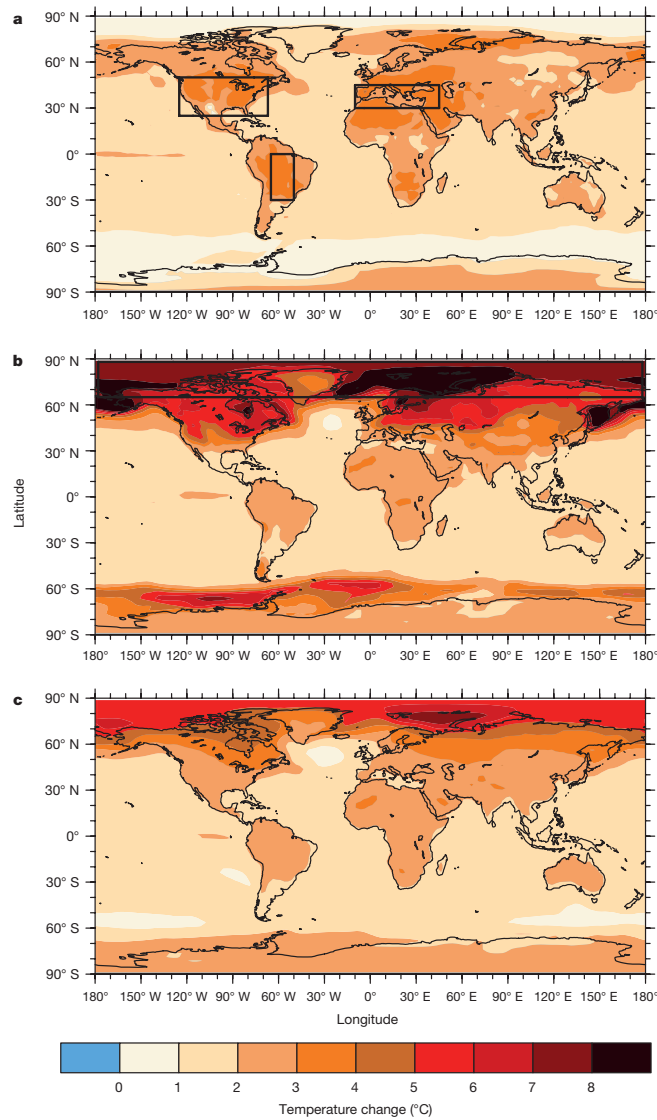


Figure 1.2 Extreme and mean temperature changes associated with a 2°C global warming target, for (a) the annual hottest daytime temperature, (b) annual coldest night-time temperature and (c) annual mean temperature. Results are based on the RCP8.5 scenario simulations from CMIP5 (source: Seneviratne et al. 2016, p. 479).

Studies comparing local extremes to local mean temperatures have predominantly focused on annual statistics calculated from daily observational data or climate model output (e.g. Kharin and Zwiers 2005; Brown et al. 2008; Donat et al. 2017). Observations and climate model simulations indicate that in recent decades, annual cold extremes have generally warmed in excess of annual mean daily minimum temperatures in Northern Hemisphere regions that are affected by snow fall and sea ice (Kharin and Zwiers 2005; Brown et al. 2008). Annual hot extremes have been warming faster than the annual mean of daily maximum temperatures for parts of Europe, the Mediterranean, southern Africa and Australia (Brown et al. 2008; Donat et al. 2017).

Climate model projections suggest that the annual hottest days are expected to warm at an accelerated rate compared with the local annual mean for many of the same regions, as well as parts of North America and Europe, which have primarily been related to land-atmosphere feedbacks as driving processes (Donat et al. 2017). However, the physical mechanisms driving the disproportionate rates of warming between extremes and mean temperatures likely differ depending on the season, and between warm and cold extremes. For example, during summer, changes in soil-moisture feedbacks and surface heat fluxes are the more likely dominant driver of accelerated warming of hot extremes (Donat et al. 2017; Vogel et al. 2017), while changes in snow and ice cover, and larger-scale changes in atmospheric circulation, are more likely to drive disproportionate rates of warming in cold extremes and decreases in temperature variability during winter months (e.g. Gregory and Mitchell 1995; Kjellström et al. 2007; Fischer et al. 2011; Kharin et al. 2007, 2013; Screen 2014). The impacts from temperature extremes are often seasonally dependent as well (Alexander et al. 2006). While Orłowsky and Seneviratne (2012) explored the scaling of regional and seasonal extremes, this was addressed in

relation to global mean temperature change and the annual regional median change in temperature. Hence, it is still largely uncertain how seasonal changes in regional extremes are changing with respect to changes in corresponding seasonal mean temperatures, and further, understanding the physical drivers influencing any disproportionate temperature changes, particularly in cold extremes, requires more attention.

Robust conclusions regarding the disproportionate changes of extremes relative to mean temperatures rely on the accuracy and reliability of the input datasets used for analysis. Due to the lack of complete daily observed temperature datasets on a global scale, the majority of studies addressing changes in temperature variability and extremes have used monthly or seasonal data (e.g. Hansen et al. 2012; Coumou and Robinson 2013; Huntingford et al. 2013; Rhines and Huybers 2013). However, the temporal aggregation of these datasets smooths out the individual events that occur on daily timescales, and it is these events that are most likely to cause the greatest impacts (Alexander and Perkins 2013). For this reason, it is imperative to use daily datasets that are both sufficiently long and continuous to detect and project changes in extremes (Klein Tank et al. 2009). *In situ*-based daily temperature datasets, such as the Hadley Centre Global Historical Climatology Network-Daily (HadGHCND) dataset (Caesar et al. 2006), which represents one of the most comprehensive global daily temperature datasets currently available, still lacks data in some regions. For this reason, reanalysis products, which assimilate *in situ*-based and satellite-based observational data into numerical weather prediction models, are commonly used for global studies (e.g. Huntingford et al. 2013; Donat et al. 2014; Kodra and Ganguly 2014; Ylhäisi and Räisänen 2014; Simmons et al. 2017). However, like climate models, reanalysis products need to be

evaluated to determine their robustness against observational products, and the impact of dataset selection depends on application.

Without drastic measures to severely limit atmospheric greenhouse gas concentrations, the IPCC has stated that “it is *virtually certain* that, in most places, there will be more hot and fewer cold temperature extremes as global mean temperature increases” (Collins et al. 2013, p. 1031). These anticipated increases will likely exacerbate the impacts from extremes. Importantly, the impacts from extremes are already unequally distributed, due to differences in socioeconomic factors that reduce a region’s capability to respond to extreme events (IPCC 2012). Disproportionate rates of warming in extreme and mean temperatures could further intensify the impacts from extremes, especially for already vulnerable regions. For example, heat-health related impacts might be exacerbated by temperature extremes warming more than mean temperatures, as a result of not being able to adequately acclimatise. Although humans are able to adapt to some changes to temperature (Gasparrini et al. 2016), the limits of these adaptation mechanisms are still an active area of research (Sherwood and Huber 2010). Natural ecosystems might also be affected by different rates of change in temperature, for instance, faster warming rates of mean temperatures compared with cold extremes during winter months may increase the longevity of disease-carrying insects that can survive through milder winters (e.g. Wolf et al. 2015; Ebi and Nealon 2016). This also highlights the seasonal dependence of impacts that disproportionate rates of warming in extremes can influence.

While many studies have investigated changes in temperature extremes, there are important limitations and gaps in the research that need to be considered and resolved. Addressing some of the sensitivities related to methodological choices by

exploring alternative methods and testing the robustness of results would reduce uncertainties and make it easier to compare studies. Then, systematically assessing several aspects of temperature distribution changes using an easily reproducible and understandable method could serve as a new standard tool for assessing changes in temperature extremes relative to mean temperatures. The implications of this type of study are far-reaching for numerous fields in climate change research.

1.3 Structure of the thesis

The overarching goal of this thesis is to provide an in-depth understanding of how daily temperature extremes have changed and might change in the future across all global land regions, specifically with regards to changes in mean temperatures. This is explored across all seasons using daily maximum and daily minimum temperature data from observations, reanalysis data and climate model output. By considering several aspects of the temperature distribution in parallel with one another, the disproportionate rates of change between different parts of the distribution become evidently clear. Methodological sensitivities underpin and limit studies regarding temperature extremes, so another key aim of this thesis is to evaluate some of these sensitivities. Each chapter has specific objectives that contribute to the overall aims of this thesis, as outlined below.

1. Investigate the impact of dataset choice for analysing statistics related to the distribution of temperature and extremes.

In *Chapter 2*, several commonly used reanalysis datasets of daily maximum and minimum temperature are compared against the HadGHCND dataset, a quasi-global, gridded *in situ*-based dataset of daily maximum and minimum temperatures. Differences and similarities between the datasets are identified using various statistical properties that are important for understanding changes in

temperature. Statistical moments, including the mean, standard deviation and skewness, are compared across the different datasets, while the Extreme Value Theory is used to compare statistics that are specific to the tails of the distribution. Lastly, trends in the tails of the distribution relative to the mean are explored for recent decades across all datasets, noting any sensitivities related to dataset choice.

2. Assess changes in seasonal temperature extremes relative to seasonal mean temperatures in observations and CMIP5 models.

Chapter 3 systematically assesses changes in local seasonal temperature extremes relative to corresponding local seasonal mean temperatures from the mid-20th century until the end of the 21st century. This follows on from the annual trends in extremes relative to the mean investigated in *Chapter 2* and uses the same observational dataset to evaluate a suite of CMIP5 models, which are then used to project future changes in both warm and cold extremes relative to mean temperatures for all seasons and all land regions across the globe.

3. Explore projected changes in seasonal cold extremes relative to seasonal mean temperatures and investigate their physical mechanisms.

In *Chapter 4*, the most robust changes identified in *Chapter 3* are assessed further to identify the possible physical mechanisms that are driving the strongest changes in temperature extremes relative to mean temperatures. Specifically, this chapter uses several individual CMIP5 models to assess the accelerated warming rates of cold extremes relative to mean temperatures in boreal winter and the shoulder seasons for Northern Hemisphere mid- to high-latitude land regions. Then, the environmental conditions on the day or just prior to the day of the extreme are explored to determine the physical processes that are plausibly contributing to the projected enhanced warming of cold extremes.

4. Examine methodological sensitivities related to the analysis of changes in temperature extremes relative to mean temperatures.

Chapter 5 explores some of the uncertainties and sensitivities related to the results presented in the previous chapters, such as sensitivities related to the choice of climatological reference period, and the methods used to define extreme temperatures. It brings together the different methodological choices used in each chapter to provide context on whether results are sensitive to the methods used to analyse changes in temperature extremes relative to the mean and summarises the robustness of conclusions in *Chapters 2 – 4* of this thesis.

Finally, the main findings and conclusions of this thesis are summarised in *Chapter 6*, which also suggests opportunities for future work relating to this field of research.

Chapter 2 is based on a published peer-reviewed article in the *Journal of Climate*, *Chapter 3* has been revised and re-submitted to the *International Journal of Climatology*, and *Chapter 4* has been submitted to *Climate Dynamics* and is currently under review. These chapters are presented as they are published or submitted, with some minor changes in formatting. Overviews preceding each chapter provide some context as to the contribution of the chapter to the thesis.

CHAPTER 2

The sensitivity of daily temperature variability and extremes to dataset choice

Chapter overview

This chapter is published as: Gross, M. H., M. G. Donat, L. V. Alexander and S. A. Sisson, 2018: The sensitivity of daily temperature variability and extremes to dataset choice. *Journal of Climate*, **31**, 1337-1359, doi:10.1175/JCLI-D-17-0243.1.

Robust conclusions regarding changes in the temperature distribution rely on the accuracy and reliability of the input datasets used. Differences between methodologies and datasets in previous studies add uncertainty when comparing and quantifying findings. This chapter assesses whether dataset choice has an impact on assessments regarding the distribution of temperature and how extremes are changing relative to mean changes. Several commonly used reanalysis datasets are compared against the gridded *in situ*-based daily temperature dataset, HadGHCND, by assessing both the entire distribution and the tails of the distribution. Empirical Probability Distribution Functions show sensitivity to the input dataset when estimating aspects such as standard deviation and skewness, with the mean showing robust results for most regions, irrespective of dataset

choice. Standard deviation is especially sensitive, with larger disagreements between datasets for some regions more than others, such as Africa and the Mediterranean region, and with larger differences in minimum temperatures compared with maximum temperatures. Estimates of extreme parameters also show sensitivity to dataset choice, particularly in the lower tails and for daily minimum temperature anomalies. Comparing changes in the means and the extremes of the temperature distributions, cold extremes in the lower tails have been warming at a faster rate than the mean of the entire distribution for much of the Northern Hemisphere extratropics, with warm extremes warming at a faster rate than the mean in some subtropical regions. Despite some sensitivity in the exact quantifications between datasets, these findings of faster rates of change are qualitatively robust regardless of the input dataset. The documented sensitivities call for caution when assessing changes in temperature variability and extremes, as dataset choice can have substantial effects on the exact quantification of results.

2.1 Introduction

Temperature extremes represent one of the most obvious impacts of climate change on society. For example, the impacts from heatwaves can range from increases in human mortality and morbidity to effects on agriculture and infrastructure (IPCC 2012). Variability within temperature can affect the probability and frequency of extreme events (Mearns et al. 1984; Katz and Brown 1992; Rahmstorf and Coumou 2011; McKinnon et al. 2016), making it critical to understand how different aspects of the temperature distribution are changing and how they might change in the future. For this reason, many studies have investigated how extremes might be affected by changes in the distribution of temperature due to climate change (e.g. Rahmstorf and Coumou 2011; Donat and Alexander 2012; Hansen et al. 2012;

Rhines and Huybers 2013; McKinnon et al. 2016). Despite the abundance of such studies, there are still some contested issues regarding changes in temperature variability. Mainly, while some studies have found increases in global temperature variability (e.g. Hansen et al. 2012), others have concluded that extremes are shifting towards hotter temperatures along with the mean, with little or no change in global variability (e.g. Brown et al. 2008; Donat and Alexander 2012; Huntingford et al. 2013; Rhines and Huybers 2013; Tingley and Huybers 2013). This suggests that while there is consensus in terms of changes in the mean temperature, there remains considerable uncertainty in changes of other aspects of the distribution.

The issues and uncertainties surrounding previous studies are likely related to the different methods and datasets used for investigating changes in the distribution of temperature, which make these studies difficult to compare and quantify. Most studies to date have used monthly or seasonal datasets (e.g. Hansen et al. 2012; Coumou and Robinson 2013; Rhines and Huybers 2013). However, the temporal aggregation of these datasets can smooth out the individual events that can occur on daily timescales. This makes it critical to use daily data to detect changes in the temperature distribution, as it is the characteristics of extremes (such as the frequency, intensity and duration) that are most likely to impact society (Alexander and Perkins 2013). This requires long-term, continuous and consistent high-quality datasets (Klein Tank et al. 2009).

The limited studies that have used daily data to investigate changes in the temperature distribution have predominantly used three main statistical methods. This includes assessing statistical moments of the probability density function (PDF), such as the mean, variance and skewness (e.g. Donat and Alexander 2012),

using Extreme Value Theory (EVT) to explicitly characterize the tails of the distribution (e.g. Kharin and Zwiers 2005; Brown et al. 2008; Christidis et al. 2011), and examining changes in different percentiles (e.g. Robeson 2004; Simolo et al. 2010; McKinnon et al. 2016). Some of these studies (e.g. Brown et al. 2008; Donat and Alexander 2012) have used HadGHCND, a quasi-global daily gridded temperature dataset (Caesar et al. 2006). However, such an *in situ*-based daily dataset still lacks data for certain regions, such as parts of South America and Africa. Reanalysis products that use assimilated observational data for the globe are therefore commonly used to investigate global changes in variability and extremes (e.g. Huntingford et al. 2013; Donat et al. 2014). Additionally, because of their global completeness and areal structure, reanalyses might be considered advantageous when evaluating climate models against historical observation-based data (e.g. Kharin et al. 2013; Sillmann et al. 2013b; Donat et al. 2016). However, it is still not clear if assessments of changes in temperature variability and extremes are sensitive to these different types of input data, a factor that might affect a study’s conclusions.

A crucial first step in gaining a more comprehensive understanding of how temperature variability and extremes are changing requires a sensitivity analysis of some of the most commonly used datasets of daily temperatures. Further, a systematic and holistic approach would require a combination of assessing both the PDFs and the tails of the distribution using EVT (Katz et al. 2013; Sardeshmukh et al. 2015). Using both methods in parallel can help address the uncertainties in assessing the tail alone, as well as inferring changes in extremes relative to the mean of the entire distribution (Sardeshmukh et al. 2015). Here, we aim to use these two approaches to compare commonly used reanalysis products with the HadGHCND daily dataset. Our aim is not to make any judgments about the quality of the

observational data itself, but rather to determine if analyses are sensitive to the input dataset. We use a mathematically consistent approach and consider both global and regional temperature distributions, to provide critical information for future studies that wish to use these types of datasets to investigate changes in temperature variability and extremes.

2.2 Data and Methods

2.2.1 Observational data

We use HadGHCND as the base dataset to compare against reanalysis products. Since it is quasi-global over land and is based only on *in situ* daily maximum and minimum temperatures, this gives mostly independent information with which to compare against other products. It is available from 1950, and at the time of analysis, ended in 2014. HadGHCND uses approximately 2500 stations that are interpolated onto a 2.5° latitude x 3.75° longitude grid using an angular distance weighting technique (Shepard 1968; Caesar et al. 2006).

2.2.2 Reanalysis data

Four reanalysis products are selected for inter-comparison: ERA-Interim (Dee et al. 2011), NCEP-DOE Reanalysis 2 (NCEP2) (Kanamitsu et al. 2002), the Japanese 55-year Reanalysis (JRA-55) (Kobayashi et al. 2015) and the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2) (Bosilovich et al. 2015; Gelaro et al. 2017). These were chosen primarily based on their prevalence in the literature for investigating temperature variability and extremes, as well as their comparability in terms of dataset length. We purposely did not include some commonly used first generation reanalyses, because for NCEP1 (Kalnay et al. 1996; Kistler et al. 2001), inhomogeneities in the

representation of warm extremes have been documented (e.g. Donat et al. 2014). Other products, such as the 20th Century Reanalysis (Compo et al. 2011), and ERA-40 (Uppala et al. 2005), were excluded as data for recent years covered by the chosen reanalyses are not available.

To test the sensitivities of analyses to the different input datasets, we focus on the common time period of 1980-2014. While this is a relatively short period of time to investigate long-term climate changes, it is sufficient for the purpose of assessing robustness between different datasets. All reanalysis products are regridded using a bilinear remapping technique and masked to the grid cell size and spatial coverage of HadGHCND. Other regridding techniques were tested, such as conservative remapping techniques (Jones 1999), and results were not sensitive to the regridding method used. We also only include grid cells that have at least 80% of temporal completeness in HadGHCND, therefore excluding data sparse regions. In addition to meeting this criterion, analysis is only performed when there are at least 50% of data available for both the first and last ten years of HadGHCND data. This ensures there are enough data points to robustly assess changes over a 35-year period. While the regridding itself is likely to add some uncertainty to the intercomparison of products, it has been shown that this effect is small compared to the structural differences between datasets (e.g. Loikith et al. 2015). We use daily maximum and minimum temperature anomalies that were calculated relative to a mean annual cycle over the entire period of investigation, that is, 1980-2014.

2.2.3 Statistical methods

Two approaches are used to assess dataset sensitivity for investigating changes in temperature variability and extremes. First, we examine PDFs for the entire temperature distribution over the period of record chosen, and second, we use

Extreme Value Theory (EVT) for direct analysis of the tails of the distribution and to understand how extremes are changing relative to the mean of the distribution.

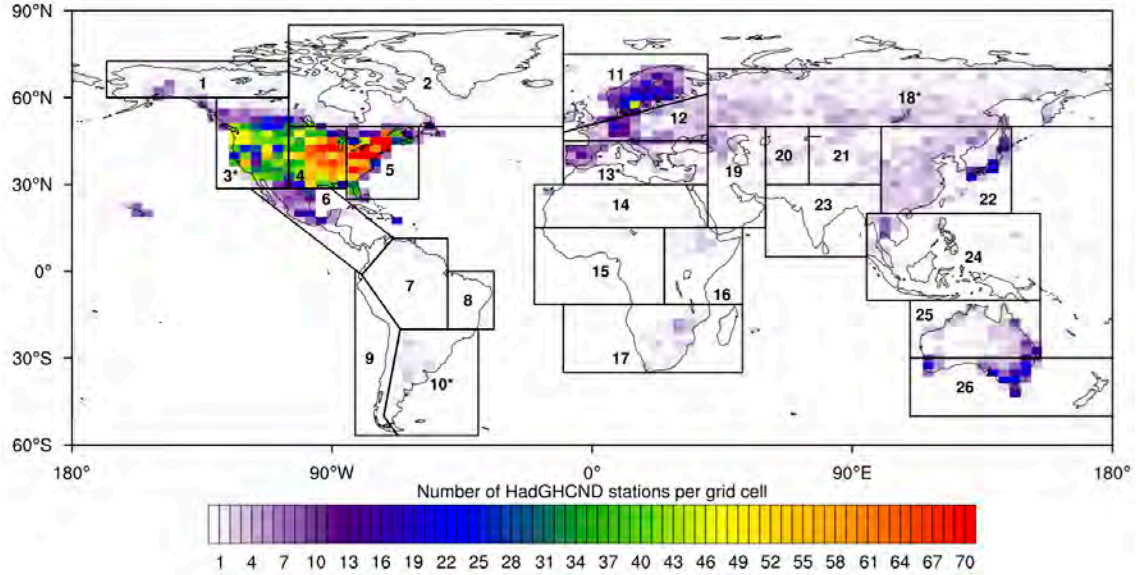


Figure 2.1 The 26 regions specified in the IPCC SREX, and the spatial distribution of stations used in HadGHCND at the time of interpolation. * indicates regions that are used for analysis of time series (i.e. Figure 2.4 – 2.13). Regions 7, 8, 9, 15 and 16 are excluded from analyses based on not fulfilling the completeness criteria specified in Methods.

We analyse PDFs for the 26 regions specified in the IPCC Special Report on Extremes (SREX) (IPCC 2012) (Figure 2.1). These were chosen for this study as they indicate regions of common climates and enabled sufficient data points for each region for reliable statistical analysis. Figure 2.1 shows the spatial distribution of all included 2500 stations used in the interpolation of HadGHCND. Furthermore, we show the number of stations per SREX region, and number of grid cells with data in each region in Table 2.1. Empirical PDFs are calculated by pooling all time steps and grid cells within a region to produce a distribution of temperature over the period 1980-2014. We note that different ways of how grid cells are spatially aggregated may affect estimates of distribution changes (Rhines and Huybers 2013; Director and Bornn 2015). The focus of our study, however, is on dataset agreement, and we treat all datasets the same. Therefore, such sensitivities

regarding spatially averaged grid cells are less relevant here. We used 140 equally spaced bins, ranging from -45°C to 45°C , to calculate the PDFs, as this choice most clearly captured the features of the dataset. We repeated our analyses for other bin width settings, but our conclusions were not sensitive to bin size choices for calculating the PDFs.

SREX regions are only included in the analysis if at least three grid cells within that region fulfill the completeness criteria specified in Section 2.2.2. While this criterion doesn't necessarily ensure spatial representativeness, it increases statistical power in terms of sample size for estimating the regional PDFs. For the time series analysis in the main text, we only show selected regions that represent different levels of data coverage and different climatic zones, those being, regions 3 (West North America), 10 (south-eastern South America), 13 (Mediterranean region), and 18 (North Asia/Russia). Figures for the remaining regions are included as supplementary material. We also include a global analysis, calculated by aggregating all grid cells that fulfill the completeness criteria.

The sample mean, standard deviation and skewness are calculated for each region and dataset. To investigate temporal changes in these quantities, time series of annual statistics are plotted over the investigation period (see Figure 2.4 – 2.8). We calculated trends using both a linear regression and Sen's trend estimator (Sen 1968), estimating trend significance using the Mann-Kendall test (Kendall 1975). Autocorrelation in the time series was accounted for by adjusting the sample size to the equivalent number of independent values, or equivalent sample size (Zwiers and von Storch 1995). Results were found to be insensitive to the test used, and so we only show linear regression trends, with significance calculated at the 5% level using a t-test.

Table 2.1 Number of HadGHCND stations within each of the 26 SREX regions, and number of grid cells with observational data per region for daily maximum and daily minimum temperature anomalies. Counts of stations only consider the number of used stations available at the time of interpolation, and so variations of stations over time are not considered. Counts of grid cells with available data per region are calculated after a criterion is applied where grid cells must have at least 80% temporal completeness, as well as at least 50% of data available in the first and last ten years of HadGHCND data.

SREX region	No. of stations per region	No. of grid cells with data per region (Tmax)	No. of grid cells with data per region (Tmin)
1. Alaska/Northwest Canada	79	77	77
2. Canada/Greenland/Iceland	142	134	114
3. West North America	1795	68	68
4. Central North America	1976	39	39
5. East North America	1435	23	23
6. Central America/Mexico	182	4	4
7. Amazon	7	0	2
8. North-East Brazil	0	0	0
9. West coast South America	1	0	0
10. South-eastern South America	25	15	8
11. North Europe	489	37	37
12. Central Europe	219	46	46
13. Mediterranean	138	39	39
14. Sahara/North Africa	16	14	15
15. West Africa	1	0	0
16. East Africa	18	0	0
17. Southern Africa	40	10	9
18. North Asia/Russia	375	245	245
19. West Asia	61	41	43
20. Central Asia	22	33	33
21. Tibetan Plateau	41	47	47
22. East Asia	350	86	86
23. South Asia	5	13	14
24. Southeast Asia	61	25	14
25. North Australia	123	50	50
26. South Australia/New Zealand	276	22	22

We use EVT to determine if analyses of the extremes are sensitive to dataset choice and assess changes in the extremes explicitly. In particular, a non-stationary point process (PP) model is fitted to the data to describe annual exceedances above and

below the upper and lower 1.5%, respectively, as in Brown et al. (2008). We additionally tested different thresholds, such as the upper and lower 1% and 2% of data, and found the results were not sensitive to the choice of threshold. This model corresponds to a Poisson process for the inter-exceedance times, with a Generalised Pareto distribution for the size of the exceedances (e.g. Coles 2001). Similar to the methods of Brown et al. (2008), we model non-stationarity of the process through a time-dependent threshold and location parameter (explicitly, a linear trend in the location parameter) over the period of investigation, as implemented in the ‘extRemes’ package in R (Gilleland and Katz 2011). Non-linearity in scale and shape parameters was also tested but found not to be present, and the fidelity of the PP model to the data was assessed using standard goodness-of-fit procedures, including qq-plots. It is clear from the qq-plots that the selected PP distribution is supported by the empirical distribution of the extremes, as the values fit close to the 1-1 line (see Figure S1.1). As in Brown et al. (2008), we calculate the difference of this location trend with the trend in the mean for all available grid cells to investigate how the magnitude of the extremes are changing relative to the mean, hereafter referred to as ‘excess trends’. Local significance of the excess trends was calculated for each grid cell using a Mann-Kendall test at the 5% level.

In addition, we plotted time series of the location, scale and shape parameters estimated from a stationary PP model, as shown in Section 2.3.3. These are calculated from decadal moving windows, starting in 1980 and ending in 2014. Trends in the PP fit parameters are similarly calculated using the same methods as those used for the moments of the entire PDF.

2.3 Results

2.3.1 Global and regional PDFs

PDFs for SREX regions that adhere to our completeness criteria, as well as global PDFs, are shown for maximum temperature anomalies (Tmax, Figure 2.2) and minimum temperature anomalies (Tmin, Figure 2.3). Dashed vertical lines of different colours represent the threshold of the top and bottom 1.5% of the data (‘extremes’) for each dataset. Plots of the respective standard deviation and skewness values are shown in supplementary material (Figure S1.1).

As expected, higher latitude regions generally have broader PDFs across all datasets, compared with lower latitudes for both maximum and minimum temperature anomalies, indicative of higher variability. For most regions, PDFs of the reanalyses are wider compared to the HadGHCND-derived PDFs for both the globe and most regions. This separation might be due to more spatial smoothing in HadGHCND, where the correlation length scale (CLS) used for interpolation is based on monthly mean temperatures (Caesar et al. 2006). This leads to search radii of several hundred to thousand kilometres for stations to include when calculating the grid box values. It is possible that this monthly CLS is larger than the correlation of daily temperature values would be, potentially causing overly strong smoothing in the HadGHCND dataset. Broader PDFs in the reanalyses are especially apparent in the Mediterranean region (region 13) and some subtropical regions, such as Northern Africa (region 14) and southern Africa (region 17). PDFs for Northern Hemisphere extratropical regions are also wider in the reanalyses compared with HadGHCND, although less so for parts of North America (e.g. regions 3-5), and North Asia/Russia (region 18). For all regions and the globe, there are fewer similarities between the PDFs of the different datasets for Tmin compared to Tmax. Skewness differences are larger between datasets for Tmin, with NCEP2

showing a more negatively skewed PDF for most regions compared with the other datasets.

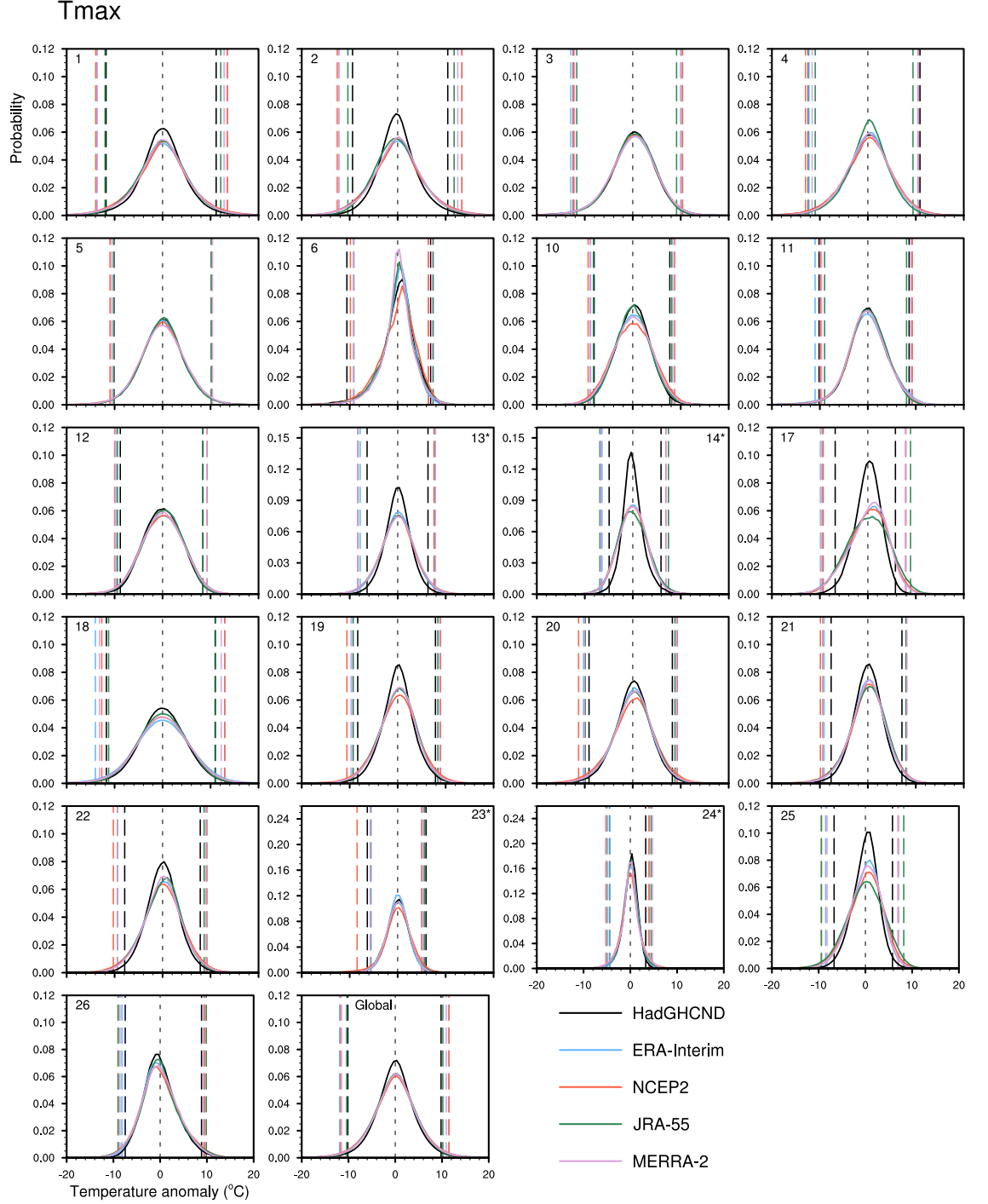


Figure 2.2 Empirical PDFs of daily maximum temperature anomalies for all SREX regions that fulfill the completeness criteria, and for the globe. The PDFs are calculated from the data that is pooled over the period 1980-2014, with different coloured PDFs representing the different datasets. Vertical long-dashed lines indicate the upper and lower 1.5% threshold. * indicates a different y-axis scale used for regions 13, 14, 23 and 24.

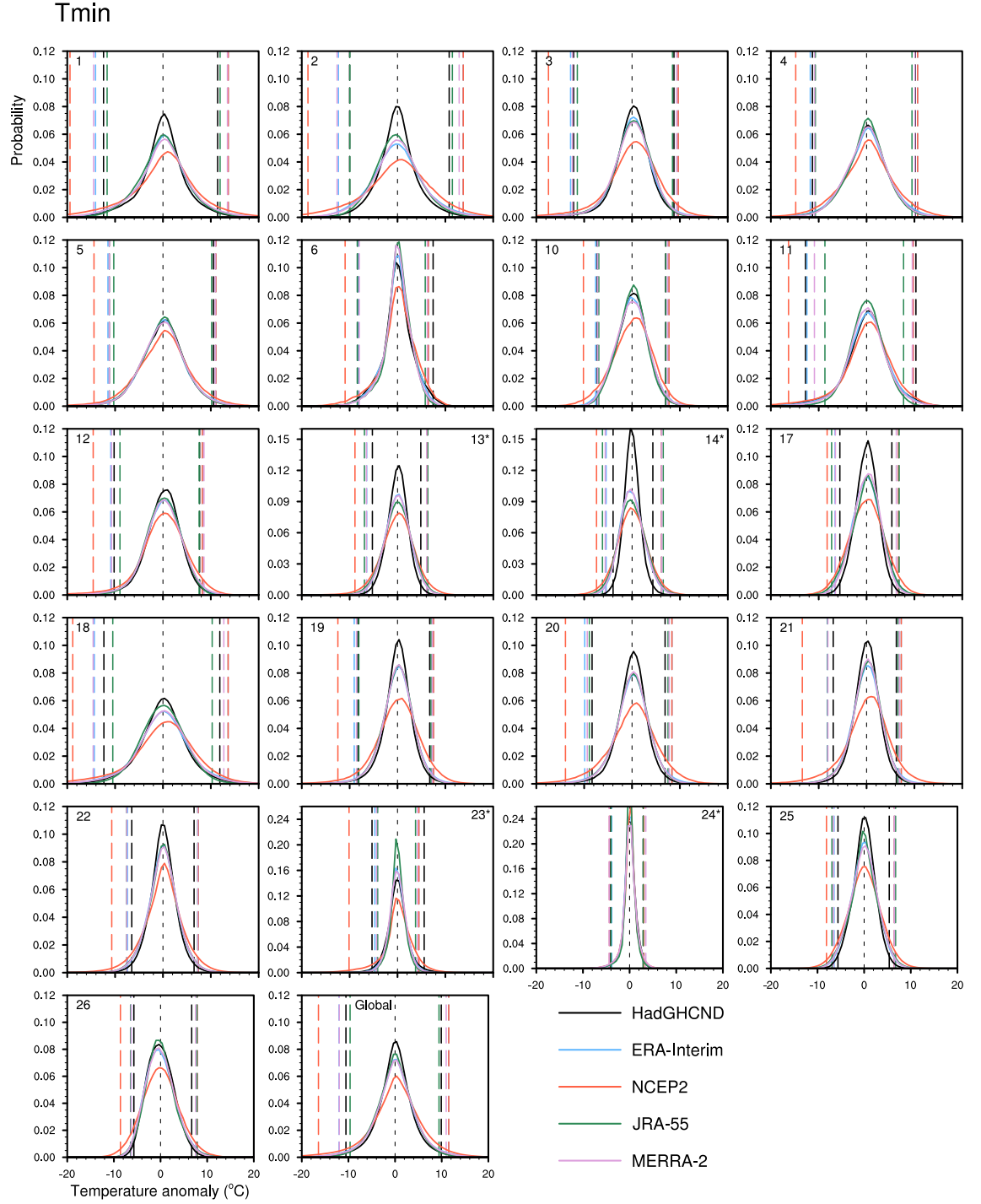


Figure 2.3 As Figure 2.2, but for daily minimum temperature anomalies.

The PDFs also show notable differences in the extremes of the distribution. That is, the 1.5% thresholds that define the tails of the distribution occur at varying temperatures depending on the dataset used, with greater differences found between datasets for Tmin compared to Tmax. For many regions, the thresholds in the

reanalysis products, occur at more extreme temperatures compared to HadGHCND. For instance, in West North America (region 3), the cold extremes threshold for T_{min} is just above -18°C for NCEP2, while it is only just below -12°C for HadGHCND. For most regions as well as the globe, NCEP2-derived PDFs consistently show a bias towards more extreme thresholds, compared with the other datasets. This is particularly apparent in the cold tails compared with the warm tails of the distribution. Though less obvious, PDFs for ERA-Interim, JRA-55 and MERRA-2 also show a bias towards more extreme temperatures thresholds compared with those shown in HadGHCND.

In addition to the higher sensitivities found in T_{min} compared to T_{max} , the tails of the PDFs show sensitivity to the input dataset, particularly in the cold tails. Overall, the NCEP2-derived PDFs are the most different to HadGHCND, however, the PDFs show sensitivity to all four reanalysis products, particularly for assessing standard deviation and skewness.

2.3.2 Time series of the statistical moments

To explore temporal changes and trends in the datasets, Figure 2.4 – 2.8 show time series of the annual mean, standard deviation and skewness for T_{max} and T_{min} for the globe and for selected regions (3, 10, 13 and 18 – see Figure 2.1). The remaining regions and their respective trends are included in supplementary material (Figures S1.2 – S1.7, Tables S1.1 – S1.3). These time series describe changes over 1980-2014. Decadal trends and their significance for each dataset are shown for each plot.

Results for the global mean are relatively robust, irrespective of the dataset used (Figure 2.4), with a significant increasing trend found in all datasets for both T_{max}

and Tmin. Aside from some small differences in NCEP2, the most notable differences occur towards the end of the time series, particularly in MERRA-2 from around 2007-2009. This discrepancy in MERRA-2 is a documented issue (Sánchez-Lugo et al. 2016; Simmons et al. 2017).

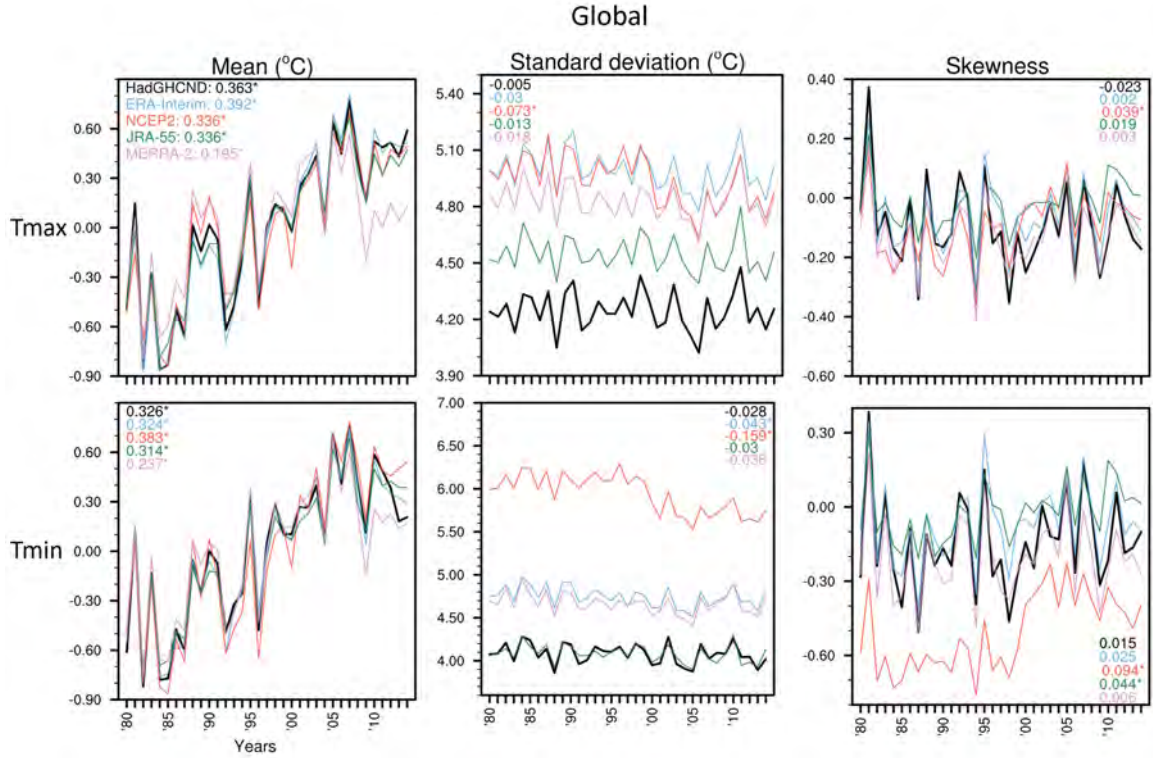


Figure 2.4 Time series of global daily maximum and minimum temperature anomalies for the first three statistical moments annually, that is, the mean, standard deviation and skewness. Decadal trends are shown for each dataset in respective colours. * indicates significant trends at the 5% level.

Unlike the mean, changes in global standard deviation and skewness (Figure 2.4) appear sensitive to the input dataset for assessing trends over the past 35 years, however, there is reasonable temporal correlation between the different time series. There is a noticeably larger spread in standard deviation across the datasets, compared with the mean or skewness. This is particularly apparent in Tmin and NCEP2, where NCEP2 shows a more substantial decreasing trend in global standard deviation in Tmin of -0.16°C per decade, compared with -0.03°C per

decade in HadGHCND. This difference in trends is likely a result of a step change in NCEP2 that occurs around 1998, that is not evident in the other datasets. This inhomogeneity is likely a signature of a step change related to only specific regions, as discussed in subsequent paragraphs. The time series for global skewness show variation and display mostly non-significant trends that are close to zero. Again, the largest differences are shown in NCEP2 for Tmin, with the discussed step change in NCEP2 likely contributing to this.

While the global mean is similar across datasets, differences become apparent for certain regions more than others. For instance, the mean for West North America (region 3 – Figure 2.5) and North Asia/Russia (region 18 – Figure 2.8) is mostly robust regardless of dataset choice, while in contrast, south-eastern South America (region 10 – Figure 2.6) and the Mediterranean (region 13 – Figure 2.7) display more sensitivity. This could be indicative of greater uncertainty within data sparse regions compared to those that are more data rich. For example, regions 3 and 18 include 1795 and 375 station observations respectively, while regions 10 and 13 only have 25 and 138 observations respectively (see Table 2.1). With regards to the Mediterranean region, most available stations are located in southern Europe, with few stations in North Africa (see Figure 2.1). Pooling the grid cells within each SREX region allow us to include region 13 in our analysis, however, the sensitivity of the Mediterranean discussed above might be a reflection of the unequal distribution of observations with data in this region.

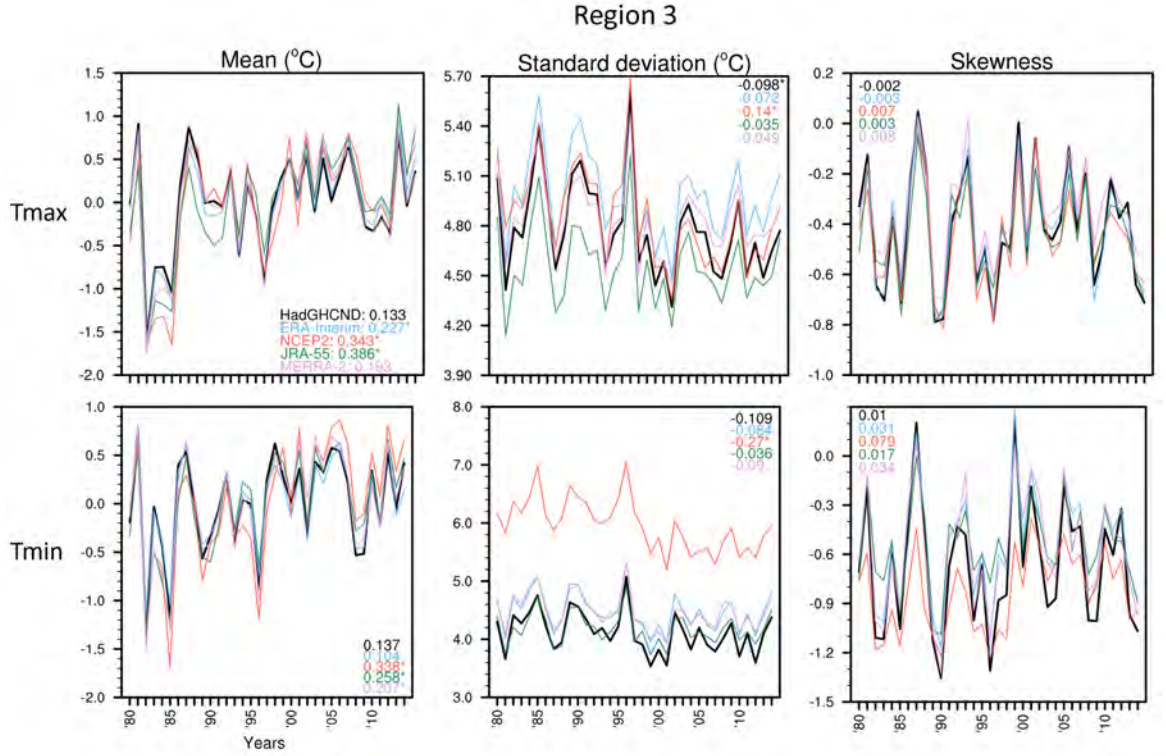


Figure 2.5 As Figure 2.4, but for SREX region 3, located in western North America.

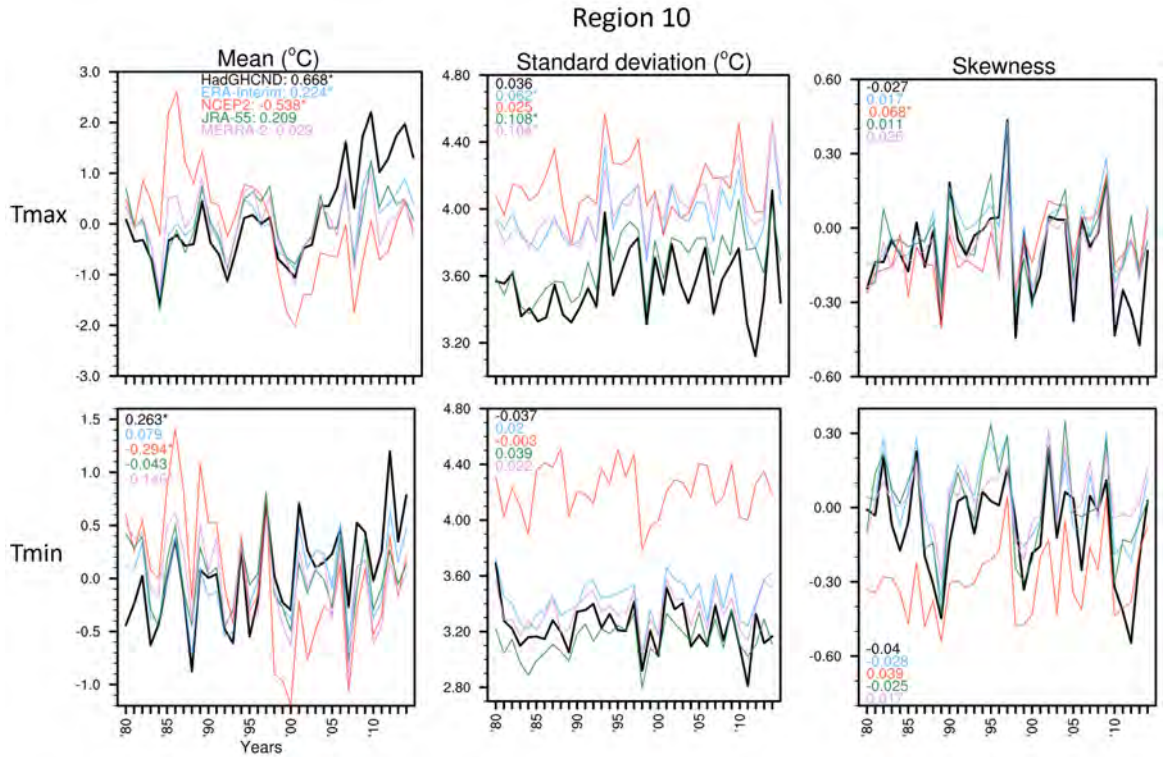


Figure 2.6 As Figure 2.4, but for SREX region 10, located in south-eastern South America.

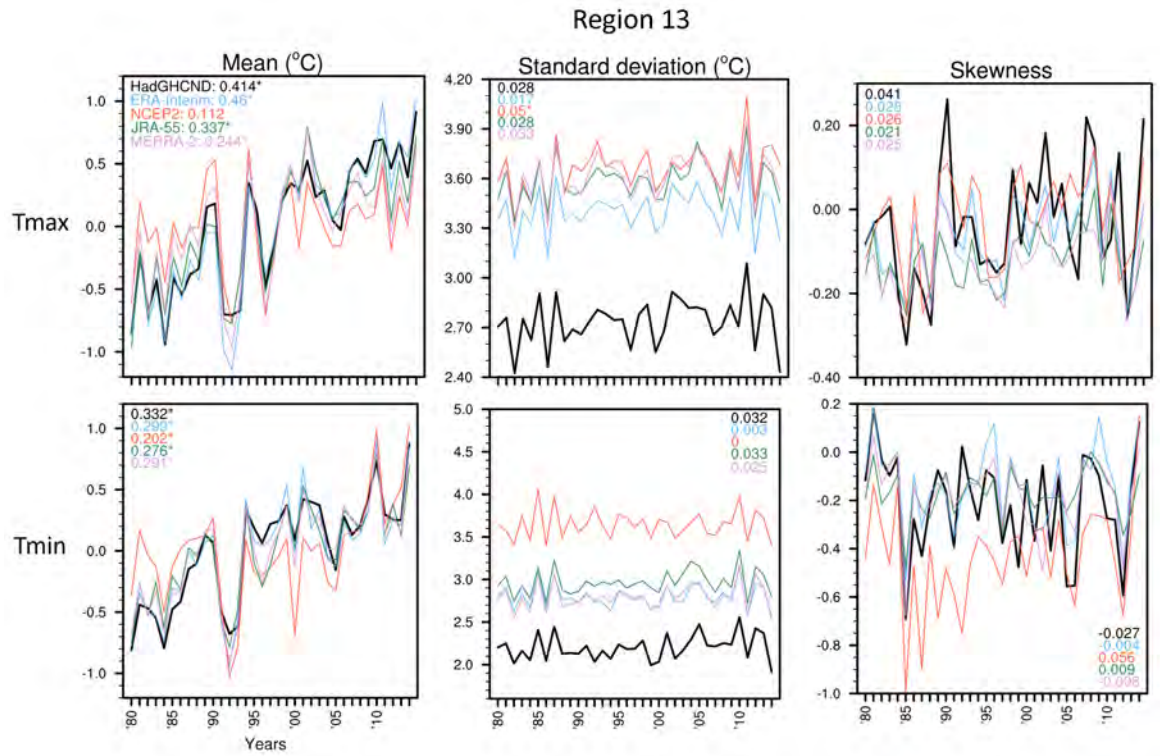


Figure 2.7 As Figure 2.4, but for SREX region 13, located in the Mediterranean region.

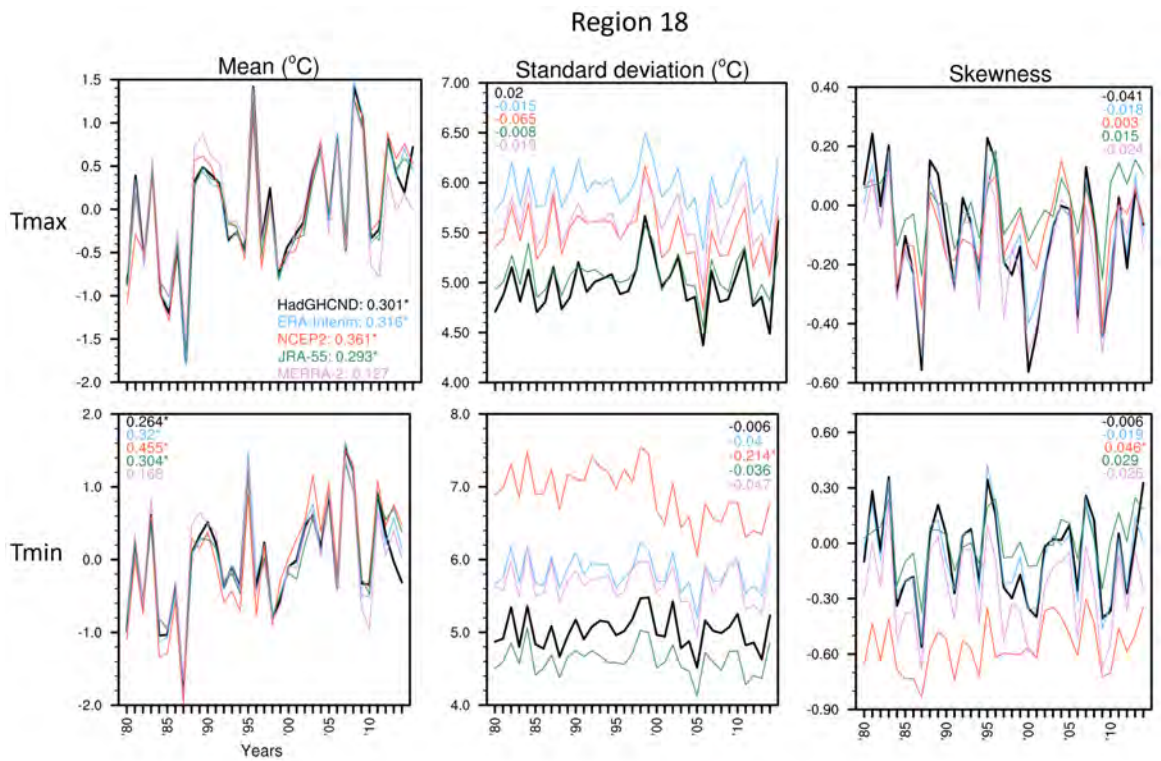


Figure 2.8 As Figure 2.4, but for SREX region 18, located in North Asia/Russia.

The differences in the mean between datasets for regions 10 and 13 are most notable in NCEP2. As for the globe, variance and skewness show a greater sensitivity to dataset choice compared with the mean. Regions 3 and 18, both located in the Northern Hemisphere extratropics, show a clear step change in NCEP2 in standard deviation in Tmin occurring around 1998. After examination of all regions, the step change is found in most regions in this latitude band, as well as in region 10. This inhomogeneity in NCEP2 is dominant enough to have a signature in the global standard deviation. We assume that this step change is artificial, but a metadata search of NCEP2 around this time has not determined the cause. Further discussion on this issue is provided in Section 2.4. While NCEP2 shows the largest differences across the regions, disagreements with HadGHCND are also found in ERA-Interim, JRA-55 and MERRA-2. For example, for Tmin, there is a decreasing trend of -0.04°C in standard deviation for region 10 in HadGHCND, while the other datasets, excluding NCEP2, indicate an increasing trend. In terms of changes in skewness, trends are mostly non-significant and centred around zero. While in region 3 skewness values are similar between all datasets for Tmax, other regions show distinctions, particularly in NCEP2 and for Tmin.

In summary, irrespective of dataset choice, changes in mean Tmax and Tmin are robust for many regions, and a significant increasing trend is found for the globe. For the other statistical moments, a higher sensitivity to dataset choice is clear. The degree of sensitivity, however, is shown to be dependent on the region, with regions of known high quality data being more similar between datasets, with the exception of NCEP2.

2.3.3 Time series of decadal extreme value fits

To investigate changes in the extremes, Figure 2.9 – 2.13 show time series of the location, scale and shape parameters that have been estimated from stationary PP fits over ten year moving windows. The extremal point process model is fitted to the high and low tails of both daily maximum and minimum temperature anomalies ($T_{\max_{\text{high}}}$, $T_{\max_{\text{low}}}$ and $T_{\min_{\text{high}}}$ and $T_{\min_{\text{low}}}$ respectively). As before, extreme values are defined as those data points that exceed a threshold of 1.5% (in the upper or lower tail). We use running decadal windows to ensure sufficient data points are available for a robust statistical analysis.

For the globe (Figure 2.9), significant increasing trends in the location parameter are shown for the high tails of both T_{\max} and T_{\min} across all datasets. Here, a trend of 0.21 per decade ($T_{\max_{\text{high}}}$) and 0.31 per decade ($T_{\min_{\text{high}}}$) is calculated for HadGHCND, with similar trends in most reanalyses. Mostly significant decreasing trends are found for the low tails of T_{\max} and T_{\min} , excluding non-significant trends in HadGHCND and MERRA-2 ($T_{\max_{\text{low}}}$). The time series of NCEP2 for $T_{\min_{\text{low}}}$ again shows a step change occurring around 1998, where the location parameter decreases substantially more than for the other datasets. Though smoother than the step change in NCEP2 shown for global variance (Figure 2.4), because of the use of decadal windows compared with annual data, this suggests the inhomogeneity in NCEP2 is related to the cold tail of the distribution of minimum temperature anomalies. Section 2.4 discusses this in more detail. Overall, the location parameter for the globe indicates increasing trends in the high tails, along with decreasing trends in the low tails. This indicates a shifting distribution towards warmer conditions, with both hot and cold extremes becoming warmer.

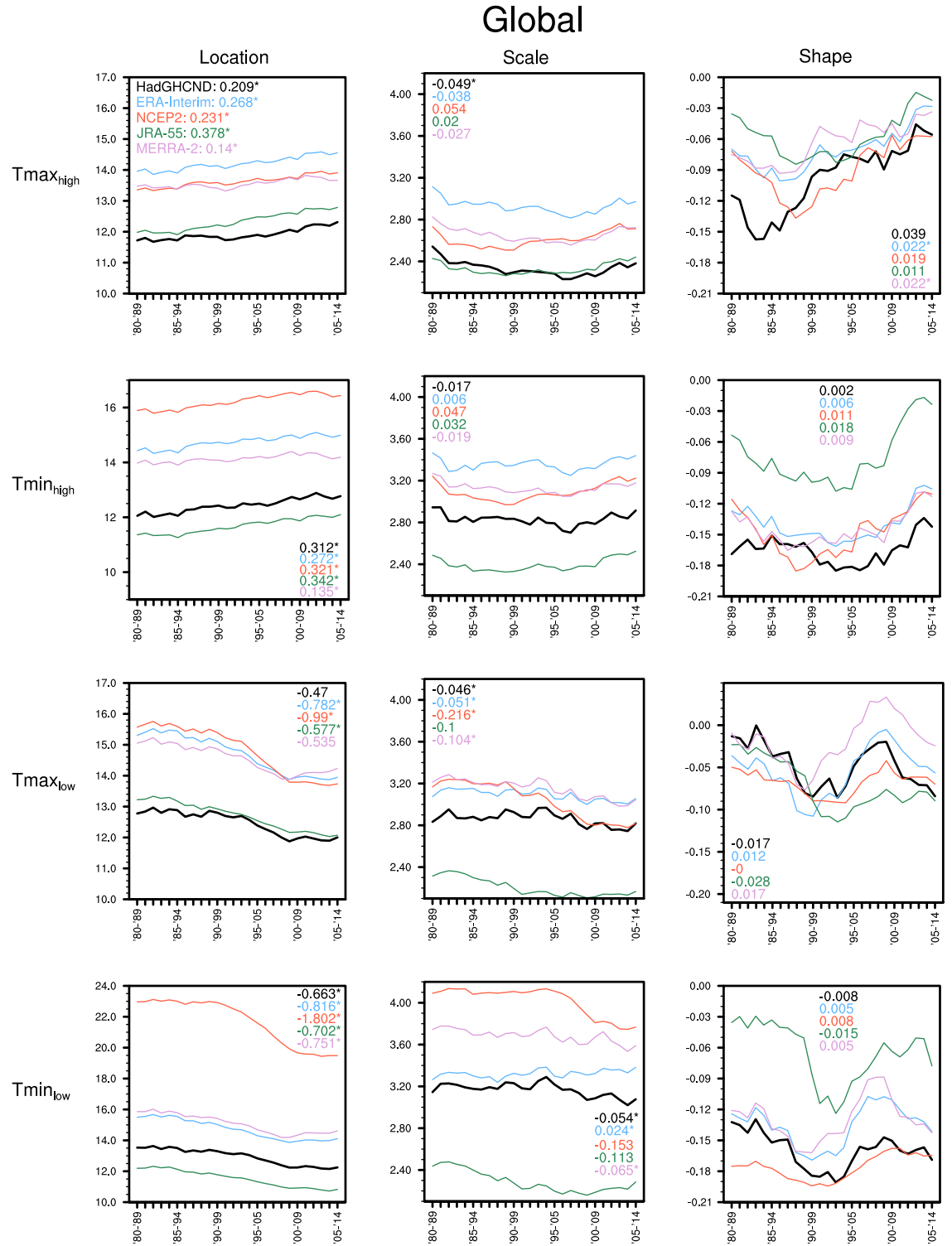


Figure 2.9 Global time series using running decadal windows of the extreme value parameters, including location, scale and shape, calculated for the high tails of maximum and minimum temperature anomalies (Tmax_{high} and Tmin_{high}, respectively), and for the low tails of maximum and minimum temperature anomalies (Tmax_{low} and Tmin_{low}, respectively). The x-axis shows years, starting with the 1980-1989 window, and ending with 2005-2014, and parameter estimates are indicated on the y-axis. Decadal trends are indicated in the respective dataset colours for each parameter and temperature variables. * indicates significant trends at the 5% level.

Trends in the global scale parameter (Figure 2.9) are close to zero for all datasets and temperature variables, although slightly decreasing significant decadal trends in HadGHCND are shown for all variables except $T_{\min_{\text{high}}}$. There is some sensitivity in scale to the input dataset, with $T_{\min_{\text{low}}}$ showing the most differences between datasets compared with the other temperature variables. As for the location parameter, NCEP2 shows a sudden decrease in scale for $T_{\min_{\text{low}}}$ starting with the 1994-2003 window, that is not apparent in the other datasets. This is also shown to a lesser extent in $T_{\max_{\text{low}}}$. Other notable differences include a steeper decreasing trend in JRA-55 and MERRA-2 compared with HadGHCND in both low tails, and a slightly increasing significant trend in ERA-Interim in $T_{\min_{\text{low}}}$. Trends in global shape are mostly non-significant and close to zero, excluding a significant increasing trend in ERA-Interim and MERRA-2 ($T_{\max_{\text{high}}}$). For all temperature variables, the reanalyses show different temporal patterns to HadGHCND. Again, this is particularly evident in T_{\min} , and in the low tails.

As shown for the statistical moments in the previous subsection, the sensitivities in the time series analyses differ between both region and temperature variable. For the location parameter, NCEP2 consistently stands out as being the most different across all regions (including those in the Supplementary Material - Figures S1.8 – S1.19, Tables S1.4 – S1.9). The step change shown in global location for $T_{\min_{\text{low}}}$ is present for both West North America (region 3 – Figure 2.10) and North Asia/Russia (region 18 – Figure 2.13), and most other Northern Hemisphere extratropical regions, but is not a feature in the time series for south-eastern South America (region 10 – Figure 2.11) and the Mediterranean (region 13 – Figure 2.12). This suggests that the apparent inhomogeneity in the global time series of NCEP2 is likely related to the cold extremes for Northern Hemisphere extratropical regions.

We note that we do not further discuss long-term trends in NCEP2 due to this inhomogeneity.

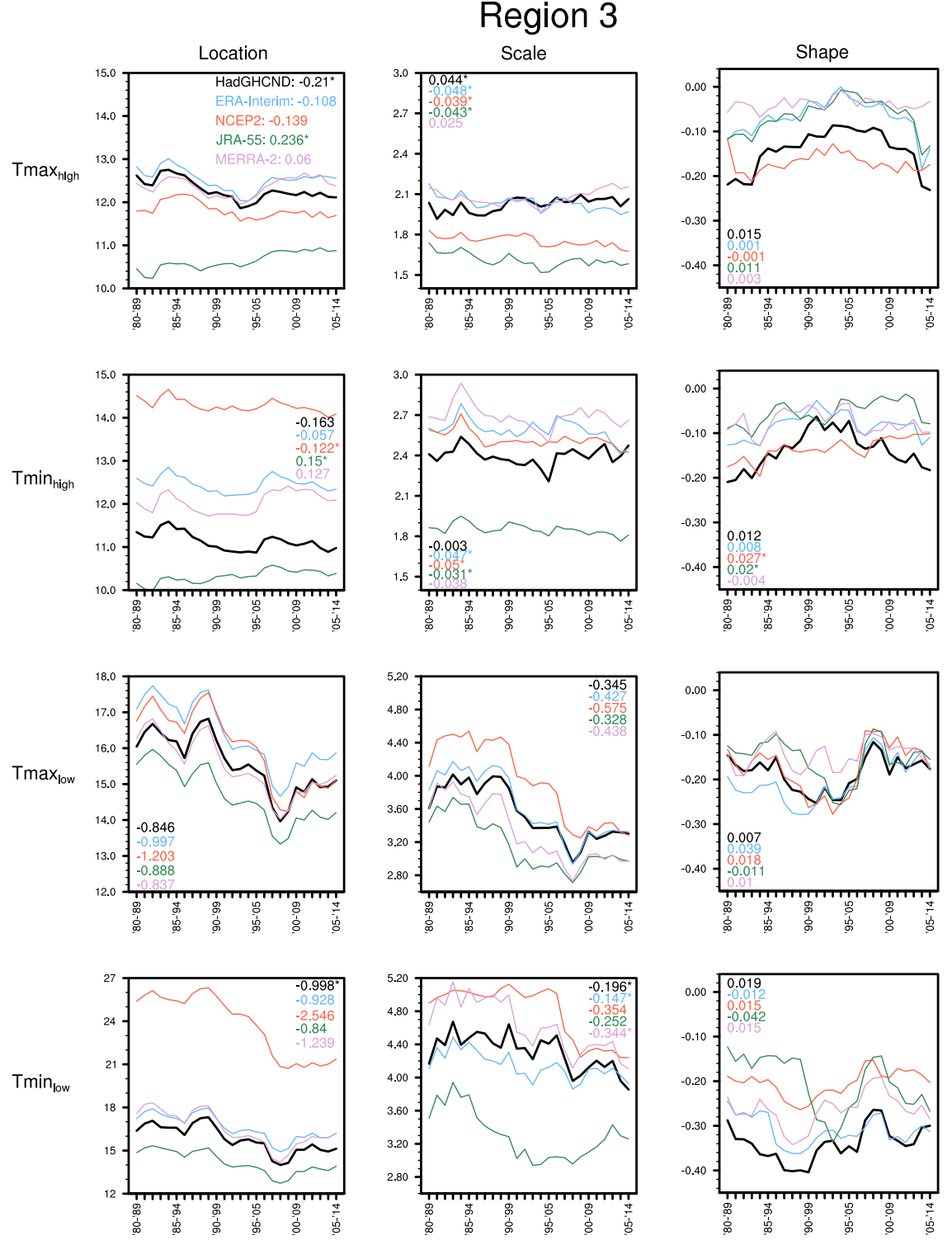


Figure 2.10 As Figure 2.9, but for SREX region 3 (west North America).

Overall, the location parameter shows sensitivity to all reanalysis products for all regions, with more data sparse regions showing the least consistency between datasets. For example, for region 10, HadGHCND shows a steeper decrease in location towards the end of the time series for $T_{\max_{\text{low}}}$, compared to the other datasets. Region 23, another data sparse region located in South Asia, shows a substantial step change in NCEP2 in location for $T_{\max_{\text{low}}}$, that is not present in the other datasets (see Supplementary Material Figure S1.10). Trends in scale and shape also differ between datasets. For example, the reanalyses show clear deviances from HadGHCND in scale and shape in the cold tails for region 10, and in more data rich regions, such as region 3 and 18. JRA-55 shows a more substantial decreasing trend in scale for $T_{\min_{\text{low}}}$, compared with the other datasets (excluding NCEP2 which is affected by the discussed inhomogeneity). Other key differences are found for region 18 for $T_{\max_{\text{low}}}$. Here, HadGHCND increases in scale from 1988-1997, during which JRA-55 shows a decreasing trend in scale. ERA-Interim and MERRA-2 capture the temporal pattern of HadGHCND here, though to a lesser extent. However, differences between these products and HadGHCND are apparent for other regions, for example, changes in shape for $T_{\min_{\text{low}}}$ in region 13. In general, the reanalyses are distinct from HadGHCND in both scale and shape.

In summary, changes in the tails of the distribution are sensitive to dataset choice. As for the analyses of the entire distribution, more overall differences in the datasets are shown for minimum temperature anomalies, with differences particularly evident in the scale and shape parameters. Analyses of changes in the extremes correspond with these findings, and more specifically, show that the largest dataset inconsistencies occur in the cold tails of the distribution of minimum temperatures.

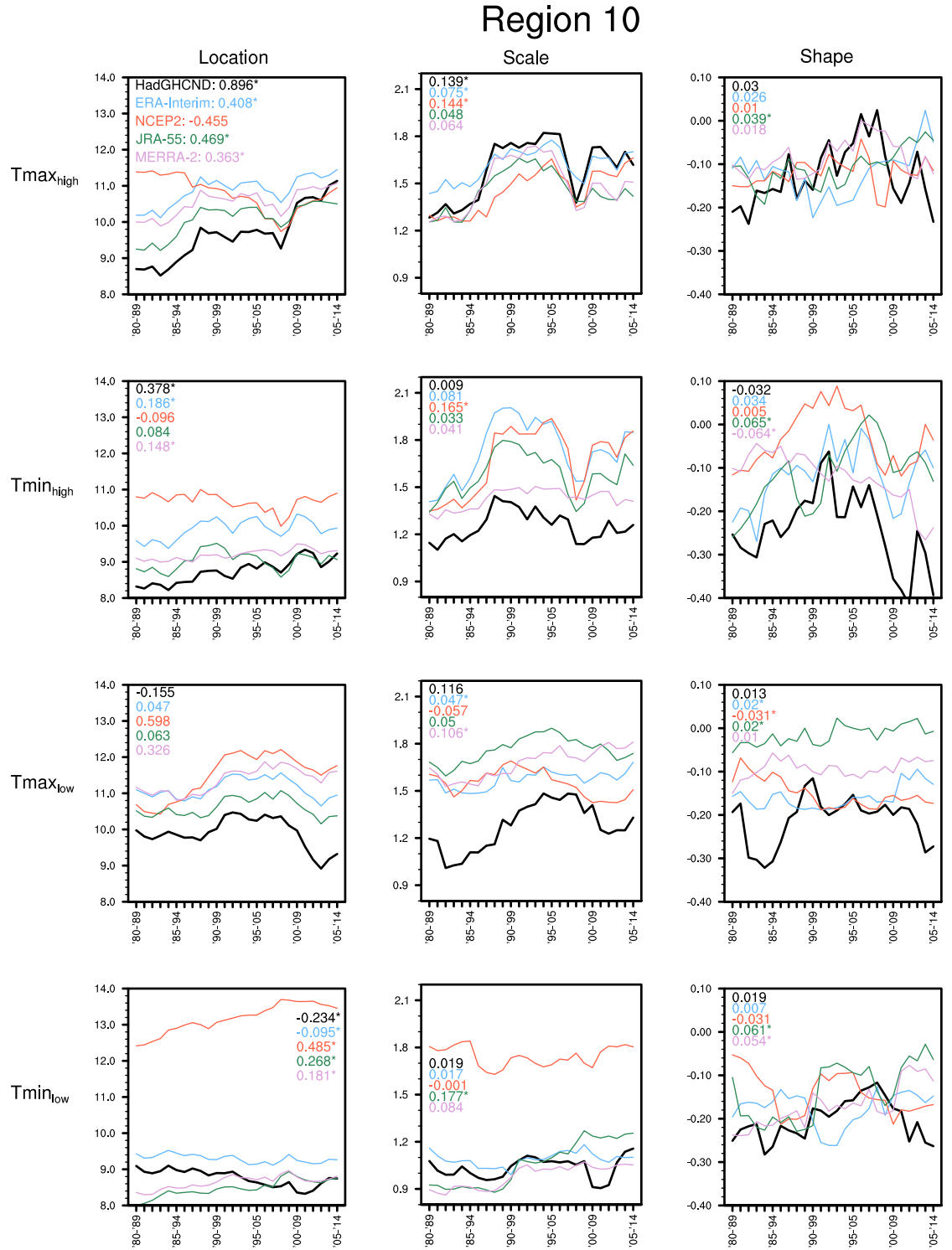


Figure 2.11 As Figure 2.9, but for SREX region 10 (south-eastern South America).

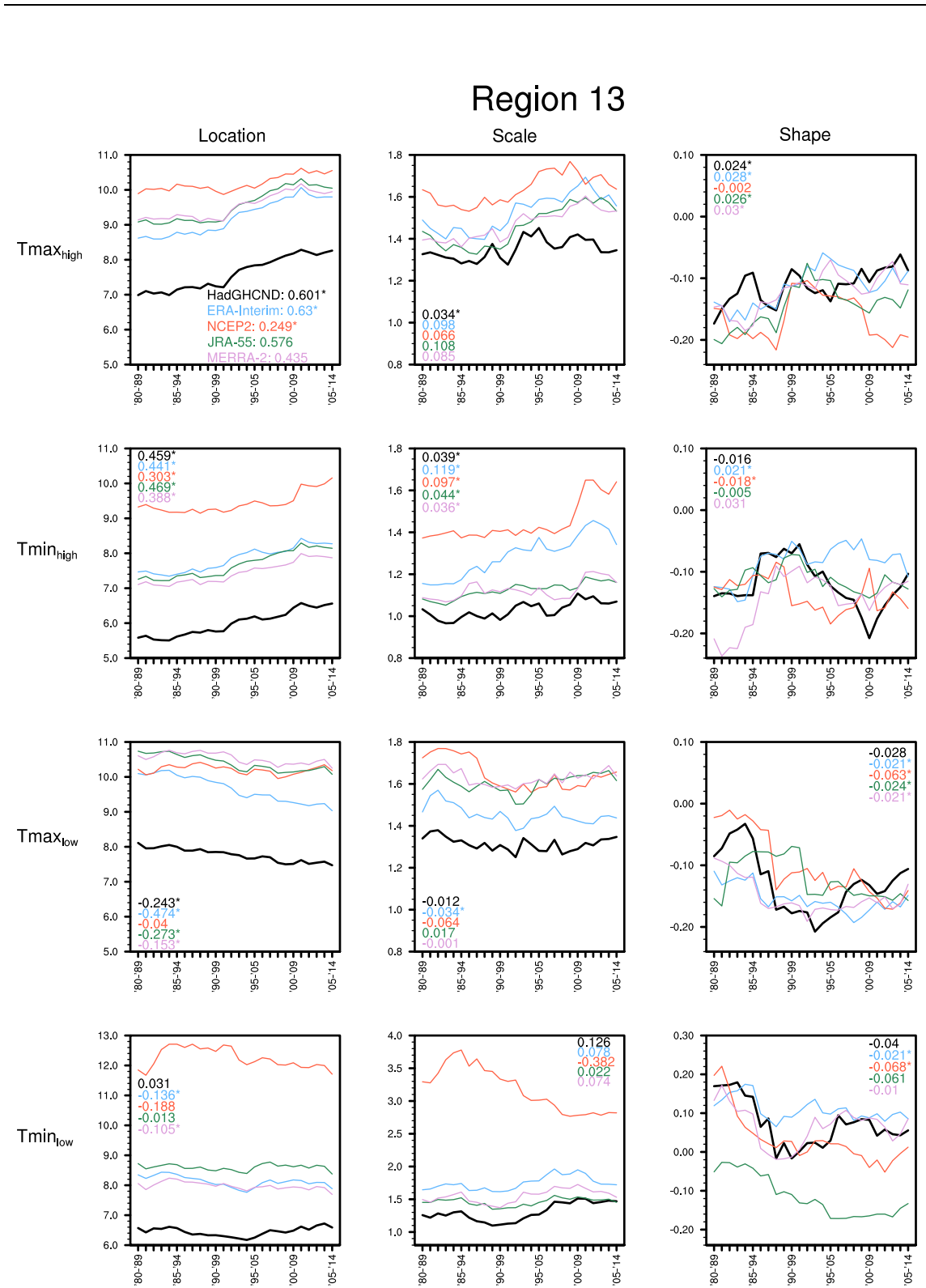


Figure 2.12 As Figure 2.9, but for SREX region 13 (Mediterranean).

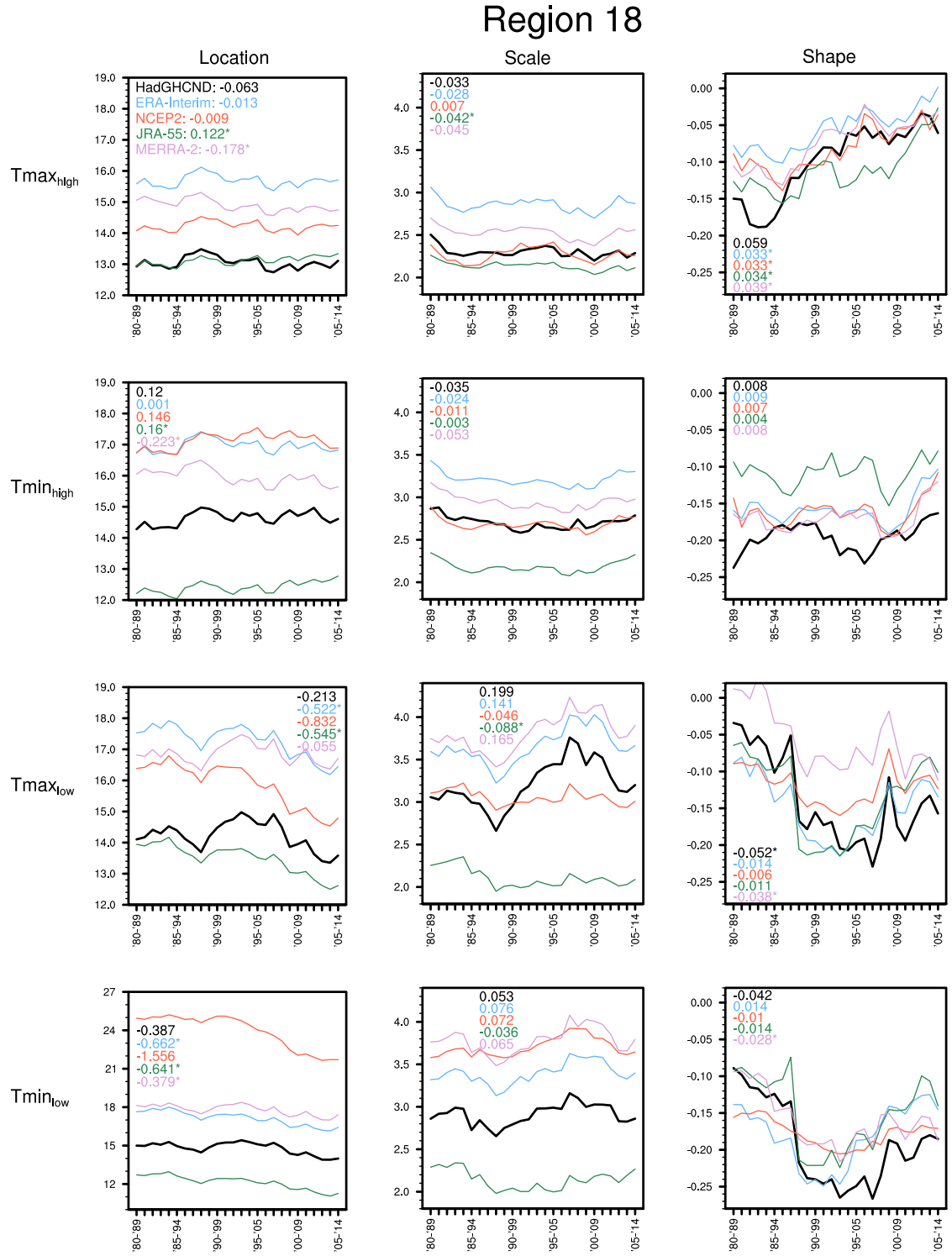


Figure 2.13 As Figure 2.9, but for SREX region 18 (North Asia/Russia).

2.3.4 Changes in extremes relative to changes in the mean

Using the above methods together, that is, jointly assessing both the entire distribution as well as the tails, can provide insight into how extremes are changing with respect to the mean. Figure 2.14 – 2.15 show the difference between the trend in the non-stationary PP fit of the location parameter and the trend in the annual mean temperature anomalies. NCEP2 is not included due to the inhomogeneity present in the time series, making it inappropriate to assess trends over a 35-year period. The trend differences (calculated per decade for the period 1980-2014) represent ‘excess trends’ and are a useful way of describing regional differences in the rates of warming between extremes and the mean (Brown et al. 2008). As in the subsequent figures, when extremes are warming faster (slower) than the mean, the excess trend is positive (negative). Stippling indicates grid boxes that are significant at the 5% level. For additional robustness, we also calculated the difference between the 98.5th quantile (1.5th quantile for the cold tails) and the mean (see Supplementary Material Figure S1.20). These results are very similar to the excess trends shown here.

For the warm extremes ($T_{\text{max}_{\text{high}}}$ - Figure 2.14), most Northern Hemisphere high latitude regions indicate that the mean has been warming faster than the extremes, irrespective of the product used. This is slightly underestimated by JRA-55. The datasets agree that warm extremes have been warming significantly faster than the mean in the Mediterranean region, south-western and East Asia, South America, Africa and parts of Australia. For some of these regions, such as the Mediterranean, the reanalyses show smaller magnitude excess trends in the extremes compared to HadGHCND, with rates of around 0.1°C to 0.2°C (reanalyses) compared to 0.4°C to 0.6°C (HadGHCND). Other regions, such as South America, show greater magnitude positive excess trends in the reanalyses compared with HadGHCND.

Differences are also found for parts of northern Canada, where a positive excess trend is shown in HadGHCND. This same area is opposite in sign for ERA-Interim, while smaller and non-significant in JRA-55 and MERRA-2.

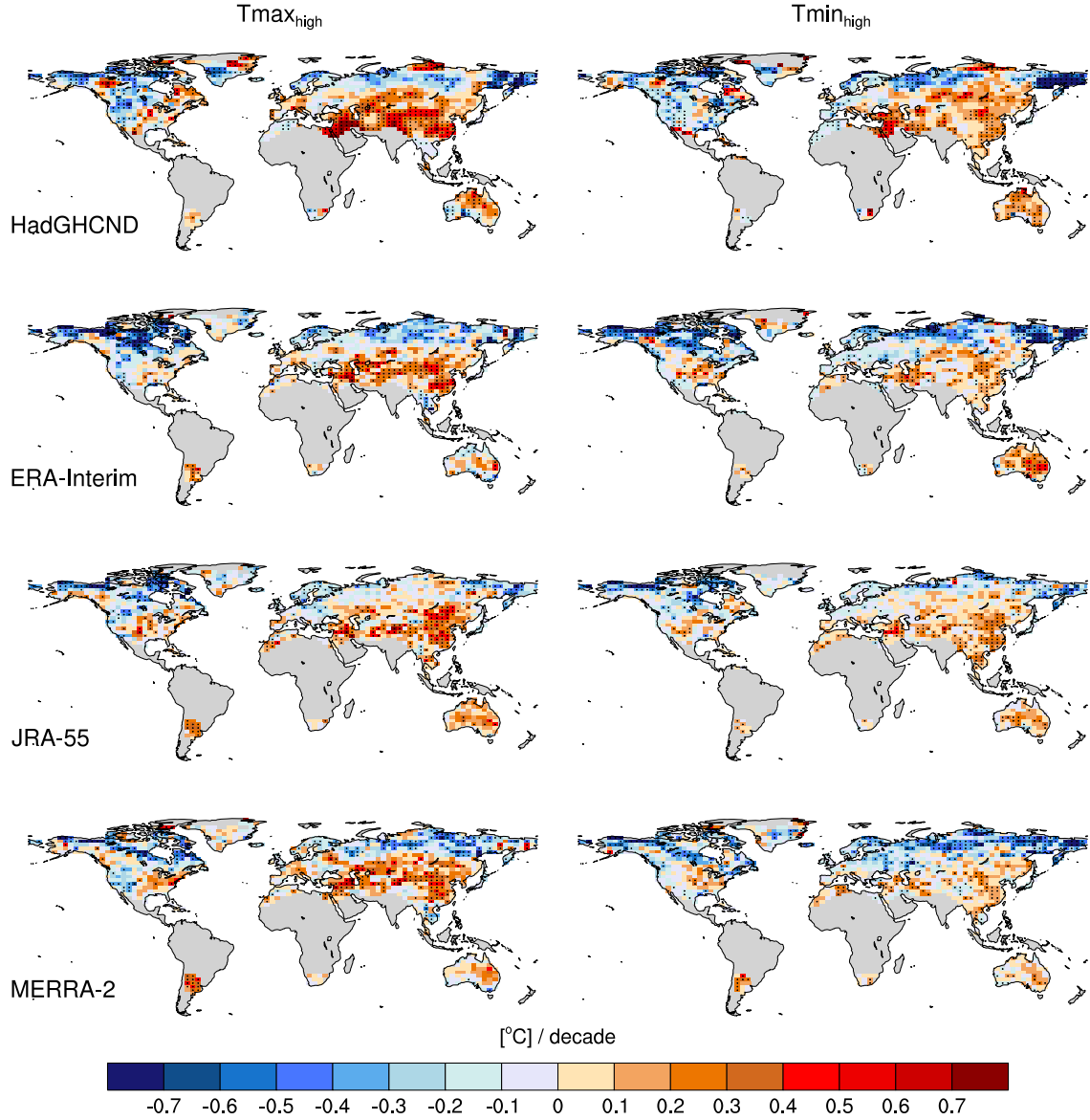


Figure 2.14 'Excess trends', that is, the difference between the trend in the location parameter and the trend in the annual mean for HadGHCND and the reanalyses, for the high tails of maximum and minimum temperature anomalies ($T_{\max_{\text{high}}}$ and $T_{\min_{\text{high}}}$, respectively). Positive excess trends represent regions where the extremes are warming faster than the mean, while negative excess trends show regions where the mean is warming faster than the extremes. Decadal trends ($^{\circ}\text{C}$) are calculated over the period 1980-2014. Stippling indicates grid boxes that are significant at the 5% level.

Excess trends in $T_{\min_{\text{high}}}$ (Figure 2.14) are similar to those of the warm extremes, in terms of both spatial patterns and sign of trend, as well as dataset agreement, though positive excess trends are generally of a smaller magnitude to those of the warm extremes. Some differences in the datasets are shown for Australia, for example, where ERA-Interim shows greater positive excess trends in parts of eastern Australia compared with HadGHCND, while MERRA-2 and JRA-55 tend to show smaller magnitude excess trends than HadGHCND for much of the country.

The datasets differ the most for excess trends in the low tails (Figure 2.15). Positive excess trends in extremes are shown for much of the Northern Hemisphere extratropics, particularly in North America. Negative excess trends are found for parts of Europe and Asia, while small negative excess trends are shown for some tropical and Southern Hemisphere regions. For $T_{\max_{\text{low}}}$, most regions with positive excess trends show larger magnitude trends in the reanalyses than for HadGHCND. Conversely, regions showing negative excess trends in HadGHCND show smaller magnitude excess trends in the reanalyses, such as in parts of Europe and West Asia. Broadly, JRA-55 shows the most differences to HadGHCND, with larger significant positive excess trends for much of the Northern Hemisphere extratropics and smaller negative trends in Europe. For example, positive excess trends greater than 0.7°C per decade are shown for some parts of Russia in JRA-55, with the same area showing negative excess trends between -0.2°C to -0.3°C in HadGHCND. ERA-Interim and MERRA-2 also differ from HadGHCND in this area, showing mixed excess trends.

For the cold extremes ($T_{\min_{\text{low}}}$ - Figure 2.15), the largest disagreements occur over Russia and the European continent. As for $T_{\max_{\text{low}}}$, JRA-55 shows greater positive

excess trends over Russia, and smaller negative excess trends over West Asia and Central Europe, compared to the other datasets.

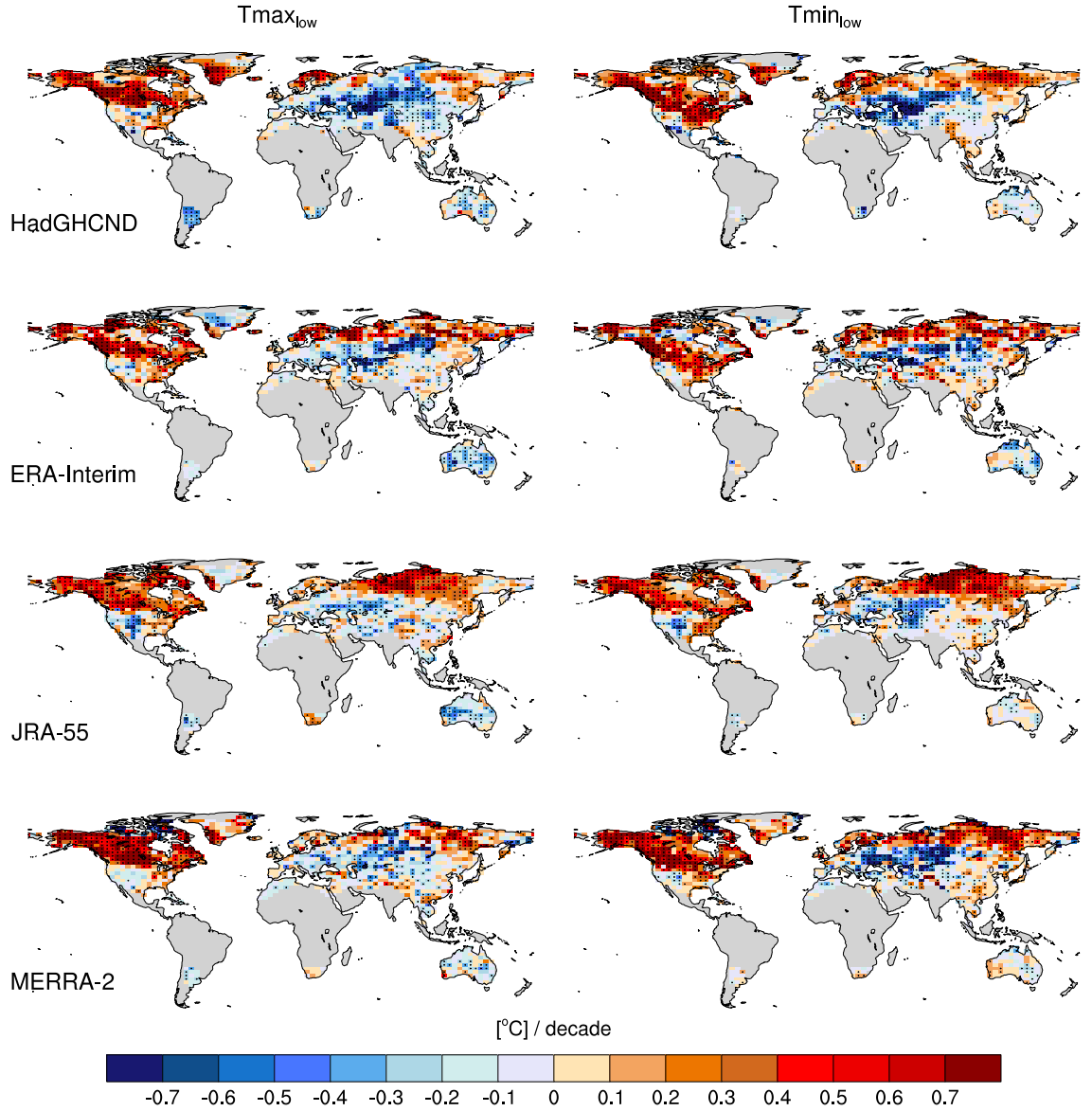


Figure 2.15 As Figure 2.14, but for the low tails of maximum and minimum temperature anomalies (Tmax_{low} and Tmin_{low}, respectively).

Overall, for the warm extremes, some subtropical regions show that the extremes have been warming at a faster rate than the mean over 1980-2014, while for the cold extremes, positive excess trends are found for much of the Northern Hemisphere extratropics. Spatially, all datasets tend to agree on this, however, the

magnitude of the excess trend differs. The warm tails of the distribution tend to be similar between datasets across the globe compared with the cold tails, which show greater sensitivity to dataset choice.

2.4 Discussion

The choice of dataset for investigating changes in temperature variability and extremes can affect conclusions regarding changes in the temperature distribution. The use of different datasets in previous work along with some differences in conclusions provides motivation for a systematic approach to determining the sensitivities of an analysis to the input datasets used. We acknowledge that non-climatic artefacts can potentially affect a spatially averaged dataset such as HadGHCND, including changes to the station network, data quality and homogeneity issues, and gridding uncertainties in terms of spatial averaging of point data, particularly for analysing extremes (Donat et al. 2014; Dunn et al. 2014; Director and Bornn 2015). In addition, HadGHCND uses a large CLS and is gridded using a relatively coarse resolution, so resultant values tend to be spatially smoothed (Caesar et al. 2006), further adding to the inherent uncertainties of using this type of dataset. However, given that it is interpolated using long-term *in situ*-based data, we are more willing to assume that there are fewer inhomogeneities than those introduced from assimilation using a highly variable network of data in the reanalyses.

Irrespective of what product is used, trends in the mean temperature are mostly robust. This suggests that changes in the mean are not particularly sensitive to dataset choice. However, some datasets show more differences in the mean than others, depending on the region, for example, in the Mediterranean region, and in more data sparse regions such as parts of South America. Differences in the mean

for these regions might be due to uncertainties in the observational data itself, rather than the ability of reanalysis data to reproduce the observations. It is also possible that the assimilation is poorly constrained by observations. Other statistical parameters aside from the mean, including the standard deviation and skewness, as well as those related explicitly to the extremes, show more sensitivity to the input dataset. As in the mean, regions with higher data uncertainty due to sparse observations, such as South America, show more disagreement between datasets than regions that are known to have high quality data, as in North America.

Much of the recent literature regarding changes in temperature variability has highlighted the importance in understanding differences in regional changes in temperature compared with the global mean. It is already clear that some regions are warming at different rates compared to the global average (e.g. Sutton et al. 2015; Seneviratne et al. 2016). Here, we show that the globally-averaged (where data are sufficiently complete) maximum temperature trend in HadGHCND is 0.36°C per decade over 1980-2014, while, for example, south-eastern South America shows a trend of 0.67°C per decade (Figure 2.4 and Figure 2.6, respectively). In addition to the mean, we also see regional differences in the trend of extremes relative to the local mean. We show, for example, that cold extremes are warming faster than the mean for much of the Northern Hemisphere, consistent with Arctic amplification (Screen 2014; Rhines et al. 2017). For the warm extremes, some subtropical regions show a faster trend in the extremes compared with the mean. Additionally, excess trends in the cold extremes are generally greater in magnitude over North America and parts of North Asia/Russia than they are for warm extremes, consistent with studies that have found stronger increases cold extremes compared with warm extremes (e.g. Donat et al. 2013b). This is robust for all

datasets, however, the magnitude of excess trend differs substantially depending on the variable and region. This is a significant finding, as regions are disproportionately impacted by extremes due to differences in socioeconomic factors that can increase a regions vulnerability to extreme events (IPCC 2012). The consequences of higher magnitude extremes can potentially further exacerbate the unequal distribution of impacts on regional scales.

Overall, ERA-Interim appears to most closely resemble HadGHCND. This is consistent with other studies. For example, Donat et al. (2014) found that ERA-Interim closely resembled trends and inter-annual variations in temperature extremes shown in interpolated observational datasets for the period 1980 – 2010, noting the higher dataset consistency in the most recent three decades, in contrast with the large disagreements between datasets in the pre-satellite era. Excluding an inhomogeneity in MERRA-2 in the global mean for the most recent years (see Figure 2.4), MERRA-2 and JRA-55 resemble mean trends in the observations reasonably well, even for more data sparse regions (see Figure 2.6). However, when assessing aspects of temperature other than the mean, there are still distinct differences between the observation-based product and the reanalyses, and so caution should be exercised if using these products to investigate changes in variability and extremes.

NCEP2 is clearly the most different to the other datasets, caused partly by the inhomogeneity shown to occur around 1998-1999, but also shown in more short-term variations, for example, in the mean for south-eastern South America (Figure 2.6). Previous studies have noted the differences in NCEP2 compared with other reanalyses (e.g. Kharin et al. 2007; Kharin et al. 2013; Sillmann et al. 2013b). Being a low-resolution, first-generation product (Kanamitsu et al. 2002; National Center

for Atmospheric Research (NCAR) 2016), this might be expected, however, perhaps not to the extent shown here. While many errors were rectified from the original NCEP1 product, some suggest a poor representation of the Southern Hemisphere in NCEP2 (NCAR 2016). Here, we show large differences in many regions around the globe. For example, the step change around 1998-1999 shown for the globe is particularly noteworthy in some Northern Hemisphere high latitude regions. Potential explanations could be related to changes in the assimilated data affecting snow cover, as these step changes are inconsistent with other datasets and are particularly notable for the cold extremes. Although snow cover data was not available in NCEP2 for the required years, we were able to investigate the water equivalent of accumulated snow depth, which is a variable that can suggest snow cover changes through the amount of snow depth (see Figure S1.22). A clear step change in the mean and standard deviation is apparent at 1998-1999 for regions 3 and 18, where a substantial decrease occurs after 1998.

The similarities or differences between the reanalysis products themselves as well as with HadGHCND are affected by their independence of *in situ*-based measurements. For example, NCEP2 does not incorporate any screen-level data to assimilate near-surface temperature (Simmons et al. 2004), unlike the other reanalyses used here. ERA-Interim and JRA-55 both interpolate between *in situ* data at the screen-level, and the atmospheric levels of the reanalysis model (Dee et al. 2011; Kobayashi et al. 2015). As such, ERA-Interim and JRA-55 are not independent of surface measurements, and so it may be expected that they are more similar to one another than to NCEP2. As in NCEP2, MERRA-2 does not assimilate using near-surface temperatures (Bosilovich et al. 2015). Despite MERRA-2 being the latest product used in this study, ERA-Interim and JRA-55 show more overall consistency between themselves as well as HadGHCND in many

instances, similarly found in other comparison studies (e.g. Simmons et al. 2017). However, short-term variations, especially in the mean for MERRA-2, show close similarities to both ERA-Interim and HadGHCND.

2.5 Conclusions

Global daily maximum and minimum temperatures are increasing. This conclusion is robust regardless of the dataset used. Regional increases in the mean are mostly robust to the dataset, however, there is a slightly greater sensitivity to the input dataset than for global assessments. Other characteristics of the temperature distribution show substantial sensitivity to the dataset used, highlighting some of the uncertainties involved in addressing changes in temperature variability. Assessing temperature extremes also displays sensitivity to the dataset choice, with results showing sometimes substantially different magnitude changes in extremes.

Irrespective of the approach used, differences between datasets regarding distribution changes are the greatest in the cold tails of the distribution and for daily minimum temperature anomalies. This suggests potentially greater uncertainties for this part of the distribution that need to be considered for future work investigating changes in the distribution of temperature. Further investigation is required to understand why greater sensitivities are shown in the cold tails. For all temperature variables, NCEP2 has the largest inconsistencies, compared with the gridded *in situ*-based observational dataset, HadGHCND. However, the ‘higher generation’ reanalysis products, that is, ERA-Interim, JRA-55 and MERRA-2, also still show distinct differences in the higher-order moments. Dataset disagreement is generally largest for regions that are more data sparse, such as south-eastern South America and southern Africa, while better agreement between datasets is

found for regions that are data rich and known to have higher quality data, for example, North America.

Despite inconsistencies in the results depending on the dataset, all products show that cold extremes have been warming at a faster rate than the mean for much of the Northern Hemisphere extratropics, while warm extremes have been warming faster for many subtropical regions. We must be able to make robust conclusions regarding the regional differences in rates of warming of extremes and the local mean, as future planning and local adaptation strategies rely on it. Future work in this field would benefit from using datasets of longer time-scales to provide more robust trend estimates.

This paper provides a first step in documenting the inconsistencies regarding changes in temperature variability and extremes and, for example, will help in making the best dataset choices for model evaluation moving forward. In addition, by understanding these preliminary data issues, more confident and robust conclusions can be made to understand changes in the characteristics of the temperature distribution, and therefore provide critical information for future planning for extremes.

CHAPTER 3

Changes in daily temperature extremes relative to the mean in CMIP5 models and observations

Chapter overview

This chapter has been accepted for publication as: Gross, M. H., M. G. Donat, and L. V. Alexander, 2019: Changes in daily temperature extremes relative to the mean in Coupled Model Intercomparison Project phase 5 models and observations. *International Journal of Climatology*, accepted.

The rates of warming in extremes and average temperatures vary both spatially and temporally, with strong geographic and seasonal differences in how local extremes are changing relative to local mean temperatures. This chapter explores the different rates of change in seasonal warm and cold extremes relative to corresponding seasonal mean temperatures, based on daily data. It builds on the final results of *Chapter 2*, where trends in annual extremes relative to the annual mean were explored, and uses the same observational dataset, HadGHCND, to evaluate a suite of CMIP5 models. The climate models are then used to assess future excess changes in seasonal warm and cold extremes relative to seasonal mean temperatures. The greatest differences in the rates of change between extremes and the mean in recent decades occur in the cold tails of the distribution for some

regions in the Northern Hemisphere extratropics during all seasons except boreal summer. For recent changes, climate model simulations have relatively low spatial correlations with each other, as well as with the observations. Future simulated changes are shown to be systematic and robust, with a clear signal in the warming of extremes relative to the mean. Boreal winter, spring and autumn months show especially pronounced excess changes in the Northern Hemisphere mid- to high-latitudes, where the models show robust agreement in cold extremes warming at least 3°C more than mean temperatures in the late 21st century compared to the mid-20th century. By comprehensively assessing past and future rates of change in seasonal extremes relative to the mean at the global scale, this chapter provides an in-depth understanding of the disproportionate rates of change that can in turn have important implications on the probability of extreme events and therefore their impacts.

3.1 Introduction

The implications from changes in daily temperature extremes have the potential to be far greater for society and the environment than increases in the mean temperature alone. Extremes occur in the tails of the temperature distribution, and changes in each of the tails, along with shifts in the mean, can affect the entire shape of the distribution which is related to the probability of extreme events (Mearns et al. 1984). Most global studies have found that changes in the tails of the distribution are primarily related to changes in the mean, with little or no change in variability and shape (e.g. Simolo et al. 2011; Huntingford et al. 2013; Rhines and Huybers 2013; Tingley and Huybers 2013; Weaver et al. 2014; McKinnon et al. 2016). However, regional-scale studies using observational data have found that there are some changes in variance and shape, depending on the

location. For example, there is evidence for changes in symmetry, or skewness, and variability of daily temperature for some mid- to high-latitude regions in the Northern Hemisphere (Kharin and Zwiers 2005; Donat and Alexander 2012; Cavanaugh and Shen 2014; Screen 2014; Loikith and Neelin 2015; McKinnon et al. 2016; Rhines et al. 2017). This highlights that changes in temperature distributions vary depending on geographic location, and importantly, policymakers and adaptation planning rely on information that is region-specific. The temperature distribution also has a strong seasonal component, as well as dependencies related to daytime and night-time temperatures (Rhines et al. 2017). A comprehensive global-scale study investigating seasonal changes in regional extremes in both daily maximum and minimum temperature distributions, in relation to changes in corresponding mean temperatures, is still lacking. This type of study would be useful in connecting previous work that has focused only on specific regions or specific aspects of the temperature distribution (e.g. Cavanaugh and Shen 2014; McKinnon et al. 2016; Rhines et al. 2017), providing a complete description of both recent and future changes in extremes and mean temperatures.

It is expected that future changes in temperature extremes by the end of the 21st century will be much larger than observed past increases if no drastic measures to limit atmospheric greenhouse gas concentrations can be implemented within the next decades. This would be associated with increases in the frequency of warm extremes, and decreases in the frequency of cold extremes (Coumou and Robinson 2013; Kharin et al. 2013; Kirtman et al. 2013), consistent with a shift in the temperature distribution towards warmer temperatures. But it has also been suggested that cold extremes are warming faster than warm extremes (e.g. Kharin and Zwiers 2005; Donat and Alexander 2012; Donat et al. 2013b), which would indeed change the shape and symmetry of the distribution along with a shift.

Studies have shown evidence for asymmetry in observed and projected temperature distributions, especially in Northern Hemisphere regions (e.g. Kodra and Ganguly 2014; Matiu et al. 2016). However, it is possible that these findings are an artefact of the methodology used to calculate temperature anomalies, causing an overestimation of the extremes and adding uncertainty to these results (Rhines and Huybers 2013; Sippel et al. 2015). Lewis and King (2017) also found changes in skewness in daily temperatures towards hot extremes for regions including Australia, Asia, Europe and North America, though climate models simulated a wide range in both sign and magnitude. Considering how both the warm and cold tails of the distribution are changing, in comparison to the mean, might help to provide a more holistic and inclusive sense of if there is more than just a shift in the distribution (Katz et al. 2013; Sardeshmukh et al. 2015; Gross et al. 2018).

Climate model simulations suggest that globally-averaged annual cold extremes warm faster than the annual mean of daily minimum, while globally-averaged annual warm extremes warm at a similar rate to the annual mean of daily maximum temperatures (Kharin and Zwiers 2005). Of course, there are considerable differences in the magnitude and rate of change of extremes and the mean depending on region. Some studies address this by assessing the rates of change in local extremes relative to average temperatures. One approach is to compare local rates of changes in extremes with global mean temperature (e.g. Sutton et al. 2015; Seneviratne et al. 2016; Vogel et al. 2017). While this is important for regional and local-scale policy-makers and stakeholders, global average temperature is a highly abstract measure (no one *experiences* a global average) and local average temperatures, like extremes, are not globally heterogeneous. An alternative, therefore, is to compare local extreme temperatures with the corresponding local mean temperature changes (e.g. Kharin and Zwiers 2005; Brown et al. 2008; Donat

et al. 2017; Gross et al. 2018), accounting for the reality that mean warming rates differ considerably between regions. Using this approach, observational and reanalysis data have shown that in recent decades, the cold tails of both daily maximum and minimum temperatures have warmed faster than the corresponding mean temperature for much of the Northern Hemisphere extratropics, while the warm tail has been warming faster than the mean for parts of Europe, South America, North America, southern Africa and Australia (Brown et al. 2008; Gross et al. 2018). Climate model simulations of future changes in extremes show robust signals of accelerated warming in the hottest annual day relative to the local annual mean for the same regions (Donat et al. 2017).

It seems clear that the temperature distribution is indeed changing shape in some regions, as a consequence of the varying rates of change in the warm tails, cold tails and average temperatures. This effect might also differ by season, as different driving mechanisms are at play in summer (e.g. changes in surface heat fluxes and soil moisture feedbacks (Donat et al. 2017; Vogel et al. 2017)) compared to winter (e.g. snow and ice cover change and temperature advection (Kharin et al. 2013; Screen 2014)). While Orłowsky and Seneviratne (2012) investigated the scaling of seasonal and regional extremes, this was explored relative to the annual regional median change in temperature, and so assessments of changes in seasonal local extremes relative to corresponding seasonal mean temperatures are still needed.

This chapter documents seasonal rates of change in local daily maximum and minimum temperature extremes with the corresponding local mean temperatures at the global scale using an observational dataset and climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5) archive (Taylor et al. 2012). Our analyses involved testing a range of statistical and methodological

sensitivities to ensure results are robust, irrespective of methodological choice. The intention here is to provide a more comprehensive and systematic approach than has been done before, by considering and documenting the different aspects of the temperature distribution for both observed and future changes.

3.2 Data and methods

3.2.1 Observational and CMIP5 datasets

We use the observational product Hadley Centre Global Historical Climatology Network-Daily (HadGHCND) (Caesar et al. 2006), a quasi-global, land only dataset of daily temperature observations that are interpolated onto a 3.75° longitude by 2.5° latitude grid. We use daily maximum (daytime) and daily minimum (night-time) HadGHCND temperature data for a 65-year period that is 1950-2014. A strict completeness criteria is applied to the HadGHCND data to exclude regions with sparse data. Specifically, grid cells are only included that have at least 80% of data available for the entire period of analysis, in addition to at least 50% of data available for both the first and last ten years of data.

We also use a suite of 26 CMIP5 that have been selected based on their availability of daily maximum and minimum temperature over the period of analysis (Table 3.1). A single simulation from each of the climate models is used, however, multiple ensemble runs from some of the models (where available) were tested to check for internal variability. Results were found to be similar irrespective of the specific model run used (not shown). Historical model simulations (1950-2005) are merged with Representative Concentration Pathway 8.5 (RCP8.5) model simulations (2006 onwards), where the years 1950-2014 are used for analysis of recent decades in comparison to observations, and 1950-2099 for the analysis of future versus past simulated changes. All models are regridded to a common $2.5^\circ \times 2.5^\circ$ grid using a

bilinear remapping technique. We acknowledge that regridding may add some uncertainty to the analysis, however, several regridding techniques, including first- and second-order remapping techniques (Jones 1999) were previously tested and shown to have little effect on the results (Loikith et al. 2015; Gross et al. 2018).

We show recent changes simulated by the CMIP5 models for all land areas, and so comparisons between HadGHCND and the model output may be inconsistent for global average time series as the observations do not have complete spatial coverage. We perform an additional analysis of spatial correlations between the models and HadGHCND, as well as between the models themselves (see Section 3.3.3). For this analysis, all CMIP5 models are regridded and masked to the grid cells available in HadGHCND, so comparisons can be made for both the recent and future analysis.

3.2.2 Methods

Daily anomalies of daily maximum and minimum temperature (T_{\max} and T_{\min} , respectively) are calculated for each dataset relative to a mean annual cycle of daily temperatures that is based on the full period of each analysis, that is, 1950-2014 for the recent period, and 1950-2099 for future changes. Because the calculation of temperature anomalies can be sensitive to the choice of reference period, potentially inflating variance in extremes (Tingley 2012; Sippel et al. 2015; Hawkins and Sutton 2016), several different climatological reference periods were tested, including the commonly chosen period 1961-1990. All of the tested anomaly calculations showed similar results (see examples in Supplementary Material Figures S2.1 – S2.4). Following the calculation of anomalies, the data are split into seasons: December, January, February (DJF), March, April, May (MAM), June, July, August (JJA) and September, October, November (SON).

Table 3.1 CMIP5 models used in this study and their modelling group.

Model	Modelling group
ACCESS1.0 ACCESS1.3	Commonwealth Scientific and Industrial Research Organisation (CSIRO) and Bureau of Meteorology (BoM), Australia (CSIRO-BOM)
BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration (BCC)
BNU-ESM	College of Global Change and Earth System Science, Beijing Normal University (GCESS)
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCMA)
CCSM4	National Center for Atmospheric Research (NCAR)
CMCC-CESM CMCC-CM	Centro Euro-Mediterraneo per I Cambiamenti Climatici (CMCC)
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique (CNRM-CERFACS)
CSIRO-Mk3.6.0	CSIRO in collaboration with Queensland Climate Change Centre of Excellence (CSIRO-QCCCE)
GFDL-CM3 GFDL-ESM2G GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory (NOAA GFDL)
HadGEM2-CC HadGEM2-ES	Met Office Hadley Centre (MOHC)
INM-CM4	Institute for Numerical Mathematics (INM)
IPSL-CM5A-LR IPSL-CM5A-MR IPSL-CM5B-LR	Institut Pierre-Simon Laplace (IPSL)
MIROC-ESM MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Atmosphere and Ocean Research Institute (AORI), and National Institute for Environmental Studies (NIES) (MIROC)
MIROC5	AORI, NIES and JAMSTEC (MIROC)
MPI-ESM-MR MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M)
MRI-CGCM3	Meteorological Research Institute, Japan (MRI)
NorESM1-M	Norwegian Climate Centre (NCC)

Extreme temperatures in this study are defined as those data points which lie above the 98.5th percentile and below the 1.5th percentile of the seasonally varying distribution. This describes the warm tails (above the threshold) and cold tails

(below the threshold) of the anomaly distributions for both daily maximum and minimum temperatures. The threshold varies with time within each season, ensuring that the probability to exceed the threshold is 1.5% on any given day within that season and in any given year throughout the time periods used to calculate changes. Moreover, because exceedances are derived using anomalies calculated from an annual cycle of daily temperatures, the climatology and thresholds are specific for each calendar day. These methods broadly follow those of Gross et al. (2018), which showed that fitting a nonstationary point process model to the data yields similar results to calculating extremes using a percentile-based approach, as done here.

Recent and future changes in daily temperature extremes relative to the seasonal mean are calculated in three steps. The first step calculates changes in extremes for each season, for daily maximum and daily minimum temperatures separately, by taking the average of all ‘extreme’ data points for each grid cell for each of the two periods used to calculate changes (i.e. period one for recent changes in 1950-1981 and period two is 1982-2014; period one for future changes is 1950-1979 and period two is 2070-2099). Similarly, the seasonal mean of daily maximum and daily minimum temperature anomalies is calculated for the two periods. The second step calculates changes in the seasonally-averaged extremes between the two periods, and similarly, changes in the seasonal means between the two periods. We then take the difference between the changes in extremes and changes in the mean, which we refer to as ‘excess changes’. For simplicity, we use the terms ‘recent excess changes’ and ‘future excess changes’ to refer to excess changes between the two periods used for the recent and future analyses, as outlined above.

To test whether the distributions of excess changes between the two periods are significantly different in HadGHCND, a two-sample Kolmogorov-Smirnov test (KS test) was used to assess significance. Statistically different distributions for each grid cell are determined at the 5% level. The pattern correlations shown in the correlation matrices (see Figure 3.7 – 3.10) are calculated using the Pearson product-moment coefficient of linear correlation. Pattern correlations of excess changes are computed as weighted to account for zonal differences in grid cell size.

3.3 Results

3.3.1 Observed and simulated excess changes in recent decades

Figure 3.1 – 3.4 show recent observed seasonal excess changes for the warm and cold tails of both Tmax and Tmin. Changes are calculated as the difference between warming rates of the local seasonal mean and corresponding extreme temperature (where ‘local’ refers to locally at the grid cell) in 1982-2014 compared to 1950-1981, as described in Section 3.2. Positive excess changes indicate regions where the extremes have been warming more than the mean, while negative excess changes show areas where the mean has warmed more than the extremes.

3.3.1.1 Recent excess changes in the warm tails

For the warm tails in HadGHCND (Figure 3.1), all seasons show a relatively heterogeneous pattern across the globe. Some of the most notable excess changes are shown during wintertime in the Northern Hemisphere (Figure 3.1a,b), with some high latitude regions in Canada, Greenland and Siberia showing excess changes below -1°C . Many locations within these regions have statistically significant changes between the two periods, indicated by stippling. The warm tail in Tmax in DJF (Figure 3.1a) also shows a ‘hot spot’ region in Europe showing excess changes of 1°C or greater, which is also apparent to a lesser magnitude in

Tmin (Figure 3.1b). Other prominent features are shown in the shoulder seasons (MAM, SON), which indicate significant positive excess changes in some areas in the Northern Hemisphere. For MAM (Figure 3.1c,d), this includes parts of eastern North America, west Asia and parts of Russia and Siberia, with the strongest of these occurring in the warm tail of Tmin (Figure 3.1d), especially in west Asia and Russia.

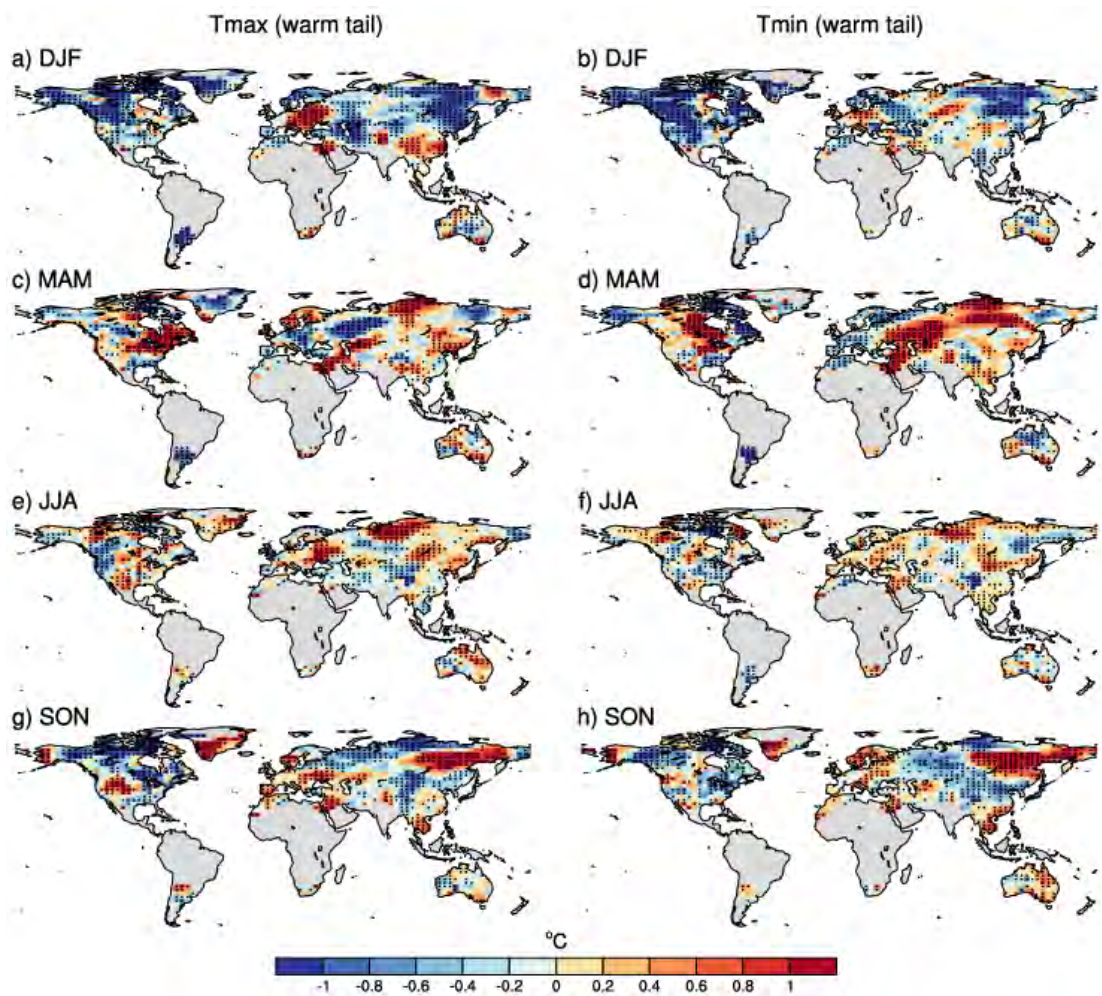


Figure 3.1 Recent observed excess changes in HadGHCND for the warm tails of both Tmax and Tmin, shown for each season (December – February (a,b), March – May (c,d), June – August (e,f), September – November (g,h)). Positive values indicate grid cells where the warm extremes have warmed more than the mean temperature, while negative values show grid cells where the mean temperature has warmed more than the extremes. Grey areas represent grid points that are either missing in HadGHCND, or that did not meet the completeness criteria outlined in Section 3.2.1. Stippling indicates grid cells that are statistically significantly different at the 5% level as assessed by a KS test.

For SON (Figure 3.1g,h), this includes parts of Siberia and North America, eastern Europe and the Mediterranean. JJA (Figure 3.1e,f) shows the weakest excess changes overall, with the only areas showing relatively strong excess changes in northern Russia and eastern Europe for Tmax (Figure 3.1e). For all seasons, spatial patterns and magnitude of excess changes between Tmax and Tmin are relatively similar, however, there are some instances with significant differences. For example, in MAM, a negative excess change exceeding -1°C is shown for eastern Europe in Tmax (Figure 3.1c), with the same region showing positive excess changes above 1°C in Tmin (Figure 3.1d). This difference suggests that on average, the mean temperature has warmed more than hot extremes during the day, while the opposite is apparent during night-time.

Figure 3.2 shows results of recent excess changes in the warm tails for the multi-model mean, calculated from the selected 26 CMIP5 models. Stippling indicates grid cells where 20 out of the 26 models agree on the sign of excess change. Results for the individual CMIP5 models are included as Supplementary Material (Figures S2.5 – S2.20 for warm tails and cold tails). As expected, the multi-model mean shows a relatively smooth spatial pattern across the globe for all seasons, due to smoothing variability across different simulations. The individual model simulations are, like HadGHCND, more spatially heterogeneous, though the models vary considerably in terms of both sign and magnitude, partly reflected in the lack of stippling in Figure 3.2. For the warm tails in Tmax, for example, some of the largest regional differences between models are shown for DJF in northern Eurasia, with some models showing negative excess changes over -1°C , while others show positive excess changes between 0.6°C and 1°C (see Figure S2.5). For Tmin, DJF also shows the largest regional differences between models (see Figure S2.10), with

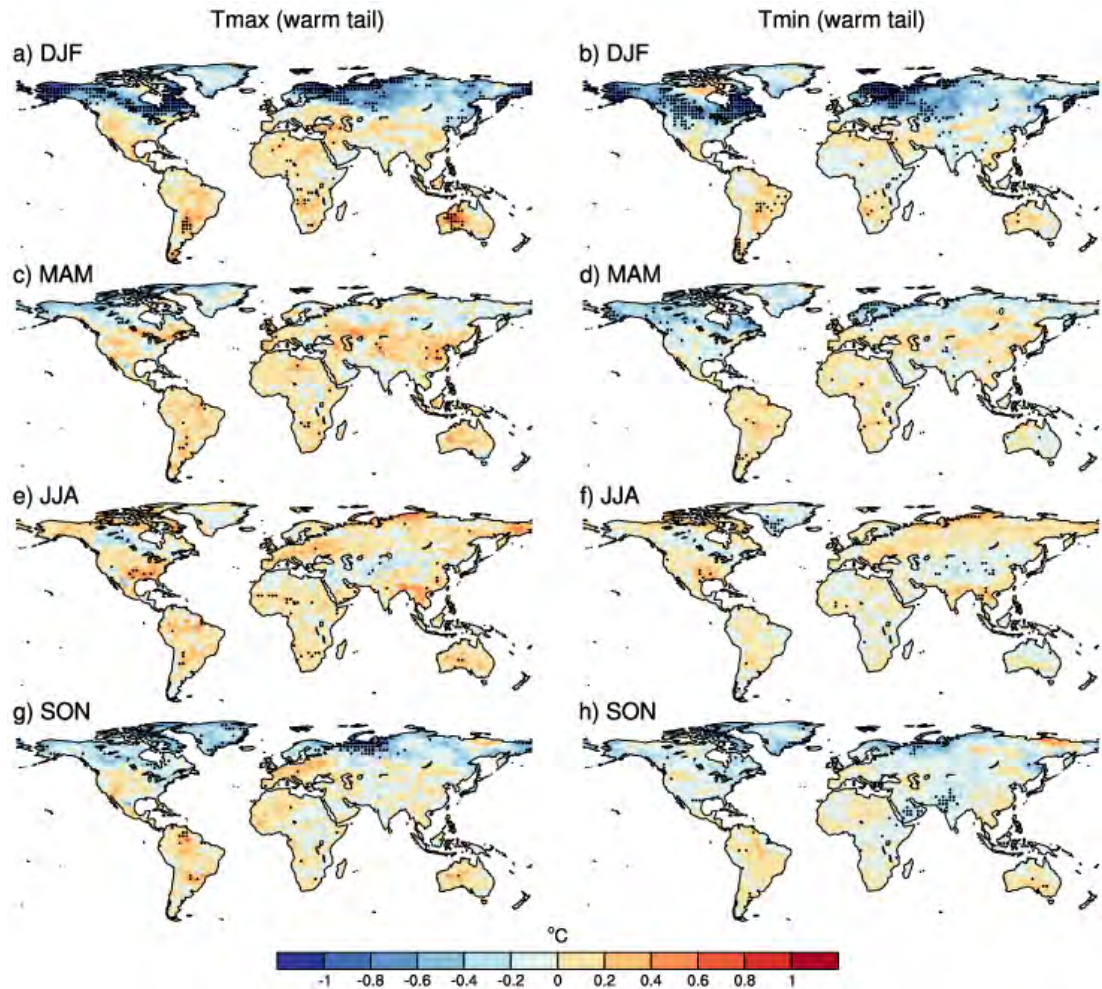


Figure 3.2 As Figure 3.1, but for the CMIP5 multi-model mean. Stippling indicates grid cells where at least 20 out of the 26 models agree on the sign of excess change.

similar magnitude differences to that of Tmax. In the multi-model mean, excess changes in the warm tails are relatively small for all seasons, with the strongest being in DJF (Figure 3.2a,b) in the Northern Hemisphere extratropics. For most of this area, excluding eastern Russia, at least 20 out of the 26 models agree that the mean has warmed more than extremes, as in HadGHCND. In some of the individual model simulations for DJF (see Figures S2.5, S2.9), the magnitude of excess change is larger and covers a greater area than in HadGHCND, while other models show a smaller or even opposite sign in magnitude of excess change. Model agreement is low for MAM (Figure 3.2c,d), JJA (Figure 3.2e,f) and SON (Figure 3.2g,h), though this is likely a consequence of excess changes being relatively small overall. As in

HadGHCND, changes in Tmax and Tmin are mostly similar both spatially and in terms of sign, though some individual models show differences in similar regions to HadGHCND, for example, around Central Europe during MAM (see Figures S2.6, S2.10).

Overall, the most prominent features in the warm tails are the negative excess changes in the Northern Hemisphere extratropics during boreal winter, shown for both Tmax and Tmin. This is evident in HadGHCND and the CMIP5 models, with high model agreement in the multi-model mean for much of this area.

3.3.1.2 Recent excess changes in the cold tails

Excess changes in the cold tails for recent decades based on the HadGHCND gridded observations are shown in Figure 3.3. As in the warm tails, Tmax and Tmin show many similarities in both spatial pattern and sign of excess change. Similarly, JJA (Figure 3.3e,f) shows the smallest excess changes in most regions, excluding around the Middle East area which shows excess changes up to 0.8°C for Tmax (Figure 3.3e). The most prominent excess changes in the cold tails are shown for some Northern Hemisphere regions in wintertime and the shoulder seasons. In DJF (Figure 3.3a,b), much of western and central Asia and the Mediterranean region show positive excess changes over 1°C in the cold tails of both Tmax and Tmin, though much of the United States, excluding parts of the west coast, shows negative excess changes between -0.6°C and -1°C. In MAM (Figure 3.3c,d), much of Eurasia and North America show significant positive excess changes, particularly for Tmin (Figure 3.3d). SON (Figure 3.3g,h) also shows strong positive excesses covering much of Eurasia, with these regions showing more significantly different grid cells in Tmin (Figure 3.3h) compared with Tmax (Figure 3.3g).

The multi-model mean again shows a relatively smooth spatial pattern in excess changes in the cold tails in recent decades (Figure 3.4), with slightly more model agreement overall compared to the warm tails. For DJF (Figure 3.4a,b), the multi-model mean does not capture the strong positive excess changes shown in HadGHCND for parts of Russia and Asia, nor the negative excess changes over much of central and eastern North America (see Figure 3.3a,b), which are mostly around 0.2°C to 0.4°C in the multi-model mean. Approximately half of the individual models show strong positive excess changes for DJF in the Northern Hemisphere, though the magnitude between the models ranges between 0.4°C and over 1°C (see Figures S2.13, S2.17). Some of the models also capture the negative excess changes shown for DJF in HadGHCND for much of the United States.

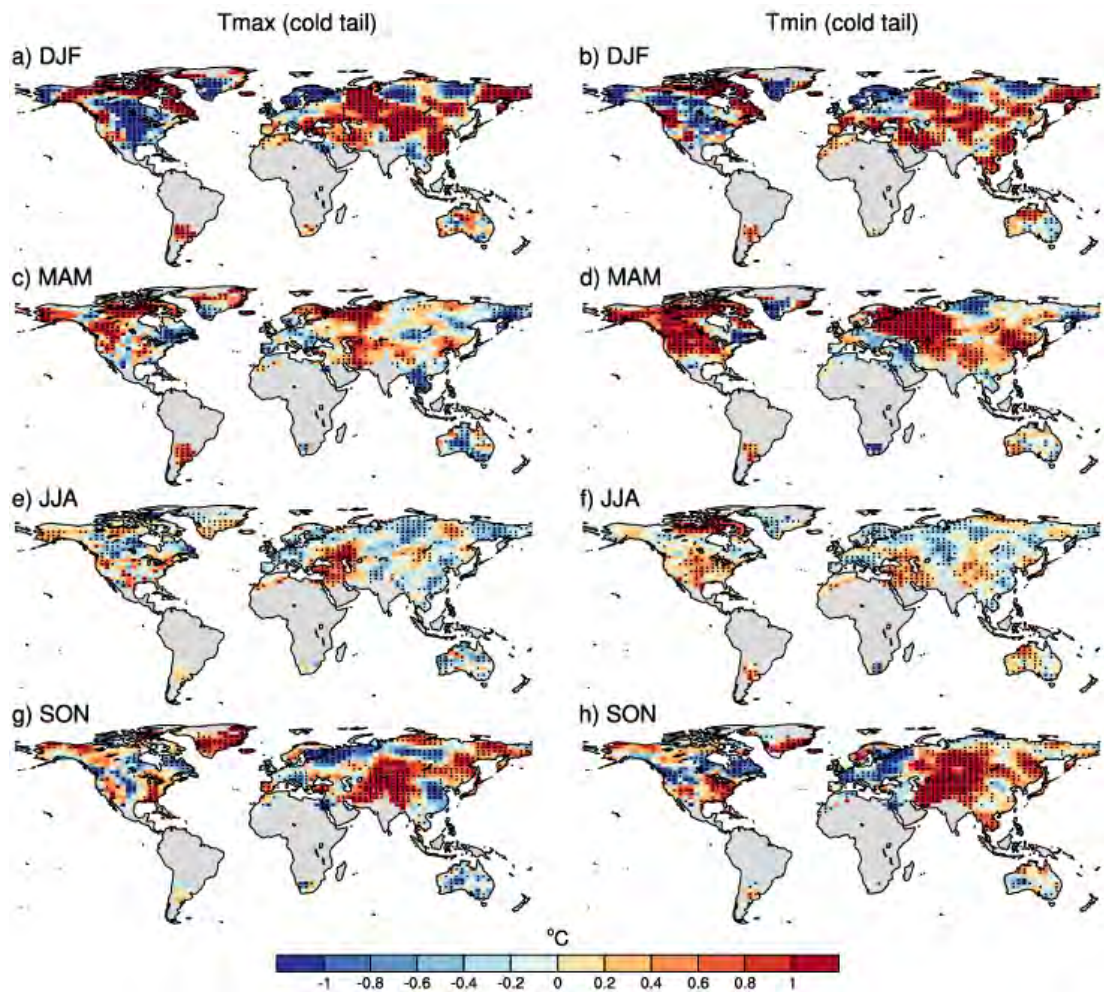


Figure 3.3 As Figure 3.1, but for the cold tails of Tmax and Tmin.

Overall, the most prominent features across all seasons in the multi-model mean are the positive excess changes in the Northern Hemisphere extratropics during MAM (Figure 3.4b,c) and SON (Figure 3.4g,h). This is especially apparent in SON, which shows moderate model agreement on the sign of change in regions showing the strongest excesses, that is, central and northern Europe and southern parts of Russia for the cold tails of both Tmax (Figure 3.4g) and Tmin (Figure 3.4h). These positive excess changes during the shoulder seasons are somewhat apparent in HadGHCND depending on the region (see Figure 3.3g,h). For example, the observations and multi-model mean agree on the sign of excess change in SON for parts of northern Europe and south-central Russia. However, the multi-model mean generally shows smaller magnitude excesses, such as an excess change of 0.6°C for south-central Russia in Tmin (Figure 3.4h), compared to 1°C or more in HadGHCND (see Figure 3.3h). In these regions during SON, many of the individual models show strong positive excesses resembling the magnitudes shown in HadGHCND, though some models show much more wide-spread changes covering much of the Northern Hemisphere extratropics (see Figures S2.16, S2.20). Some models also show strong negative excess changes between -0.2°C to -1°C in parts of Australia. This sign in excess agrees with HadGHCND and is reflected in the multi-model mean. For MAM (Figure 3.4c,d), the observations and multi-model mean agree on positive excess changes for much of central Europe and North America, particularly for Tmin (Figure 3.4d), though as in SON, excess changes in the multi-model mean are smaller than those shown in HadGHCND. As in HadGHCND, and as shown for the warm tails, JJA (Figure 3.4e,f) shows the smallest excess changes overall in the multi-model mean, with the individual models generally showing excess changes between -0.4°C and 0.4°C (see Figures S2.15, S2.19). Though it is not apparent in the multi-model mean, some individual models (see Figure S2.15)

capture the positive excess changes in Tmax during JJA in the Middle East shown in HadGHCND (see Figure 3.3e).

In recent decades, extremes in the cold tails of the distribution have warmed more than mean temperatures for much of the Northern Hemisphere in the shoulder seasons. This positive excess change is a common feature in HadGHCND, the individual CMIP5 models and the multi-model mean.

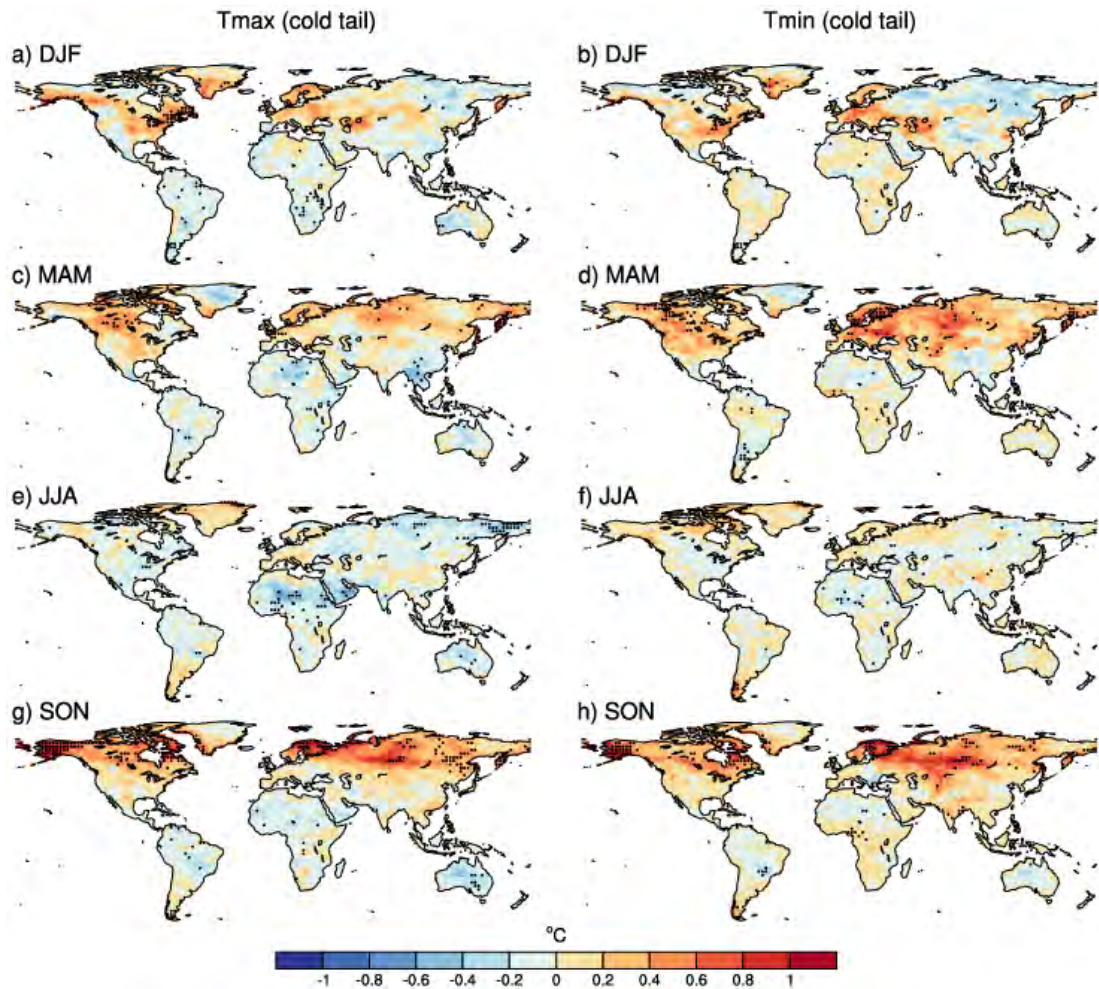


Figure 3.4 As Figure 3.2, but for the cold tails of Tmax and Tmin.

3.3.2 Projected excess changes between future and past decades

Figure 3.5 – 3.6 show future excess changes in the CMIP5 multi-model mean between the end of the 21st century and the mid-20th century, with results from the individual models shown in Figures S2.21 – S2.36. As for the recent excess changes (see Figure 3.2, Figure 3.4), spatial pattern and sign of excess change between Tmax and Tmin are similar. In particular, for the warm tails (Figure 3.5), the mid-to high-latitudes in the Northern Hemisphere show negative excess changes between -1°C and -7°C for all seasons except JJA in both Tmax and Tmin. For the same regions, positive excess changes between 1°C and 7°C are shown for the cold tails (Figure 3.6) in both Tmax and Tmin. In the Southern Hemisphere, low magnitude changes for both the warm and cold tails are projected, however, some differences are apparent such as over South America where the cold tail of Tmax mostly shows projected excess changes around -0.5°C, while Tmin shows positive excess changes around 0.5°C for the same regions.

3.3.2.1 Future excess changes in the warm tails

For the warm tails (Figure 3.5), DJF shows high model agreement in terms of sign for much of the globe for both Tmax (Figure 3.5a) and Tmin (Figure 3.5b), especially for those areas which show the largest excess changes. This includes much of the Northern Hemisphere extratropics, which show negative excess changes between -3°C and -7°C for much of Canada and eastern Russia. Though most of the individual models capture this (see Figures S2.21, S2.25), some show smaller negative excess changes closer to -1°C and -2°C shown across eastern Russia. Here, the IPSL models especially stand out, showing positive excess changes over Siberia. For DJF in the Southern Hemisphere, there is also high model agreement for many regions, with hot extremes in the austral summertime projected to warm more than the mean for much of South America and southern parts of Africa. The magnitudes

of positive excess warming in the individual models range between 0.5°C and over 5°C for regions such as central to northern South America (see Figures S2.21, S2.25). Both MAM (Figure 3.5c,d) and SON (Figure 3.5g,h) show similarities to DJF, with similar regions showing model agreement and the same sign in change, although with a smaller magnitude. JJA (Figure 3.5b,c) shows the smallest magnitude in excess changes overall, with good model agreement for regions showing the largest excesses, including southern Australia and southern Asia, parts of South America and Africa, and Europe, which all show positive excess changes in the range of 0.5°C to 1.5°C .

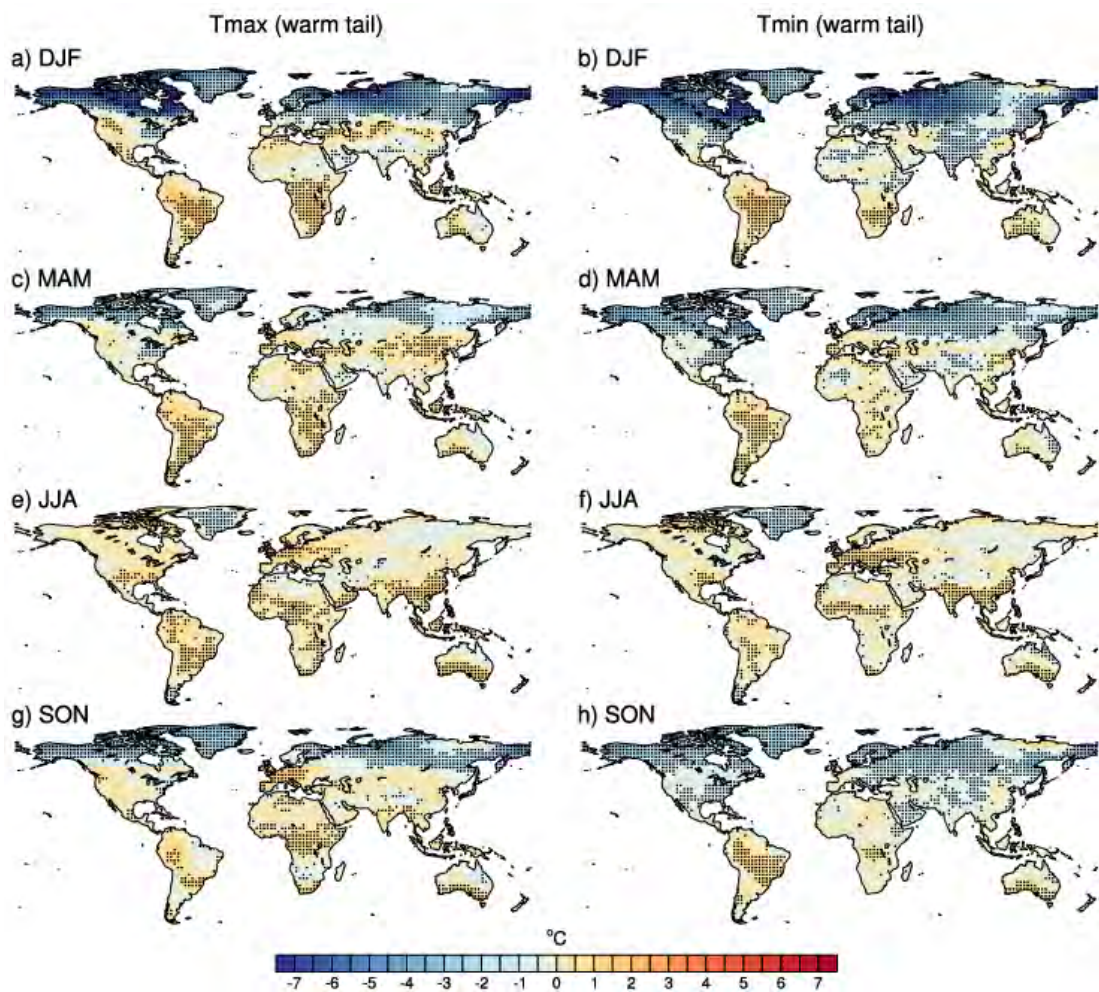


Figure 3.5 Future excess changes in the warm tails of Tmax (left column) and Tmin (right column) in the CMIP5 multi-model mean for each season (December – February (a,b), March – May (c,d), June – August (e,f), September – November (g,h)). Stippling indicates grid cells where at least 20 out of the 26 models agree on the sign of change.

Projections of future excess changes indicate that mean temperatures will increase more than warm extremes for much of the Northern Hemisphere, for all seasons except JJA. This is a robust signal across the suite of 26 CMIP5 models. Southern Hemisphere land regions generally show that extremes in the warm tails of both Tmax and Tmin are predicted to warm more than the corresponding seasonal mean temperatures.

3.3.2.2 Future excess changes in the cold tails

The cold tails generally show the opposite sign of excess changes in the multi-model mean compared to the warm tails for all seasons across the globe (Figure 3.6). Northern Hemisphere regions, particularly in the mid- to high-latitudes, show strong positive excesses and high model agreement for DJF (Figure 3.6a,b), MAM (Figure 3.6c,d) and SON (Figure 3.6g,h). There is also high model agreement for Southern Hemisphere regions, particularly for Tmax, which show mostly negative excess changes. In both shoulder seasons, the Northern Hemisphere extratropics show excesses in cold extremes of at least 1.5°C, with some locations in Canada, Alaska, northern Europe, central-western Asia and Russia exceeding 5°C. This is robust across most of the models, with few showing excess changes below 1.5°C, especially for SON (see Figures S2.32, S2.36). These positive excess changes extend further south for Tmin than Tmax, into parts of India. For JJA (Figure 3.6e,f), there is strong model agreement for Tmax (Figure 3.6e) for much of the globe, which shows relatively small negative excess changes, excluding Greenland and southern South America. Though most of the models agree on this, some models show larger negative excess changes around -2.5°C for northern Africa, while others show positive excess changes between 0.5°C and 2.5°C over southern Russia and central Asia (see Figure S2.31). For all seasons except JJA, Tmax and Tmin tend

to agree on spatial pattern and sign of change for much of the globe, excluding some Southern Hemisphere locations which show differences in sign for regions showing small excess changes, such as parts of central South America and Africa. All seasons show robust positive excess changes in Greenland for both Tmax and Tmin.

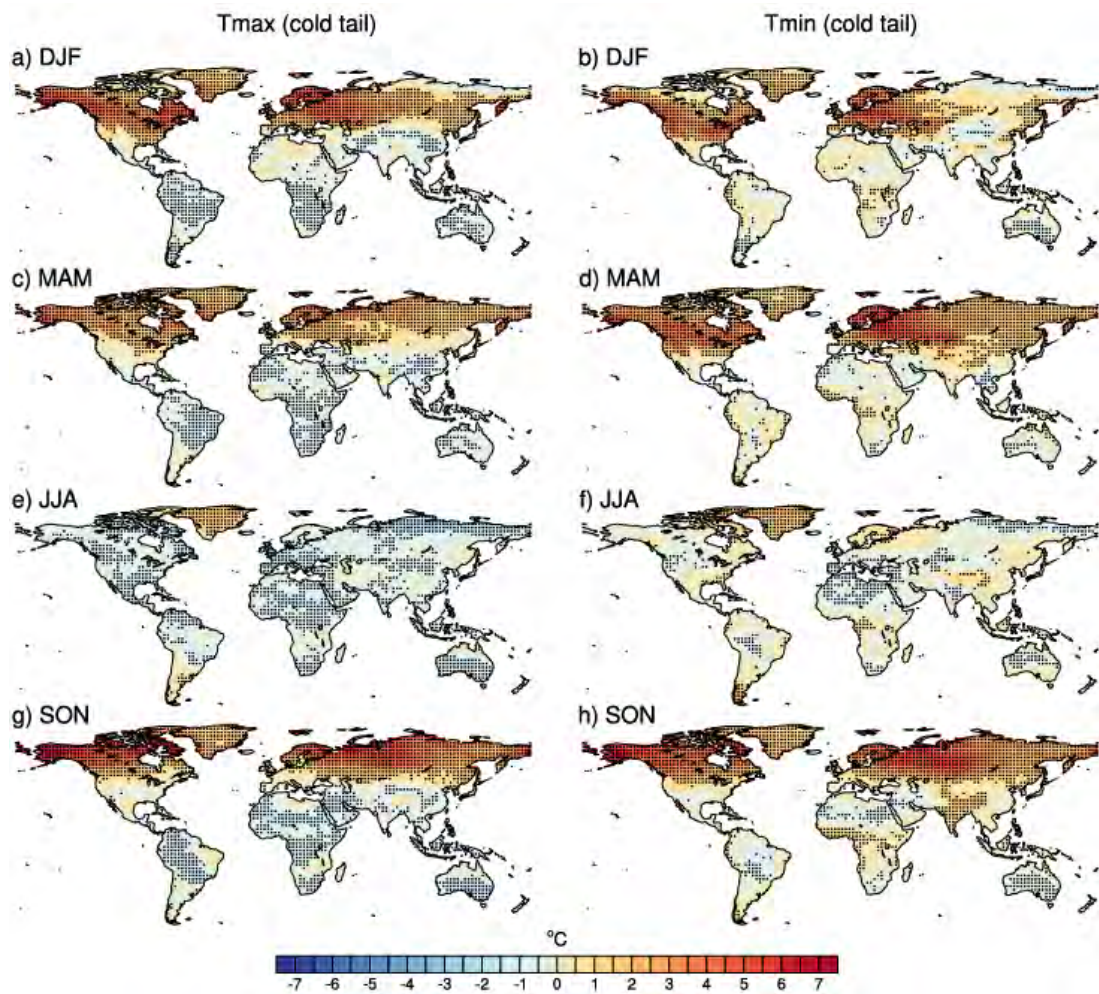


Figure 3.6 As Figure 3.5, but for the cold tails of Tmax and Tmin.

In the cold tails of the distribution, future changes generally show the opposite signal to those shown for the warm tails (Section 3.3.2.1), that is, a strong warming of cold extremes compared with the mean in the Northern Hemisphere (excluding JJA) in both Tmax and Tmin, and slightly more warming in the mean compared

with cold extremes in some Southern Hemisphere regions. Notably, for both the warm and cold tails, all seasons except boreal summer show strong inter-hemispheric differences in excess changes. Previous work has noted the inter-hemispheric differences in surface temperature after the pre-industrial era has increased due to sea ice and snow melt in the Northern Hemisphere and will continue to increase into the future due to the land-sea warming contrast and unequal spatial impacts from greenhouse gas emissions (Friedman, 2013; Blunden and Arndt 2014; Hutchinson et al. 2015).

3.3.3 Spatial correlations of recent and future excess changes

In order to better evaluate how excess changes are represented across the different models, correlation matrices for global spatial patterns of excess changes in the warm and cold tails of Tmax and Tmin are shown for all seasons in Figure 3.7 – 3.10 (DJF, MAM, JJA, and SON respectively). Here, the models are all regridded and masked to HadGHCND, so that comparisons to observations can be made for the patterns of past changes. Pattern correlations for excess changes in recent decades are generally low for all seasons and temperature variables. Importantly, the low correlations are apparent not only between the models and the observations, but between the models themselves. This suggests a general spatial noise for the globe, along with the relatively low signal of excess changes in the CMIP5 models in the analysis of recent changes.

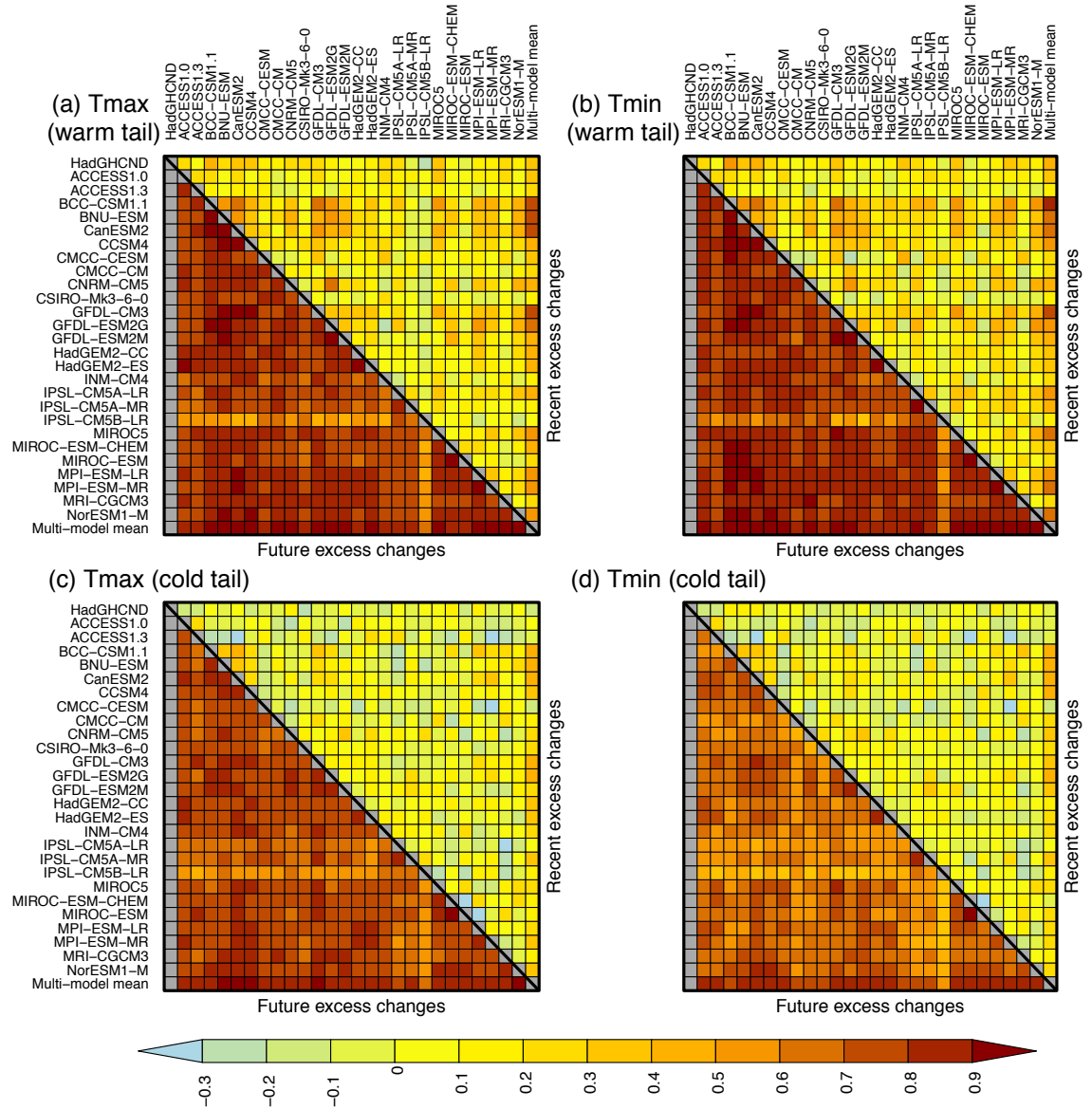


Figure 3.7 Globally-averaged pattern correlations for December – February, for each temperature variable: Tmax (warm tail) (a), Tmin (warm tail) (b), Tmax (cold tail) (c), and Tmin (cold tail) (d). Correlations for recent excess changes are shown above the diagonal, and future excess changes are shown below the diagonal.

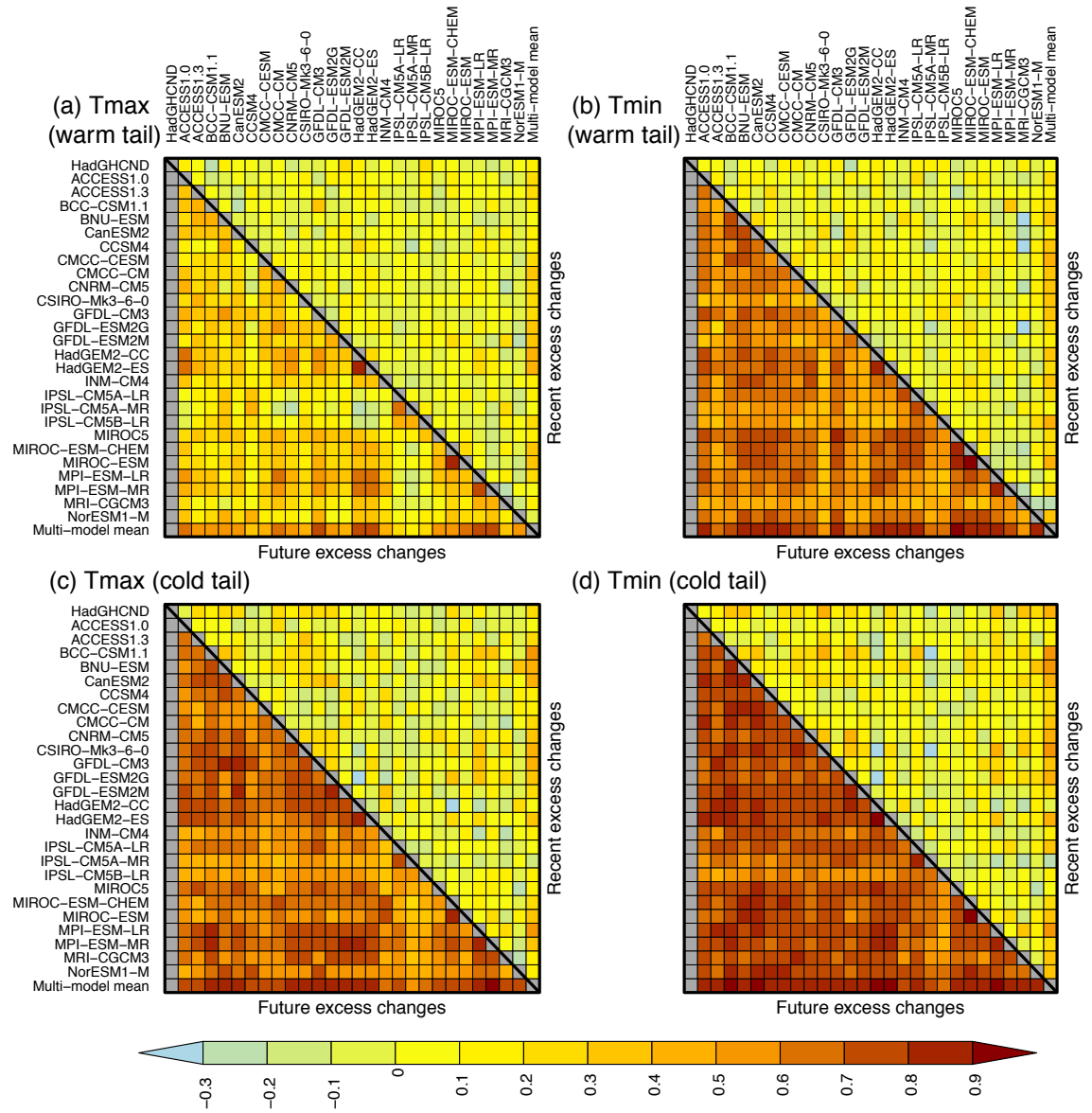


Figure 3.8 As Figure 3.7, but for March – May.

Irrespective of the season and temperature variable, excess changes in the future projections are much more spatially correlated than recent excess changes. JJA (Figure 3.9) shows comparatively lower spatial correlations for future changes compared with other seasons, though the models are still more correlated with each other than they are for the analysis of past changes. Across all temperature variables, future excess changes in DJF (Figure 3.7) show the highest pattern correlations, with the highest overall being in the warm tails. For MAM (Figure 3.8), spatial correlations are the highest for patterns of future changes in the cold

tails of the distribution, as is the case for SON (Figure 3.10). In comparison to the low correlations and noise shown for the recent excess changes, the high pattern correlations for future excess changes indicate a systematic change in extremes relative to the mean, especially for DJF and in the cold tails of the shoulder seasons

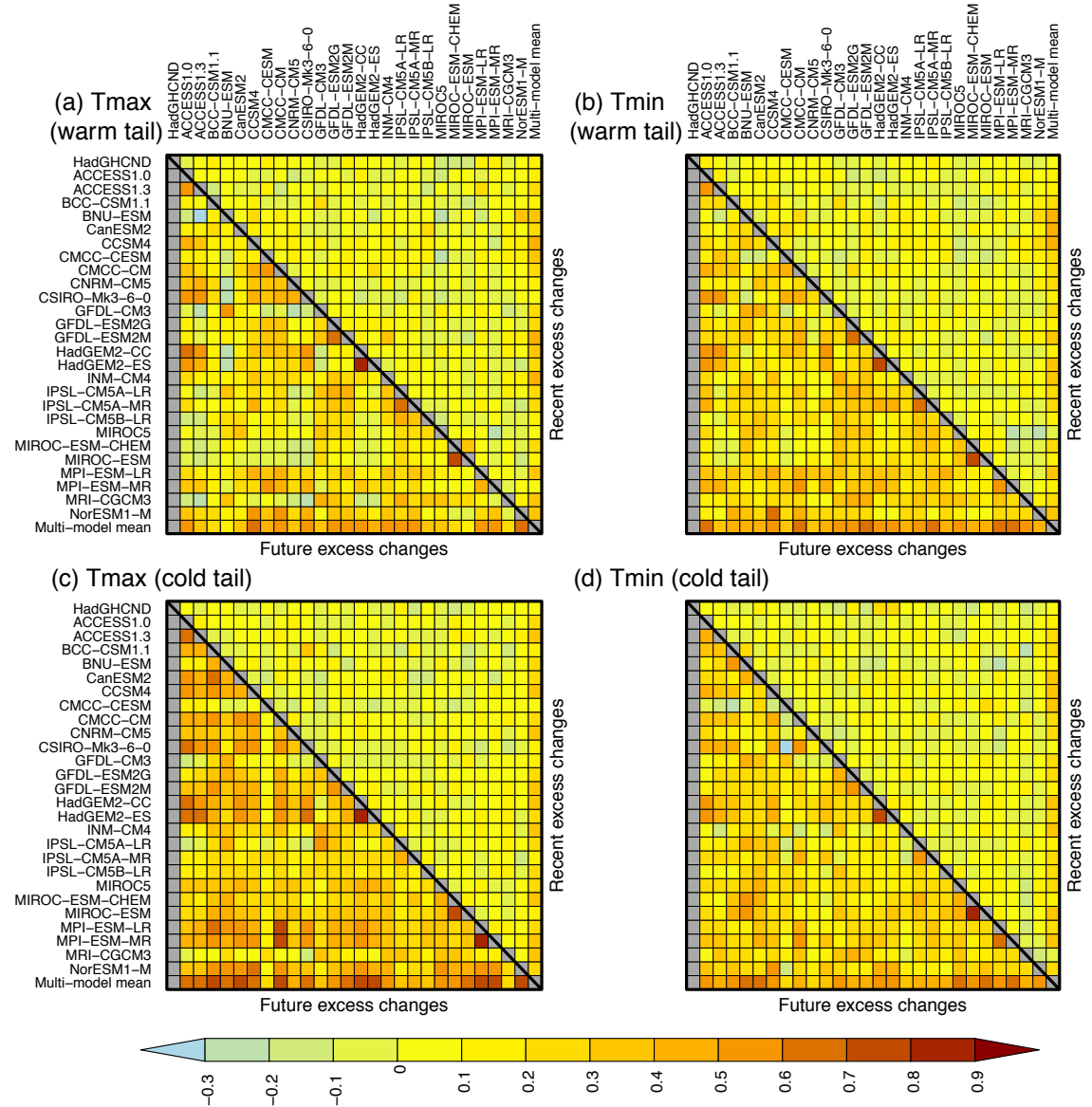


Figure 3.9 As Figure 3.7, but for June – August.

under stronger greenhouse gas forcing. The magnitude of warming in future changes compared with recent changes is reflected in the high correlations shown for future changes, in all seasons except JJA. Moreover, models that show greater warming

also show greater sensitivity. For example, the IPSL-CM5B-LR model, which shows lower correlation with other models compared with those that are more highly correlated with each other, is generally cooler than other models and shows smaller magnitude excess changes. Further, for the warm tail in Tmax in MAM (Figure 3.8a), relatively low correlations are shown for the future compared with Tmin and with the cold tails. This is likely a function of relatively small magnitude future excess changes across the majority of the models.

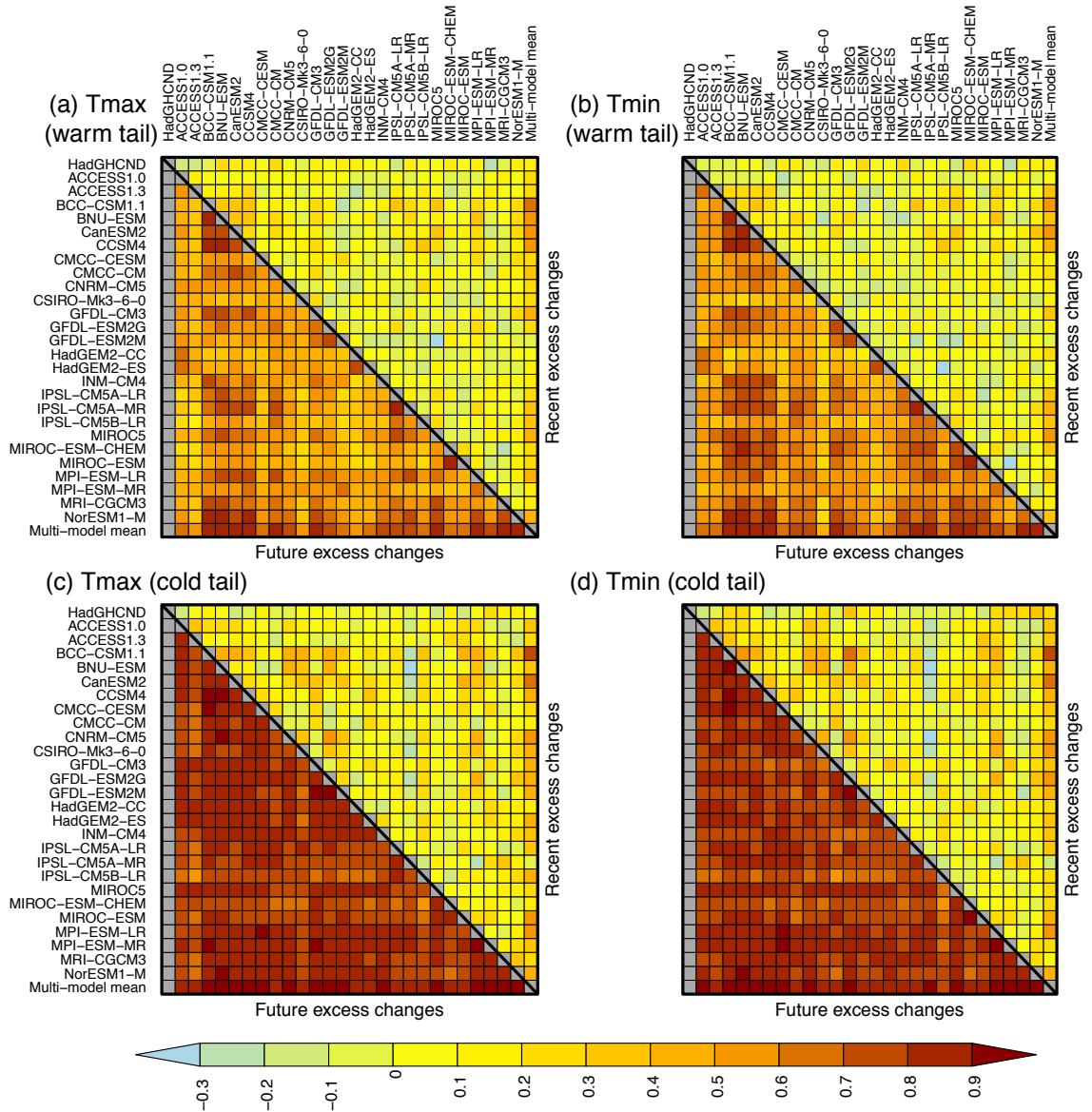


Figure 3.10 As Figure 3.7, but for September – November.

3.4 Discussion and conclusions

We have documented how seasonal extremes of daily temperatures in both tails of the distribution are changing, and might change in the future, relative to seasonal local mean temperatures for global land areas. Our findings indicate that warm extremes in all seasons are projected to warm more than the corresponding mean temperatures for much of the tropics and Southern Hemisphere regions. In these areas, the CMIP5 multi-model mean shows strong agreement for accelerated warming of warm extremes in much of South America, southern Africa and southern Australia. This is especially true in DJF, MAM and JJA, which also show excess warming in the warm tails in parts of the United States, the Mediterranean and Eurasia. It should be noted that some of these regions, such as parts of South America and southern Africa, do not have sufficiently complete data available in the observational dataset used in this study, and so simulated magnitudes of accelerated warming of warm extremes in these areas cannot be evaluated against observations here and need to be interpreted with caution. Further, in our analysis of recent decades, excess changes in the multi-model mean are sometimes opposite in sign to those shown in the observations in regions where data are available, such as in the United States during DJF for Tmax. This adds some uncertainty to projections of warm extremes relative to the mean. Some individual model simulations, however, do resemble the excess changes shown in the observations, in both sign and magnitude, and projections for these models show strong positive excess changes. The aforementioned ‘hot spot’ regions are also identified as areas of accelerated warming of hot extremes by Donat et al. (2017) and seem robust when comparing local hot extremes against both the local annual mean and local seasonal mean. The opposite sign of excess change, where the rate of change in the mean is greater than that of the warm extremes, is projected for high-latitude areas in the Northern Hemisphere during all seasons except JJA. These regions have

previously been shown to have the strongest mean warming signal during boreal winter (Screen and Simmonds 2013). The increase of warm extremes relative to the mean in summertime Northern Hemisphere regions is relatively small, compared with other seasons.

Southern Hemisphere regions generally show smaller magnitude excess changes overall compared with the Northern Hemisphere, especially for future changes. Hutchinson et al. (2015) looked at simulations of equilibrated control runs and 2xCO₂ warming experiments to show the asymmetry is caused primarily by ocean circulation being affected by greater melting of sea ice in the Arctic. Although outside of the scope of the present study, it would be interesting to similarly explore equilibrated simulations to understand whether the inter-hemispheric asymmetry in excess changes is related to model simulations being further from equilibrium in the Southern Hemisphere, or whether the asymmetry is due to the land-ocean contrast and enhanced sea ice loss in the Northern Hemisphere.

There is strong evidence in both observations and climate models that cold extremes are warming at a faster rate than warm extremes, changing the symmetry of the distribution (e.g. Kharin and Zwiers 2005; Donat et al. 2013b). We show that, for all seasons except JJA, the magnitude of excess change in regions where the cold tail is projected to warm more than the mean, such as in the Northern Hemisphere mid- to high-latitudes, is greater than that of the negative excess changes in the warm tails in the same regions. This is consistent with the distribution becoming narrower. Evidence for a decrease in wintertime variability in these regions has previously been discussed (Screen 2014; Holmes et al. 2016; Rhines et al. 2017). Our findings further suggest a decrease in variability in DJF in these areas, but also during the shoulder seasons. Perhaps the most striking

features of all are the strong excess changes in the cold tails in the shoulder seasons, shown for many of these mid- to high-latitude Northern Hemisphere regions. Regions that experience snow and ice-related feedbacks, such as these areas with strong excess changes, are more associated with the disproportionate warming of cold extremes to warm extremes (Kharin et al. 2013). This strengthens the argument for an asymmetrical, or skewed distribution in the Northern Hemisphere mid- to high-latitudes, with more warming in the cold tails relative to the mean, and less warming of warm extremes relative to the mean.

The possible drivers of the warming in the cold tails are likely different to those affecting the warm tails. The physical processes relating to the rates of warming in extremes relative to the mean have been explored mostly for hot extremes (e.g. Donat et al. 2017; Vogel et al. 2017). These changes have been largely explained by changes in the land surface heat fluxes and soil moisture-temperature feedbacks amplifying the intensity of hot extremes. Recent work has shown that the strength of these land-atmosphere interactions differs between models, and those models with moisture-temperature relationships more similar to observations over the past decades tend to simulate smaller warming of future heat extremes compared with models showing unrealistically strong land-atmosphere interactions (Donat et al. 2018; Vogel et al. 2018). The mechanisms behind the warming rates of cold extremes have been less studied, though some have attributed the accelerated warming of cold extremes relative to warm extremes in Northern Hemisphere regions to Arctic amplification and the reduction of cold air advection (e.g. Screen 2014).

Future work should continue to investigate the possible physical processes that are driving the excess warming of extremes relative to the mean. Based on this study,

this process-driven work could focus on the regions and seasons that show the most robust signals, that is, cold extremes in mid- to high-latitude Northern Hemisphere regions during the shoulder seasons and DJF. One possibility, aside from reductions in cold air temperature advection, is that changes in snow cover and albedo are playing a role in the warming of cold extremes relative to average conditions in these regions and seasons. Exploring regional differences and similarities between individual models can also provide insight into the possible physical mechanisms driving excess changes. Improving our understanding of why these differences occur, by understanding differences in the physical parametrizations within the models could lead to a more process-based evaluation such as linking changes in snow and ice cover with the amplified warming of cold extremes simulated in the models used in this study. Additionally, departures from normality may be linked with changes in extremes that are proportional to mean changes. Previous work has found that in the U.S., extreme and mean temperature changes in winter can be explained by non-normal intra-seasonal variability which drives the enhanced warming of wintertime extremes (Huybers et al. 2014). It would be useful to investigate excess changes in temperature with regards to non-normality for other regions showing the largest excess changes that are identified in this study, such as the Northern Hemisphere extratropics. Explaining what is driving excess changes could then lead to higher confidence in projected future changes and enable more effective planning and adaptation tailored to each region's needs.

Anticipated increases in temperature are coupled with greater potential impacts from extremes. Much of the existing literature on the impacts from climate change has focused on heat extremes, arguably because of the more perceivable effects from events such as heatwaves, for example. However, there are also consequences associated with the warming of cold extremes, and this is especially important if

the cold tail is warming faster than the warm tail. While there may be some advantages associated with the warming of cold extremes, such as lower human mortality rates from cold spells (Smith et al. 2014; Wolf et al. 2015), there are also detrimental consequences. For example, the warming of cold extremes can lead to more disease-carrying insects that can survive through milder winters (Smith et al. 2014; Wolf et al. 2015; Ebi and Nealon 2016). Under a warming climate, mosquitoes might expand and shift in geographical location (Ebi and Nealon 2016). Coupled with the accelerated rates of warming of cold extremes relative to average conditions during boreal winter, this puts more regions at risk, including North America, much of Europe and central Asia. In addition, the warming of cold extremes can lead to longer exposure to pollen from earlier flowering in the seasons (Wolf et al. 2015), which can then trigger allergies in those that suffer from respiratory illnesses. This could be a potential consequence of positive excess changes in cold extremes during the shoulder seasons. So, as with warm extremes, the need to understand how cold extremes are changing is crucially important to society.

The existing literature has used a range of methods to explore changes in extremes relative to the mean, making the uncertainties surrounding reported changes in variability and shape of the distribution difficult to compare and quantify. While some have defined extremes as the hottest or coldest day of the year (e.g. Seneviratne et al. 2016; Donat et al. 2017; Vogel et al. 2017), others have used globally or seasonally-averaged extremes (e.g. Kharin and Zwiers 2005) or annual extremes defined by a percentile (e.g. Brown et al. 2008; Gross et al. 2018). Some of these studies have then compared extremes to global annual average temperatures. Because of the strong seasonal and regional differences also in mean warming, in this study, we use seasonal averages of local mean temperatures, to

compare with the corresponding seasonal local extreme temperature. Another outstanding issue raised in the literature relates to the methods in which temperature anomalies are calculated. Showing temperature anomalies is useful because it more easily communicates how unusual the temperature is to ‘normal’, based on a long enough reference period (Hawkins and Sutton 2016). Normalising temperature relative to the local mean relative to a sufficiently long reference period is the most commonly used approach, but this can lead to inflated results of variability and extremes (Sippel et al. 2015). Here, daily temperature anomalies were calculated by subtracting the mean annual cycle of daily values from the raw daily data of each respective dataset, and this was tested across different reference periods. Results were found to be similar regardless of whether a shorter or longer period was used to derive the percentiles (see Figures S2.1 – S2.4). This ensured that the results of excess changes shown here are robust to such methodological choices. By testing and documenting possible sensitivities and finding no outstanding issues, the methods and results presented here can stand as a comprehensive framework for assessing a number of different aspects of the temperature distribution.

Differences in the local rates of change in the warm and cold extremes relative to the corresponding mean temperature can provide information on the changing shape, or skewness, of the distribution. This is meaningful with regards to climate change impacts because changes in the variability or shape of the distribution can have greater implications than shifts in the mean alone (Katz and Brown 1992; Schär et al. 2004). This study adds to the existing literature by systematically assessing characteristics of the tails of the distribution in parallel with the mean. Though the aim here is to provide a comprehensive documentation on several different aspects of the temperature distribution, some limitations remain. Firstly,

we have only presented projections made from the CMIP5 RCP scenario with the strongest greenhouse gas forcing, and it would be useful to know the range of excess changes across different RCP scenarios to determine uncertainties. In addition, despite selecting CMIP5 models from different groups to enhance independence, we did not assess model skill, aside from pattern correlations against HadGHCND, and so all model results need to be interpreted with caution. Despite finding relatively low agreement between the models and the observations for the recent analyses, many of the individual models resemble the observations more closely than the multi-model mean. The findings for future excess changes are systematic and robust, and our discussion of the multi-model mean reflects those of the individual models.

Changes in different characteristics of the temperature distribution, including changes in the tails of the distribution relative to the mean, vary significantly depending on the season and the geographic location. For many Northern Hemisphere regions, particularly those in the mid- to high-latitudes and during boreal winter, autumn and spring, average temperatures are projected to warm at a faster rate than warm extremes, while the opposite is true for the Southern Hemisphere. The greatest difference in warming rates between cold extremes and the mean are again shown for mid- to high-latitudes in the Northern Hemisphere, for all seasons except boreal summer. In these regions, cold extremes are projected to warm at a substantially faster rate than the mean, especially in Alaska, Canada, central Europe and western Russia. For these regions that show substantial differences in the rates of warming between extremes and average temperatures, the symmetry of the temperature distribution is changing. This can affect the probability of extreme events, potentially leading to significant impacts from the warming of cold extremes, as well as those from exacerbated warm extremes. By

systematically and comprehensively documenting how these aspects of the distribution are changing relative to each other, this study provides crucial information for future work aiming to understand the underlying physical processes that are driving the disproportionate rates of warming between extreme and mean temperatures.

CHAPTER 4

Enhanced warming of seasonal cold extremes relative to the mean in the Northern Hemisphere extratropics

Chapter overview

This chapter is currently under review for publication as: Gross, M. H., M. G. Donat, L. V. Alexander, and S. C. Sherwood, 2019: Enhanced warming of seasonal cold extremes relative to the mean in the Northern Hemisphere extratropics. *Climate Dynamics*, submitted.

In *Chapter 3*, we showed that seasonal cold extremes are projected to warm faster than seasonal mean temperatures for much of the Northern Hemisphere mid- to high-latitudes in all seasons except boreal summer. In this chapter, we use a novel approach to examine the environmental conditions on or prior to the day in which the cold extreme occurs to improve our understanding of the physical mechanisms that might contribute to the accelerated warming of cold extremes relative to mean temperatures. Using a selection of Global Climate Models, regions showing the largest disproportionate rates of warming are identified and explored using several different climate variables, including horizontal temperature advection, snow cover, and surface albedo. During winter months, accelerated warming rates of cold extremes relative to mean temperatures are projected for much of North America,

Europe and Eurasia for the late 21st century, compared with the mid-20th century. This is shown to be largely driven by reductions in cold air temperature advection, suggested to be a likely consequence of Arctic amplification. In spring and autumn months, cold extremes are expected to warm more than mean temperatures for most of the Northern Hemisphere extratropics, particularly Alaska, northern Canada and northern Eurasia. In the shoulder seasons, feedbacks related to projected decreases in snow cover and associated reductions in surface albedo are suggested as the main contributor influencing the enhanced warming of cold extremes relative to the mean. In addition, we show that the anomalously coldest day of the season in Northern Hemisphere high-latitude regions is projected to occur later in spring and earlier in autumn months, suggesting a flattening of the seasonal cycle in these regions. Overall, this chapter provides insight into the physical processes that might be driving the enhanced warming of seasonal cold extremes relative to the seasonal mean. This increases our overall confidence in the simulated excess changes and could help to develop constraints in future research to decrease the uncertainty in climate model projections.

4.1 Introduction

Daily temperature extremes are expected to continue to warm, along with increases in the mean, as a consequence of increasing greenhouse gases in the atmosphere. The rates of warming of extremes and average temperatures are, however, not uniform, and differ depending on the season and region. Disproportionate rates of warming for different parts of the temperature distribution imply a change in the shape of the distribution. This is significant because it effects the probability of extreme events (Mearns et al. 1984), which can cause widespread impacts on society

and the environment, more so than changes in the mean temperature alone (Intergovernmental Panel on Climate Change (IPCC) 2012).

Both observational data and climate model simulations suggest that cold extremes are warming faster than warm extremes for much of the globe (e.g. Kharin and Zwiers 2005; Donat and Alexander 2012; Donat et al. 2013b). Studies have also shown that in recent decades, cold extremes have been warming at a faster rate than local mean temperatures for some regions in the Northern Hemisphere (Brown et al. 2008; Gross et al. 2018). The enhanced warming of cold extremes in these regions, relative to both the mean temperature and warm extremes, is indicative of decreasing variability in these regions during boreal winter (Screen 2014; Ylhäisi and Räisänen 2014; Schneider et al. 2015; Rhines et al. 2017). Climate model projections suggest this decrease in variability due to the accelerated warming of the coldest days will continue (Holmes et al. 2016), with some regions in the mid- to high-latitudes projected to increase over 5°C more than mean temperatures by the late 21st century (Gross et al. submitted). These disproportionate rates of warming suggest that changes in cold extremes are driven by mechanisms other than increases in local mean temperatures alone. A better understanding of the physical drivers related to the projected rates of enhanced warming of cold extremes is therefore crucial for assessing the probability and potential impacts of future changes in cold extremes.

The physical mechanisms driving the accelerated warming rates of cold extremes differ depending on the region and season. For land regions in the Northern Hemisphere mid- to high-latitudes, the warming of cold extremes and the associated decreases in temperature variability during winter months are consistent with reductions in cold air temperature advection that are a consequence of Arctic

amplification (Screen 2014; Schneider et al. 2015; Holmes et al. 2016; Rhines et al. 2017). Arctic amplification, a phenomenon describing the enhanced warming of the Arctic relative to lower latitudes (Serreze and Francis 2006), has been suggested as one of the dominant causes of the observed and projected reductions in the severity of extremely cold days during winter in the Northern Hemisphere mid- to high-latitudes (Screen 2014; Schneider et al. 2015; Holmes et al. 2016; Rhines et al. 2017). This effect on cold extremes from Arctic amplification is shown to be a consequence of northerly winds from the Arctic bringing warmer than usual air to more southerly regions on the coldest days, which are warming faster than the warm days, therefore reducing the sub-seasonal temperature variability (Screen 2014; Holmes et al. 2016). Though it seems relatively clear that changes in temperature advection are linked with decreases in temperature variability in many mid- to high-latitude Northern Hemisphere regions, there is still uncertainty as to its role in driving the enhanced warming of seasonal cold extremes relative to the corresponding seasonal mean. It is more likely that multiple factors are influencing the seasonal and regional differences in rates of warming.

Aside from changes in atmospheric circulation patterns and thermal advection that may be altering cold extremes, variations in surface fluxes affecting the overall surface energy budget have strong links with surface temperatures and extremes. In particular, changes in snow cover play an important role in altering surface temperature in Northern Hemisphere regions that experience snowfall (e.g. Cohen and Rind 1991; Mote 2008; Diro et al. 2018). The high reflectivity and thermal emissivity of snow, compared to other natural surfaces, increases the surface albedo, lowers the absorbed shortwave radiation at the surface, and increases shortwave radiation reflected at the surface (Cohen and Rind 1991). The surface albedo feedback from snow cover is more likely to influence winter months and early spring

in the Northern Hemisphere, where snow accumulation is at its highest (He et al. 2014; Thackeray et al. 2015; Diro et al. 2018). However, the effect of snow cover on surface temperature is suggested to be strongest during spring when snow melt is at its highest, leading to increases in latent heat at the surface (Cohen and Rind 1991; Dutra et al. 2011; Xu and Dirmeyer 2011; Diro et al. 2018). The snow-temperature relationship is also effected by the snowpack, due to melting snow and consequent increases in latent heat, and vegetation cover, which acts to limit the role of snow cover and snow melt (Chapin III et al. 2005; Mote 2008).

Climate model simulations have shown differences in the regions with the strongest snow-temperature relationship, with some studies looking at North America finding the strongest links between temperature and snow cover over parts of eastern North America (e.g. Xu and Dirmeyer 2011), and others suggesting northwestern U.S. and southern Canada (e.g. Dutra et al. 2011). Many of the uncertainties related to biases within climate model simulations are related to the land cover parameterizations within the models, such as how the models capture the masking effect of vegetation on snow cover (Lorantý et al. 2014; Qu and Hall 2014). Evaluating the differences between climate model simulations of snow cover, surface albedo and their influences may help to improve future projections of warming.

This paper is structured by first evaluating a selection of climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5) archive (Taylor et al. 2012) against a gridded observational dataset in terms of their ability to capture recent warming rates of seasonal cold extremes relative to the seasonal mean. This is followed by discussing predicted future changes in the suite of climate models used. Next, the possible physical mechanisms driving the enhanced warming of cold extremes relative to seasonal means are explored. The investigated variables are

chosen based on evidence that has been suggested by prior studies, as previously discussed. We follow an approach similar to Donat et al. (2017), assessing conditions on the day on which the cold extreme occurs, or the average of days prior to day of the extreme.

4.2 Data and Methods

4.2.1 Observational and CMIP5 data

We use the Hadley Centre Global Historical Climatology Network-Daily (HadGHCND) dataset (Caesar et al. 2006) to evaluate climate model simulations for the period 1950-2014. HadGHCND is a land only, daily gridded dataset of daily maximum and minimum temperatures from ground stations, for which daily mean temperatures are calculated by taking the average of each daily maximum and minimum temperature value for each grid cell.

The HadGHCND data are used to evaluate six individual CMIP5 models (see Table 4.1), which were selected based on their data availability for all of the daily climate variables being investigated. While we only show a single simulation from each model (r1i1p1), multiple ensemble runs were analysed (where available) to determine model robustness and assess internal climate variability within the models. Results of multiple ensemble runs were found to be highly correlated in both spatial pattern and magnitude of simulated changes (not shown), indicating that sensitivity of the results to internal variability within the models is small. Historical model simulations (1950-2005) are merged with Representative Concentration Pathway 8.5 (RCP8.5) simulations (2006 onwards) to assess changes between the mid-20th century and early 21st century (1950-2014), as well as between the mid-20th century and late 21st century (1950-2099). For analysis of recent decades, a bilinear remapping technique is used to re-grid all models to the grid cell

size of HadGHCND, that is, 2.5° latitude x 3.75° longitude, and masked to only cover land regions where sufficient observational data are available. We define ‘sufficient’ as being grid cells with at least 80% of daily data available over 1950-2014, as well as at least 50% of data available for the first and last ten years of observational data. For analysis of future projected changes, all models are regridded to a common grid size of 2.5° latitude x 2.5° longitude to enable inter-model comparison.

Table 4.1 List of CMIP5 models used in this study and their institution.

Model	Modelling group
CanESM2	Canadian Centre for Climate Modelling and Analysis (CCCMA)
CNRM-CM5	Centre National de Recherches Météorologiques / Centre Européen de Recherche et Formation Avancée en Calcul Scientifique (CNRM-CERFACS)
CSIRO-Mk3-6-0	CSIRO in collaboration with Queensland Climate Change Centre of Excellence (CSIRO-QCCCE)
INM-CM4	Institute for Numerical Mathematics (INM)
MPI-ESM-LR	Max Planck Institute for Meteorology (MPI-M)
MPI-ESM-MR	Max Planck Institute for Meteorology (MPI-M)

4.2.2 Methods

For each model simulation as well as HadGHCND, daily temperature anomalies are calculated relative to a mean annual cycle of daily mean temperatures based on the entire period of analysis (1950-2014 for analysis of recent changes, and 1950-2099 for analysis of future changes). The data are then split into seasons, December to February (DJF), March to May (MAM) and September to November (SON), and all analyses are only applied to Northern Hemisphere land areas north of 30°N . For this study, we do not include boreal summer in our analysis as it was previously found to have only small changes in cold extremes relative to the mean that were less robust across a suite of CMIP5 models (Gross et al. submitted).

For each grid cell in each dataset, the seasonal minima of daily temperature anomalies are calculated annually for 1950-2014 and 1950-2099 separately, accounting for the differences in base period selection. The seasonal minima are then averaged over two periods, 1950-1981 and 1982-2014 for analysis of recent changes, and 1950-1979 and 2070-2099 for analysis of future changes, to calculate changes in the anomalously coldest days. Changes in seasonal mean temperature is similarly computed from daily mean temperature data. The difference between changes in the seasonal minima and changes in the seasonal mean is then calculated, hereafter referred to as ‘excess changes’. ‘Recent excess changes’ refer to excess changes between the mid-20th century and early 21st century, while ‘future excess changes’ refer to excess changes between the mid-20th century and late 21st century. Local significance of future excess changes is assessed by a Kolmogorov-Smirnov test (KS-test) at the 5% level.

To investigate the possible drivers of the enhanced warming of seasonal cold extremes relative to the mean in the mid- to high-latitude Northern Hemisphere regions, we assess several variables available at the daily time-scale in the selected CMIP5 models. This includes daily data for snow cover (CMIP5 variable name *snc*), snow amount (*snw*) and upwelling and downwelling longwave and shortwave radiation fluxes at the surface (*rlus*, *rlds*, *rsus*, *rsds*). We also assess surface albedo, calculated as the ratio of upwelling shortwave radiation and downwelling shortwave radiation, and horizontal temperature advection, which is derived for each model using the equation:

$$Tadv = u \times \left(\frac{dT}{dlon} \right) + v \times \left(\frac{dT}{dlat} \right)$$

where u and v are the zonal and meridional wind components (uas and vas , respectively), and dT is the local absolute daily mean temperature gradient in the zonal and meridional direction. For surface albedo, there are some instances in high-latitude regions where values are unrealistically large, as a result of low incoming shortwave radiation values that affect the calculation of surface albedo. In any instance where surface albedo values are outside of the physically reasonable 0 to 1 range, values are set to missing. Several other daily variables were also assessed, such as surface heat fluxes and cloud cover, but were found of low relevance as potential drivers of cold extremes in the seasons and regions being examined.

The analysis of the physical mechanisms related to the enhanced warming of cold extremes is limited to future changes, where the signal is stronger than for recent changes, and therefore shows a more robust identification of relationships. For each of the variables assessed, except temperature advection, data are evaluated on the specific day when the seasonal minima occurs. For temperature advection, a three-day average prior to the day the cold extreme occurs is used. This is because it is likely that larger changes in circulation would have more of an influence on temperature in the days leading up to the event, rather than the day of the event. Days leading up to the cold event were also examined for the other variables, but results showed no clear difference compared to using values on the exact day of the event. Excess changes are also calculated for each variable in much the same way as temperature, by taking the difference between the value of the variable on the days of the temperature extreme (or three-day average prior for temperature advection) and the seasonal mean of the variable. In essence, this removes the mean from the analysis and shows regions that experience increases or decreases in conditions related entirely to the days on which the cold extremes occur.

Results of the physical relationships are presented in two ways: one by showing maps of the variables as is shown for excess changes in temperature (to infer on the similarity of spatial patterns), and the other by showing scatter plots of correlations between each variable and future excess changes in cold extremes. The former is included in supplementary material due to the number of figures while the latter are included within the main body of the manuscript. For the scatter plots, annual ‘excess’ values for the two time periods used for the future analysis are calculated as the difference between the variable value on the day the seasonal minima occurs and the seasonal mean of the respective variable. Weighted area-averages of the annual excess values are then calculated for all grid boxes within a selected region that adhere to a specified condition that only includes grid cells that have a statistically significant future excess change exceeding 1°C. Two regions are assessed, one covering North America, and the other covering much of northern Eurasia, and are common to all models (see Supplementary Material Figure S3.1). Regressions are calculated using total least squares regression, with correlation coefficients computed using Spearman’s rank correlation.

4.3 Results

4.3.1 Recent changes in cold extremes relative to the mean

Excess changes between recent decades in seasonal cold extremes relative to corresponding mean temperatures are shown for HadGHCND and individual CMIP5 models for DJF (Figure 4.1), MAM (Figure 4.2) and SON (Figure 4.3). In all figures, and subsequent maps of excess changes, positive values indicate regions where cold extremes have warmed more than the mean, while negative values indicate regions where the mean has warmed more than the cold extremes. Global pattern correlations are shown in the top right above each map, indicating the spatial similarity between each model with HadGHCND.

Pattern correlations between the observations and the individual CMIP5 models are low, indicating differences in the spatial patterns, however, this may be due to the relatively low magnitude of excess change and high spatial noise across the observations and models over the past 60 years. During winter (Figure 4.1), the observations and most of the individual models agree that cold extremes have warmed more than the mean for parts of Canada and Alaska, as well as parts of western Europe, and central Asia, although the exact locations of these excess changes differ (as reflected in the low pattern correlations). Spring months (Figure 4.2) show slightly higher pattern correlations between observations and the models, with all models, excluding INM-CM4, showing even higher correlations in autumn (Figure 4.3). Compared with winter, MAM and SON generally show stronger and more widespread positive excess changes over Europe and Eurasia. For most of the models, the strongest excess changes overall are shown for parts of Russia/Eurasia, especially during the shoulder seasons, except for CanESM2 which shows particularly strong positive excess changes in DJF in central/eastern Europe. In the observations, the strongest excess changes are shown for much of Eurasia in SON and DJF, but mostly central Europe and North America during MAM.

The observations and the CMIP5 models suggest the same general pattern of excess changes, with the most prominent positive excess changes in recent decades occurring in the northern continental interiors for all seasons shown. This motivates us to assess projections of enhanced warming of cold extremes in the selected six climate models over the Northern Hemisphere.

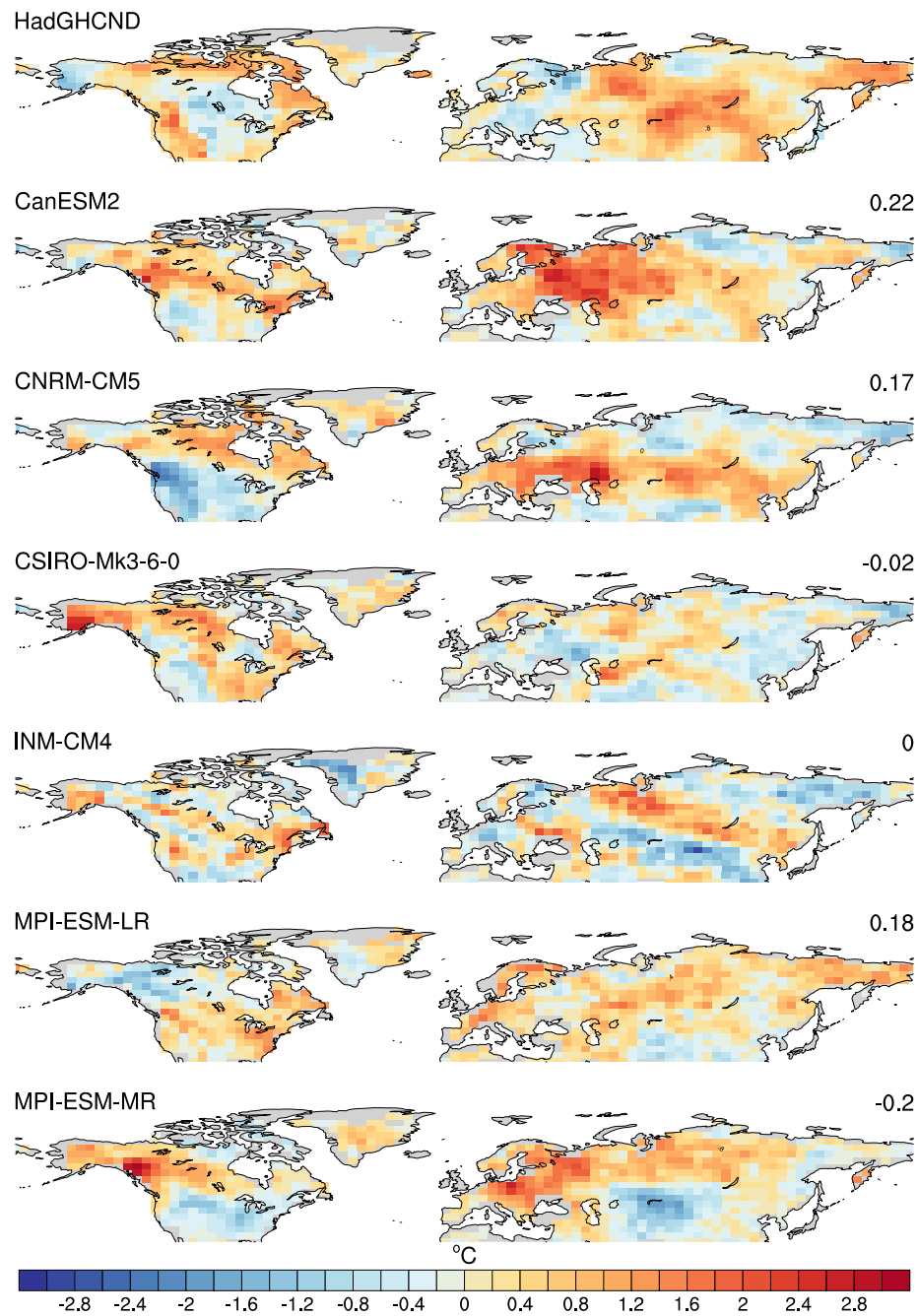


Figure 4.1 Recent excess changes (1982-2014 – 1950-1981) in cold extremes (seasonal minima – seasonal mean) for December – February. The number in the top right above the maps for each model represents the pattern correlation with that model and HadGHCND.

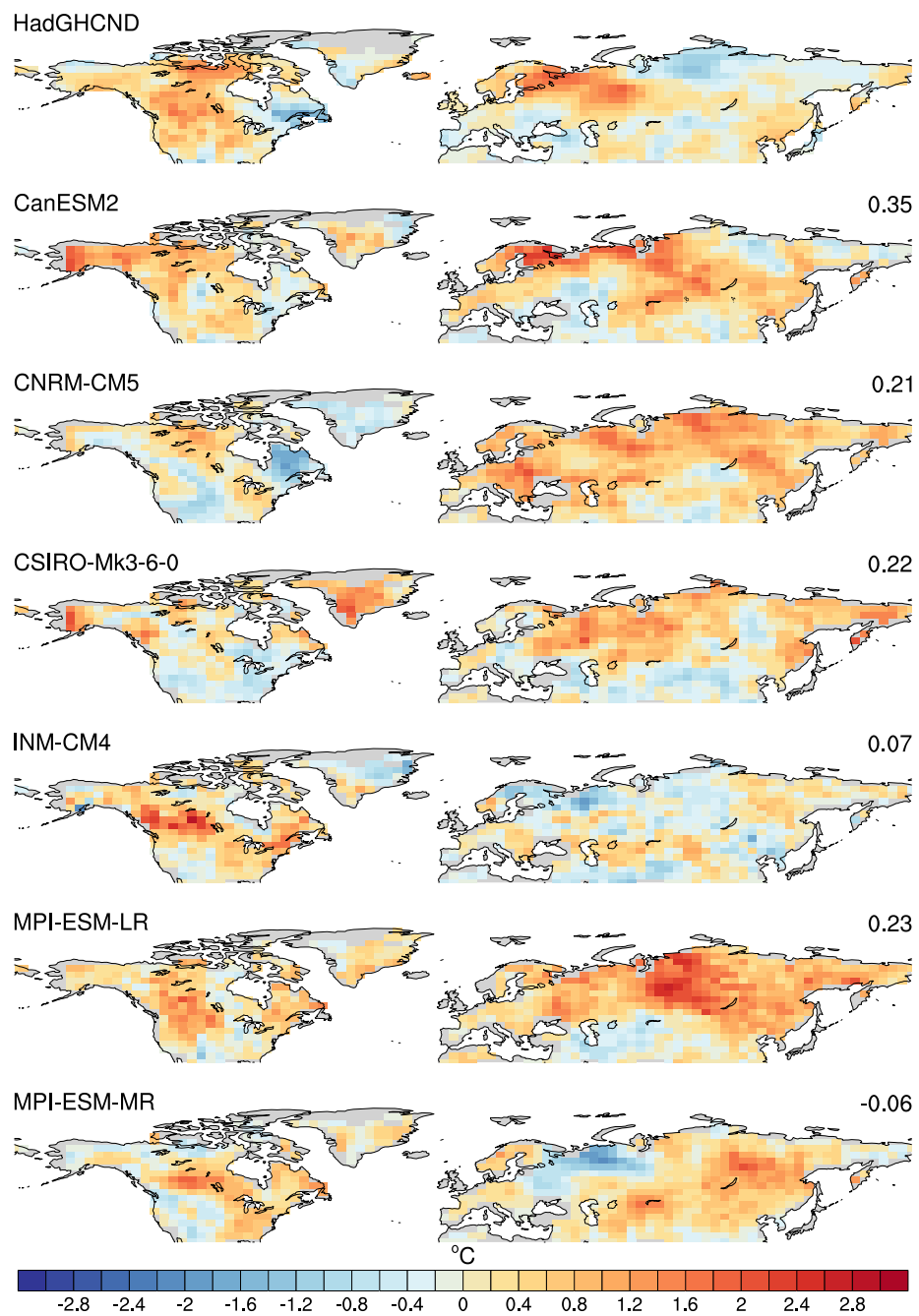


Figure 4.2 As Figure 4.1, but for March – May.

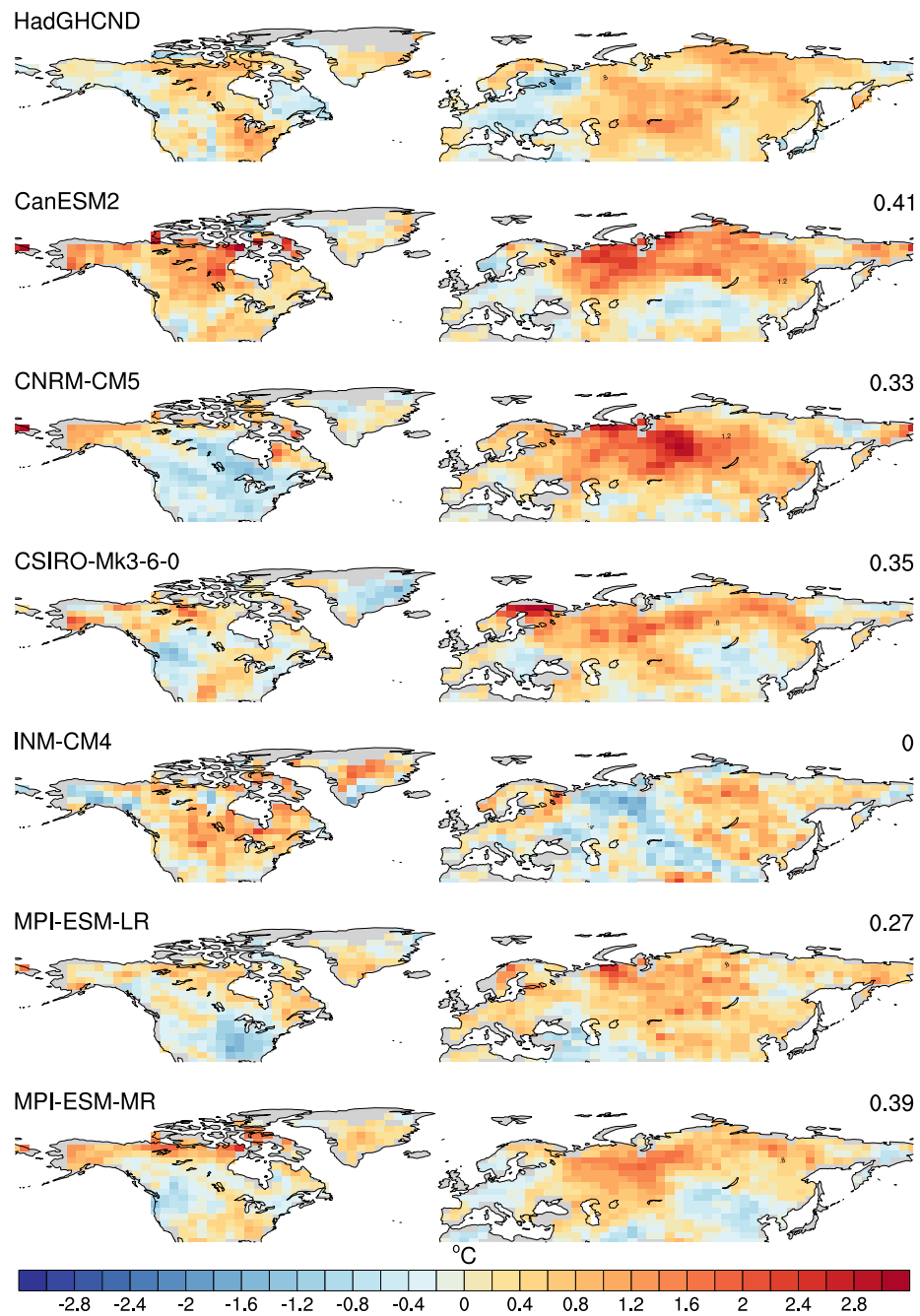


Figure 4.3 As Figure 4.1, but for September – November.

4.3.2 Projected excess changes in cold extremes between future and past decades

Projections of excess changes in cold extremes comparing the mid-20th century with the late 21st century are shown for each of the six individual CMIP5 models for DJF (Figure 4.4), MAM (Figure 4.5) and SON (Figure 4.6). Stippling indicates

grid cells where changes are statistically significant, as assessed by a KS-test. Much of the Northern Hemisphere mid- to high-latitudes indicate that cold extremes are projected to warm significantly more than mean temperatures. For winter (Figure 4.4), CanESM2 shows the most prominent excess changes, covering much of North America, eastern and central Europe, and northern Eurasia. While not quite as strong, all other models show mostly significant positive excess changes in cold extremes for many of the same regions, particularly Alaska, eastern Canada, and eastern and central Europe. There is some variation over northern Russia/Siberia, with CNRM-CM5 showing significant negative excess changes around -1.5°C , opposed to relatively strong positive excess changes around 3°C shown in CanESM2, MPI-ESM-LR and MPI-ESM-MR. As in the recent excess changes, spring and autumn generally show a more widespread and systematic pattern of positive excess changes over Eurasia, Canada and Alaska. In MAM (Figure 4.5), the models show some differences in the southern parts of the U.S. and Eurasia, which mostly show non-statistically significant negative excess changes. There is relatively good agreement between models for the strong excess changes over much of the Eurasian and Russian region during MAM, except for INM-CM4 which shows smaller and mostly non-significant positive excess changes for the northern half of Russia. In terms of model agreement, SON (Figure 4.6) is similar to MAM, with the greatest model agreement over Alaska, Canada and central/eastern Europe and northern Eurasia, excluding non-significant small excess changes in northern Russia for CSIRO-Mk3-6-0. For Alaska, northern Canada and northern Eurasia, positive excess changes are generally largest for SON compared with the other seasons, except CanESM2 which shows particularly large excess changes in DJF.

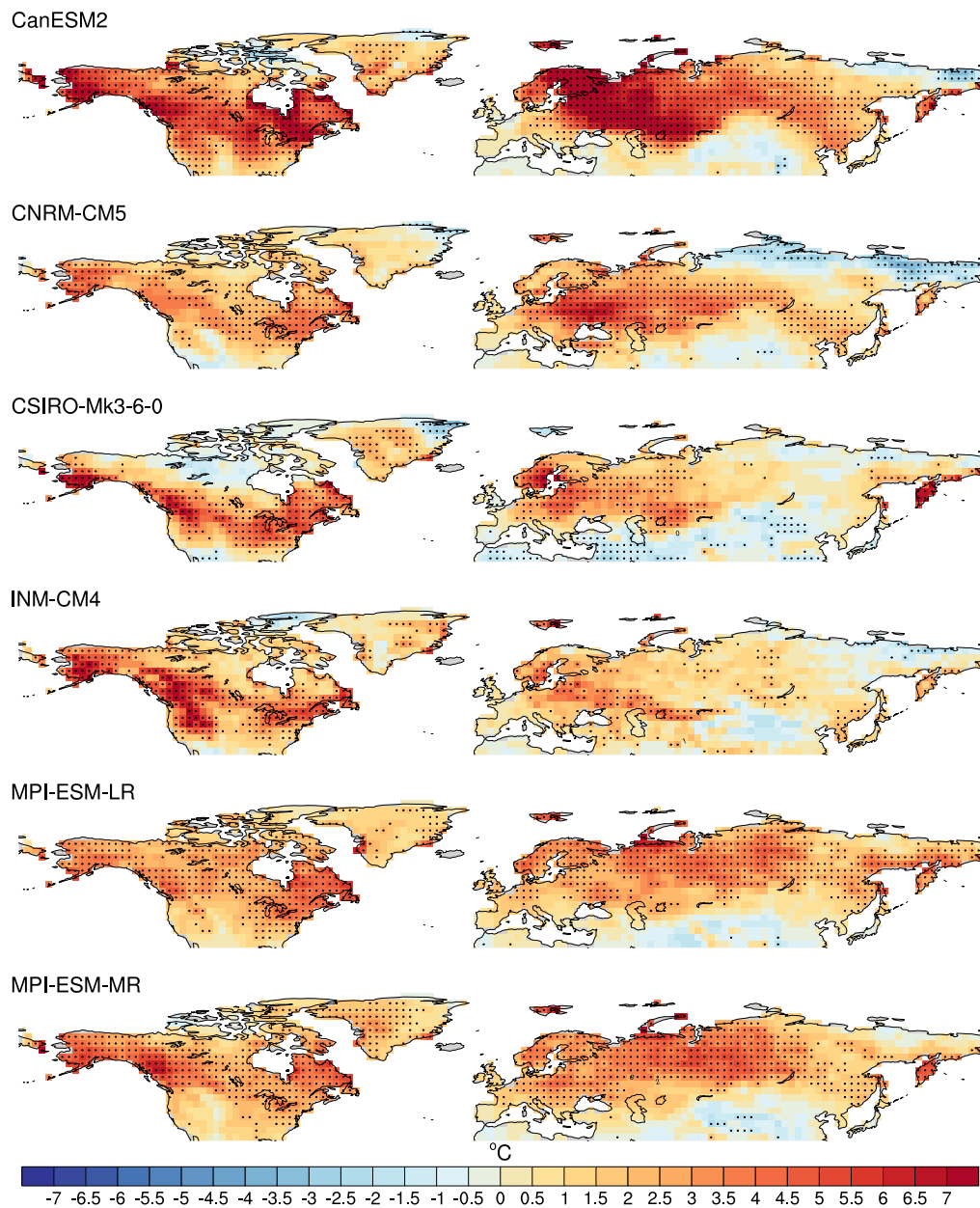


Figure 4.4 Future excess changes (2070-2099 – 1950-1979) in cold extremes (seasonal minima – seasonal mean) for December – February. Stippling indicates grid cells that are significant at the 5% level as assessed by a KS-test.

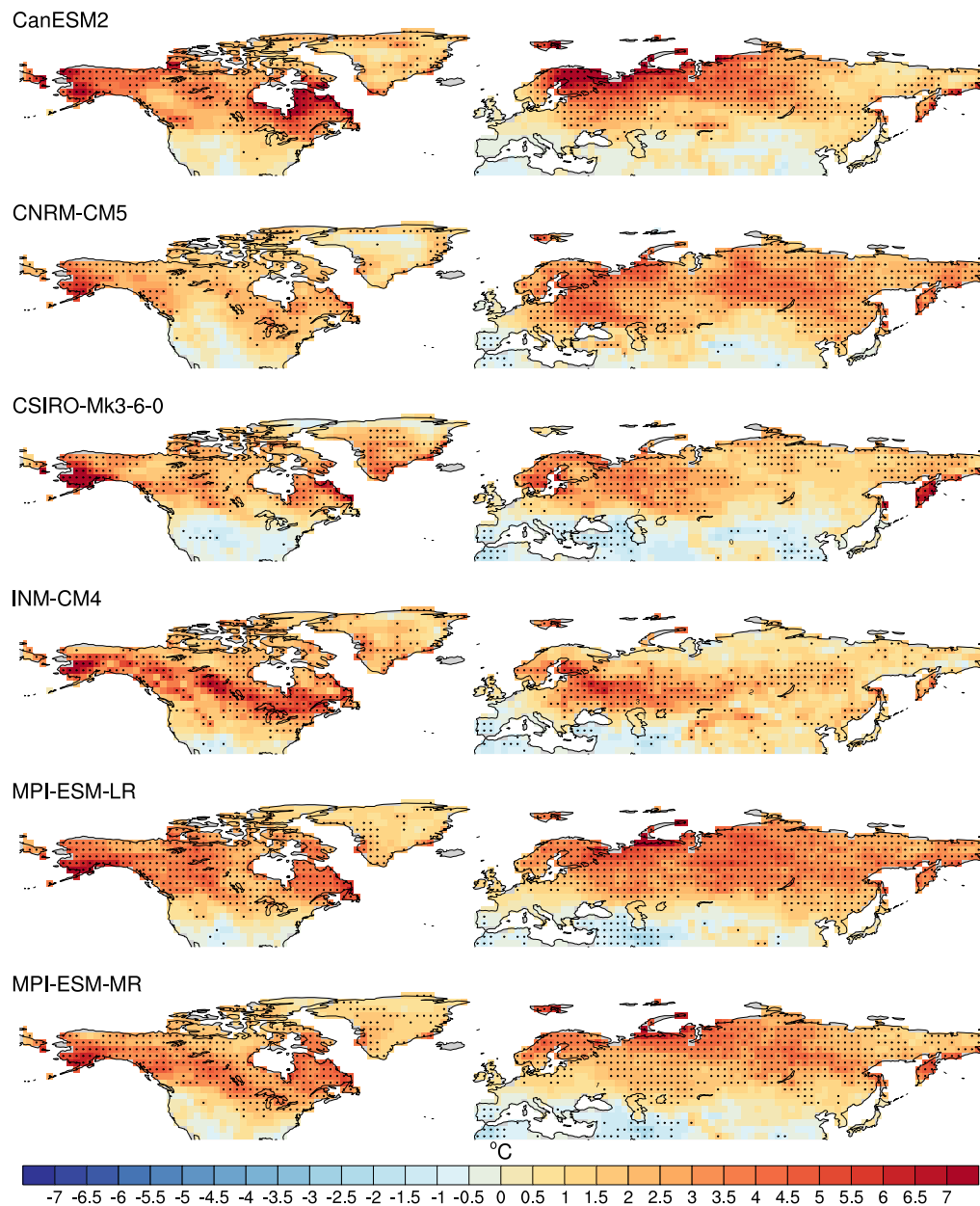


Figure 4.5 As Figure 4.4, but for March – May.

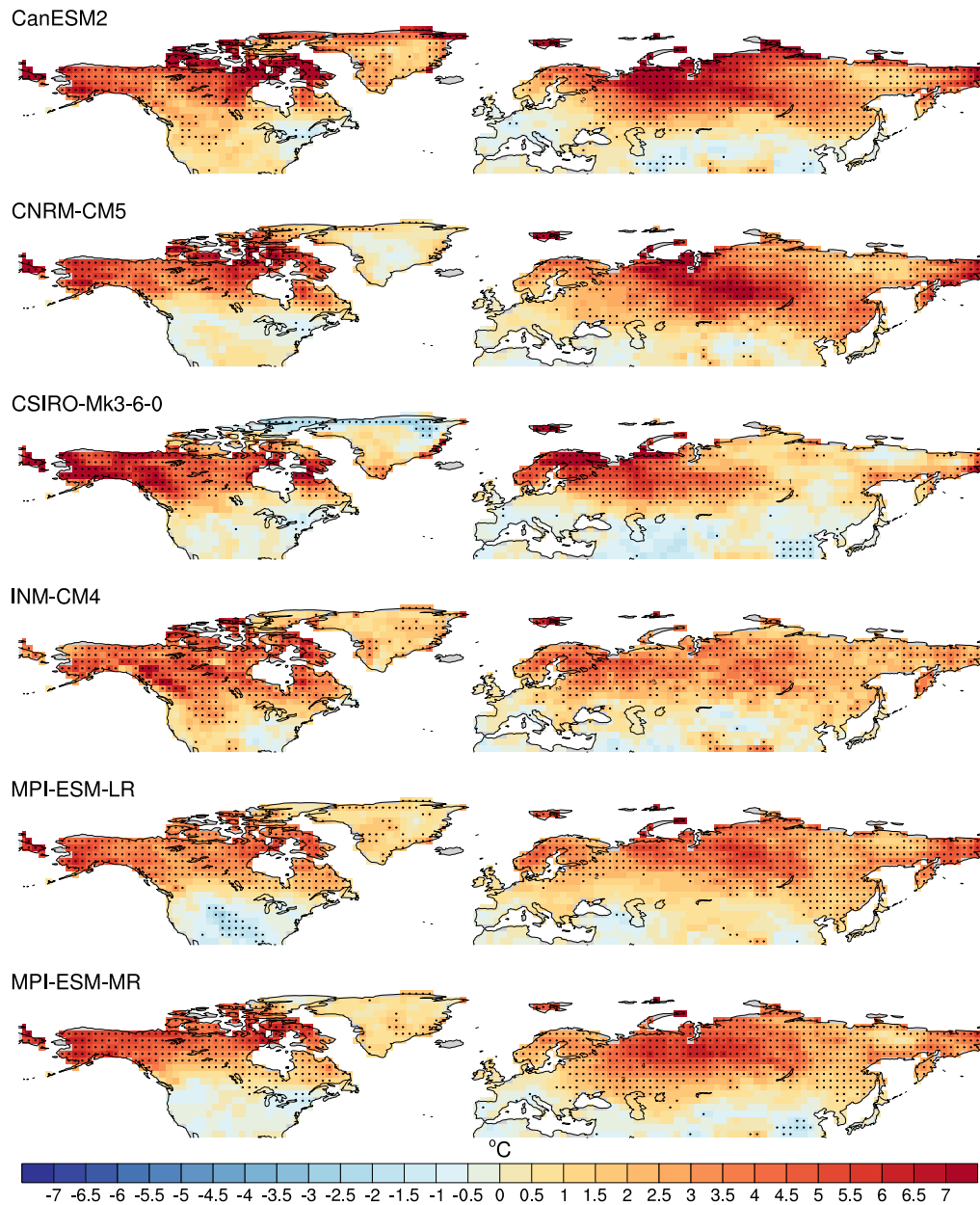


Figure 4.6 As Figure 4.4, but for September – November.

Overall, future excess changes are robust and systematic across the models in winter, spring and autumn, with many mid- to high-latitude regions projecting enhanced warming in cold extremes relative to the mean. To explore the possible physical mechanisms driving the enhanced warming of cold extremes, we focus on the regions that show the most robust signals. The strongest excess changes across all the shown seasons are over northern Eurasia and northern North America. This

is relatively consistent with the largest recent excess changes occurring in the northern continental interiors in observations and CMIP5 models, though is much more widespread and systematic in the predicted patterns.

4.3.3 Projected changes in cold air temperature advection prior to cold extremes

Due to the evidence suggesting Arctic amplification, and consequent changes in thermal advection, as a main driver of decreasing temperature variability in Northern Hemisphere regions (e.g. Screen and Simmonds 2010; Screen 2014; Schneider et al. 2015; Holmes et al. 2016; Rhines et al. 2017), we first consider projections of changes in cold air temperature advection for the three days prior to the cold event. Figure 4.7 shows future changes in actual and excess cold air temperature advection in the CanESM2 model. Results for the remaining models are included as Supplementary Material (Figures S3.2 – S3.3), with any key differences between the models highlighted below.

The most notable features occur for DJF, which shows reductions in cold air temperature advection for much of North America as well as Eurasia. This is evident for changes in both actual and excess cold air temperature advection, which suggests the changes are related to a change in the overall mean state of cold air temperature advection, rather than changes associated with the days directly prior to the day the cold extreme occurs. This is similarly true for the other models (see Figures S3.2 – S3.3), however, the magnitude of reduction in cold air temperature advection is generally larger in CanESM2 compared with the other models. This is reflected in future excess changes in cold extremes, where CanESM2 is generally warmer than the other models for DJF. For much of the central and eastern U.S., reductions in cold air temperature advection of at least -2°C are projected for the

late 21st century, with the same areas showing the largest positive excess changes in cold extremes. This is indicative of reductions in cold air temperature advection, both related to the day of the extreme as well as changes in the mean-state, being a dominant driver of the enhanced warming of cold extremes relative to the mean during boreal winter (see Figure 4.4). Similarly, the greatest reductions in cold air temperature advection over the European continent occur in Scandinavia and Eurasia, which also show high magnitude excess warming in cold extremes during DJF. This same pattern is evident across all of the selected climate models. Though some reductions in cold air temperature advection are shown scattered over North America and Eurasia for MAM and SON, the spatial pattern does not match with the seasonal future excess changes in cold extremes like DJF does.

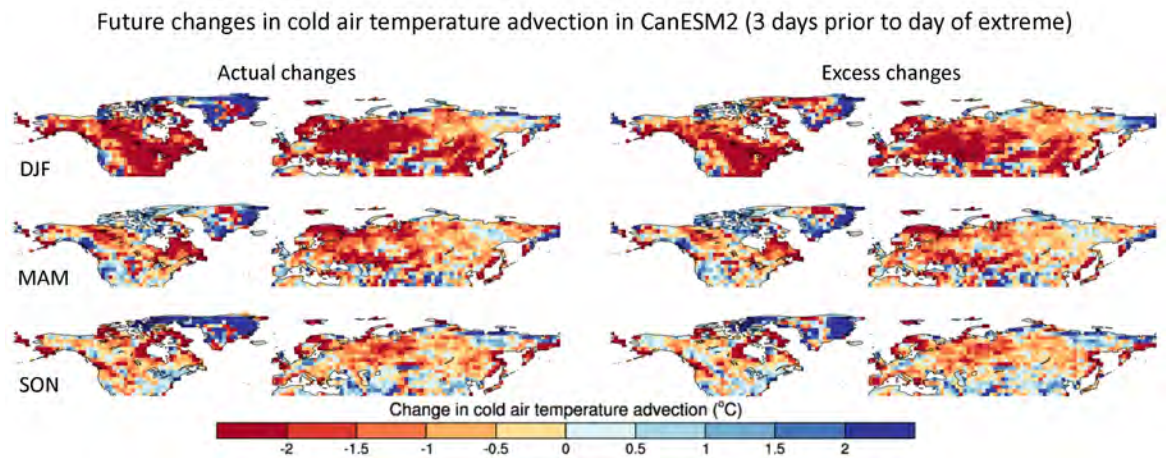


Figure 4.7 Future changes (2070-2099 – 1950-1979) in actual (left column) and excess (right column) cold air temperature advection for December – February (top row), March – May (middle row) and September – November (bottom row) in CanESM2. Values are calculated using the average cold air temperature advection for the three days prior to the day the annual seasonal minimum occurs, with negative values indicating reductions in cold air advection, and positive values indicating increases. All six of the selected climate models used in this study are shown in Figures S3.2 – S3.3.

Based on these results, it is evident that a reduction in cold air temperature advection is driving the projected excess changes in cold extremes over much of North America and Eurasia during winter. Both shoulder seasons, however, have a

less clear signal with generally smaller, spatially scattered reductions in cold air temperature advection, pointing to other mechanisms as a more dominant driver of the enhanced warming of cold extremes in spring and autumn months.

4.3.4 Projected changes in snow cover and surface albedo associated with cold extremes

Many of the grid cells showing significantly strong excess changes are located in regions that experience high snow fall. Snow cover and associated surface albedo feedbacks therefore play a major role in temperature variability in these regions, but it is not clear if this relationship extends to the accelerated warming of cold extremes relative to local mean temperatures, and the seasonal influence remains uncertain. The subsequent results show scatter plots of excess changes in cold extremes and snow cover (Figure 4.8) and surface albedo (Figure 4.9) for the North American region and northern Eurasian region outlined in Figure S3.1. The different coloured symbols and regression lines represent each of the six CMIP5 models. As outlined in Section 4.2, these changes are calculated for the exact day when the cold extreme occurs. Projections of changes in actual and excess snow cover and surface albedo for the days of the cold extreme are included as Supplementary Material (Figures S3.4 – S3.5 and S3.8 – S3.9, respectively). For additional information on the snow-temperature relationship, future changes in snow amount are also included in Supplementary Material (see Figures S3.6 – S3.7).

For both regions, mostly significant negative correlations between snow cover and excess cold extremes are shown for all seasons, aside from excess snow cover in DJF. For the North America region (Figure 4.8a-f), while all models show significant negative correlations of at least -0.74 for actual snow cover in DJF (Figure 4.8a), all models except CSIRO-Mk3-6-0 show significantly positive

correlations for excess snow cover in DJF (Figure 4.8b). From Figures S3.4 – S3.5, parts of North America, particularly southern Alaska, southern Canada and along the north-western coast of the U.S., show projected decreases in actual snow cover, but slight increases in excess snow cover. This suggests that the feedback between snow cover and the projected enhanced warming of cold extremes is related to overall reductions in the mean snow cover over winter, rather than decreases in snow cover specifically on the day of the extreme only. For MAM and SON, negative correlations are stronger for both actual snow cover (Figure 4.8c,e) and excess snow cover (Figure 4.8d,f), compared to DJF, and are the strongest for SON. Again, this is reflected in the maps, where actual snow cover is projected to decrease around 40% for much of Alaska and northern Canada during autumn months, while decreasing somewhat less and slightly further south during spring (Figure S3.4). Smaller decreases are projected for excess snow cover for many of the same regions, excluding Alaska which shows mostly small positive increases during spring (Figure S3.5).

Northern Eurasia (Figure 4.8g-l) shows similar correlations to that of North America. The overall largest correlations between snow cover and excess cold extremes occur in SON (Figure 4.8k,l), with some models, for example CanESM2 and CNRM-CM5, showing correlations as high as -0.91 between actual snow cover and excess cold extremes, and decreases over 45% for parts of western Russia and Scandinavia (Figure S3.4). Correlations are slightly lower for actual snow cover in MAM (Figure 4.8i), with the largest decreases in snow cover occurring in Scandinavia and central/eastern Europe, and no substantial changes in northern Russia (Figure S3.4). This is reflected in projected changes in snow amount (Figure S3.6), with increases shown for regions showing no changes in snow cover. Correlations with excess snow cover in MAM are substantially smaller for most

models, with parts of northern Russia showing small increases in snow cover (Figure S3.5) and snow amount (Figure S3.7).

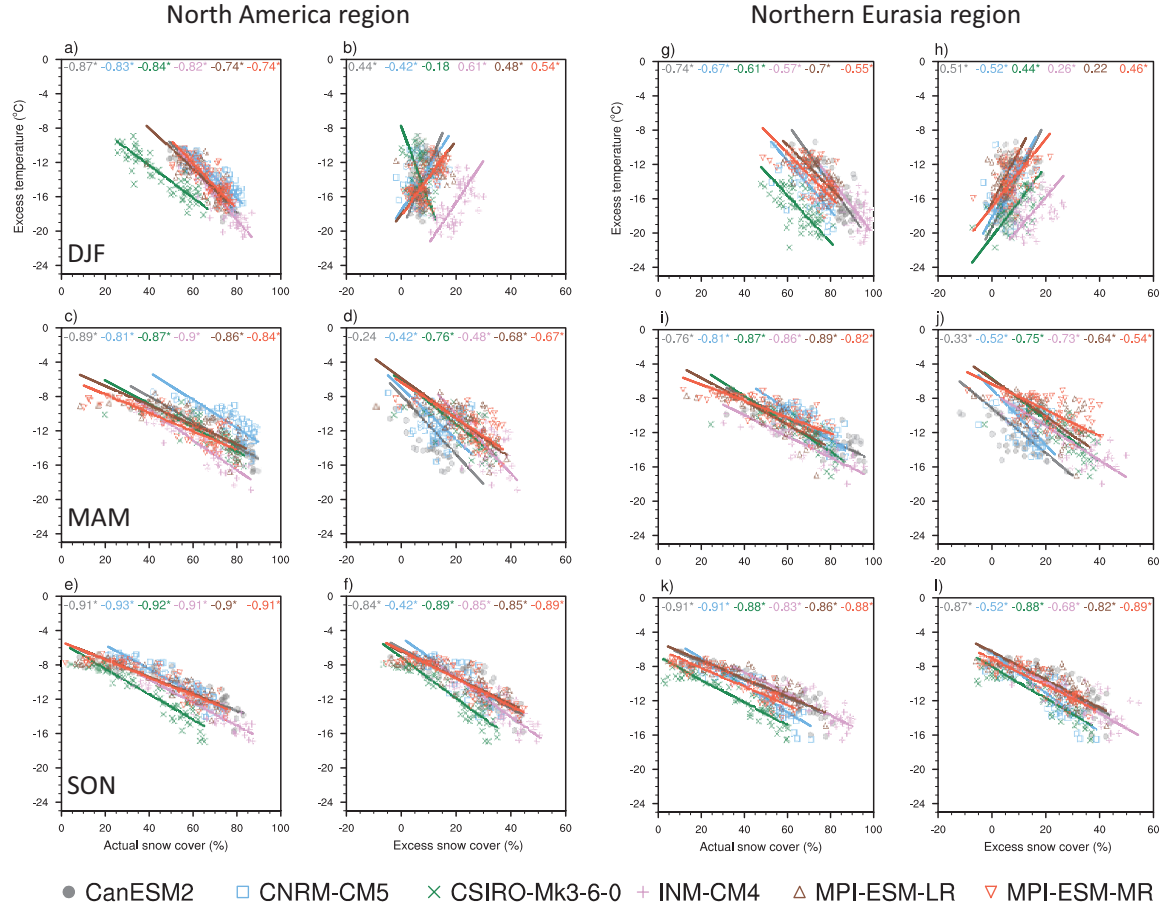


Figure 4.8 Scatter plots showing annual values of excess temperatures in cold extremes for each season (seasonal minima – seasonal mean) on the y-axis, and annual values in each season of actual snow cover, calculated on the day the cold extreme occurs, on the x-axis (1st and 3rd column). The 2nd and 4th column show values of excess snow cover on the x-axis (i.e. snow cover on the day of the extreme – mean annual snow cover). Each row represents a different season: December – February in the 1st row, March – May in the 2nd row, and September – November in the 3rd row. Each value is an area-average of two regions (see Figure S3.1): North America (a-f) and northern Eurasia (g-l). The straight lines indicate the regression slope for each model calculated using total least squares regression. Correlation coefficients are shown at the top of each panel, with the different colours indicating the model. * indicates significance at the 5% level.

Decreases in snow cover imply that reductions in surface albedo are a likely factor contributing to the enhanced warming of cold extremes relative to the mean. Correlations between surface albedo and excess temperature (Figure 4.9) indeed show strong similarities with those of snow cover, with the largest negative

correlations shown for SON for both North America (Figure 4.9a-f) and northern Eurasia (Figure 4.9g-l). As for snow cover, the strongest overall decreases are shown for actual changes in surface albedo over Alaska, northern Canada and Eurasia during autumn months (Figure S3.8). Differences in the magnitude and sign between actual surface albedo and excess surface albedo are also clear. For example, mostly positive correlations with excess surface albedo are shown for DJF for both regions (Figure 4.9b,h). During boreal winter in high-latitude regions, solar insolation is low, and so it is expected that surface albedo, which is calculated using shortwave radiation fluxes, is less of a factor in driving excess changes in temperature during these months.

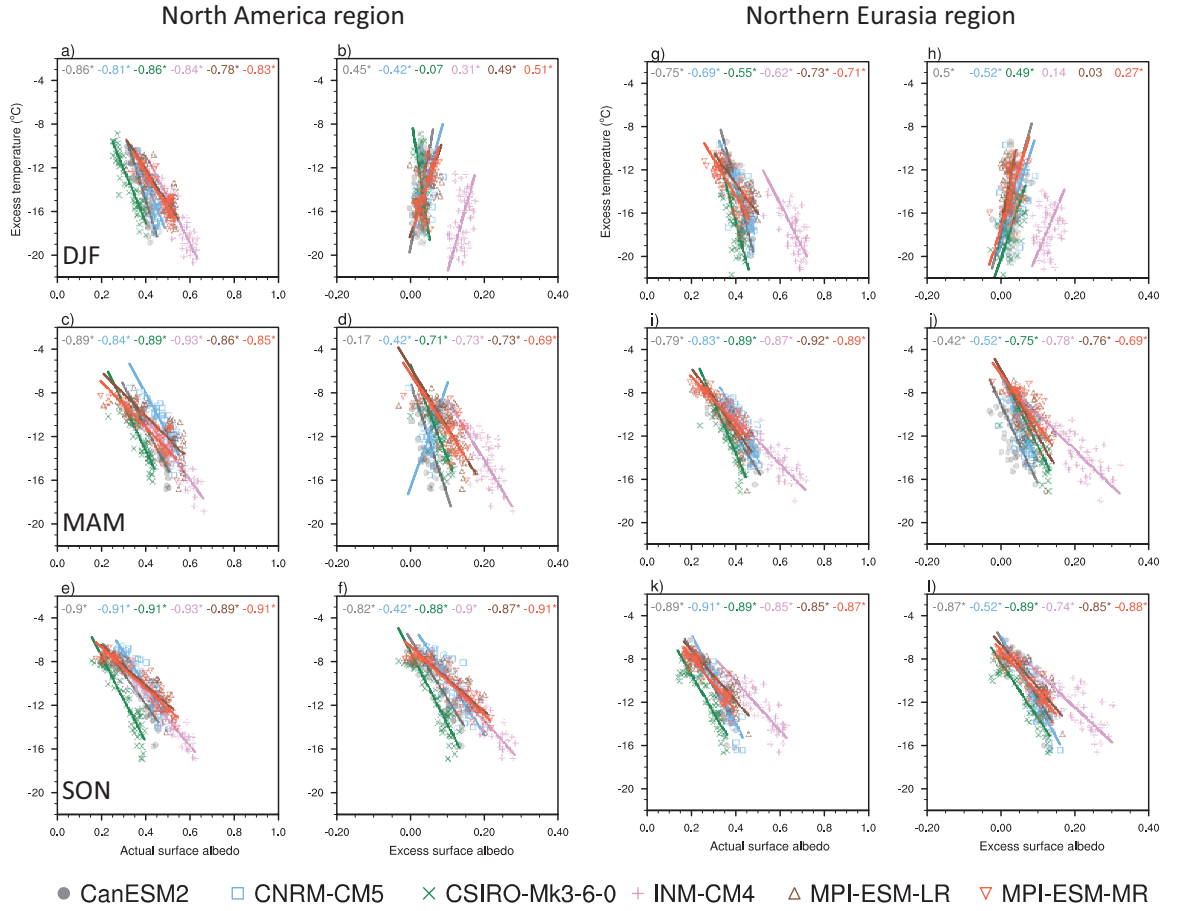


Figure 4.9 As Figure 4.8, but for surface albedo.

There is a clear relationship between decreases in snow cover, associated lower albedo and the enhanced warming of cold extremes for many regions in the Northern Hemisphere mid- to high-latitudes. While negative correlations are shown for actual snow cover and excess cold extremes during winter for both North America and Eurasia, projected decreases in actual snow cover, as shown in the maps in Figure S3.4, are generally much smaller than they are for both shoulder seasons, especially autumn months which show the overall largest decreases and highest correlations with excess temperatures in cold extremes. Much of this relationship between snow cover, surface albedo and excess temperatures in cold extremes is a consequence of overall decreases in the mean-state of both snow cover and surface albedo, rather than decreases in snow cover specifically on the day the cold extreme occurs. This is consistent across the selection of CMIP5 models used in this study.

4.3.5 Projected changes in the timing of anomalously cold days

The projected enhanced warming of cold extremes relative to the mean found for many mid- to high-latitude Northern Hemisphere regions, in all seasons except boreal summer, is clearly related to excess heat at the surface that acts to decrease the severity of the anomalously coldest days of the season. During spring and autumn, much of this is likely a consequence of less snow and lower albedo, leading to increases in absorbed shortwave radiation at the surface and consequently amplifying the warming of cold extremes, creating a positive feedback. In addition to these relationships presented above, we also analysed an increase in net radiation on the days of the cold extremes in both shoulder seasons, with increases in incoming shortwave radiation being the largest contributor (not shown). These increases are, however, largely attributable to temporal shifts in the occurrence of the largest negative temperature anomalies in the shoulder seasons.

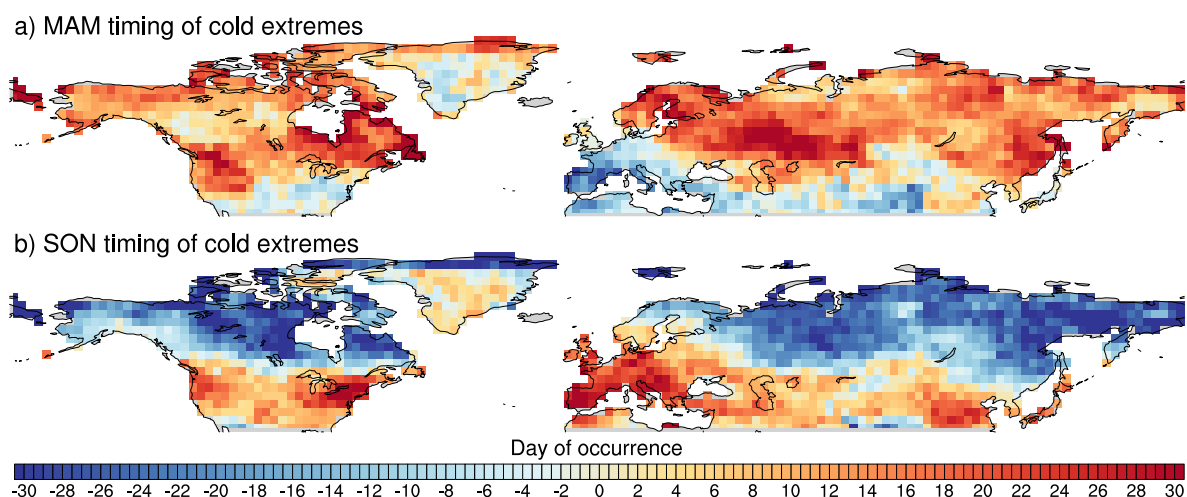


Figure 4.10 Changes in the timing of the anomalously coldest day in the season between 2070–2099 and 1950–1979 for March – May (a) and September – November (b), as shown for the CanESM2 model. All models and seasons, excluding boreal summer, are shown in Figure S3.10.

Figure 4.10 shows the projected change in the timing of the seasonal minimum of daily anomalies during MAM and SON, as simulated in CanESM2 (see Figure S3.10 for all models and seasons). For much of the Northern Hemisphere mid- to high latitudes, excluding the most southerly parts, the Mediterranean region and parts of Greenland, the anomalously coldest days are projected to occur later in the season during MAM (Figure 4.10a). In some regions, such as central-western Europe and eastern Canada, the anomalously coldest spring days are projected to occur more than 30 days later in the late 21st century, compared to those simulated in the mid-20th century. For SON (Figure 4.10b), the anomalously coldest days are projected to shift to earlier in the season for most high latitude regions in the Northern Hemisphere. For example, the anomalously coldest days are projected to occur over 30 days earlier than they did in the mid-20th century much some regions in northern Canada and northern Eurasia. This change in the timing of cold extremes suggests an overall flattening of the seasonal cycle in these extratropical Northern Hemisphere regions. Coupled with the cold extremes warming at a faster

rate than average temperatures, this suggests these regions will experience generally a longer duration of warmer months and a shorter duration of colder months.

4.4 Discussion and conclusions

Cold extremes are projected to warm more than seasonal average temperatures for much of the Northern Hemisphere mid- to high-latitude regions, for all seasons except boreal summer. Though these projected changes differ slightly in magnitude and spatial pattern depending on the CMIP5 model used, the most prominent excess changes are robust across the selection of models. These changes are likely related to projected changes in horizontal temperature advection, snow cover and surface albedo feedbacks. The season in which the excess changes in cold extremes occur largely dictates which physical mechanisms are at play.

Decreases in snow cover and lower surface albedo are more associated with excess changes in cold extremes during spring and autumn months. Due to low solar insolation in winter months, and subsequently only small effects from changes in shortwave radiation and surface albedo, reductions in cold air temperature advection in the days leading up to the extreme event is the dominant driver during boreal winter. This latter finding is likely a consequence of Arctic amplification and is in agreement with previous studies linking the warming of cold days in winter months with warmer than usual air being brought from the Arctic to lower latitudes (e.g. Screen 2014; Schneider et al. 2015; Holmes et al. 2016; Rhines et al. 2017).

In contrast, Arctic warming and associated sea ice loss has been argued to result in more persistent severe cold air outbreaks over continental regions in the mid-latitudes during boreal winter (e.g. Kodra et al. 2011; Cohen et al. 2014, 2018; Francis and Vavrus 2015; Zhang et al. 2016). Recent cold snaps in the United

States and Eurasia, such as those observed in the boreal winter of 2012/2013, can largely be explained by a southward shift in the jet stream and a weakening of the stratospheric polar vortex (Francis and Vavrus 2015; Zhang et al. 2016; Cohen et al. 2018; Kretschmer et al. 2018). Though some argue these events are likely transient and related to atmospheric decadal variability (e.g. Barnes and Screen 2015; Ayarzagüena and Screen 2016; Sun et al. 2016), others suggest that severe cold snaps in the Northern Hemisphere mid-latitudes might persist in response to continued Arctic warming (e.g. Kodra et al. 2011; Francis and Vavrus 2012, 2015; Cohen et al. 2014). While there is some disagreement between models and observations in how they simulate the observed cold outbreaks (e.g. Cohen et al. 2013; Sun et al. 2016), there is robust model agreement that mid-latitude cold extremes are projected to decrease in severity (Screen 2014; Barnes and Screen 2015; Screen et al. 2015a,b; Ayarzagüena and Screen 2016). Some have also suggested that cold air outbreaks will decrease in duration and frequency (e.g. Screen et al. 2015a,b), however, this remains unclear and requires further work (e.g. Ayarzagüena and Screen 2016). Though the results in this study cannot infer anything regarding the frequency and duration of cold spells, it is evident that cold extremes are projected to warm in excess of increasing mean temperatures over North America and Eurasia during boreal winter by the end of the 21st century.

Arctic amplification and associated thermal advection is also suggested to be a particularly strong driver of decreased cold extremes in autumn months (e.g. Screen 2014; Holmes et al. 2016). While significant negative correlations between cold air temperature advection and excess cold extremes are shown for all models for both North America and Eurasia during autumn (see Figure 4.7), reductions in cold air temperature advection are substantially less coherent than they are for winter (see Figures S3.2 – S3.3). This highlights that area-averaged correlations are sometimes

misleading in terms of exploring the processes driving temperature extremes over a large region. While Arctic amplification and associated reductions in cold air temperature advection may be having somewhat of an impact on the warming of cold extremes during the shoulder seasons, other physical mechanisms likely have a greater influence on changes in spring and autumn cold extremes.

For both shoulder seasons, ‘hot spots’ of enhanced warming of cold extremes relative to the mean are shown for much of Alaska, Canada and northern Eurasia (Figures 4.5 – 4.6). During autumn, snow cover shows an exceptionally similar spatial pattern to excess changes in cold extremes for all models (see Figure S3.4 and Figure 4.6, respectively), with the largest excess changes in cold extremes matching regions showing the largest decreases in snow cover. Spatial similarities between snow cover and excess changes in cold extremes during spring are less obvious than they are for autumn, with slightly lower correlations, though the largest decreases in snow cover are still associated with significant excess changes in cold extremes (see Figure S3.4 and Figure 4.5). Previous work has suggested that spring has the strongest snow-temperature relationship, largely due to increases in latent heat from snowmelt (e.g. Dutra et al. 2011; Xu and Dirmeyer 2011; Diro et al. 2018). Many of the regions showing the strongest relationship between projected snow cover and the projected amplification of warming cold extremes, such as northwestern U.S., southern and northeast Canada and the Rocky Mountains, are in agreement with historical simulations of the snow-temperature association during winter and spring months (Dutra et al. 2011; Diro et al. 2018). While some high-latitude regions in northern Canada and northern Russia show projected increases in snow amount during spring (see Figures S3.6 – S3.7), with the same regions and seasons showing no substantial changes in snow cover (see Figures S3.4 – S3.5), correlations between snow and excess temperature in autumn are generally larger.

This infers that even if springtime is associated with a stronger snow-temperature relationship, due to increases in snowmelt, decreases in snow cover have more of an influence on warming anomalously cold days in autumn months.

A change in surface albedo feedback, as a result of a change in snow cover, is more likely to influence cold days in winter and early spring due to peak snow accumulation during this time. While results presented here show projections of decreasing albedo for many regions in North America and Europe, autumn shows the largest decreases in surface albedo (see Figure S3.8), which is closely related to the projected decreases in snow cover. We note that our calculation of surface albedo, as simply the ratio between absorbed and reflected shortwave radiation, may not be capturing certain aspects that are important to snow affected areas. For example, the boreal forest is a region with extensive snow fall and dense vegetation cover, and the varying land parameterizations within the climate models may not necessarily be capturing the snow that is intercepted by trees in the canopy (Lorant et al. 2014; Thackeray et al. 2015). This then has important implications on surface albedo and therefore surface temperature.

Biases in climate model simulations of snow-albedo feedbacks have been found over the boreal forest region, with significant underestimations compared with observations, especially during periods where snowmelt is high, such as in early March (Fletcher et al. 2012; Lorant et al. 2014; Qu and Hall 2014; Thackeray et al. 2014, 2015). However, biases in the models are reduced over larger study regions (Thackeray et al. 2015), with area-averaging over large regions also likely to suppress any biases. Biases may also simply be a consequence of temperature, with cold biases having more snow, and warm biases leading to more snowmelt. The ability of climate models to capture snow-albedo feedbacks is also complicated by

factors relating to snow type and the ageing of snow, which can also influence surface temperatures (Thackeray et al. 2015; Diro et al. 2018). Improving the ability of climate models to capture realistic changes in snow cover and surface albedo would enable more accurate projections of future cold extremes. Biases in the representation of physical relationships may control the simulation of long-term changes in cold extremes. Given the availability of suitable observations of relevant land variables, an evaluation of the land-atmosphere relationships as outlined here may serve to develop process-based constraints to reduce the uncertainty in future projections, similar to previous approaches focussing on the processes driving hot extremes in summer (Donat et al. 2018).

While our findings are consistent with the theory that less snow cover and associated reductions in surface albedo lead to anomalously warmer temperatures on cold days, it is unclear whether these variables are driving the enhanced warming of cold extremes, or vice versa. It is true, however, that the positive feedback between snow cover, surface albedo and surface temperature exacerbate the warming of cold extremes. It would be useful for future studies to run climate model simulations with and without snow cover prescribed to quantify the specific impact on simulated cold extremes, enabling more confident conclusions regarding snow cover and albedo as a driver of enhanced cold extremes.

Similar to albedo, radiative fluxes are strongly influenced by changes in the surface which affects the overall surface energy budget. Decreases in snow cover which lead to lower albedo will result in increased absorption of incoming shortwave radiation for regions and seasons with enough solar insolation. While we did find some increases in incoming shortwave radiation on the days when the coldest anomalies occur, this is more an artefact of the timing in which the cold extreme occurs. For

high-latitude regions, the seasonal minimum in spring time is projected to occur later in the season, with the autumn minima projected to occur earlier in the season, suggesting an overall flattening of the seasonal cycle. Changes in the annual cycle of surface temperature have been detected before, with a shift to earlier seasons by 1.7 days shown for extratropical land regions in recent decades (Stine et al. 2009). Further, changes in the seasonal cycle have been projected by CMIP3 models too, with colder temperatures occurring later and warmer temperatures occurring earlier, reducing the amplitude of the seasonal cycle (Dwyer et al. 2012). These shifts are argued to be a consequence of anthropogenic climate change driving sea ice loss (Dwyer et al. 2012; Santer et al. 2018), but have also been linked with changes in the Northern Annular Mode (Stine et al. 2009).

The day the cold extreme occurs is also associated with less snow, albeit largely due to decreases in mean snow cover, so it can be inferred that the snow will likely melt more as well, as it is occurring closer to summer months than normal. This describes another positive feedback within the system, with snowmelt leading to increases in latent heat which in turn heat the surface. This highlights the fact that multiple factors within the surface energy budget are contributing to an overall greater heating at the surface, thus influencing the decrease in the severity of cold days relative to mean warming during spring and autumn months. The relationship between excess changes in cold extremes and decreases in the mean state of snow cover warrants further investigation, namely to explore how snow cover and associated albedo changes can predict excess changes in cold extremes relative to the mean. The first step towards this would require investigating mean trends in observational snow cover data to determine if this can predict recent excess changes using observational temperature data, and then similarly evaluating the ability of climate models to represent this relationship.

The enhanced warming of seasonal cold extremes relative to seasonal mean temperature is projected for much of the Northern Hemisphere mid- to high-latitudes. The main findings of this paper show that the possible drivers of this enhanced warming depend on the season. Reductions in cold air temperature advection, likely as a consequence of Arctic amplification, are the most probable driver of enhanced cold extremes during winter months. For boreal autumn and spring, snow cover and surface albedo feedbacks are the most likely contributors affecting the accelerated warming rates of cold extremes. These findings are robust across the selection of CMIP5 models used in this study. While observational data were used to evaluate simulations of excess temperature in recent decades, the climate variables are only explored as future changes, with model agreement suggesting how robust changes are. Future work in understanding the physical mechanisms driving cold extremes would benefit in evaluating observational data of snow cover against model simulations. Further, regional climate models would provide finer-scale information that would be more useful for future planning.

CHAPTER 5

Methodological sensitivities related to the analysis of daily temperature extremes

Chapter overview

The findings presented in *Chapters 2 – 4* of this thesis are subject to specific methodological choices which can result in parametric and structural uncertainties that have the potential to impact conclusions. This chapter discusses some of the uncertainties related to the different methodological choices used for assessing changes in daily temperature extremes. Alternative methodological choices are explored and compared to see whether these choices impact the conclusions drawn from the findings of the previous chapters of the thesis. Specifically, the sensitivities explored are related to the choice of climatological reference period used for the calculation of temperature anomalies, different methodologies used to define temperatures that are considered ‘extreme’, the use of absolute temperatures as opposed to temperature anomalies, and uncertainties related to different levels of radiative forcing used in climate models. Generally, the conclusions drawn from this thesis remain robust regardless of methodological choice. However, some differences are apparent when anomalies are calculated relative to a moving window climatology, as opposed to a fixed climatological reference period. Excess changes calculated using a moving window

reflect changes in variability, because a moving climatology removes the effect of mean changes. Using a moving window as the climatological reference period therefore rephrases the initial research question to relate excess changes entirely to changes in extremes and variability, emphasising the need to clearly communicate what the results are actually showing. In addition, local excess changes are shown to be proportional to the amount of global warming, as reflected through comparisons of a high and mid-greenhouse gas emissions scenario.

5.1 Introduction

Uncertainties stemming from methodological choices are inevitable in climate change studies. The study of extremes, in particular, have inherent uncertainties due to the fact that errors show up as ‘extreme’ and more complex statistical methodologies are generally required for analysis. Further, the methods used in previous studies vary widely, making results difficult to compare and quantify. Evaluating the potential sensitivities to alternative methods not only adds to the robustness of conclusions, but also allows for a better understanding of the more appropriate methodological choices for specific research questions. While many studies touch on the various uncertainties related to assessing changes in temperature extremes (e.g. Klein Tank et al. 2009; Katz et al. 2013; Kirtman et al. 2013; Sippel et al. 2015; Hawkins and Sutton 2016), including sensitivities related to the choice of dataset (as discussed in *Chapter 2*), there is a distinct lack of documentation providing systematic assessments of the different types of sensitivities that might affect conclusions. In this chapter, some of the key potential sensitivities relating to the results presented in *Chapters 2 – 4* are explored in more detail.

So far in this thesis, changes in temperature extremes have been evaluated as anomalies. Anomalies are the change in temperature relative to a baseline, or climatological reference period, and are more commonly used in analyses of changes in temperature extremes compared with actual, or absolute temperature values. One reason for this is that anomalies are spatially coherent across large areas, allowing the development of global timeseries which can then be used to identify trends in temperature (Tingley 2012; Schmidt 2014). The varying number of stations over time, and their location and elevation, can result in inhomogeneities in absolute temperatures, with these factors being less important for anomalies (Jones et al. 1999; Jones et al. 2009; World Meteorological Organization (WMO) 2017). Anomalies are also an effective communication tool, describing how unusual the temperature is relative to average conditions. In this sense, it is immediately clear how ‘extreme’ a particular event is. Absolute temperature values, particularly when used for assessing global temperature, do not necessarily represent the risk to society (Schmidt 2014). Despite the advantages of using temperature anomalies, it is still useful to understand how absolute temperatures are changing for local-scale studies, where absolute thresholds can indicate the temperature at which society and the environment is at risk (Klein Tank et al. 2009). Further, some have argued that absolute temperatures are more appropriate to use for assessing changes in temperature extremes and variability because there is no need to calculate values relative to a chosen climatological reference period (e.g. Huntingford et al. 2013). The reference period used to calculate temperature anomalies is an important choice that can potentially impact results.

The choice of climatological reference period first and foremost depends on the research question being asked, and the reason for choosing a particular reference

period should be communicated clearly to avoid misinterpretation of results. Studies using anomalies relative to a reference period should explore the robustness of findings to alternative choices of reference period (Hawkins and Sutton 2016). The WMO outlines that a sufficiently long reference period is necessary to capture the variability of temperature, suggesting that a fixed 30-year period is generally suitable, with the most commonly used period being 1961-1990 for long-term assessments of changes (WMO 2017). The reference period is used to normalise the data relative to the mean and variability of the period, but this method has been shown to cause overestimations of extremes (e.g. Tingley 2012; Huntingford et al. 2013; Rhines and Huybers 2013; Sippel et al. 2015; Hawkins and Sutton 2016). For this reason, temperature anomalies in this thesis are calculated relative to the entire period of analysis, therefore not excluding variability and mean changes outside of the reference period and eliminating the potential for inflating extremes. Another option is to calculate anomalies relative to a moving window, which only describes changes other than the mean, such as changes related to variability. In this chapter, the sensitivity of results of excess changes to the choice of climatological reference period is explored further.

Another potential source of uncertainty relating to the analysis of temperature extremes stems from the methods used to define extremes and calculate exceedances. The methods used to define extremes vary widely between studies, with some using specialised statistics based on the Extreme Value Theory (EVT) (e.g. Kharin and Zwiers 2005; Brown et al. 2008; Christidis et al. 2011; Katz et al. 2013), and others using simpler methods such as using fixed values, quantiles or percentiles to define an extreme threshold (e.g. Simolo et al. 2010; Orłowsky and Seneviratne 2012; Donat et al. 2014; McKinnon et al. 2016). Different methods have previously resulted in different conclusions (Katz et al. 2013), for example,

some have suggested temperature is becoming more skewed in the warm tails (e.g. Donat and Alexander 2012; Hansen et al. 2012; Lewis and King 2017), while others using more complex techniques, such as those based on EVT (e.g. Brown et al. 2008), have concluded that changes in temperature are mostly related to a uniform shift in the tails of the distribution, with mostly no significant change in the shape parameter. As with climatological reference periods, testing alternative methods for calculating extremes is critical for robust conclusions.

Projections of climate change have extensive uncertainties, especially in terms of future forcings of greenhouse gases, which are dependent on how climate change is mitigated. Output from climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5) archive (Taylor et al. 2012) includes several different scenarios with different levels of radiative forcings to project the potential consequences from possible development pathways. These scenarios need to be plausible in order to provide useful information for the end-user. While alternative scenarios increase the spread of results, adding to the uncertainty of projections, this provides a range of potential outcomes for policymakers and planners who can then use this information to limit the risk of negative impacts from climate change (Moss et al. 2010). Aside from model scenario, internal variability within climate models is another possible source of uncertainty. In *Chapter 4* of this thesis, however, we show that long-term changes from several climate models were not especially sensitive to internal variability for the type of analysis being performed.

In this chapter, several sources of uncertainty are explored, comparing results shown in previous chapters to alternative methodological choices. The chapter is split into parametric uncertainties, namely, how parameter choice within a chosen

methodological framework affects results (e.g. reference period used), and structural uncertainties relating to methodological choices (e.g. interpolation scheme used). It is expected that structural uncertainties will be larger than parametric uncertainties (Dunn et al. 2014). By way of example, the comparisons in this chapter are based on climate model simulations from one ensemble member from the CanESM2 model, though it is noted that results remain consistent across a wide range of CMIP5 models and ensemble members. Only future changes in cold extremes are shown for September through November (SON), though conclusions remain true for other seasons, or are otherwise noted in the text.

5.2 Parametric uncertainties – choice of reference period

Due to the potential for overestimating extremes from defining anomalies based on standardising temperature data using a fixed 30-year period (Huntingford et al. 2013; Sippel et al. 2015), the results presented so far in the thesis are calculated relative to the entire period of analysis. In this sense, a longer, fixed reference period includes data that would otherwise be left out of the analysis if using a fixed 30-year climatology, removing the chance of inflated variance and extremes in the analysis of data outside the reference period. Because there is a possibility that the analysis is sensitive to the choice of reference period, it is important to explore alternative reference periods to gauge the robustness of results of both temperature and other climate variables.

Figure 5.1 compares future excess changes in cold extremes based on the climatological reference period, where results from an alternative reference period are subtracted from excess changes calculated relative to a mean annual cycle over the entire period of investigation, that is, 1950-2099 for future changes

(actual magnitudes of future excess changes is shown in Figure 5.1a). Figure 5.1b shows the difference in results between the longer, fixed reference period with using a climatological reference period of 1961-1990, while Figure 5.1c shows the difference when using a 30-year moving window, which for this particular analysis, uses climatologies of 1950-1979 for past decades, and 2070-2099 for future decades.

Differences in future excess changes depending on the climatological reference period used are relatively small for both instances, with the largest changes being in regions showing the largest excess changes, that is, in the high-latitudes in the Northern Hemisphere. Differences between the two fixed climatologies (Figure 5.1b) are small, with generally only 0°C to 0.5°C difference for most of the globe. Some high-latitude regions in Russia, Canada and Greenland show some larger differences exceeding 5°C , corresponding with regions where excess changes calculated from the 1950-2099 reference period are particularly large (Figure 5.1a). For the 30-year moving climatology, again, the largest differences are shown for high-latitude regions in Canada and Russia, which are regions with the largest excess changes in cold extremes in SON. Some regions show small negative differences, on the order of -0.5°C , for example, over eastern/central-Europe and southern Greenland, where excess changes calculated using the 30-year moving window are slightly larger than they are for the fixed, long-term climatology, though this is generally only the case for regions showing relatively small magnitude excess changes that are less than 2°C . Though not shown, DJF shows virtually the same excess changes, irrespective of choice of reference period, with SON (shown here) showing the greatest differences from reference period choice compared with all seasons. Similarly, results for recent excess changes are robust

regardless of the reference period (see Supplementary Material Figures S2.1 – S2.2 for example from *Chapter 3*).

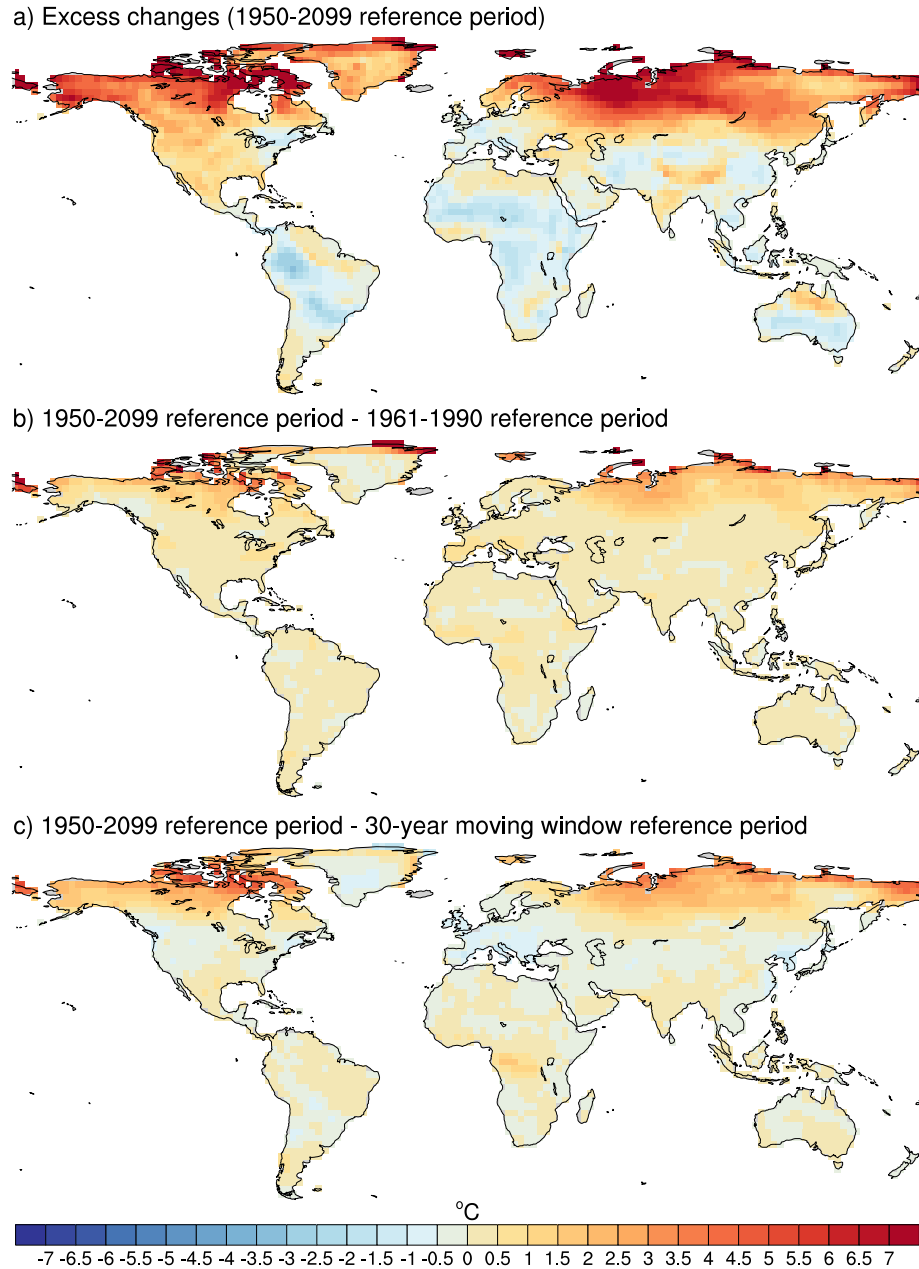


Figure 5.1 Differences in future excess changes in cold extremes depending on the climatological reference period used to calculate daily temperature anomalies. (a) shows the actual future excess changes calculated from the 1950-2099 fixed climatological reference period. The second and third row shows the differences between (a) and results calculated using the 1961-1990 reference period (b), and those relative to a 30-year moving window (c), which in this analysis, uses 1950-1979 for past decades, and 2070-2099 for future decades. Results are only shown for SON using the CanESM2 model.

Chapter 4 of the thesis investigated several different climate variables that might be driving the enhanced warming of cold extremes in the Northern Hemisphere mid- to high-latitudes. For this reason, it is also important to understand whether the choice of reference period has an impact on variables other than temperature. Figure 5.2 shows changes in snow cover on the day the cold extreme occurs (a-c) and changes in the timing of cold extreme days (d-f). As in Figure 5.1, actual changes calculated relative to the long-term, fixed reference period for each variable are shown in Figure 5.2a,d respectively, with differences calculated by subtracting alternative reference period choices from the 1950-2099 climatological reference period (Figure 5.2b,c for changes in snow cover and Figure 5.2e,f for changes in timing). Because these results are directly related to those of *Chapter 4*, only the mid- to high-latitudes in the Northern Hemisphere are shown, where the influence of reference period is greatest.

For changes in snow cover (Figure 5.2a-c), that is, using values on the day the seasonal cold extreme occurs, any differences between reference period choice are due to differences in the day the cold extreme occurs. For both instances of alternative reference period choice, high-latitude regions in Russia, Alaska and Canada show the largest differences, much like those shown in Figure 5.1b-c, with no substantial differences for most mid-latitude regions. When using a shorter, fixed climatology (Figure 5.2b), changes in snow cover in those regions showing the most difference are between 10% and 20% greater than they are for the 1950-2099 reference period. For the two fixed climatologies, the decreases in snow cover on the day of the extreme mainly reflect a decrease in the overall seasonal mean snow cover, as concluded in *Chapter 4*. Some negative differences are apparent when using a specific climatological reference period that moves through time (Figure 5.2c), which reflects changes in snow cover variability rather than

changes in the seasonal mean snow cover. These differences, mainly over Siberia and northern Canada, indicate smaller magnitude decreases in snow cover on the day of the extreme when using a moving window climatology, compared to when using a fixed reference period. Although not shown, only slight differences are apparent for MAM, with no apparent differences between reference periods found for DJF. Generally, the main conclusions drawn from *Chapter 4* regarding decreases in seasonal mean snow cover being a major contributor of the enhanced warming of cold extremes in the shoulder seasons remain robust irrespective of reference period choice.

In *Chapter 4*, a shift in the timing of cold extremes is shown, with anomalously cold days occurring later in high-latitude regions in the Northern Hemisphere spring months, and earlier in autumn months. If the day on which the extreme occurs is affected by choice of reference period, as suggested by the slight differences in snow cover depending on the reference period, then a shift in the timing of cold extremes would be impacted. While the shifts using the 1961-1990 reference period are only slightly larger than they are for those calculated using the 1950-2099 climatology (Figure 5.2e), differences when using a 30-year moving window climatology are substantial in some regions (Figure 5.2f). In the latter instance, substantially smaller shifts in the timing of cold extremes are projected in high-latitude regions, with cold extremes occurring only around 10 days earlier, compared with over 30 days earlier when using the 1950-2099 reference period. Central/southern Europe show much larger positive shifts (i.e. cold days are occurring later) compared with results from the 30-year moving window, which shows differences sometimes over 30 days in some areas. These large differences are related to the change in mean being removed when using a 30-year moving climatology, whereas a fixed climatology is affected by changes in the mean

seasonal cycle. Rather than adding uncertainty to results, this in fact strengthens the argument put forward in *Chapter 4* that a flattening of the seasonal cycle is projected for the late 21st century, because it is related to changes in the mean seasonal cycle which is specific for the time of year, as reflected through using fixed climatologies.

In general, conclusions are qualitatively robust when using a fixed climatological reference period, whether short or long. High-latitude regions in the Northern Hemisphere show the greatest sensitivity reference period choice, where the day in which the extreme occurs differs depending on this choice. Importantly, these results are specific to the particular methodology used here, where averages over 30-year periods are taken to calculate changes, removing the chance of a bias in results due to inflated variance using a 30-year fixed climatology. When using a 30-year moving climatology, two different means are included, which effectively removes influences related to changes in the mean temperature, and so all changes shown for this choice of reference period are likely related to changes in variability or higher-order moments. Any differences in the 30-year moving climatology results compared with the fixed climatologies is not so much of a sensitivity, but rather answering a different question entirely. That is, the comparison of a fixed climatology, long or short-term, versus a 30-year moving climatology, helps to separate changes that are due to changes in the mean from changes that are outside of this. In this thesis, it seems most appropriate to use the longer-term fixed climatology, as this highlights entire changes, which includes changes in the mean.

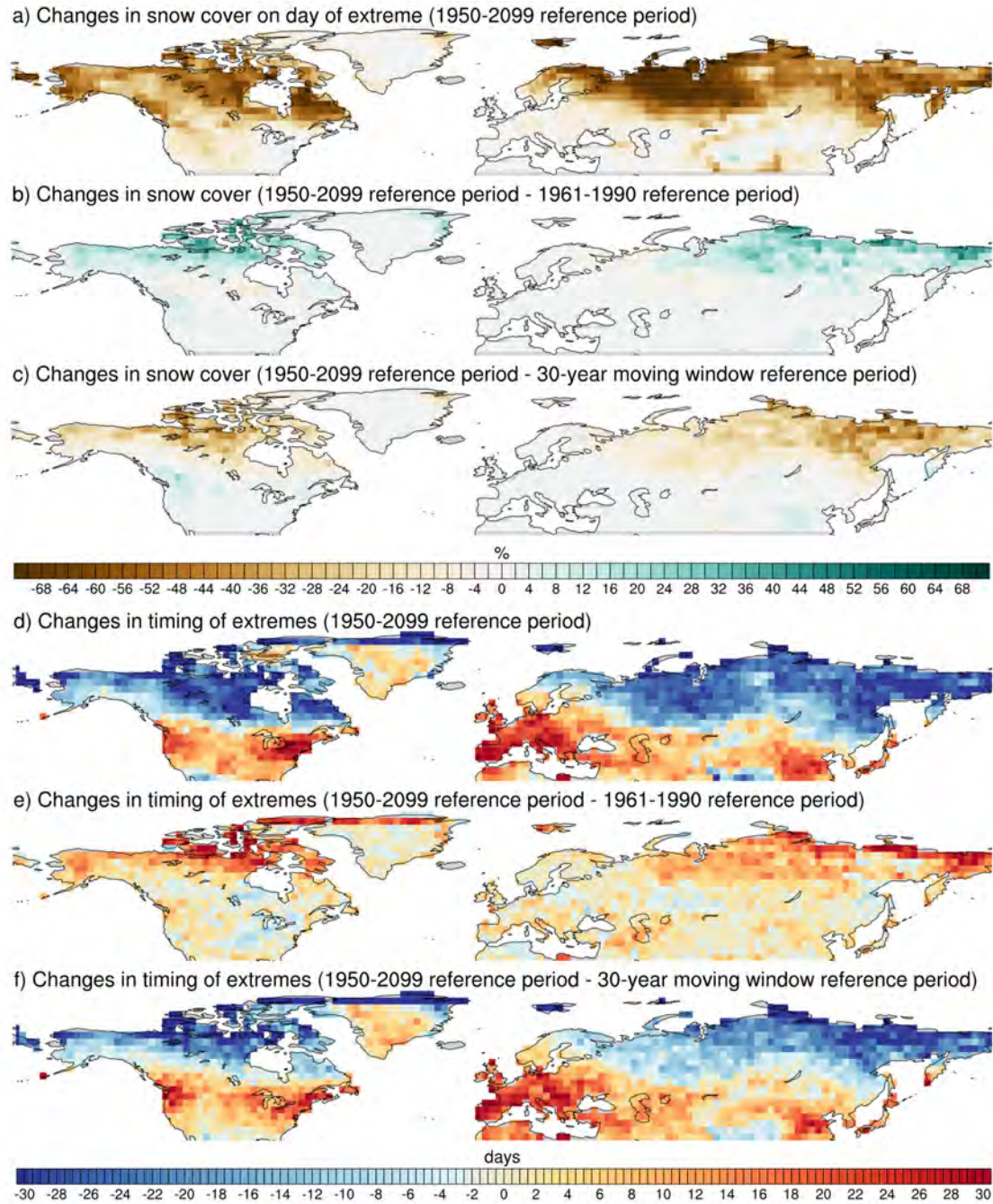


Figure 5.2 Similar to Figure 5.1, but showing differences in projected changes in snow cover on the day the cold extreme occurs (a-c) and the projected changes in the timing of cold extremes (d-f) depending on the reference period used. (a) shows actual changes in snow cover on the day the cold extreme occurs, while (d) shows actual changes in the timing of cold extremes based on the 1950-2099 climatological reference period. (b) and (e) show the difference between results using the 1950-2099 reference period and 1961-1990 reference period, and (c) and (f) show differences with results calculated relative to a 30-year moving window.

5.3 Structural uncertainties

5.3.1 Absolute temperatures

The robustness of results of changes in excess temperature can be further evaluated by assessing excess changes in absolute temperatures (Figure 5.3). Figure 5.3b shows the difference between future seasonal excess changes in cold extremes calculated from anomalies and those calculated from absolute temperatures. For reference, Figure 5.3a shows the actual magnitude future excess changes calculated from anomalies (as also shown in Figure 5.1a). Because the method for calculating excess changes takes temporal averages, it is unsurprising that results from absolute temperatures match closely with those using anomalies, with only small differences mostly between -1°C to 1°C shown for most of the globe, aside from the northmost region in Greenland which shows some larger differences. Excess changes in regions showing the greatest positive values, that is, mid- to high-latitude Northern Hemisphere regions, are of a slightly higher magnitude when using absolute temperatures, for example, over north-east Canada and central-eastern Europe and Russia. Other regions showing relatively low magnitude excess changes show small differences in sign between absolute temperatures and anomalies, for example, negative excess changes are shown over eastern U.S. in absolute temperatures in SON, compared with small positive excess changes in the anomalies. Overall, the most prominent future excess changes in cold extremes are robust, regardless of whether anomalies or absolute temperatures are used in the calculation.

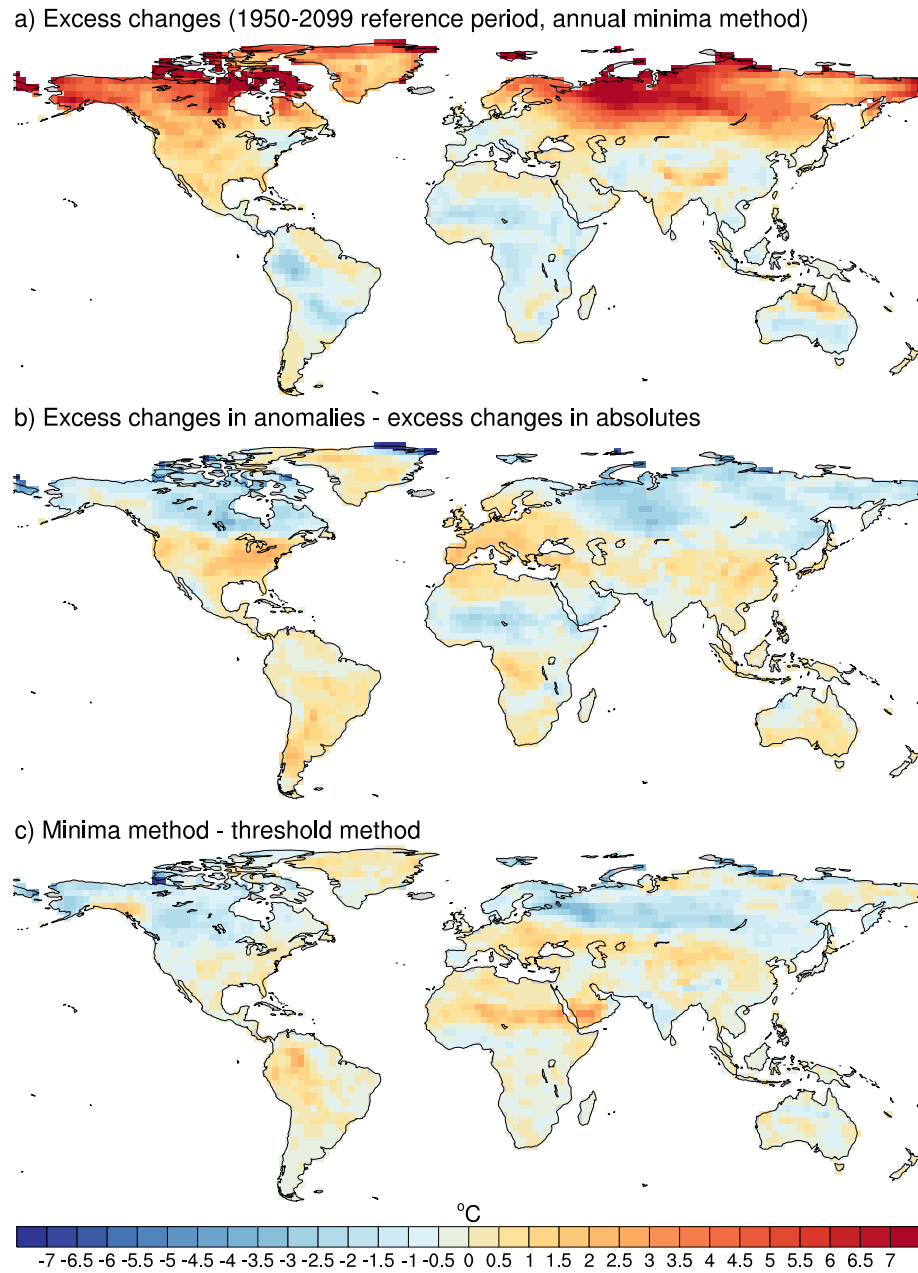


Figure 5.3 Differences between future excess changes in SON in anomalies and other methodological choices. (a) shows the actual magnitude of future excess changes using the 1950-2099 reference period and annual minima method to calculate seasonal exceedances, (b) shows the difference between this and using absolute temperatures to calculate excess changes, and (c) shows the difference between (a) and calculating exceedances using the seasonally-varying threshold approach to calculate cold extremes that fall below 1.5% of all data points within the time period.

5.3.2 Methods for defining temperature extremes

Throughout the thesis, three different methods are used to calculate temperature extremes. In *Chapter 2*, we used the EVT to define extremes, where a non-

stationary point-process model was fitted to the data to describe annual exceedances above (below) the upper (lower) 1.5% of all data points. Within that study, an alternative method was tested, which selects all data points that exceed the 98.5th quantile (1.5th quantile for the cold tails) per year, and then subtracts those values from the annual mean. These results yielded little difference compared with results calculated using the EVT (see Figure S1.20). Because results are mostly insensitive to this choice of method, excess changes in *Chapter 3* were calculated using the simpler and less computationally-expensive threshold approach, which defines exceedances as those data points which lie above the 98.5th percentile and below the 1.5th percentile of the seasonally varying distribution throughout each time period. Assessing only cold extremes, results shown in *Chapter 4* are calculated using a block minima approach, selecting the annual minimum temperature per season. Aside from further assessing sensitivities related to the choice of method, this was a simpler approach for investigating potential drivers of excess changes in temperature on the coldest day of the season, per year, compared to the former approach where extremes can occur in any given year throughout the season.

Figure 5.3c shows a comparison between the block minima approach and the 1.5% threshold approach. Overall, there is little sensitivity to the choice of method used to define temperature extremes, with most differences being between -1°C and 1°C. Slightly higher magnitude excess changes are shown for the threshold approach, for example, the latitude band around the equator in Africa which shows higher magnitude negative excess changes compared with the block minima approach. This could reflect extremes that are part of the same event within the same year, which are excluded when using an annual block minima. As with the choice of reference period for assessing excess changes in temperature,

the choice of method used to define extremes does not have a substantial impact on results for much of the globe, adding further confidence to the overarching findings of excess temperature changes especially for those areas showing the greatest excess changes.

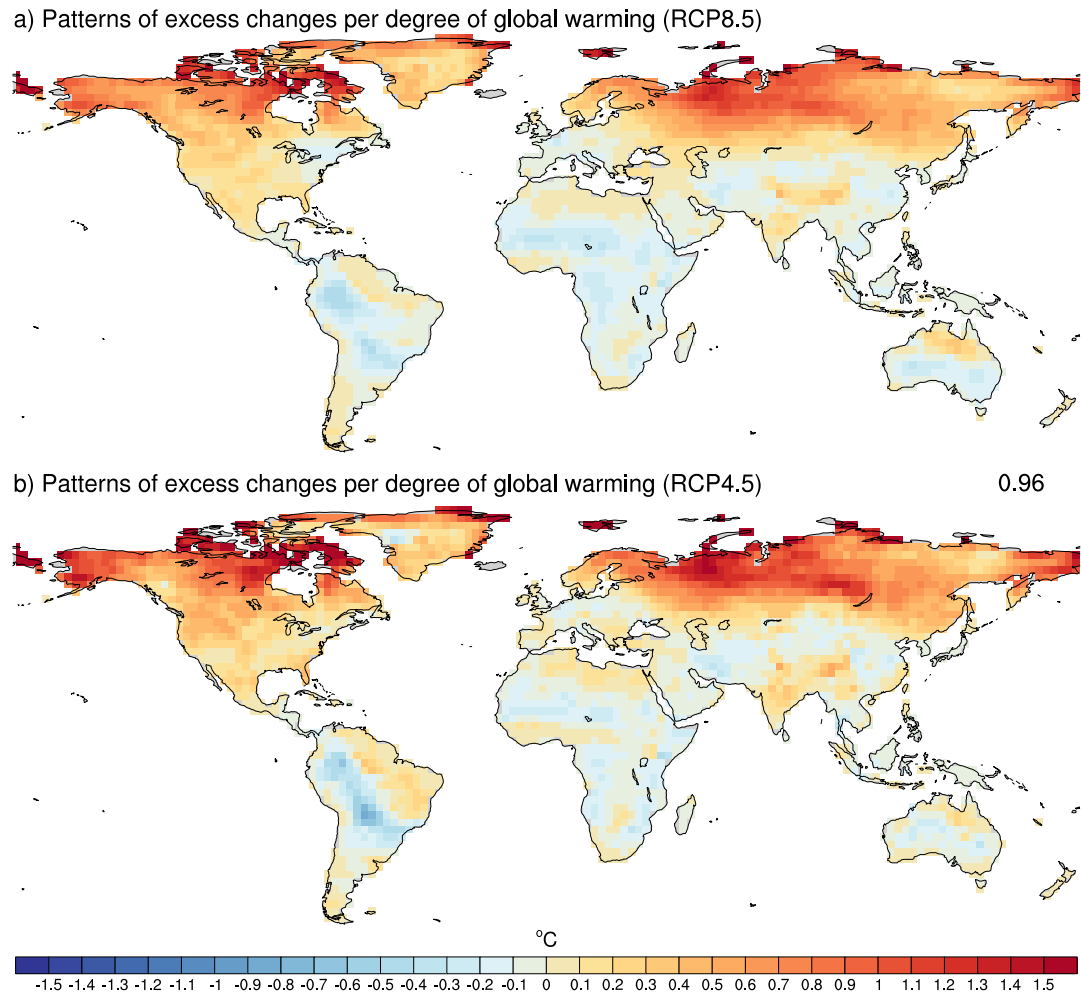


Figure 5.4 Comparison of RCP8.5 (a) and RCP4.5 (b) emissions scenarios, shown as future excess changes in cold extremes per degree of global warming. The number on the right-hand side of (b) represents the global pattern correlation with (a). As in previous figures, all results are from one ensemble member from CanESM2 for SON.

5.4 Uncertainty due to different levels of radiative forcing

All results of future projections presented so far in the thesis are based on the RCP8.5, “business-as-usual” scenario, or highest emissions scenario. It would be

useful to know the magnitude of excess changes for alternative scenarios with weaker greenhouse gas forcings, especially in terms of attributing projections to greenhouse gases and providing further information regarding, for instance, the benefits from mitigation for policymakers, stakeholders and the public. Therefore, by way of example, Figure 5.4 shows excess changes per degree of global warming for RCP8.5 (a) and RCP4.5 (b) with the latter representing a mid-range mitigation emissions scenario (Taylor et al. 2012). These results are calculated using simple pattern scaling (e.g. Tebaldi and Arblaster 2014), dividing excess changes for both emissions scenarios by the respective change in global mean temperature. Comparing the two figures, it is clear that the local responses are very similar between both scenarios, suggesting that future excess changes in cold extremes are indeed a consequence of, and mostly proportional to the amount of global warming. This is further emphasised in the high pattern correlation of 0.96 which describes how well the spatial patterns for the globe match between the two scenarios.

5.5 Discussion and conclusions

This chapter serves to explore some of the potential sensitivities and uncertainties of the results shown in *Chapters 2 – 4* of this thesis. Firstly, it highlights that the conclusions of the previous chapters generally remain robust, irrespective of methodological choice in defining extremes, and for the most part, regardless of choosing a long or short reference period. A moving window reference period does show some differences compared with using a fixed climatological reference period, however, this is less of an issue related to sensitivities, and rather points to a different question altogether. This emphasises the importance of clearly communicating the research question as well why the associated methods and

parameters are appropriate for the particular study, for example, why a specific reference period is chosen and what it explicitly means for the results.

When communicating how extreme temperatures are changing, anomalies are particularly useful to emphasise how unusual a temperature is compared with what is normal, especially if comparing between regions that experience different climates. However, it is still useful to know how extreme temperatures are changing in an absolute sense, in terms of exceeding absolute thresholds that pose a risk to society. For example, in Melbourne, Australia, the health of people aged 65 years and over is at risk when daily temperatures exceed 30°C (Nicholls et al. 2008). This type of information is crucial for making informed decisions that aim to reduce the risk of negative impacts from warming temperatures, however, absolute temperature thresholds are locally specific and cannot be used in a universal sense, as anomalies can. In the results presented here, it is shown that excess changes in absolute temperatures are similar in magnitude to those of anomalies, further adding to the robustness of the findings related to the rate of warming of extreme temperatures relative to mean temperatures.

Chapters 3 and 4 only show projections from the highest emissions scenario available in CMIP5, because in these chapters, we were interested in analysing results with the strongest signals in order to identify common patterns of change and relationships between different variables. By providing information of projections of other possible future scenarios, the range of possible outcomes can be used to understand how greenhouse gases might impact excess temperatures in the future. As shown here, the excess changes in temperature are generally proportional to the amount of global warming, with the mid-emissions scenario still projecting cold extremes to warm at least 3°C more than the mean for many

extratropical Northern Hemisphere regions, compared to greater than 5°C in the ‘business-as-usual’ scenario that is associated with a stronger global warming signal. A newly updated set of emissions scenarios are included as part of the newest model intercomparison in CMIP6, with a greater focus on societal development (Eyring et al. 2016a; O’Neill et al. 2016). Future work would need to explore these new emissions scenarios and how this affects projections of temperature extremes.

There are still other uncertainties related to the analyses, for example, those related to regridding datasets and spatially aggregating data. *Chapter 2* touches on these aspects, stating that different remapping techniques are unlikely to cause any substantial differences, with previous studies showing only very small uncertainties related to the choice of regridding method (e.g. Loikith et al. 2015). However, this is likely to depend on the chosen metric, and extremes are usually far more sensitive to interpolation choice than means (e.g. Avila et al. 2015). There are also uncertainties related to the gridded observational product itself, that is, HadGHCND, some of which are discussed briefly in *Chapter 2*, such as the possible artefacts that can result from spatially aggregated datasets (e.g. Rhines and Huybers 2013; Director and Bornn 2015).

Even though alternative methodological choices do not have a substantial impact on results for the specific type of analyses used throughout this thesis, it is still essential to test whether results are sensitive to these choices as this may depend on the type of analysis being performed. A greater understanding of how methodological choices affect results can reduce the uncertainties regarding future changes in climate, thereby providing more accurate and reliable information for the end-user.

CHAPTER 6

Conclusions and future work

6.1 Summary of findings

The overarching goal of this thesis was to provide an in-depth understanding of how different aspects of the temperature distribution are changing. In particular, analyses focused on how temperature extremes are changing relative to mean temperatures for all seasons and all global land areas, in both recent and future decades. This is a simple measure to illustrate disproportionate warming rates for different parts of the distribution, especially for the tails that are arguably most relevant in terms of impacts, in comparison to mean temperatures, a frequently used and therefore easily understood measure. Understanding the different rates of change between extremes and average temperatures was achieved by addressing several smaller aims which were presented in four chapters in this thesis.

In *Chapter 2*, several reanalysis datasets were evaluated against a quasi-global gridded observational dataset of daily maximum and minimum temperature, that is then used in subsequent chapters to evaluate a selection of Global Climate Models (GCMs). *Chapters 3* and *4* then focused on changes in seasonal temperature extremes relative to changes in seasonal mean temperatures in

observations and climate models. *Chapter 4* also included an investigation into the physical mechanisms related to the most robust changes in extremes relative to mean temperatures. Lastly, *Chapter 5* provides insight into the various sensitivities related to method choices that were not considered in earlier thesis chapters. This chapter presents the four main outcomes of the aims presented in *Chapter 1* and summarises the key findings of each chapter and overarching outcomes of this thesis, including opportunities for future research. The main conclusions from this thesis are:

1. When assessing daily temperature variability and extremes, dataset choice has an impact on analysing extremes and statistical moments other than the mean, but general spatial and temporal patterns in trends remain qualitatively robust.

In *Chapter 2*, several commonly used reanalysis datasets of daily maximum and minimum temperature were compared to the Hadley Centre Global Historical Climatology Network-Daily (HadGHCND) dataset (Caesar et al. 2006) to evaluate whether assessments of daily temperature variability and extremes are sensitive to the input dataset. This chapter also introduced a method for investigating changes in daily temperature extremes relative to daily mean temperatures. There are some limitations to this study, namely those related to the spatial coverage of the HadGHCND dataset, which lacks data for parts of South America, Africa, southern Asia and the Middle East. Some of the largest disagreements between datasets are shown for regions that are more data sparse, so information on these regions needs to be taken with caution. The key findings from this chapter highlighted that:

- General trends in daily mean maximum and minimum temperature anomalies from 1980-2014 are similar irrespective of dataset choice.
- Other statistics, such as standard deviation and skewness, show differences depending on the dataset, particularly for daily minimum temperature anomalies.
- Using the Extreme Value Theory (EVT) to assess the tails of the distribution specifically, where a point-process model was fitted to the data to describe extremes that exceed the top and bottom 1.5%, the lower tails in daily minimum temperature are most sensitive to dataset choice.
- In recent decades, annual cold extremes have been warming at a faster rate than annual mean temperatures for much of the Northern Hemisphere extratropics, while annual warm extremes have been warming faster than the mean in some tropical and subtropical regions, including parts of Australia, Asia, the Mediterranean and southern South America.
- While the uncertainties are sometimes large for the exact quantification of changes in temperature variability and extremes, the qualitative conclusions regarding general spatial patterns in the sign of change, especially for regions showing the strongest trends, remain robust across the range of datasets.

2. In recent decades, the greatest differences in the rates of change between extreme and mean temperatures occur in parts of North America and Eurasia, where cold extremes are warming faster than the mean, particularly in autumn and spring. Based on the CMIP5 models, this pattern becomes more pronounced in the future, with cold extremes warming faster than the mean for much of the Northern Hemisphere extratropics, in all seasons except boreal summer.

Chapter 3 assessed recent and future projected changes in seasonal warm and cold extremes relative to seasonal mean temperatures, referred to as ‘excess changes’. This follows on from the annual excess trends investigated in *Chapter 2*. The same observational dataset, HadGHCND, is assessed alongside a suite of climate models from the Coupled Model Intercomparison Project phase 5 (CMIP5) archive (Taylor et al. 2012), which were then used to analyse possible future changes in seasonal extremes relative to the seasonal mean. The findings in this study are important because they show that disproportionate rates of change are evident for many regions across the globe, which have an impact on the probability and frequency of extreme events. Further, the fact that some seasons show larger disproportionate rates of warming compared to others could potentially exacerbate the seasonal impacts from extremes. Other than the finding that the most striking excess changes occur in the cold tails relative to the mean, this chapter has a greater focus on cold extremes because previous work has generally given more attention to the disproportionate changes in hot extremes (e.g. Donat et al. 2017; Vogel et al. 2017). The main conclusions from this chapter are:

- In recent decades, the cold tails show the largest difference in warming relative to the mean, shown for some mid- to high-latitude Northern Hemisphere regions in all seasons except boreal summer. However, pattern correlations between CMIP5 models and the observational dataset are generally low for assessing recent excess changes. Low pattern correlations are also apparent between the CMIP5 models themselves, suggesting that recent excess changes are relatively spatially incoherent for much of the globe.

- Future simulated excess changes in the CMIP5 models are systematic and robust with high model agreement. The strongest excess changes are shown for many mid- to high-latitude regions in the Northern Hemisphere, where cold extremes are projected to warm faster than the mean.
- These projected disproportionate changes in the cold tails are apparent for all seasons except boreal summer, with cold extremes warming at least 3°C, or 50% more than mean temperatures by the end of the 21st century, and sometimes more than 5°C faster than the mean for some regions in Alaska, Canada and northern Eurasia.

3. Different physical mechanisms dominate the enhanced warming of cold extremes in the Northern Hemisphere extratropics for different seasons: reductions in cold air temperature advection dominate the accelerated warming of cold extremes during winter months, while decreases in snow cover and related albedo changes are likely the major contributors during autumn and spring.

Chapter 4 focused on the most robust findings of the previous chapter, that is, future excess changes in seasonal cold extremes in the Northern Hemisphere mid- to high-latitudes, and went further to investigate the physical mechanisms that are plausibly contributing to these projected changes. This was done by focusing on the environmental conditions on or prior to the anomalously coldest day of the season. The main findings are:

- During boreal winter, reductions in cold air temperature advection prior to the days when the cold extremes occur is the dominant driver of the enhanced warming of cold extremes relative to mean temperatures in North America, parts of Europe and Eurasia.

-
- During the shoulder seasons, particularly boreal autumn, decreases in snow cover and associated reductions in surface albedo on the anomalously coldest day of the season play a major role in the enhanced warming of cold extremes relative to daily mean temperature in the Northern Hemisphere extratropics.
 - The anomalously coldest days in the season are projected to shift in timing by the late 21st century, occurring earlier in autumn months, and later in spring months in high-latitude Northern Hemisphere regions. This suggests an overall flattening of the seasonal cycle in these regions and contributes to providing more incoming shortwave radiation on the days with the largest negative temperature anomalies.

4. Conclusions regarding changes in temperature extremes relative to the mean remain robust irrespective of methodological choice.

Within each chapter, there are various sensitivities related to methodological choices that are not examined in much detail, other than those related to dataset choice which are discussed thoroughly in *Chapter 2*. For instance, *Chapters 2, 3* and *4* all use slightly different methods to define extremes. *Chapter 2* characterises extremes by fitting a point-process model to the data, which describes exceedances above or below a non-stationary 1.5% threshold, while *Chapter 3* describes extremes as the average of data points exceeding a 1.5% threshold that can occur in any given year throughout the season, and *Chapter 4* identifies extremes as the seasonal block minimum. For this reason, *Chapter 5* explored some of these sensitivities by comparing alternative methodological choices. The major points from this chapter found:

- Results of future simulated excess changes in cold extremes show little sensitivity to the choice of fixed reference period used to calculate temperature anomalies.
- When using a 30-year moving window to calculate anomalies, excess changes are qualitatively similar to those calculated using a fixed reference period, however, the day the cold extreme occurs is affected. Much smaller shifts in the timing of cold extremes are projected for the Northern Hemisphere extratropics, compared to more substantial shifts when using a fixed reference period. When using a moving climatology as a reference period, changes in the mean are removed, so any changes shown in this instance reflect changes in variability and extremes.
- The methods used to define extremes, for example, using a threshold approach versus taking the seasonal block minimum, are robust for looking at future excess changes, with spatial patterns, sign of excess change, and general magnitudes being similar irrespective of the method. Similarly, using absolute temperatures instead of temperature anomalies result in similar magnitude excess changes for the globe, with the overall conclusions remaining the same.
- Two CMIP5 future emissions scenarios were compared: a high emissions scenario and a mid-emissions scenario. Results suggest that future excess changes in cold extremes are proportional to the magnitude of global warming.

While each chapter addresses individual aims, the overarching findings provide insight into the most robust disproportionate rates of change between extreme and mean temperatures. In combination, these different studies highlight that cold extremes in many mid- and high-latitude Northern Hemisphere regions are

observed and expected to warm disproportionately faster than regional increases in mean temperatures. This finding is robust across a range of different observational datasets and climate models and is true irrespective of the methods used to define extreme temperatures. While this is evident for annual cold extremes, autumn and winter months contribute the most to this excess, with spring months also showing widespread enhanced warming of cold extremes relative to the mean. For these regions in the Northern Hemisphere mid- to high-latitudes, the rates of warming of cold extremes are faster than both mean temperatures and hot extremes, suggesting overall variability in these regions is projected to decrease by the late 21st century compared with the mid-20th century.

6.2 Future work

This thesis creates several opportunities for future research. In *Chapter 4*, model output for several different climate variables was assessed to explore the possible drivers of the enhanced warming of cold extremes relative to mean temperatures. This work could be expanded on in several ways. Firstly, only simulations of future changes in snow cover, surface albedo and horizontal temperature advection are assessed, so this work is limited in that there is no evaluation of the climate models against observational data. Assessing observational snow cover data, for example, against CMIP5 simulations could reduce uncertainties in future projections of snow cover changes. Further, only existing simulations of climate variables were analysed, where the relationships between the variables were explored to make inferences about the physical processes driving the enhanced warming of the anomalously coldest days. Running specific model simulations with prescribed boundary conditions in variables such as snow cover and surface albedo would more robustly attribute the excess changes in cold extremes to specific drivers.

Aside from assessing global pattern correlations between the CMIP5 models and observations in *Chapters 3* and *4*, model skill, in terms of how well the models resemble the observations, was not evaluated. The Perkins skill score (Perkins et al. 2007) is one application used to measure the agreement, or overlap between modelled and observed distributions, however, in more recent years, the scientific community is moving towards more process-oriented diagnostics (Eyring et al. 2019). Future research could work towards developing a process-based constraint for future projections based on the results of *Chapter 4*. Given the availability of suitable observational data for the variables investigated, it would be useful to identify models for which the identified relationships or physical processes between different variables are more realistic. Further, new tools, such as the Earth System Model Evaluation Tool (ESMValTool) (Eyring et al. 2016b), are being developed for rapid climate model evaluation to address biases in the models. Implementing these tools could greatly increase confidence in the model-based results in this thesis. Additionally, the methods used in this thesis could be applied to the new generation of CMIP6 models (Eyring et al. 2016a) and associated Scenario Model Intercomparison Project (ScenarioMIP) future emissions scenarios, which extend beyond the current Representative Pathway Scenarios (RCPs) by updating trends in recent greenhouse gas emissions as well as integrating socioeconomic factors into the scenarios (O'Neill et al. 2016).

Another avenue of future work would be to expand the study presented in *Chapter 2*. For example, exploring the reasons behind the differences in the datasets would be useful to improve reanalysis data, particularly to understand why greater sensitivities are shown in the cold tails compared with the warm tails, and in daily minimum temperatures compared with daily maximum

temperatures. In addition, the methods used in *Chapter 2* could be easily applied to evaluate sensitivities in newly developed datasets of daily temperature.

6.3 Concluding remarks

The research presented in this thesis comprehensively assesses changes in several aspects of the distribution of daily temperature, with a specific focus on how extremes are changing in relation to average temperatures. Aside from assessing changes in the warm tails of the distribution relative to mean temperatures, there is a strong focus in this thesis on the accelerated warming of cold extremes and the potential drivers influencing this. Cold extremes have generally been a less prominent focus in existing literature, however, like hot extremes, the impacts from warming cold extremes affect numerous sectors. These impacts will likely be exacerbated with the anticipated disproportionate warming rates of cold extremes relative to the rest of the temperature distribution.

The findings presented in this thesis contribute to the understanding of changes in extreme temperatures by taking a more holistic approach than is generally taken, by considering how they are changing in parallel with changes in the mean. This is assessed systematically in both observations and climate model simulations across all seasons and at the global scale. In addition, this thesis adds to the current knowledge by exploring the environmental conditions on the day of the extreme and the physical relationships that are contributing to the projected disproportionate rates of warming in cold extremes. The methodologies and statistical analyses applied in this thesis are simple, yet robust measures that have been evaluated against alternative methods to add confidence to the conclusions. This reduces the potential sensitivities surrounding methodological choices in future studies of extremes, and the approach taken here can easily be

reproduced to not only further analyse changes in temperature extremes but also investigate other climate variables in a similar way. Addressing the disproportionate rates of warming in extremes and mean temperatures is crucial so that policymakers and future planners have the information necessary to reduce the negative impacts that can arise from exacerbated temperature extremes due to climate change.

REFERENCES

- Alexander, L. V., and Coauthors, 2006: Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research*, **111**, D05109, doi:10.1029/2005JD006290.
- Alexander, L., and S. Perkins, 2013: Debate heating up over changes in climate variability. *Environmental Research Letters*, **8**, 41001, doi:10.1088/1748-9326/8/4/041001.
- Avila, F. B., S. Dong, K. P. Menang, J. Rajczak, M. Renom, M. G. Donat, and L. V. Alexander, 2015: Systematic investigation of gridding-related scaling effects on annual statistics of daily temperature and precipitation maxima: A case study for south-east Australia. *Weather and Climate Extremes*, **9**, 6–16, doi:10.1016/j.wace.2015.06.003.
- Ayarzagüena, B., and J. A. Screen, 2016: Future Arctic sea ice loss reduces severity of cold air outbreaks in midlatitudes. *Geophysical Research Letters*, **43**, 2801–2809, doi:10.1002/2016GL068092.
- Barnes, E. A., and J. A. Screen, 2015: The impact of Arctic warming on the midlatitude jet-stream: Can it? Has it? Will it? *Wiley Interdisciplinary Reviews: Climate Change*, **6**, 277–286, doi:10.1002/wcc.337.
- Blunden, J., and D. S. Arndt, 2014: State of the climate in 2013. *Bulletin of the American Meteorological Society*, **95**, S1–S279, doi:10.1175/2014BAMSSStateoftheClimate.1.
- Bosilovich, M. G., and Coauthors, 2015: MERRA-2: Initial evaluation of the climate. Technical Report Series on Global Modeling and Data Assimilation,

-
- Vol. 43, NASA Tech. Rep. NASA/TM-2015-104606, 139 pp. [Available online at <https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich803.pdf>].
- Brown, S. J., J. Caesar, and C. A. T. Ferro, 2008: Global changes in extreme daily temperature since 1950. *Journal of Geophysical Research*, **113**, D05115, doi:10.1029/2006JD008091.
- Caesar, J., L. Alexander, and R. Vose, 2006: Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *Journal of Geophysical Research: Atmospheres*, **111**, D05101, doi:10.1029/2005JD006280.
- Cattiaux, J., H. Douville, R. Schoetter, S. Parey, and P. Yiou, 2015: Projected increase in diurnal and interdiurnal variations of European summer temperatures. *Geophysical Research Letters*, **42**, 899–907, doi:10.1002/2014GL062531.
- Cavanaugh, N. R., and S. S. P. Shen, 2014: Northern hemisphere climatology and trends of statistical moments documented from GHCN-Daily surface air temperature station data from 1950 to 2010. *Journal of Climate*, **27**, 5396–5410, doi:10.1175/JCLI-D-13-00470.1.
- Chapin III, F., and Coauthors, 2005: Role of Land-Surface Changes in Arctic Summer Warming. *Science*, **310**, 657–660, doi:10.1126/science.1117368.
- Christidis, N., P. A. Stott, and S. J. Brown, 2011: The role of human activity in the recent warming of extremely warm daytime temperatures. *Journal of Climate*, **24**, 1922–1930, doi:10.1175/2011JCLI4150.1.
- Cohen, J., and D. Rind, 1991: The effect of snow cover on the climate. *Journal of Climate*, **4**, 689–706.
- Cohen, J., J. Jones, J. C. Furtado, and E. Tziperman, 2013: Warm Arctic, cold continents: A common pattern related to Arctic sea ice melt, snow advance,

- and extreme winter weather. *Oceanography*, **26**, 150–160,
doi:10.5670/oceanog.2013.70.
- Cohen, J., and Coauthors, 2014: Recent Arctic amplification and extreme mid-latitude weather. *Nature Geoscience*, **7**, 627–637, doi:10.1038/ngeo2234.
- Cohen, J., K. Pfeiffer, and J. A. Francis, 2018: Warm Arctic episodes linked with increased frequency of extreme winter weather in the United States. *Nature Communications*, **9**, 1–12, doi:10.1038/s41467-018-02992-9.
- Coles, S, 2001: *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics. Springer, 208 pp.
- Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis. Contributing of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker et al., Eds., IPCC, Cambridge University Press, 1029–1136.
- Compo, G. P., and Coauthors, 2011: The twentieth century reanalysis project. *Quarterly Journal of the Royal Meteorological Society*, **137**, 1–28, doi:10.1002/qj.776.
- Coumou, D., and A. Robinson, 2013: Historic and future increases in the global land area affected by monthly heat extremes. *Environmental Research Letters*, **8**, 034018, doi:10.1088/1748-9326/8/3/034018.
- Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, **137**, 553–597, doi:10.1002/qj.828.
- Director, H., and L. Bornn, 2015: Connecting point-level and gridded moments in the analysis of climate data. *Journal of Climate*, **28**, 3496–3510, doi:10.1175/JCLI-D-14-00571.1.

-
- Diro, G. T., L. Sushama, and O. Huziy, 2018: Snow-atmosphere coupling and its impact on temperature variability and extremes over North America. *Climate Dynamics*, **50**, 2993–3007, doi:10.1007/s00382-017-3788-5.
- Donat, M. G., and L. V. Alexander, 2012: The shifting probability distribution of global daytime and night-time temperatures. *Geophysical Research Letters*, **39**, L14707, doi:10.1029/2012GL052459.
- Donat, M.G., L. V. Alexander, H. Yang, I. Durre, R. Vose, and J. Caesar, 2013a: Global land-based datasets for monitoring climatic extremes. *Bulletin of the American Meteorological Society*, **94**, 997–1006, doi:10.1175/BAMS-D-12-00109.1.
- Donat, M. G., and Coauthors, 2013b: Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset. *Journal of Geophysical Research: Atmospheres*, **118**, 2098–2118, doi:10.1002/jgrd.50150.
- Donat, M. G., J. Sillmann, S. Wild, L. V. Alexander, T. Lippmann, and F. W. Zwiers, 2014: Consistency of temperature and precipitation extremes across various global gridded in situ and reanalysis dataset. *Journal of Climate*, **27**, 5019–5035, doi:10.1175/JCLI-D-13-00405.1.
- Donat, M. G., L. V. Alexander, N. Herold, and A. J. Dittus, 2016: Temperature and precipitation extremes in century-long gridded observations, reanalyses, and atmospheric model simulations. *Journal of Geophysical Research: Atmospheres*, **121**, 11,174–11,189, doi:10.1002/2016JD025480.
- Donat, M. G., A. J. Pitman, and S. I. Seneviratne, 2017: Regional warming of hot extremes accelerated by surface energy fluxes. *Geophysical Research Letters*, **44**, 7011–7019, doi:10.1002/2017GL073733.
- Donat, M. G., A. J. Pitman, and O. Angélil, 2018: Understanding and reducing future uncertainty in midlatitude daily heat extremes via land surface feedback

- constraints. *Geophysical Research Letters*, **45**, 10627–10636, doi:10.1029/2018GL079128.
- Dunn, R. J. H., M. G. Donat, and L. V. Alexander, 2014: Investigating uncertainties in global gridded data sets of climate extremes. *Climate of the Past*, **10**, 2171–2199, doi:10.5194/cp-10-2171-2014.
- Dutra, E., C. Schär, P. Viterbo, and P. M. A. Miranda, 2011: Land-atmosphere coupling associated with snow cover. *Geophysical Research Letters*, **38**, L15707, doi:10.1029/2011GL048435.
- Dwyer, J. G., M. Biasutti, and A. H. Sobel, 2012: Projected changes in the seasonal cycle of surface temperature. *Journal of Climate*, **25**, 6359–6374, doi:10.1175/JCLI-D-11-00741.1.
- Ebi, K. L., and J. Nealon, 2016: Dengue in a changing climate. *Environmental Research*, **151**, 115–123, doi:10.1016/j.envres.2016.07.026.
- Eyring, V., S. Bony, G. A. Meehl, C. A. Senior, B. Stevens, R. J. Stouffer, and K. E. Taylor, 2016a: Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geoscientific Model Development*, **9**, 1937–1958, doi:10.5194/gmd-9-1937-2016.
- Eyring, V., and Coauthors, 2016b: ESMValTool (v1.0)-a community diagnostic and performance metrics tool for routine evaluation of Earth system models in CMIP. *Geoscientific Model Development*, **9**, 1747–1802, doi:10.5194/gmd-9-1747-2016.
- Eyring, V., and Coauthors, 2019: Taking climate model evaluation to the next level. *Nature Climate Change*, **9**, 102–110, doi:10.1038/s41558-018-0355-y.
- Fischer, E. M., and C. Schär, 2009: Future changes in daily summer temperature variability: Driving processes and role for temperature extremes. *Climate Dynamics*, **33**, 917–935, doi:10.1007/s00382-008-0473-8.

-
- Fischer, E. M., D. M. Lawrence, and B. M. Sanderson, 2011: Quantifying uncertainties in projections of extremes-a perturbed land surface parameter experiment. *Climate Dynamics*, **37**, 1381–1398, doi:10.1007/s00382-010-0915-y.
- Fletcher, C. G., H. Zhao, P. J. Kushner, and R. Fernandes, 2012: Using models and satellite observations to evaluate the strength of snow albedo feedback. *Journal of Geophysical Research: Atmospheres*, **117**, D11117, doi:10.1029/2012JD017724.
- Francis, J. A., and S. J. Vavrus, 2012: Evidence linking Arctic amplification to extreme weather in mid-latitudes. *Geophysical Research Letters*, **39**, L06801, doi:10.1029/2012GL051000.
- Francis, J. A., and S. J. Vavrus, 2015: Evidence for a wavier jet stream in response to rapid Arctic warming. *Environmental Research Letters*, **10**, 014005, doi:10.1088/1748-9326/10/1/014005.
- Friedman, A. R., Y.-T. Hwang, J. C. H. Chiang, and D. M. W. Frierson, 2013: Interhemispheric temperature asymmetry over the twentieth century and in future projections. *Journal of Climate*, **26**, 5419–5433, doi:10.1175/JCLI-D-12-00525.1.
- Gasparrini, A., and Coauthors, 2016: Changes in Susceptibility to Heat during the Summer: A Multicountry Analysis. *American Journal of Epidemiology*, **183**, 1027–1036, doi:10.1093/aje/kwv260.
- Gelaro, R., and Coauthors, 2017: The Modern-ERA Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *Journal of Climate*, **30**, 5419–5454, doi:10.1175/JCLI-D-16-0758.1.
- Gilleland, E., and R. W. Katz, 2011: New software to analyze how extremes change over time. *EOS Transactions American Geophysical Union*, **92**, 13–20, doi:10.1029/2011EO020001.

- Gregory, J. M., and J. F. B. Mitchell, 1995: Simulation of daily variability of surface temperature and precipitation over Europe in the current and 2 x CO₂ climates using the UKMO climate model. *Quarterly Journal of the Royal Meteorological Society*, **121**, 1451–1476.
- Gross, M. H., M. G. Donat, L. V. Alexander, and S. A. Sisson, 2018: The sensitivity of daily temperature variability and extremes to dataset choice. *Journal of Climate*, **31**, 1337–1359, doi:10.1175/JCLI-D-17-0243.1.
- Hartmann, D. L., and Coauthors, 2013: Observations: Atmosphere and surface. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker et al., Eds., IPCC, Cambridge University Press, 159-254.
- Hansen, J., M. Sato, and R. Ruedy, 2012: Perception of climate change. *Proceedings of the National Academy of Sciences*, **109**, E2415-E2423, doi:10.1073/pnas.1205276109.
- Hawkins, E., and R. Sutton, 2016: Connecting climate model projections of global temperature change with the real world. *Bulletin of the American Meteorological Society*, **97**, 963–980, doi:10.1175/BAMS-D-14-00154.2.
- He, T., S. Liang, and D.-X. Song, 2014: Analysis of global land surface albedo climatology and spatial-temporal variation during 1981-2010 from multiple satellite products. *Journal of Geophysical Research: Atmospheres*, **119**, 10,281–10,298, doi:10.1002/2014JD021667.
- Holmes, C. R., T. Woollings, E. Hawkins, and H. de Vries, 2016: Robust future changes in temperature variability under greenhouse gas forcing and the relationship with thermal advection. *Journal of Climate*, **29**, 2221–2236, doi:10.1175/JCLI-D-14-00735.1.

-
- Huntingford, C., P. D. Jones, V. N. Livina, T. M. Lenton, and P. M. Cox, 2013: No increase in global temperature variability despite changing regional patterns. *Nature*, **500**, 327-330, doi:10.1038/nature12310.
- Hutchinson, D. K., M. H. England, A. M. Hogg, and K. Snow, 2015: Interhemispheric asymmetry of warming in an eddy-permitting coupled sector model. *Journal of Climate*, **28**, 7385-7406, doi:10.1175/JCLI-D-15-0014.1.
- Huybers, P., K. A. McKinnon, A. Rhines, and M. Tingley, 2014: U.S. daily temperatures: The meaning of extremes in the context of nonnormality. *Journal of Climate*, **27**, 7368-7384, doi:10.1175/JCLI-D-14-00216.1.
- Intergovernmental Panel on Climate Change (IPCC), 2012: Summary for policymakers. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*, C. B. Field et al., Eds., IPCC, Cambridge University Press, 3-21.
- Jones, D. A., W. Wang, and R. Fawcett, 2009: High-quality spatial climate datasets for Australia. *Australian Meteorological and Oceanographic Journal*, **58**, 233-248.
- Jones, P. D., M. New, D. E. Parker, S. Martin, and I. G. Rigor, 1999: Surface air temperature and its changes over the past 150 years. *Review of Geophysics*, **37**, 173-199.
- Jones, P. W., 1999: First- and second-order conservative remapping schemes for grids in spherical coordinates. *Monthly Weather Review*, **127**, 2204-2010.
- Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77**, 437-471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.
- Kanamitsu, M., W. Abisuzaki, J. Woollen, S. -K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP-DOE AMIP-II Reanalysis (R-2). *Bulletin of the*

- American Meteorological Society*, **83**, 1631-1643, doi:10.1175/BAMS-83-11-1631.
- Katz, R., and B. Brown, 1992: Extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, **21**, 289-302, doi:10.1007/BF00139728.
- Katz, R. W., P. F. Craigmile, P. Guttorp, M. Haran, B. Sanso, and M. L. Stein, 2013: Uncertainty analysis in climate change assessments. *Nature Climate Change*, **3**, 769-771, doi:10.1038/nclimate1980.
- Kendall, M. G, 1975: *Rank Correlation Methods*. Charles Griffin, 272 pp.
- Kharin, V. V., and F. W. Zwiers, 2000: Changes in the extremes in an ensemble of transient climate simulations with a coupled atmosphere-ocean GCM. *Journal of Climate*, **13**, 3760-3788, doi:10.1016/j.proenv.2016.07.054.
- Kharin, V. V., and F. W. Zwiers, 2005: Estimating extremes in transient climate change simulations. *Journal of Climate*, **18**, 1156-1173, doi:10.1174/JCLI3320.1.
- Kharin, V. V., F. W. Zwiers, X. Zhang, and G. C. Hegerl, 2007: Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations. *Journal of Climate*, **20**, 1419-1444, doi:10.1175/JCLI4066.1.
- Kharin, V. V., F. W. Zwiers, X. Zhang, and M. Wehner, 2013: Changes in temperature and precipitation extremes in the CMIP5 ensemble. *Climatic Change*, **119**, 345-357, doi:10.1007/s10584-013-0705-8.
- Kirtman, B., and Coauthors, 2013: Near-term climate change: Projections and predictability. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T. F. Stocker et al., Eds., IPCC, Cambridge University Press, 953-1028.

-
- Kistler, R., and Coauthors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, **82**, 247-267.
- Kjellström, E., L. Bärring, D. Jacob, R. Jones, G. Lenderink, and C. Schär, 2007: Modelling daily temperature extremes: Recent climate and future changes over Europe. *Climatic Change*, **81**, 249–265, doi:10.1007/s10584-006-9220-5.
- Klein Tank, A. M. G., F. W. Zwiers, and X. Zhang, 2009: Guidelines on analysis of extremes in a changing climate in support of informed decisions for adaptation. WMO, WCDMP-72, WMO-TD 1500, 56pp.
- Kobayashi, S., and Coauthors, 2015: The JRA-55 reanalysis: General specifications and basic characteristics. *Journal of the Meteorological Society of Japan*, **93**, 5-48, doi:10.2151/jmsj.2015-001.
- Kodra, E., K. Steinhaeuser, and A. R. Ganguly, 2011: Persisting cold extremes under 21st-century warming scenarios. *Geophysical Research Letters*, **38**, L08705, doi:10.1029/2011GL047103.
- Kodra, E., and A. R. Ganguly, 2014: Asymmetry of projected increases in extreme temperature distributions. *Scientific Reports*, **4**, 1-8, doi:10.1038/srep05884.
- Kretschmer, M., D. Coumou, L. Agel, M. Barlow, E. Tziperman, and J. Da. Cohen, 2018: More-persistent weak stratospheric polar vortex States linked to cold extremes. *Bulletin of the American Meteorological Society*, **99**, 49–60, doi:10.1175/BAMS-D-16-0259.1.
- Lavell, A., and Coauthors, 2012: Climate change: New dimensions in disaster risk, exposure, vulnerability, and resilience. *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change*, C. B. Field et al., Eds., IPCC, Cambridge University Press, 25-64.

- Lewis, S. C., and A. D. King, 2017: Evolution of mean, variance and extremes in 21st century temperatures. *Weather and Climate Extremes*, **15**, 1–10, doi:10.1016/j.wace.2016.11.002.
- Loikith, P. C., and J. D. Neelin, 2015: Short-tailed temperature distributions over North America and implications for future changes in extremes. *Geophysical Research Letters*, **42**, 1–9, doi:10.1002/2015GL065602.
- Loikith, P.C., and Coauthors, 2015: Surface temperature probability distributions in the NARCCAP Hindcast Experiment: Evaluation methodology, metrics, and results. *Journal of Climate*, **28**, 978–997, doi:10.1175/JCLI-D-13-00457.1.
- Lorantý, M. M., L. T. Berner, S. J. Goetz, Y. Jin, and J. T. Randerson, 2014: Vegetation controls on northern high latitude snow-albedo feedback: Observations and CMIP5 model simulations. *Global Change Biology*, **20**, 594–606, doi:10.1111/gcb.12391.
- Matiu, M., D. P. Ankerst, and A. Menzel, 2016: Asymmetric trends in seasonal temperature variability in instrumental records from ten stations in Switzerland, Germany and the UK from 1864 to 2012. *International Journal of Climatology*, **36**, 13–27, doi:10.1002/joc.4326.
- McKinnon, K. A., A. Rhines, M. P. Tingley, and P. Huybers, 2016: The changing shape of Northern Hemisphere summer temperature distributions. *Journal of Geophysical Research: Atmospheres*, **121**, 8849–8868, doi:10.1002/2016JD025292.
- Mearns, L. O., R. W. Katz, and S. H. Schneider, 1984: Extreme high temperature events: Changes in their probabilities with changes in mean temperature. *Journal of Climate and Applied Meteorology*, **23**, 1601–1613.
- Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, doi:10.1038/nature08823.

-
- Mote, T. L., 2008: On the role of snow cover in depressing air temperature. *Journal of Applied Meteorology and Climatology*, **47**, 2008–2022, doi:10.1175/2007JAMC1823.1.
- National Center for Atmospheric Research (NCAR), 2016: The climate data guide: NCEP Reanalysis (R2). Accessed 26 October 2016. [Available online at <https://climatedataguide.ucar.edu/climate-data/ncep-reanalysis-r2>].
- Nicholls, N., C. Skinner, M. Loughnan, and N. Tapper, 2008: A simple heat alert system for Melbourne, Australia. *International Journal of Biometeorology*, **52**, 375–384, doi:10.1007/s00484-007-0132-5.
- O’Neill, B. C., and Coauthors, 2016: The Scenario Model Intercomparison Project (ScenarioMIP) for CMIP6. *Geoscientific Model Development*, **9**, 3461–3482, doi:10.5194/gmd-9-3461-2016.
- Orlowsky, B., and S. I. Seneviratne, 2012: Global changes in extreme events: Regional and seasonal dimension. *Climatic Change*, **110**, 669–696, doi:10.1007/s10584-011-0122-9.
- Perkins, S. E., A. J. Pitman, N. J. Holbrook, and J. McAneney, 2007: Evaluation of the AR4 climate models’ simulated daily maximum temperature, minimum temperature, and precipitation over Australia using probability density functions. *Journal of Climate*, **20**, 4356–4376, doi:10.1175/JCLI4253.1.
- Qu, X., and A. Hall, 2014: On the persistent spread in snow-albedo feedback. *Climate Dynamics*, **42**, 69–81, doi:10.1007/s00382-013-1774-0.
- Rahmstorf, S., and D. Coumou, 2011: Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, **108**, 17905–17909, doi:10.1073/pnas.1201491109.
- Rahmstorf, S., G. Foster, and N. Cahill, 2017: Global temperature evolution: Recent trends and some pitfalls. *Environmental Research Letters*, **12**, doi:10.1088/1748-9326/aa6825.

REFERENCES

- Rhines, A., and P. Huybers, 2013: Frequent summer temperature extremes reflect changes in the mean, not the variance. *Proceedings of the National Academy of Sciences*, **110**, E546, doi:10.1073/pnas.1218748110.
- Rhines, A., K. A. McKinnon, M. P. Tingley, and P. Huybers, 2017: Seasonally resolved distributional trends of North American temperatures show contraction of winter variability. *Journal of Climate*, **30**, 1139-1157, doi:10.1175/JCLI-D-16-0363.1.
- Robeson, S. M., 2004: Trends in time-varying percentiles of daily minimum and maximum temperature over North America. *Geophysical Research Letters*, **31**, 4379-4384, doi:10.1029/20003GL019019.
- Sánchez-Lugo, A., C. Morice, and P. Berrisford, 2016: Surface temperature [in “State of the Climate in 2015”]. *Bulletin of the American Meteorological Society*, **97**, S12-S13.
- Santer, B. D., and Coauthors, 2018: Human influence on the seasonal cycle of tropospheric temperature. *Science*, **361**, eaas8806, doi:10.1126/science.aas8806.
- Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2015: Need for caution in interpreting extreme weather statistics. *Journal of Climate*, **28**, 9166-9187, doi:10.1175/JCLI-D-15-0020.1.
- Schär, C., P. L. Vidale, D. Luthi, C. Frei, C. Haberli, M. A. Liniger, and C. Appenzeller, 2004: The role of increasing temperature variability in European summer heatwaves. *Nature*, **427**, 332-336, doi:10.1038/nature02300.
- Schmidt, G, 2014: Real Climate: Absolute temperatures and relative anomalies. Accessed 28 October 2018. [Available online at <http://www.realclimate.org/index.php/archives/2014/12/absolute-temperatures-and-relative-anomalies/>].

-
- Schneider, T., T. Bischoff, and H. Plotka, 2015: Physics of changes in synoptic midlatitude temperature variability. *Journal of Climate*, **28**, 2312–2331, doi:10.1175/JCLI-D-14-00632.1.
- Screen, J. A., and I. Simmonds, 2010: The central role of diminishing sea ice in recent Arctic temperature amplification. *Nature*, **464**, 1334–1337, doi:10.1038/nature09051.
- Screen, J. A., and I. Simmonds, 2013: Exploring links between Arctic amplification and mid-latitude weather. *Geophysical Research Letters*, **40**, 959–964, doi:10.1002/grl.50174.
- Screen, J. A., 2014: Arctic amplification decreases temperature variance in northern mid-to high-latitudes. *Nature Climate Change*, **4**, 577–582, doi:10.1038/nclimate2268.
- Screen, J. A., C. Deser, and L. Sun, 2015a: Reduced risk of North American cold extremes due to continued arctic sea ice loss. *Bulletin of the American Meteorological Society*, **96**, 1489–1503, doi:10.1175/BAMS-D-14-00185.1.
- Screen, J. A., C. Deser, and L. Sun, 2015b: Projected changes in regional climate extremes arising from Arctic sea ice loss. *Environmental Research Letters*, **10**, 84006, doi:10.1088/1748-9326/10/8/084006.
- Sen, P. K, 1968: Estimates of the regression coefficient based on Kendall’s Tau. *Journal of the American Statistical Association*, **63**, 1379–1389, doi:10.1080/01621459.1968.10480934.
- Seneviratne, S. I., M. G. Donat, A. J. Pitman, R. Knutti, and R. L. Wilby, 2016: Allowable CO₂ emissions based on regional and impact-related climate targets. *Nature*, **529**, 477–483, doi:10.1038/nature16542.
- Serreze, M. C., and J. A. Francis, 2006: The Arctic amplification debate. *Climatic Change*, **76**, 241–264, doi:10.1007/s10584-005-9017-y.

- Shepard, D., 1968: A two-dimensional interpolation function for irregularly-spaced data. *Proceedings of the 1968 23rd ACM National Conference*, ACM Publ. P-68, Brandon/Systems Press, 517-524.
- Sherwood, S. C., and M. Huber, 2010: An adaptability limit to climate change due to heat stress. *Proceedings of the National Academy of Sciences*, **107**, 9552–9555, doi:10.1073/pnas.0913352107.
- Sillmann, J., V. V. Kharin, F. W. Zwiers, X. Zhang, and D. Bronaugh, 2013a: Climate extremes indices in the CMIP5 multimodel ensemble: Part 2. Future climate projections. *Journal of Geophysical Research: Atmospheres*, **118**, 2473–2493, doi:10.1002/jgrd.50188.
- Sillmann, J., V. V. Kharin, X. Zhang, F. W. Zwiers, and D. Bronaugh, 2013b: Climate extremes indices in the CMIP5 multimodel ensemble: Part 1. Model evaluation in the present climate. *Journal of Geophysical Research: Atmospheres*, **118**, 1716-1733, doi:10.1002/jgrd.50203.
- Simmons, A. J., and Coauthors, 2004: Comparison of trends and low-frequency variability in CRU, ERA-40, and NCEP/NCAR analysis of surface air temperature. *Journal of Geophysical Research*, **109**, D24115, doi:10.1029/2004JD005306.
- Simmons, A. J., P. Berrisford, D. P. Dee, H. Hersbach, S. Hirahara, and J.-N. Thépaut, 2017: A reassessment of temperature variations and trends from global reanalyses and monthly surface climatological datasets. *Quarterly Journal of the Royal Meteorological Society*, **143**, 101-119, doi:10.1002/qj.2949.
- Simolo, C., M. Brunetti, M. Maugeri, T. Nanni, and A. Speranza, 2010: Understanding climate change-induced variations in daily temperature distributions over Italy. *Journal of Geophysical Research: Atmospheres*, **115**, D22110, doi:10.1029/2010JD014088.

-
- Simolo, C., M. Brunetti, M. Maugeri, T. Nanni, and A. Speranza, 2011: Evolution of extreme temperatures in a warming climate. *Geophysical Research Letters*, **38**, L16701, doi:10.1029/2011GL048437.
- Sippel, S., J. Zscheischler, M. Heimann, F. E. L. Otto, J. Peters, and M. D. Mahecha, 2015: Quantifying changes in climate variability and extremes : Pitfalls and their overcoming. *Geophysical Research Letters*, **42**, 1–9, doi:10.1002/2015GL066307.
- Smith, K. R., and Coauthors, 2014: Human health: Impacts, adaptation, and co-benefits. *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*, C. B. Field et al., Eds., IPCC, Cambridge University Press, 709–754.
- Sun, L., J. Perlwitz, and M. Hoerling, 2016: What caused the recent “Warm Arctic, Cold Continents” trend pattern in winter temperatures? *Geophysical Research Letters*, **43**, 5345–5352, doi:10.1002/2016GL069024.
- Sutton, R., E. Suckling, and E. Hawkins, 2015: What does global mean temperature tell us about local climate? *Philosophical Transactions of the Royal Society A*, **373**, 20140426, doi:10.1098/rsta.2014.0426.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An Overview of CMIP5 and the Experiment Design. *Bulletin of the American Meteorological Society*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Tebaldi, C., and J. M. Arblaster, 2014: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Climatic Change*, **122**, 459–471, doi:10.1007/s10584-013-1032-9.
- Thackeray, C. W., C. G. Fletcher, and C. Derksen, 2014: The influence of canopy snow parameterizations on snow albedo feedback in boreal forest regions.

- Journal of Geophysical Research: Atmospheres*, **119**, 9810–9821, doi:10.1002/2014Jd021858.
- Thackeray, C. W., C. G. Fletcher, and C. Derksen, 2015: Quantifying the skill of CMIP5 models in simulating seasonal albedo and snow cover evolution. *Journal of Geophysical Research: Atmospheres*, **120**, 5831–5849, doi:10.1002/2015JD023325.
- Tingley, M. P., 2012: A bayesian ANOVA scheme for calculating climate anomalies, with applications to the instrumental temperature record. *Journal of Climate*, **25**, 777–791, doi:10.1175/JCLI-D-11-00008.1.
- Tingley, M., and P. Huybers, 2013: Recent temperature extremes at high northern latitudes unprecedented in the past 600 years. *Nature*, **496**, 201–205, doi:10.1038/nature11969.
- Uppala, S. M., and Coauthors, 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131**, 2961–3012, doi:10.1256/qj.04.176
- Vogel, M., R. Orth, F. Cheruy, S. Hagemann, R. Lorenz, B. van den Hurk, and S. I. Seneviratne, 2017: Regional amplification of projected changes in extreme temperatures strongly controlled by soil moisture-temperature feedbacks. *Geophysical Research Letters*, **44**, 1511–1519, doi:10.1002/2016GL071235.
- Vogel, M., J. Zscheischler, and S. I. Seneviratne, 2018: Varying soil moisture-atmosphere feedbacks explain divergent temperature extremes and precipitation projections in central Europe. *Earth System Dynamics*, **9**, 1107–1125, doi:10.5194/esd-9-1107-2018.
- Weaver, S. J., A. Kumar, and M. Chen, 2014: Recent increases in extreme temperature occurrence over land. *Geophysical Research Letters*, **41**, 4669–4675, doi:10.1002/2014GL060300.

-
- Wolf, T., K. Lyne, G. Martinez, and V. Kendrovski, 2015: The Health Effects of Climate Change in the WHO European Region. *Climate*, **3**, 901–936, doi:10.3390/cli3040901.
- World Meteorological Organization (WMO), 2017: WMO Guidelines on the calculation of climate normals. WMO No. 1203, 29 pp. [Available online at https://library.wmo.int/doc_num.php?explnum_id=4166].
- Xu, L., and P. Dirmeyer, 2011: Snow-atmosphere coupling strength in a global atmospheric model. *Geophysical Research Letters*, **38**, L13401, doi:10.1029/2011GL048049.
- Ylhäisi, J. S., and J. Räisänen, 2014: Twenty-first century changes in daily temperature variability in CMIP3 climate models. *International Journal of Climatology*, **34**, 1414–1428, doi:10.1002/joc.3773.
- Zhang, J., W. Tian, M. P. Chipperfield, F. Xie, and J. Huang, 2016: Persistent shift of the Arctic polar vortex towards the Eurasian continent in recent decades. *Nature Climate Change*, **6**, 1094–1099, doi:10.1038/nclimate3136.
- Zwiers, F. W., and H. von Storch, 1995: Taking serial correlation into account in tests of the mean. *Journal of Climate*, **8**, 336–351.

S.1 SUPPLEMENTARY MATERIAL FOR CHAPTER 2

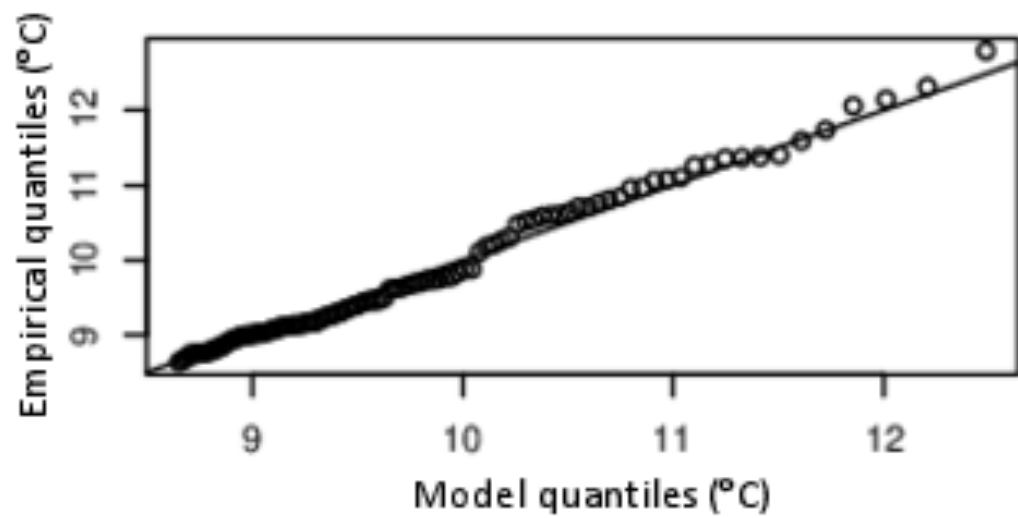


Figure S1.1 Example of a quantile-quantile plot using HadGHCND data showing the goodness-of-fit for empirical and model distributions. If the data lie close to the 1-1 diagonal line, it indicates that the data are a good fit to the PP distribution used to define extremes.

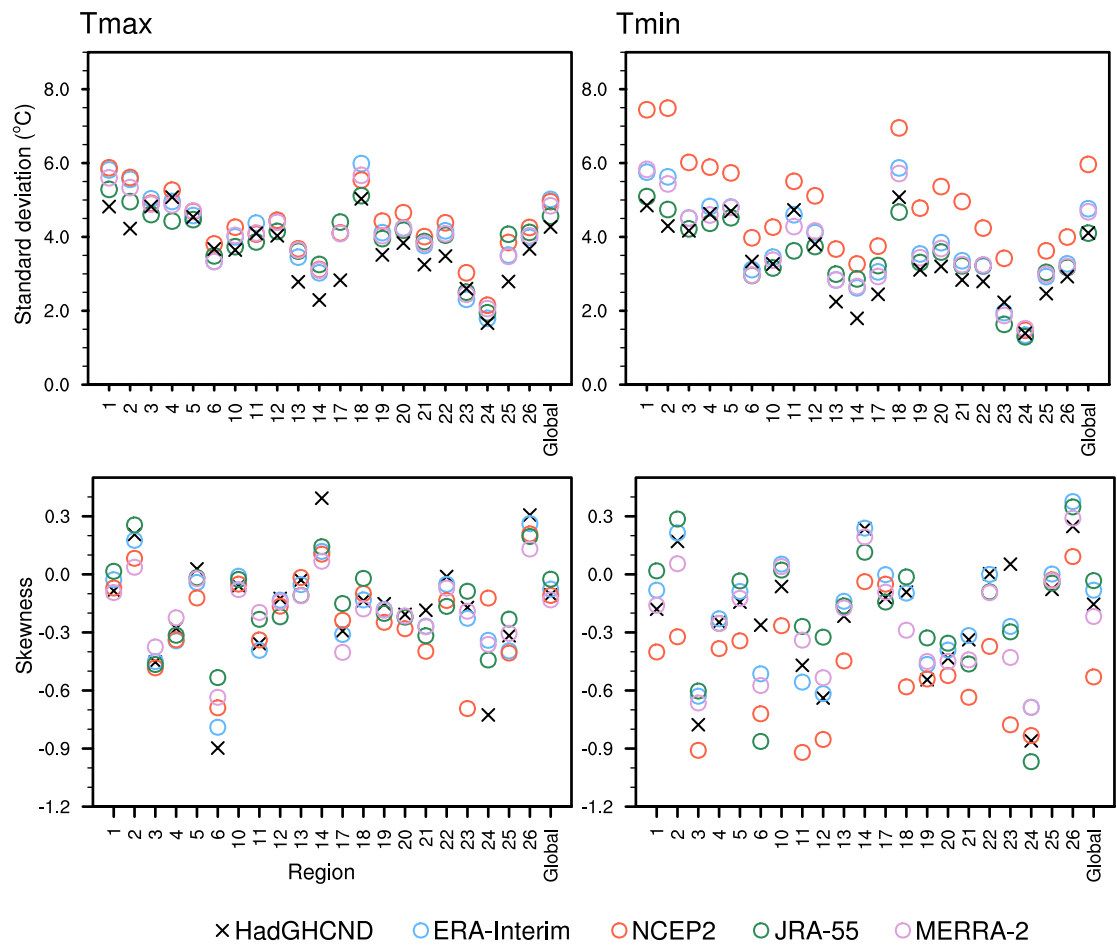


Figure S1.2 Parameter estimates for standard deviation (°C) and skewness, for both daily maximum and daily minimum temperature anomalies. Estimates are shown for SREX regions that fulfill the completeness criteria, and for the globe, and are calculated by pooling data for each region over the period 1980-2014.

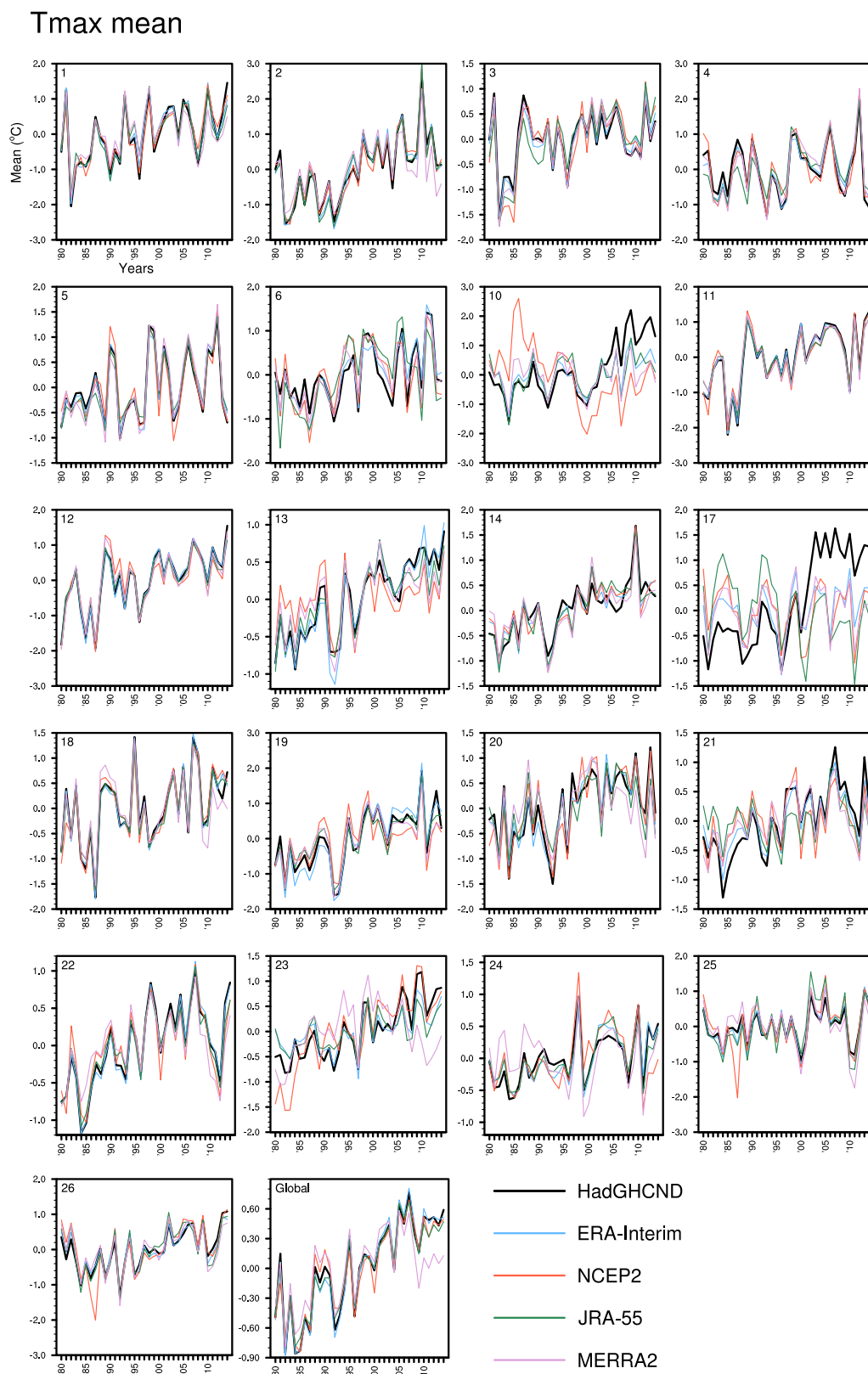


Figure S1.3 Time series plots of the annual mean for daily maximum temperature anomalies (°C) for each region with sufficient data, as well as the globe.

Tmin mean

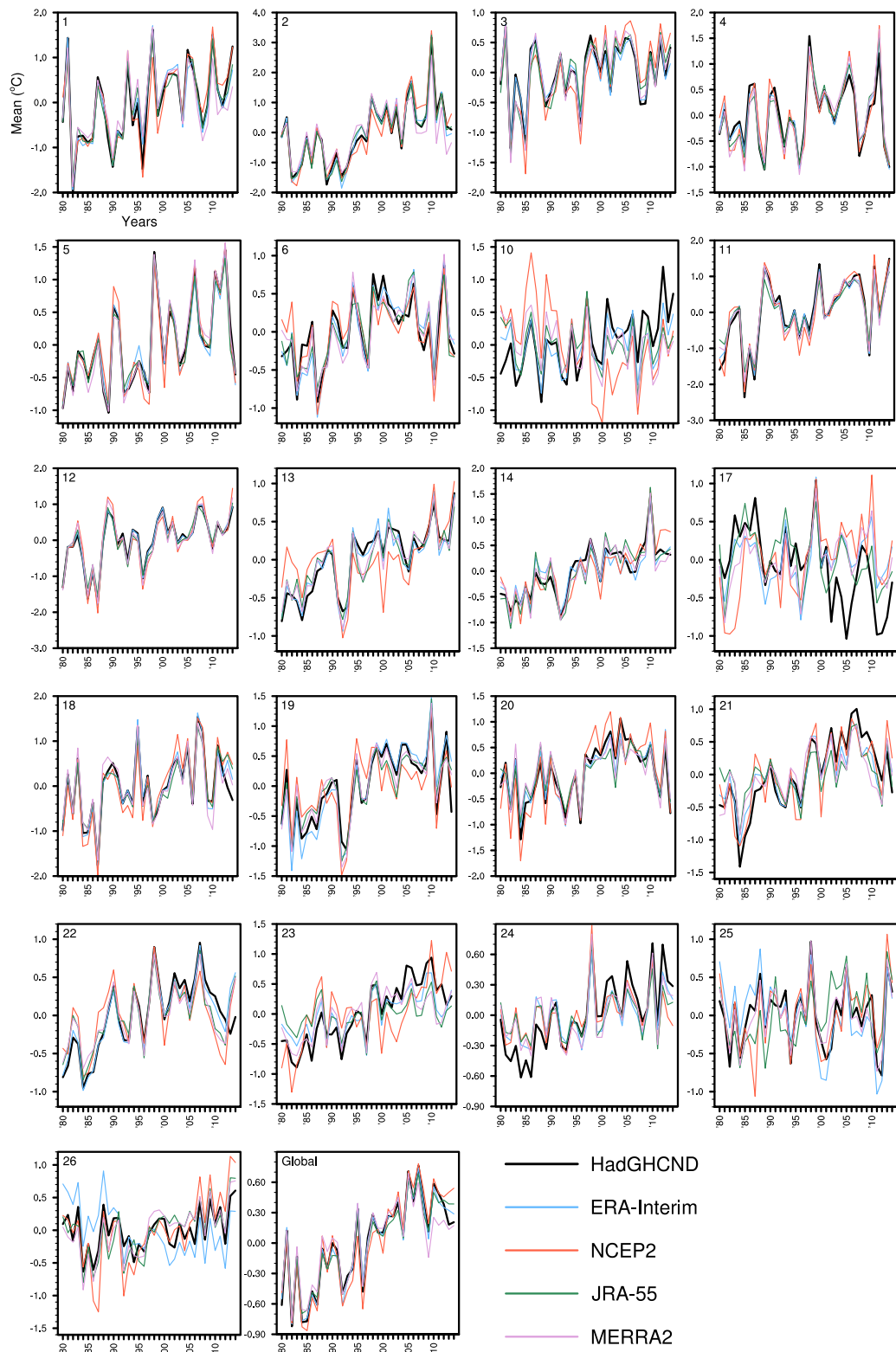


Figure S1.4 As Figure S1.3, but for the annual mean of daily minimum temperature anomalies (°C).

Tmax standard deviation

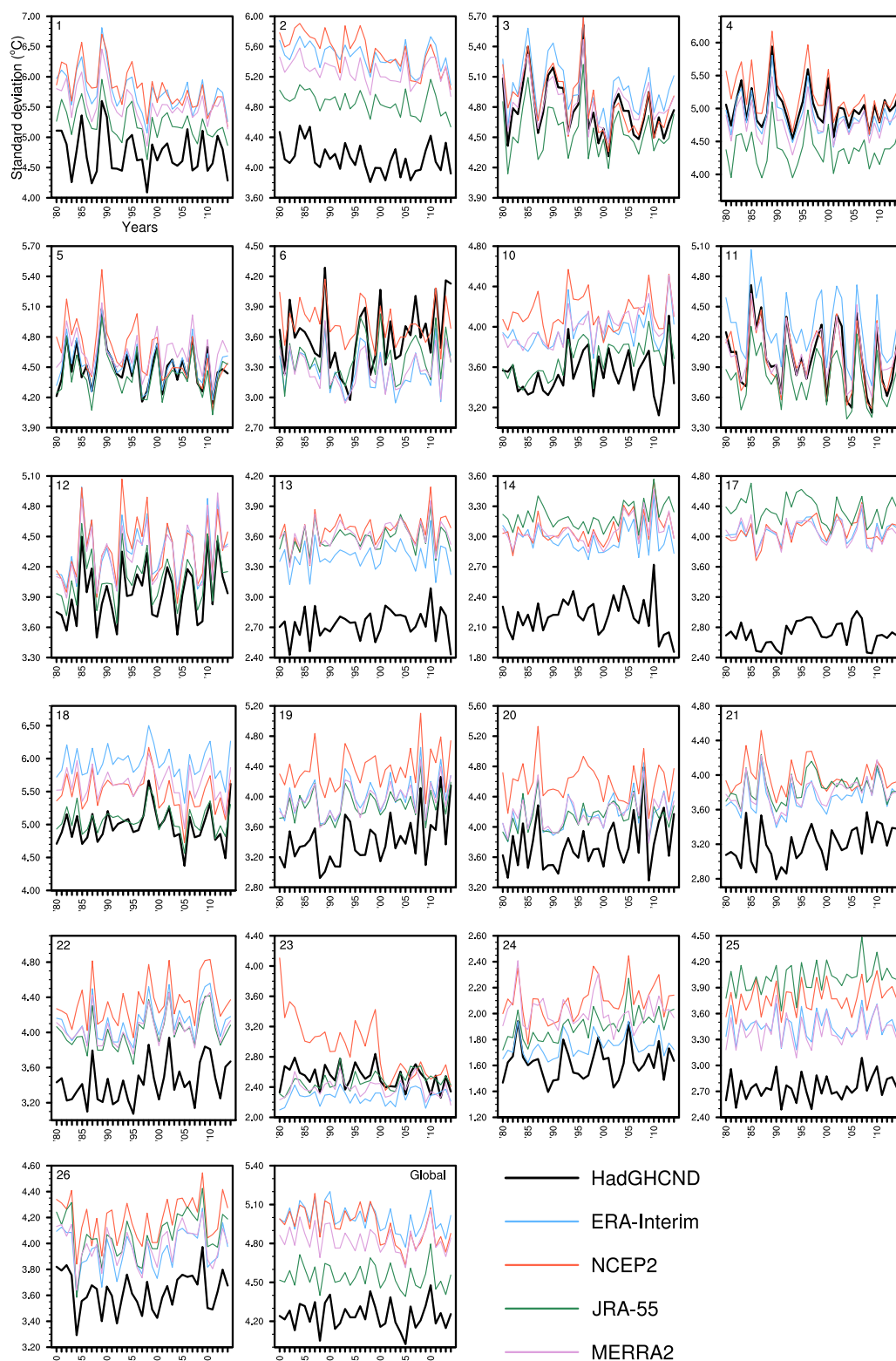


Figure S1.5 As Figure S1.3, but for the annual standard deviation of daily maximum temperature anomalies (°C).

Tmin standard deviation

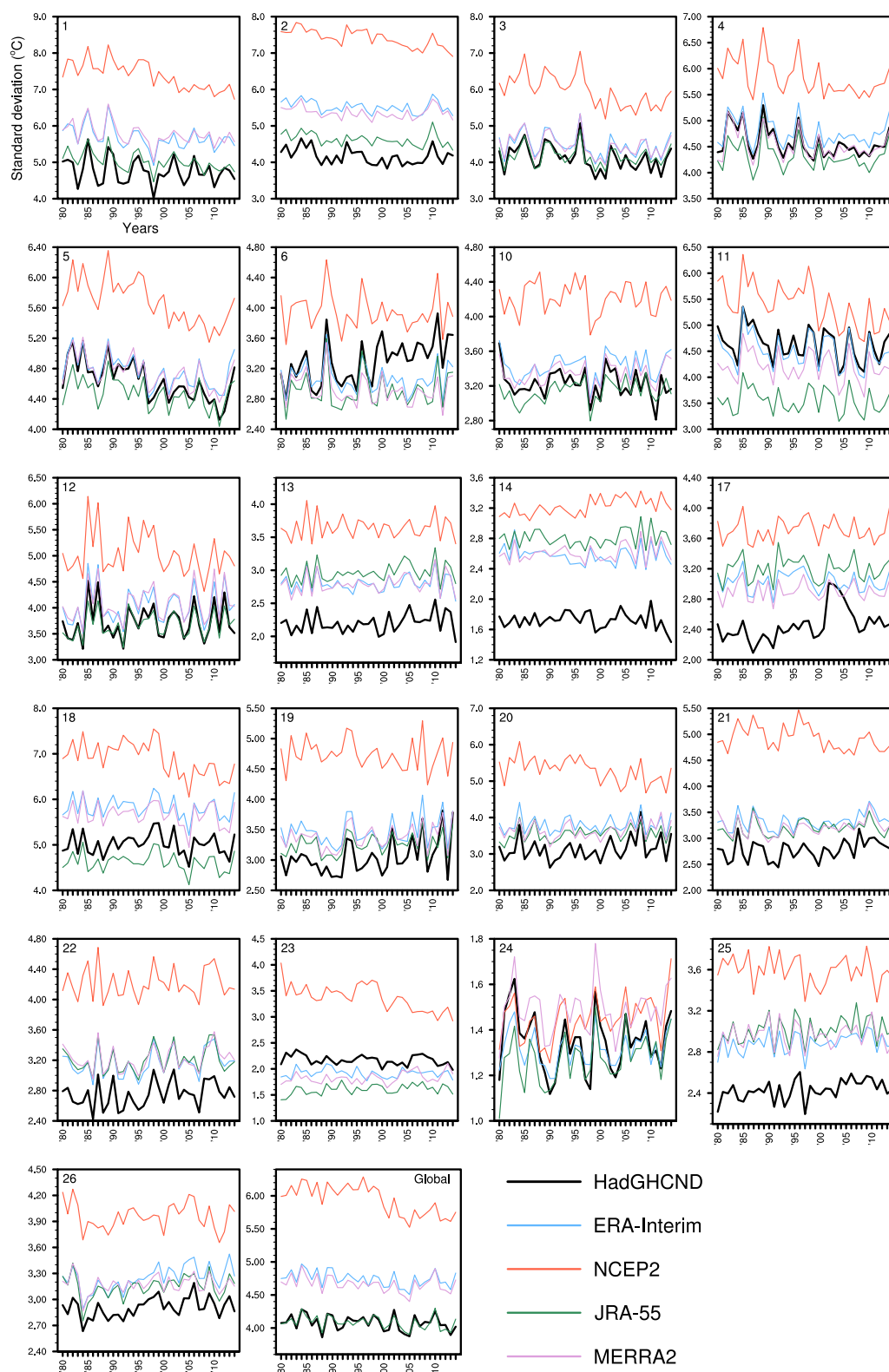


Figure S1.6 As Figure S1.3, but for the annual standard deviation of daily minimum temperature anomalies (°C).

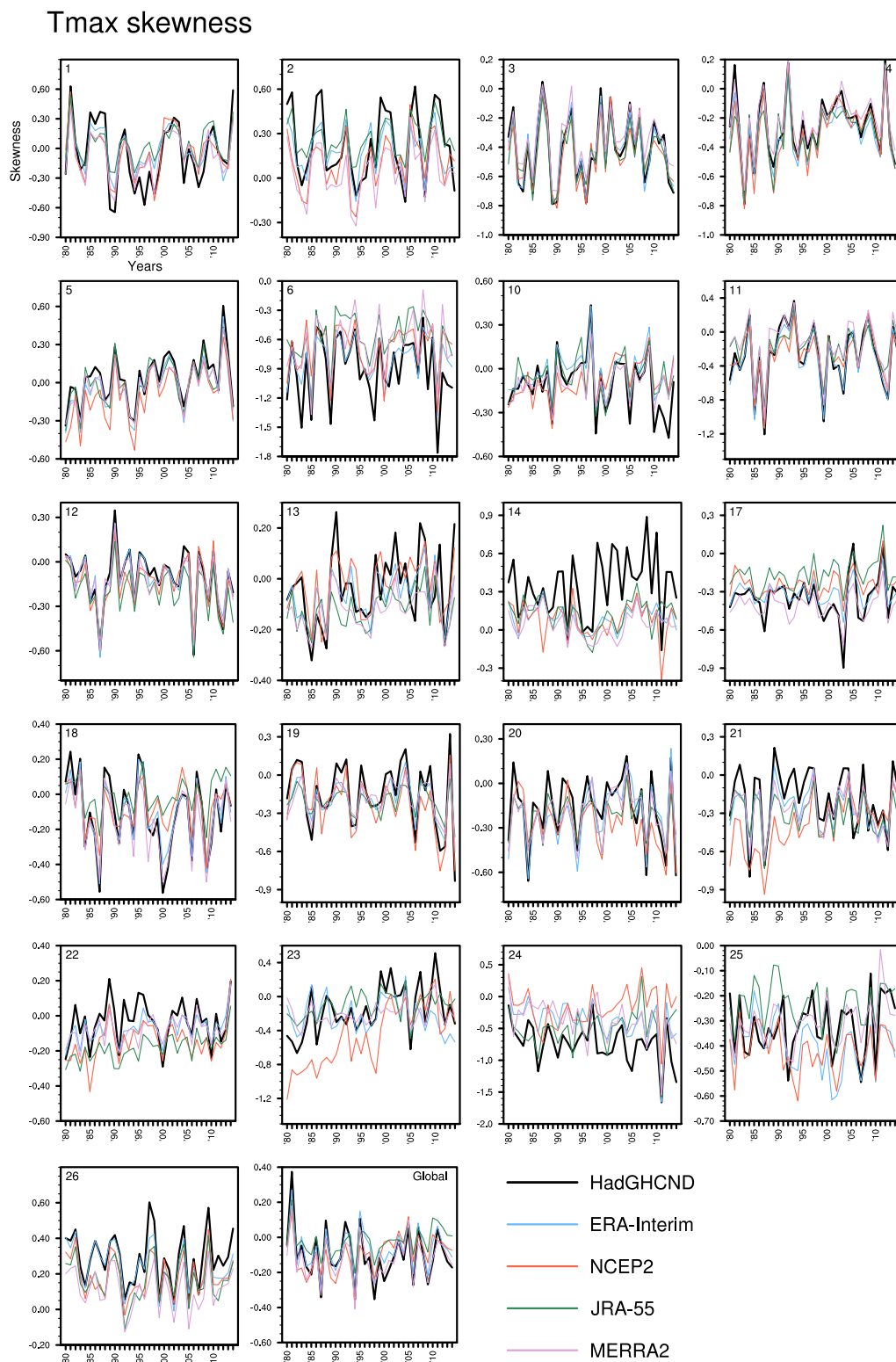


Figure S1.7 As Figure S1.3, but for the annual skewness of daily maximum temperature anomalies ($^{\circ}\text{C}$).

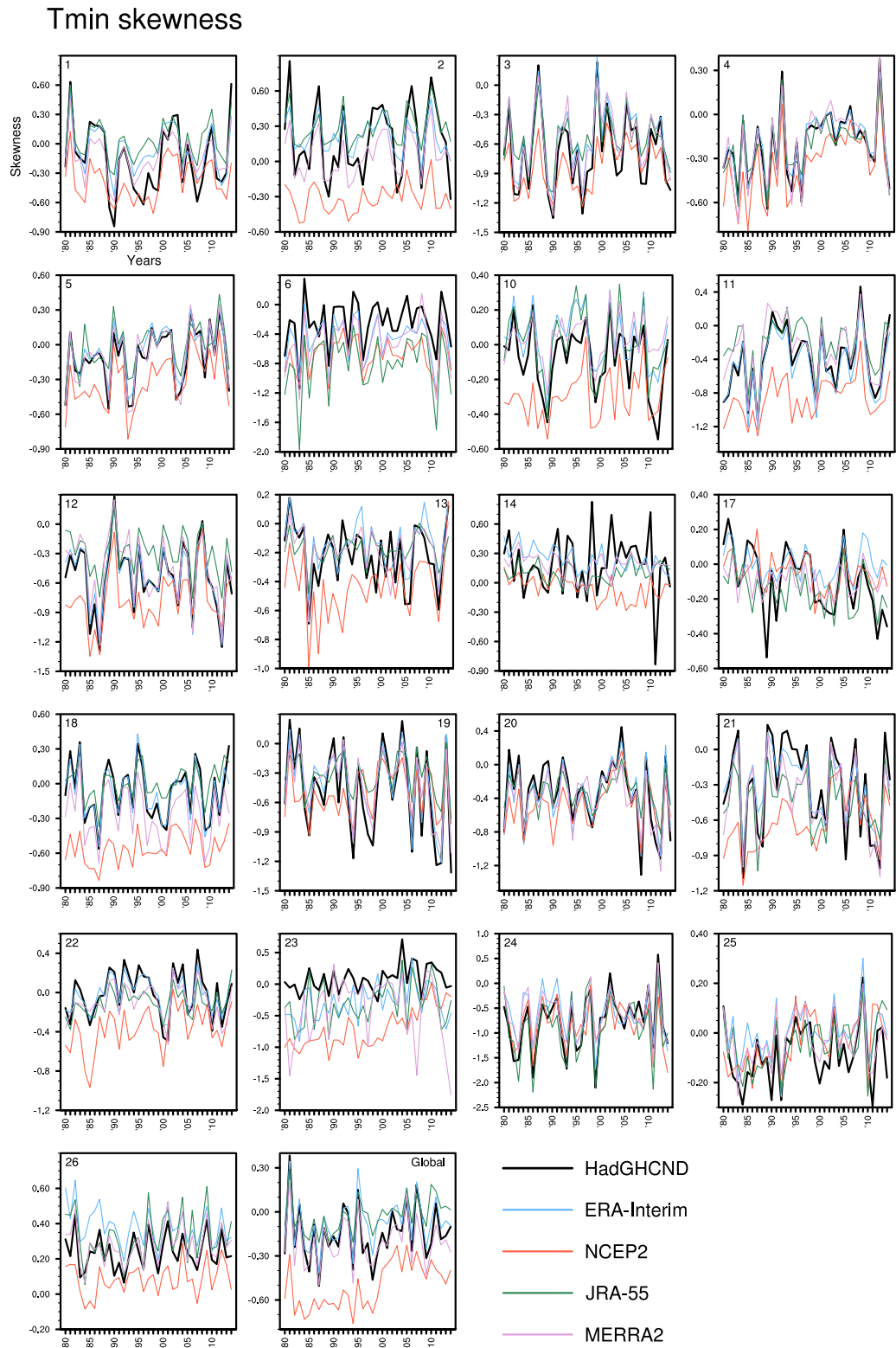


Figure S1.8 As Figure S1.3, but for the annual skewness of daily minimum temperature anomalies (°C).

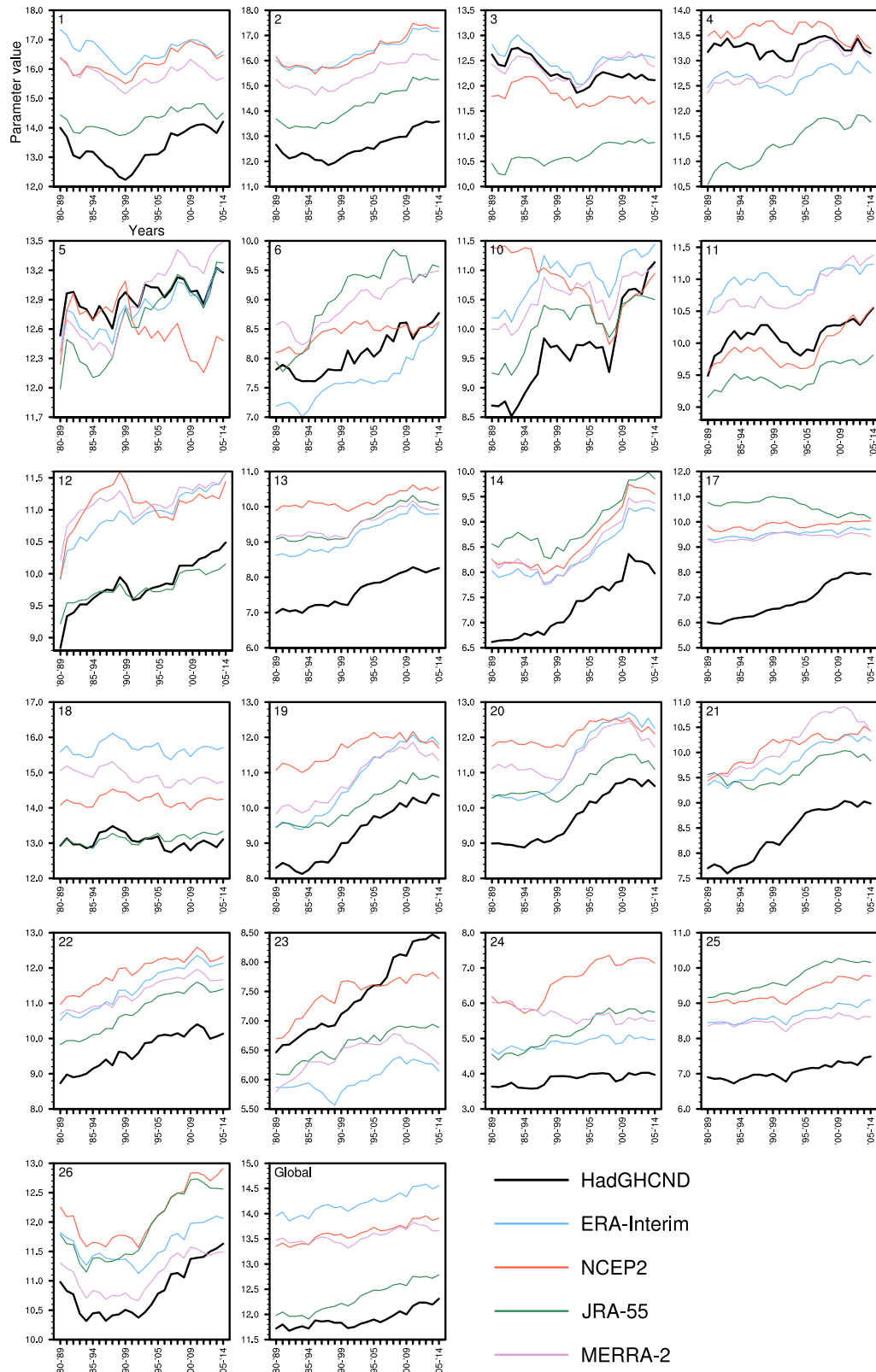
Tmax_{high} location

Figure S1.9 Time series plots of the location parameter of the high tail of maximum temperature anomalies for each region and the globe, based on a PP model fit using running decadal windows from 1980 to 2014.

Tmin_{high} location

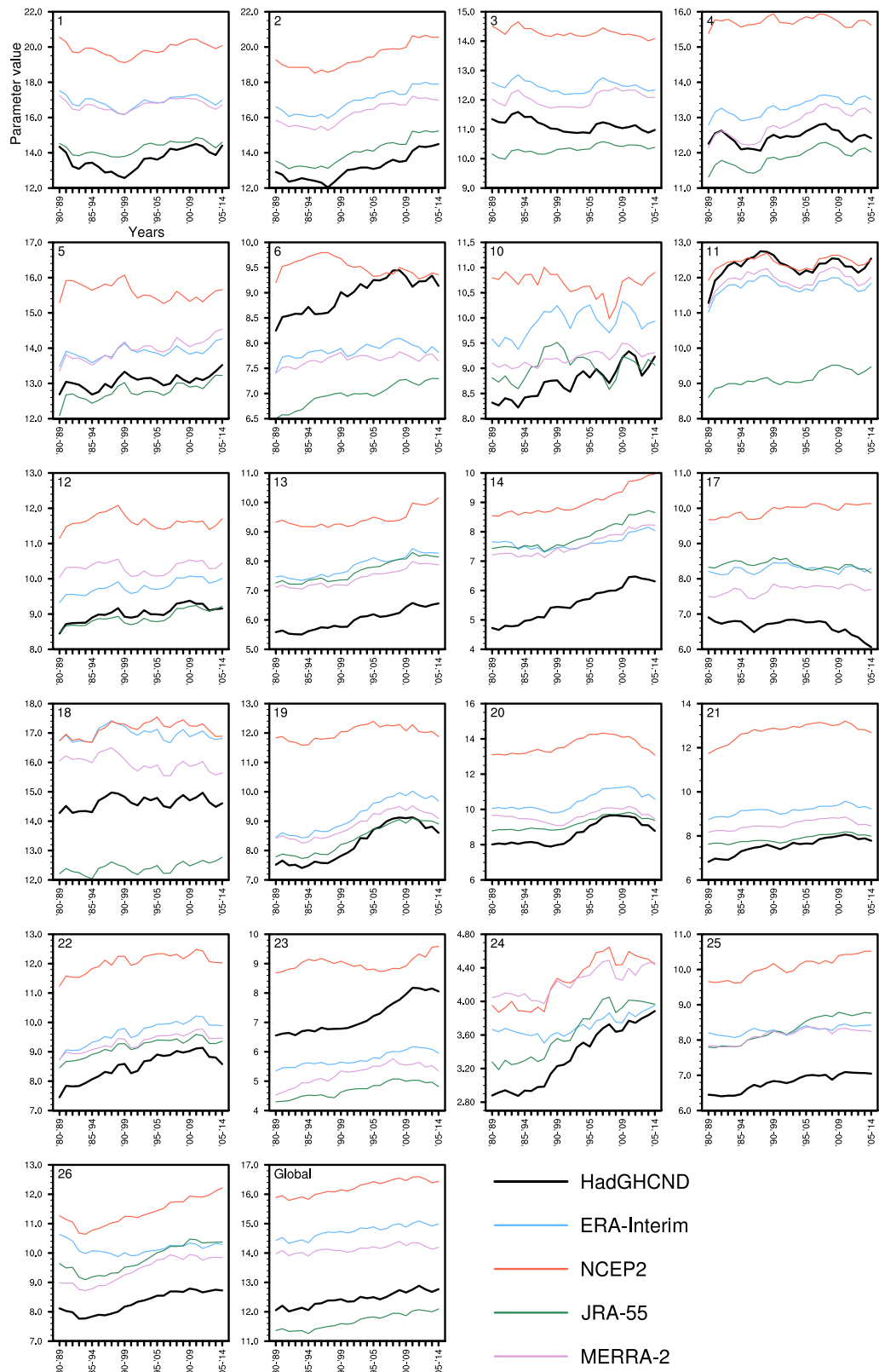


Figure S1.10 As Figure S1.9, but for the high tail of minimum temperature anomalies.

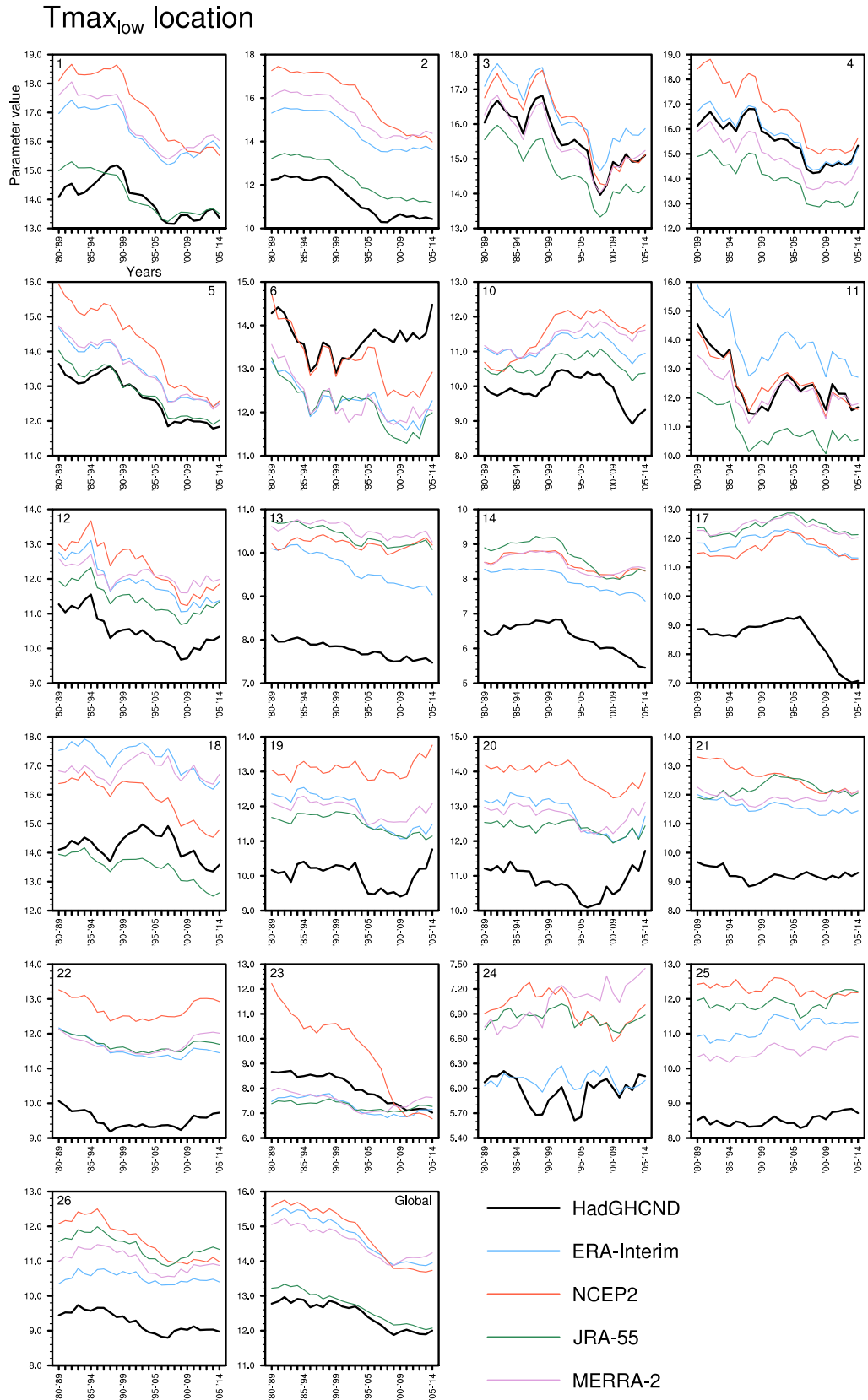


Figure S1.11 As Figure S1.9, but for the low tail of maximum temperature anomalies.

Tmin_{low} location

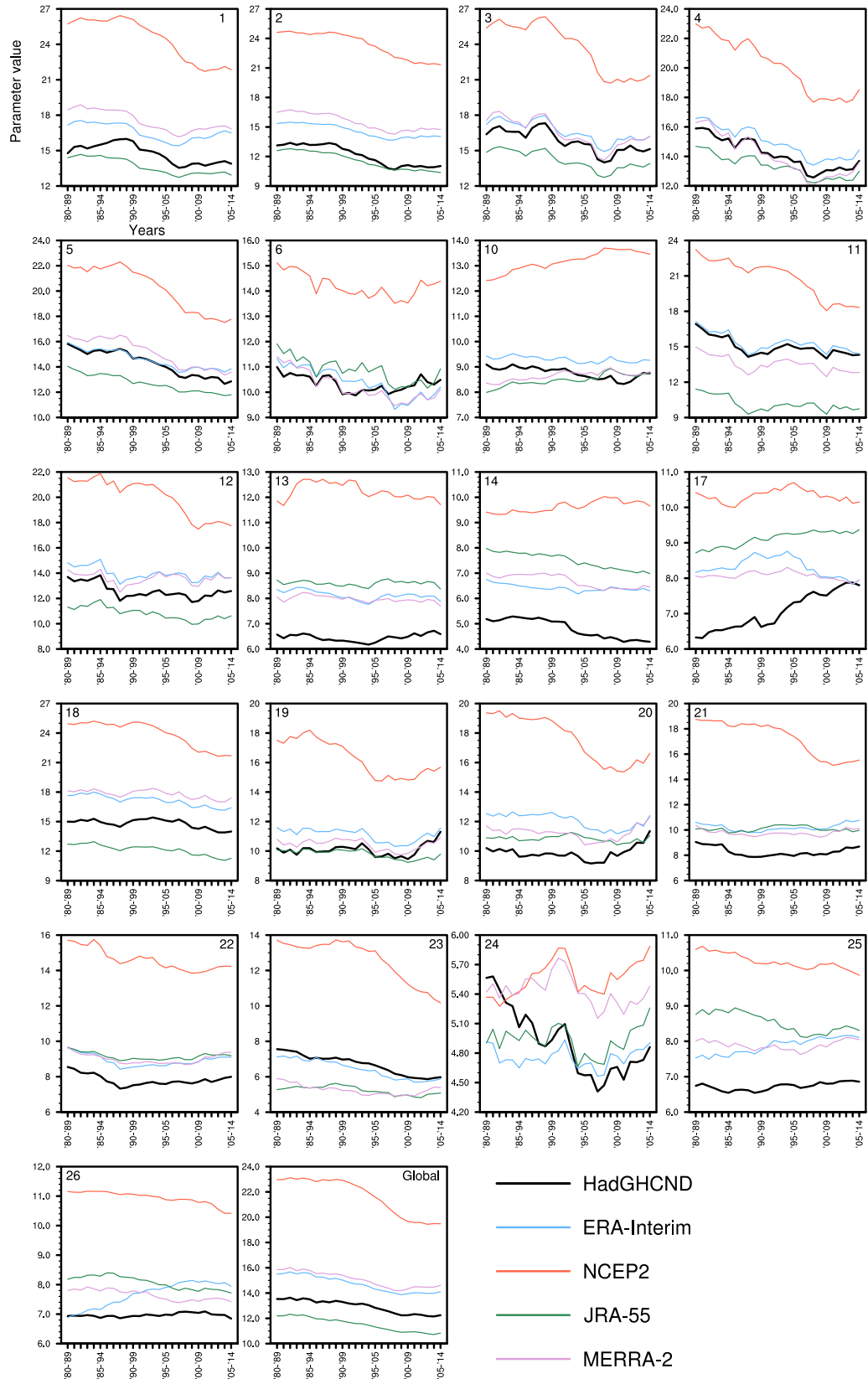


Figure S1.12 As Figure S1.9, but for the low tail of minimum temperature anomalies.

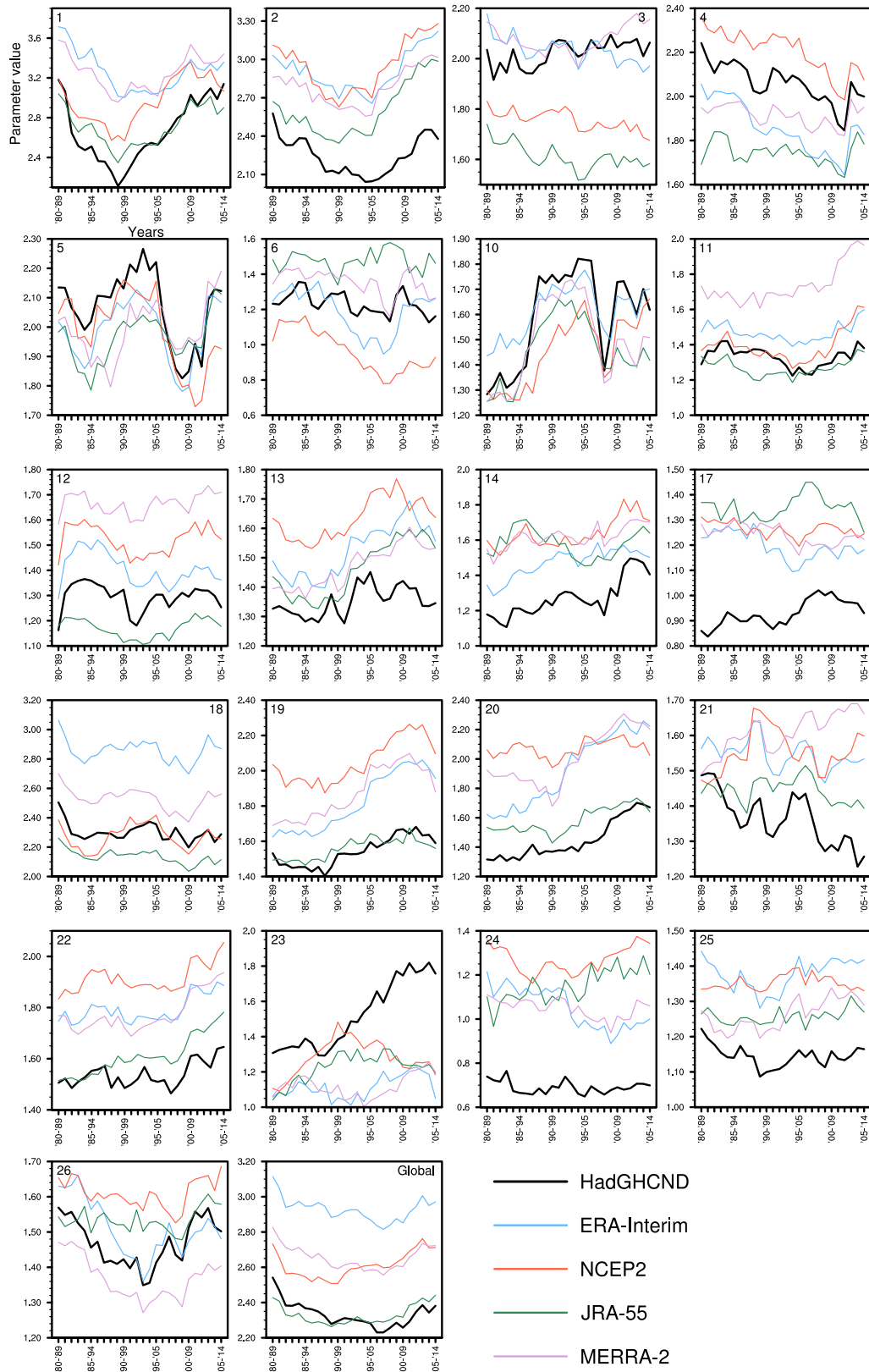
Tmax_{high} scale

Figure S1.13 Time series plots in the scale parameter of the high tail of maximum temperature anomalies for each region and the globe, based on a PP model fit using running decadal windows from 1980 to 2014.

Tmin_{high} scale

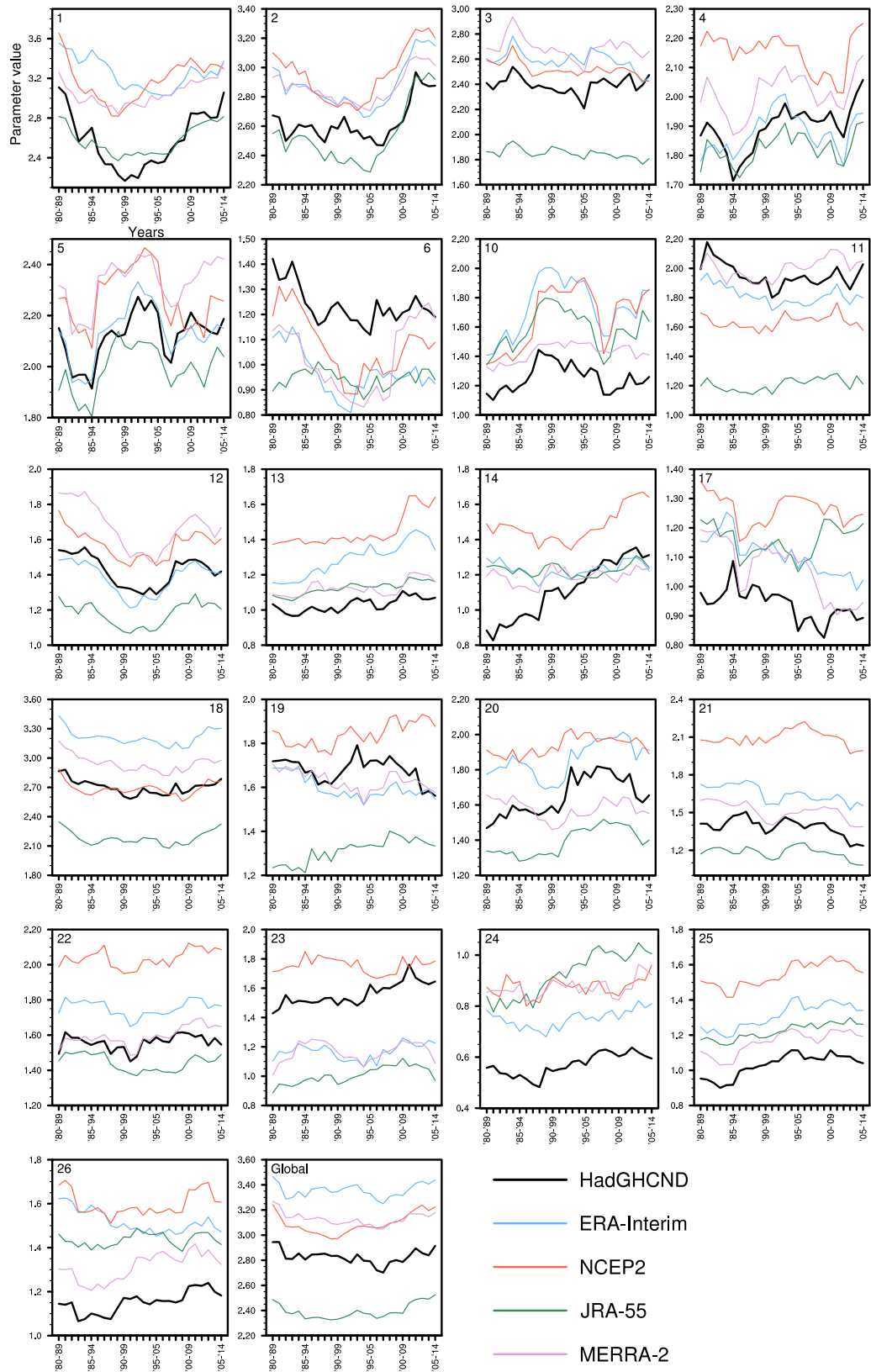


Figure S1.14 As Figure S1.13, but for the high tail of minimum temperature anomalies.

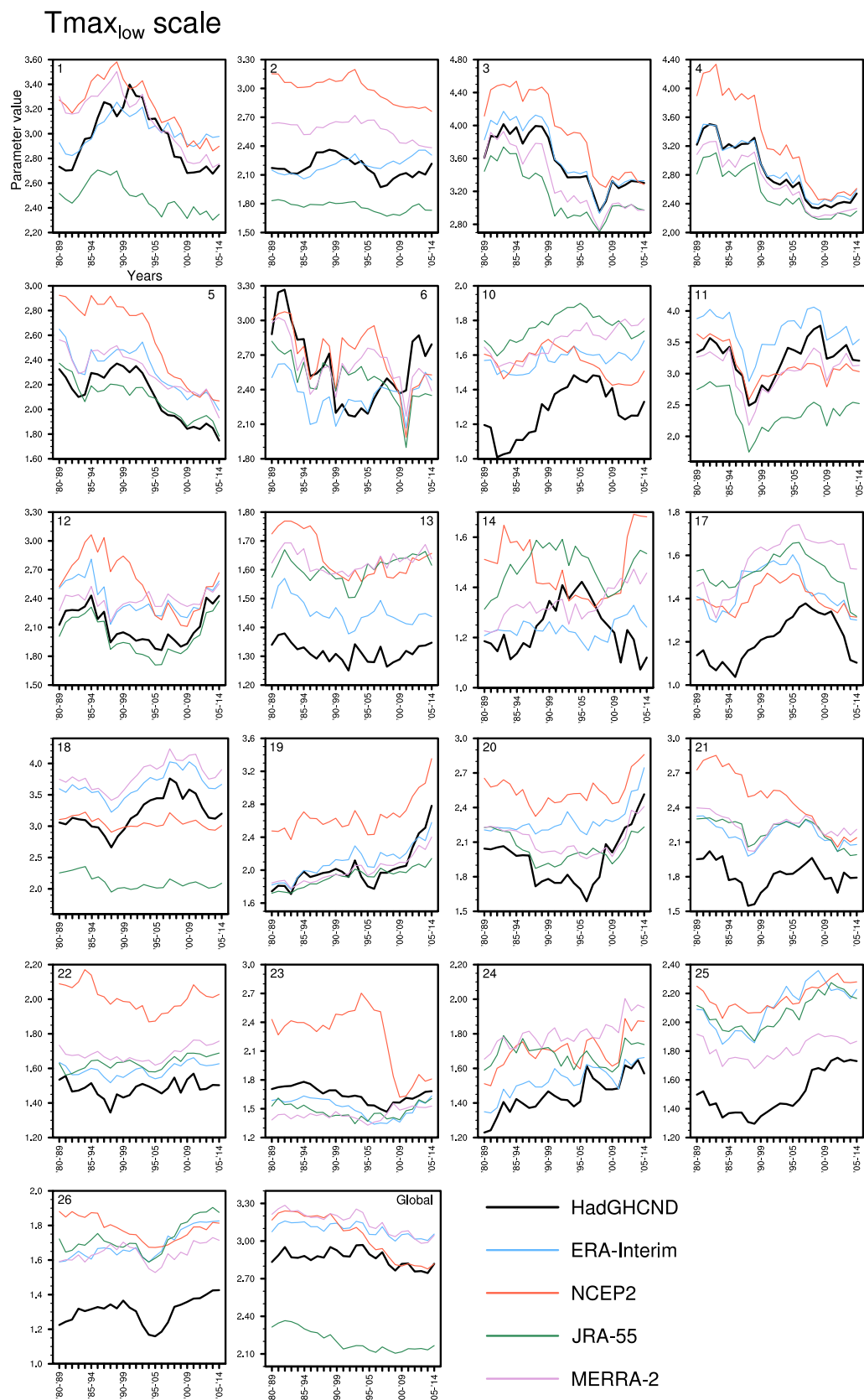


Figure S1.15 As Figure S1.13, but for the low tail of maximum temperature anomalies.

Tmin_{low} scale

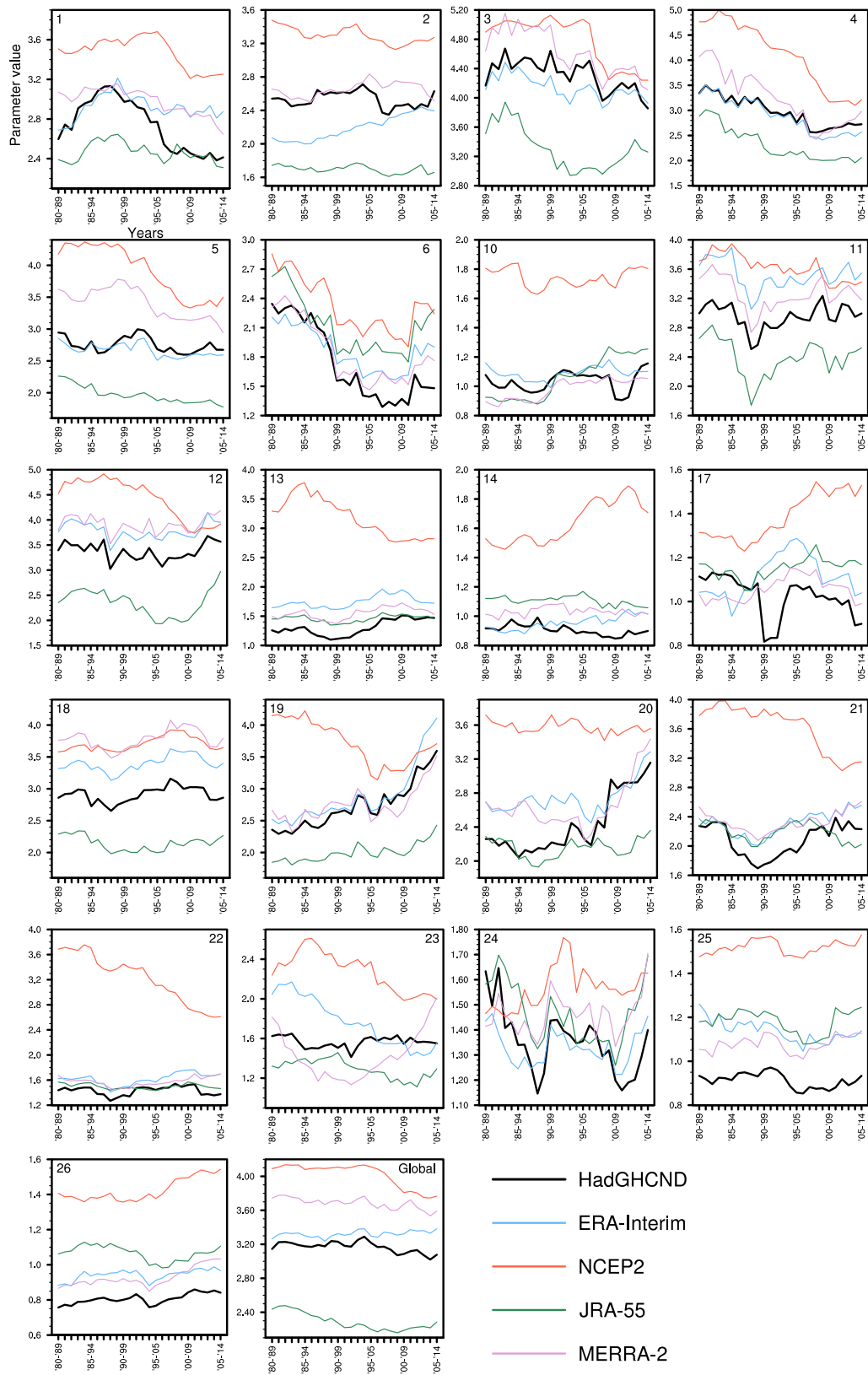


Figure S1.16 As Figure S1.13, but for the low tail of minimum temperature anomalies.

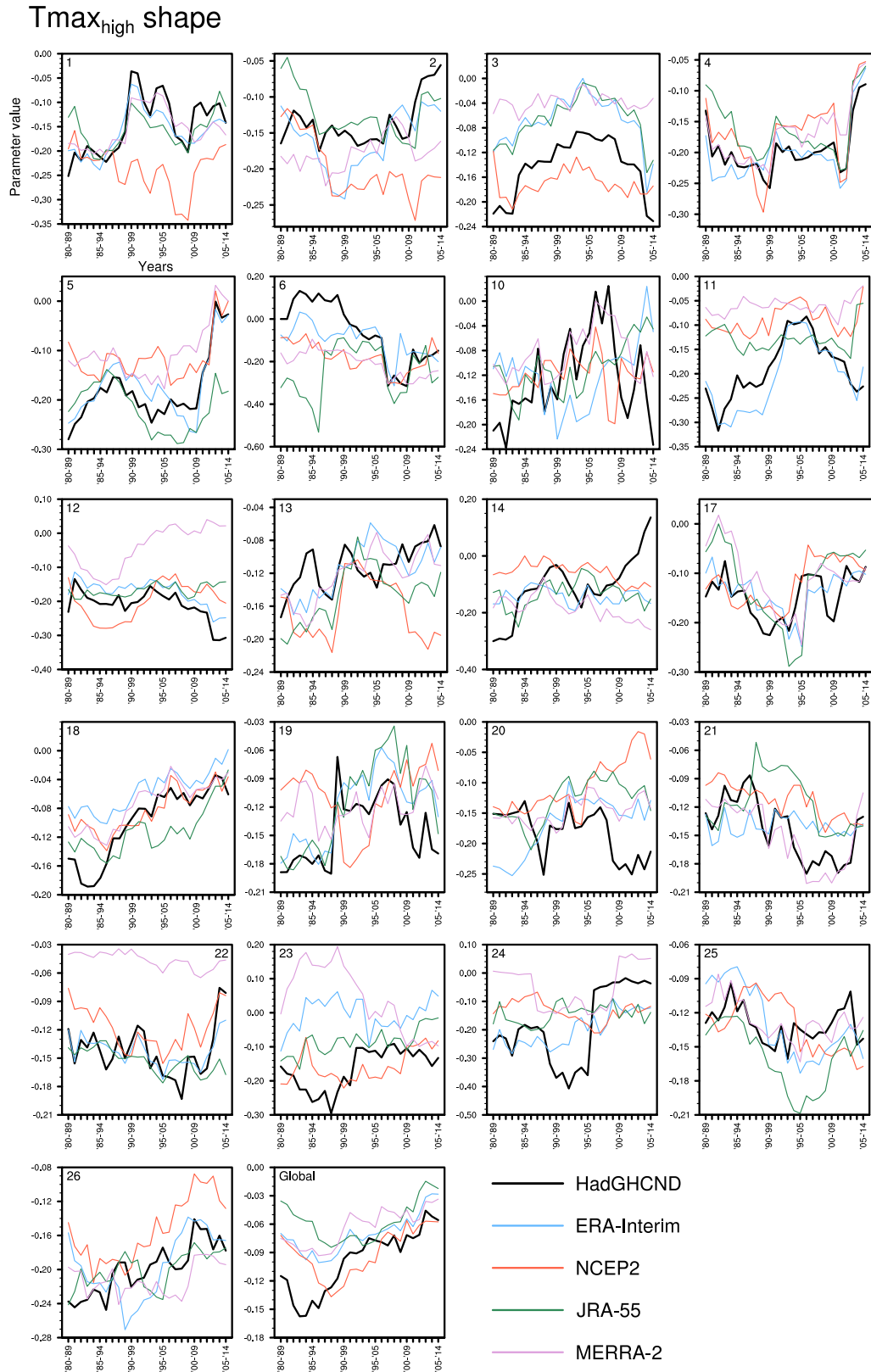


Figure S1.17 Time series plots in the shape parameter of the high tail of maximum temperature anomalies for each region and the globe, based on a PP model fit using running decadal windows from 1980 to 2014.

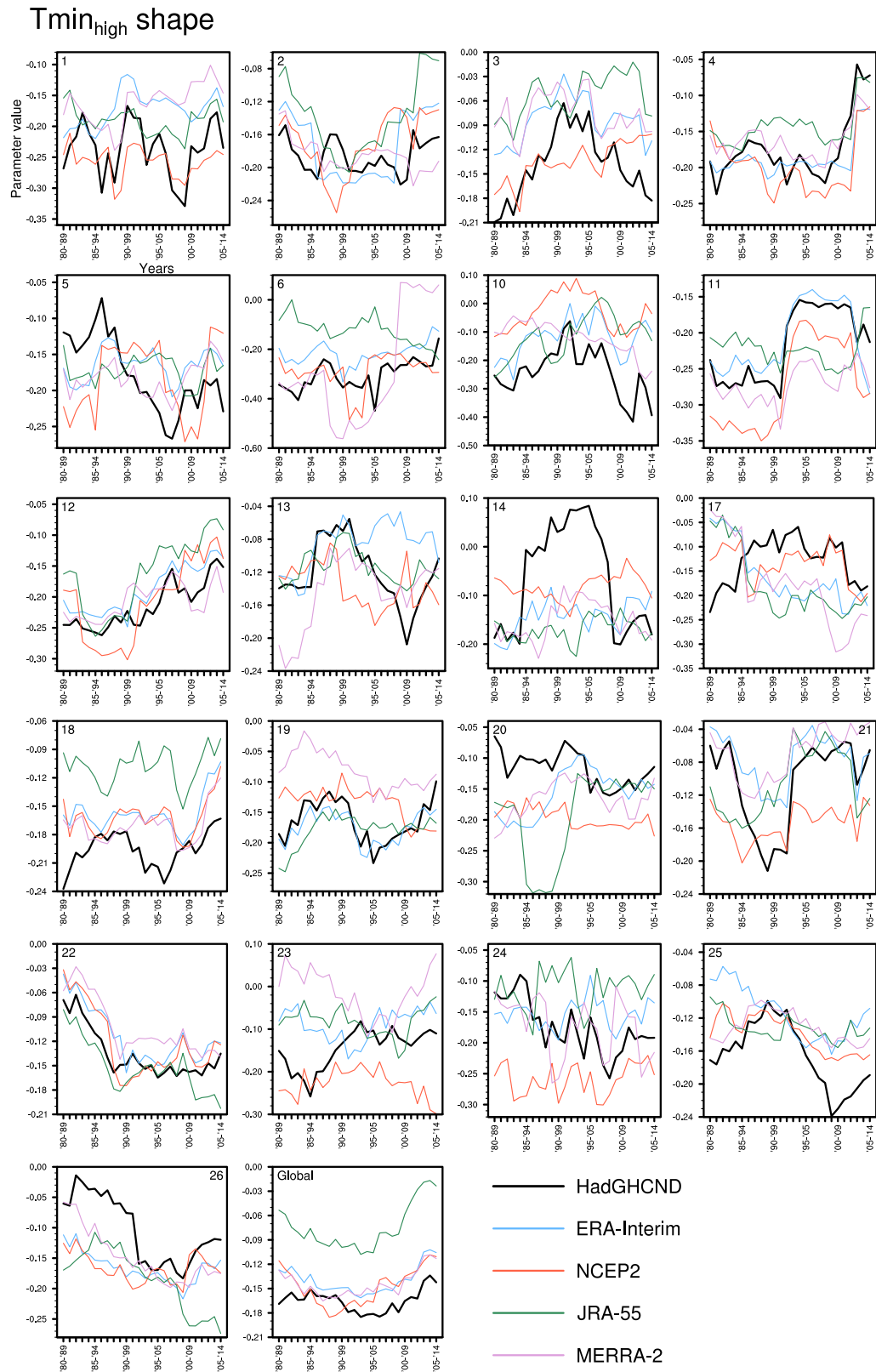


Figure S1.18 As Figure S1.17, but for the high tail of minimum temperature anomalies.

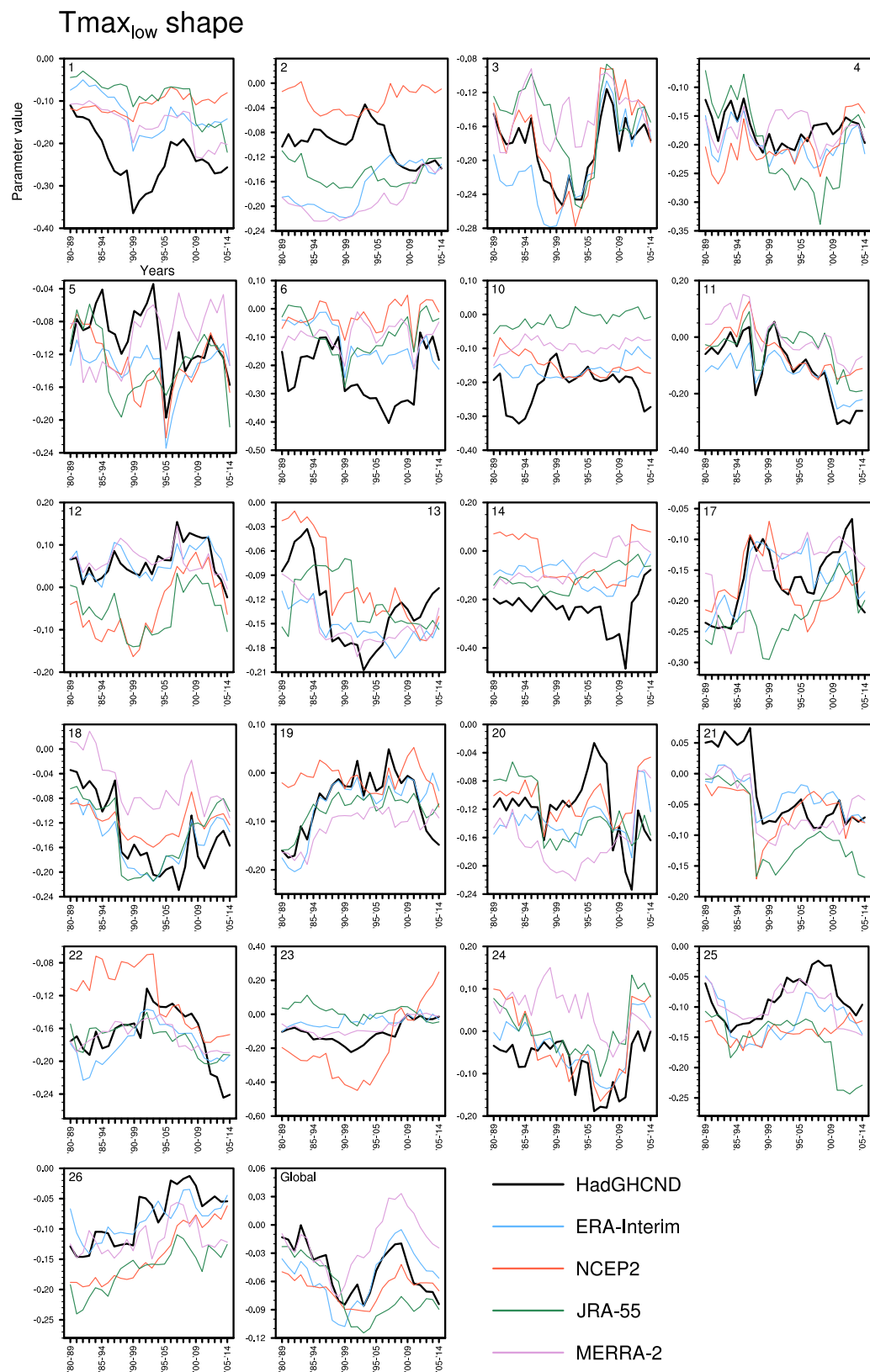


Figure S1.19 As Figure S1.17, but for the low tail of maximum temperature anomalies.

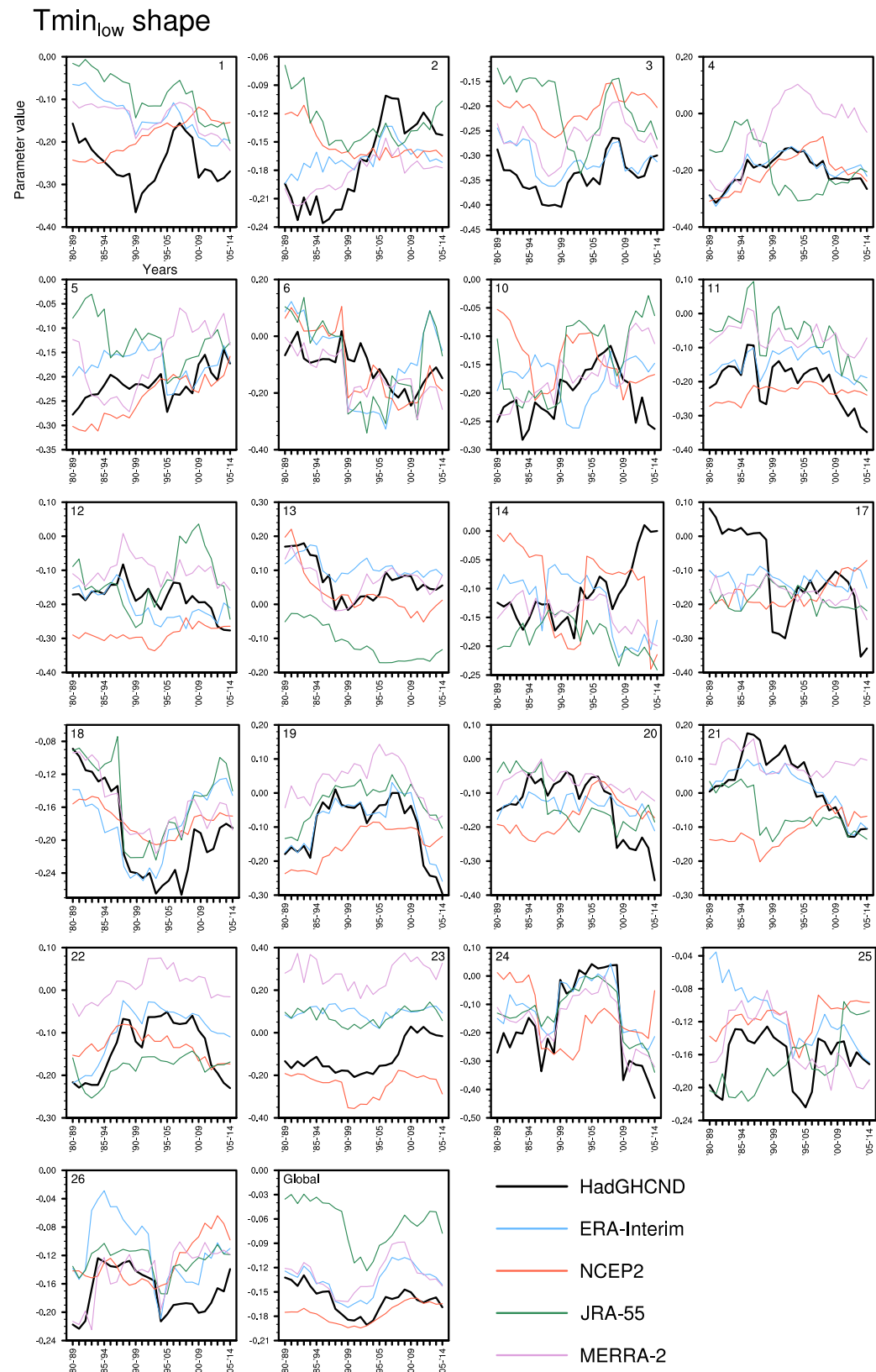


Figure S1.20 As Figure S1.17, but for the low tail of minimum temperature anomalies.

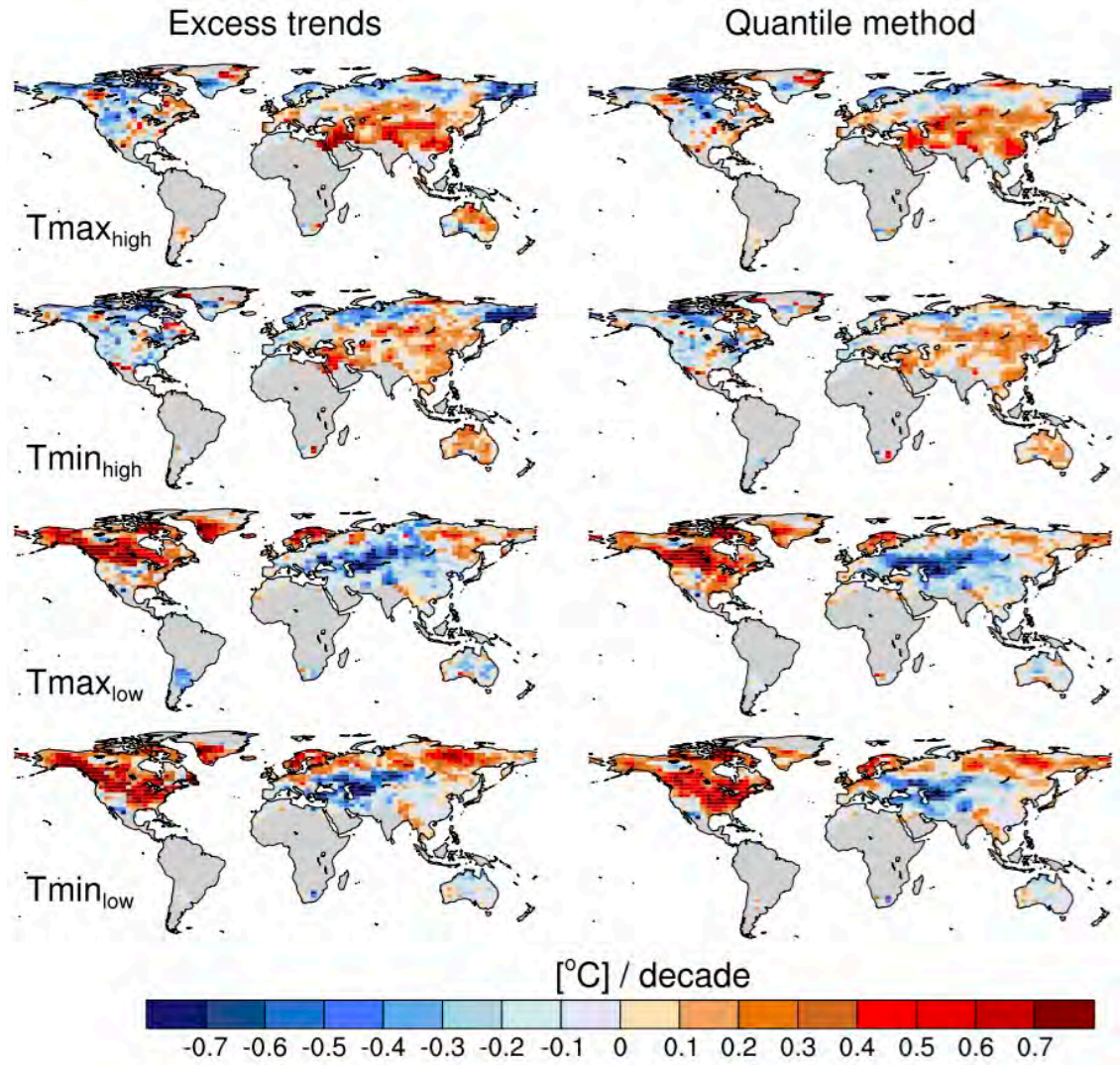


Figure S1.21 Comparison of two different methods used to calculate excess trends. The left column shows results calculated using a non-stationary PP model fit to the data to describe exceedances above (below for the cold tails) the upper (lower) 1.5% of all data points, while the right column shows results calculated as the difference between the 98.5th quantile (1.5th quantile for the cold tails) and the mean.

Water equivalent of accumulated snow depth in NCEP2

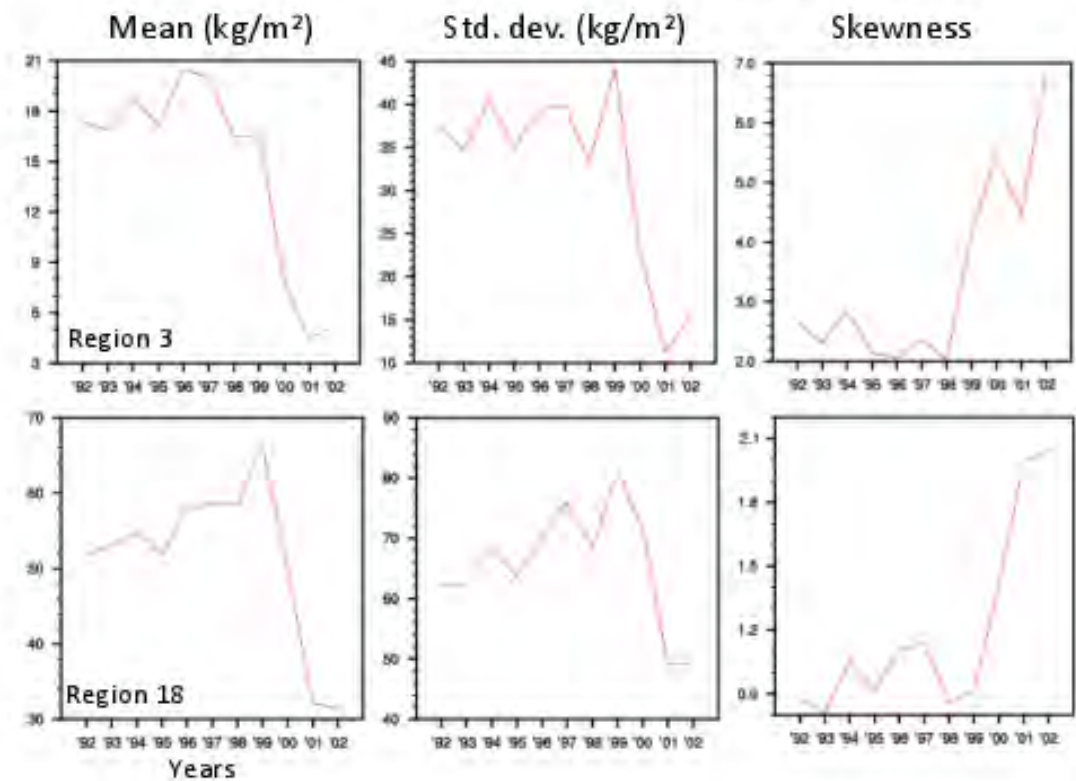


Figure S1.22 Time series of the mean, standard deviation and skewness for the water equivalent of accumulated snow depth in NCEP2 over the period 1992-2002. The first row shows time series for region 3 (western North America) and the bottom row are time series for region 18 (North Asia/Russia). A step change in 1998-1999 is consistent with the step change in NCEP2 $T_{min_{low}}$ apparent for these regions (see Figures 2.5 and 2.8).

S1. SUPPLEMENTARY MATERIAL FOR CHAPTER 2

Table S1.1 Trends in the mean (°C) per decade for each region and the globe over 1980-2014. Bold font indicates statistical significance at the 5% level.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin
1	0.40	0.40	0.38	0.35	0.43	0.44	0.33	0.37	0.16	0.20
2	0.57	0.63	0.62	0.63	0.58	0.74	0.61	0.65	0.32	0.39
3	0.13	0.14	0.23	0.10	0.34	0.34	0.39	0.26	0.19	0.21
4	0.04	0.08	0.20	0.13	0.06	0.13	0.31	0.21	0.24	0.24
5	0.19	0.35	0.23	0.28	0.16	0.29	0.28	0.30	0.31	0.41
6	0.27	0.16	0.40	0.27	0.24	0.06	0.37	0.19	0.29	0.21
10	0.67	0.26	0.22	0.08	-0.54	-0.29	0.21	-0.04	0.03	-0.15
11	0.50	0.51	0.47	0.43	0.43	0.51	0.38	0.39	0.37	0.40
12	0.53	0.37	0.56	0.39	0.39	0.37	0.43	0.36	0.42	0.34
13	0.41	0.33	0.46	0.30	0.11	0.20	0.34	0.28	0.24	0.29
14	0.37	0.36	0.38	0.29	0.37	0.42	0.42	0.41	0.29	0.33
17	0.72	-0.26	0.15	0.05	-0.04	0.21	-0.27	-0.11	0.05	0.08
18	0.30	0.26	0.32	0.32	0.36	0.46	0.29	0.30	0.13	0.17
19	0.51	0.35	0.70	0.52	0.28	0.20	0.45	0.34	0.37	0.35
20	0.38	0.24	0.41	0.22	0.42	0.40	0.21	0.15	0.10	0.14
21	0.39	0.36	0.25	0.25	0.14	0.32	0.06	0.12	0.17	0.19
22	0.36	0.29	0.37	0.31	0.24	0.10	0.29	0.19	0.15	0.17
23	0.45	0.42	0.24	0.23	0.60	0.35	0.20	0.09	0.13	0.23
24	0.21	0.25	0.24	0.11	0.16	0.09	0.19	0.08	-0.01	0.11
25	0.12	0.02	0.16	-0.17	0.15	0.12	0.19	0.17	-0.11	0.02
26	0.32	0.11	0.24	-0.19	0.32	0.29	0.27	0.22	0.18	0.24
Global	0.36	0.33	0.39	0.32	0.34	0.38	0.34	0.31	0.19	0.24

Table S1.2 As Table S1.1, but for standard deviation (°C per decade).

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin
1	-0.07	-0.07	-0.16	-0.14	-0.21	-0.29	-0.12	-0.13	-0.12	-0.13
2	-0.07	-0.08	-0.08	-0.06	-0.17	-0.17	-0.07	-0.07	-0.05	-0.03
3	-0.10	-0.11	-0.07	-0.08	-0.14	-0.27	-0.04	-0.04	-0.05	-0.09
4	-0.03	-0.10	-0.05	-0.05	-0.12	-0.17	0.01	-0.04	-0.02	-0.06
5	-0.03	-0.17	-0.03	-0.10	-0.16	-0.20	-0.02	-0.08	-0.02	-0.11
6	0.08	0.16	-0.01	-0.02	-0.04	-0.03	0.03	0.01	-0.01	-0.04
10	0.04	-0.04	0.06	0.02	0.02	0.00	0.11	0.04	0.10	0.02
11	-0.08	-0.09	-0.09	-0.08	-0.07	-0.24	-0.05	-0.04	-0.04	-0.05
12	0.07	0.02	0.06	0.04	0.04	-0.09	0.04	0.04	0.06	0.05
13	0.03	0.03	0.02	0.00	0.05	0.00	0.03	0.03	0.03	0.02
14	0.00	-0.02	0.00	-0.02	0.06	0.07	0.04	0.02	0.05	0.02
17	0.02	0.10	0.02	-0.01	0.05	0.01	-0.03	0.00	-0.01	0.01
18	0.02	-0.01	-0.01	-0.04	-0.06	-0.21	-0.01	-0.04	-0.02	-0.05
19	0.16	0.12	0.09	0.05	0.07	-0.03	0.06	0.05	0.08	0.06
20	0.11	0.08	0.08	0.02	-0.01	-0.18	0.06	0.05	0.06	0.05
21	0.09	0.06	0.05	0.04	-0.02	-0.08	0.03	0.03	0.05	0.02
22	0.09	0.04	0.07	0.04	0.08	0.01	0.07	0.03	0.05	0.03
23	-0.04	-0.03	0.02	-0.01	-0.30	-0.17	0.05	0.04	-0.01	0.04
24	0.01	-0.01	0.02	0.01	0.05	0.04	0.07	0.04	-0.01	0.01
25	0.03	0.04	0.01	0.03	0.05	-0.03	0.06	0.04	0.02	0.00
26	0.01	0.04	0.02	0.07	0.03	-0.01	0.04	0.03	0.01	0.01
Global	0.00	-0.03	-0.03	-0.04	-0.07	-0.16	-0.01	-0.03	-0.02	-0.04

Table S1.3 As Table S1.1, but for skewness (trend per decade).

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin	Tmax	Tmin
1	0.00	-0.02	-0.02	-0.04	0.02	0.03	0.01	0.00	0.02	0.02
2	0.00	0.02	-0.02	-0.02	0.04	0.03	0.01	0.03	0.02	0.03
3	0.00	0.01	0.00	0.03	0.01	0.08	0.00	0.02	0.01	0.03
4	0.03	0.07	0.03	0.08	0.07	0.12	0.05	0.05	0.04	0.09
5	0.08	0.07	0.08	0.06	0.10	0.10	0.07	0.07	0.07	0.08
6	-0.01	0.04	0.02	0.00	0.02	0.02	0.04	0.05	0.06	0.04
10	-0.03	-0.04	0.02	-0.03	0.07	0.04	0.01	-0.03	0.03	0.02
11	0.03	0.08	0.03	0.05	0.04	0.09	0.01	0.03	0.00	0.01
12	-0.03	-0.03	-0.03	-0.05	-0.01	0.07	0.04	-0.03	-0.03	-0.08
13	0.04	-0.03	0.03	0.00	0.03	0.06	0.02	0.01	0.03	-0.01
14	0.06	0.00	0.00	-0.08	-0.02	-0.06	0.00	0.01	-0.01	-0.02
17	0.01	-0.09	0.01	-0.01	0.20	-0.04	0.3	-0.04	0.00	-0.01
18	-0.04	-0.01	-0.02	-0.02	0.00	0.05	0.02	0.03	-0.02	-0.03
19	-0.05	-0.12	-0.02	-0.11	-0.09	-0.03	-0.01	-0.02	-0.01	-0.11
20	-0.02	-0.09	0.05	0.01	-0.05	0.06	0.00	0.00	0.03	-0.06
21	-0.02	-0.06	0.00	-0.08	0.13	0.12	-0.02	-0.02	0.00	-0.05
22	0.01	0.01	0.03	0.01	0.04	0.09	0.05	0.04	0.03	0.00
23	0.13	0.07	-0.04	0.09	0.29	0.25	0.06	0.10	0.03	-0.02
24	-0.12	0.16	-0.14	-0.06	0.05	-0.06	0.05	0.07	-0.15	-0.01
25	0.02	0.02	-0.03	0.02	0.00	0.06	0.00	0.04	0.03	0.01
26	-0.01	0.01	-0.04	-0.03	0.00	0.02	0.00	0.02	-0.01	0.02
Global	-0.02	0.01	0.00	0.02	0.04	0.09	0.02	0.04	0.00	0.01

Table S1.4 Decadal trends in the location parameter for the high tails of maximum and minimum temperature anomalies (Tmax_H and Tmin_H, respectively) for each region and the globe, based on a PP model fit using running decadal windows from 1980 to 2014. Bold font indicates statistical significance at the 5% level.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	Tmax _H	Tmin _H	Tmax _H	Tmin _H	Tmax _H	Tmin _H	Tmax _H	Tmin _H	Tmax _H	Tmin _H
1	0.43	0.44	-0.04	0.05	0.36	0.14	0.31	0.31	-0.01	0.10
2	0.59	0.84	0.71	0.84	0.76	0.82	0.91	0.94	0.62	0.79
3	-0.21	-0.16	-0.11	-0.06	-0.14	-0.12	0.24	0.15	0.06	0.13
4	0.02	0.10	0.11	0.23	-0.07	0.03	0.52	0.28	0.38	0.45
5	0.15	0.17	0.24	0.14	-0.19	-0.15	0.44	0.25	0.46	0.32
6	0.44	0.41	0.50	0.12	0.18	-0.12	0.75	0.29	0.49	0.09
10	0.90	0.38	0.41	0.19	-0.46	-0.10	0.47	0.08	0.36	0.15
11	0.20	0.10	0.18	0.08	0.28	0.06	0.19	0.26	0.36	0.10
12	0.44	0.24	0.46	0.22	0.18	-0.03	0.26	0.24	0.27	0.04
13	0.60	0.46	0.63	0.44	0.25	0.30	0.58	0.47	0.43	0.39
14	0.74	0.77	0.61	0.21	0.72	0.56	0.60	0.56	0.61	0.47
17	0.93	-0.19	0.17	0.03	0.13	0.19	-0.26	-0.07	0.12	0.11
18	-0.06	0.12	-0.01	0.00	-0.01	0.15	0.12	0.16	-0.18	-0.22
19	1.02	0.78	1.13	0.72	0.41	0.20	0.71	0.64	0.87	0.52
20	0.94	0.72	1.21	0.53	0.30	0.35	0.53	0.43	0.58	0.20
21	0.67	0.44	0.47	0.22	0.38	0.39	0.28	0.22	0.60	0.21
22	0.63	0.55	0.76	0.47	0.55	0.31	0.76	0.34	0.51	0.32
23	0.86	0.73	0.24	0.30	0.40	0.15	0.36	0.31	0.25	0.40
24	0.17	0.46	0.16	0.12	0.68	0.33	0.64	0.39	-0.22	0.19
25	0.26	0.30	0.27	0.12	0.35	0.38	0.50	0.47	0.11	0.22
26	0.43	0.43	0.23	0.01	0.47	0.54	0.62	0.55	0.28	0.53
Global	0.21	0.31	0.27	0.27	0.23	0.32	0.38	0.34	0.14	0.13

Table S1.5 As Table S1.4, but for the low tails of maximum and minimum temperature anomalies (T_{\max_L} and T_{\min_L} , respectively) of the location parameter.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}
1	-0.60	-0.85	-0.87	-0.66	-1.46	-2.18	-0.86	-0.81	-1.01	-0.98
2	-1.03	-1.27	-1.01	-0.80	-1.59	-1.62	-1.11	-1.14	-1.07	-1.03
3	-0.85	-1.00	-1.00	-0.93	-1.20	-2.55	-0.89	-0.84	-0.84	-1.24
4	-0.92	-1.27	-1.10	-1.26	-1.64	-2.36	-0.95	-0.92	-1.05	-1.65
5	-0.78	-1.27	-0.92	-0.93	-1.44	-2.16	-0.84	-0.89	-0.98	-1.41
6	0.02	-0.19	-0.43	-0.67	-0.68	-0.37	-0.56	-0.52	-0.50	-0.65
10	-0.15	-0.23	0.05	-0.10	0.60	0.48	0.06	0.27	0.33	0.18
11	-0.75	-0.73	-0.83	-0.71	-0.75	-2.01	-0.56	-0.54	-0.40	-0.61
12	-0.56	-0.53	-0.70	-0.41	-0.79	-1.78	-0.46	-0.55	-0.25	-0.11
13	-0.24	0.03	-0.47	-0.14	-0.04	-0.19	-0.27	-0.01	-0.15	-0.10
14	-0.44	-0.47	-0.38	-0.12	-0.26	0.24	-0.47	-0.41	-0.23	-0.29
17	-0.58	0.67	-0.13	-0.14	0.08	0.02	0.06	0.26	-0.03	-0.05
18	-0.21	-0.39	-0.52	-0.66	-0.83	-1.56	-0.54	-0.64	-0.05	-0.38
19	-0.14	0.11	-0.58	-0.32	0.11	-1.37	-0.28	-0.30	-0.19	-0.17
20	-0.15	0.13	-0.53	-0.48	-0.31	-1.85	-0.15	-0.08	-0.19	-0.06
21	-0.11	-0.13	-0.21	0.10	-0.54	-1.68	0.09	0.03	0.01	0.00
22	-0.10	-0.17	-0.25	-0.14	-0.08	-0.70	-0.13	-0.10	0.01	-0.08
23	-0.76	-0.72	-0.36	-0.68	-2.15	-1.36	-0.15	-0.24	-0.25	-0.26
24	-0.01	-0.36	-0.01	0.00	-0.13	0.12	-0.04	0.00	0.26	-0.08
25	0.13	0.09	0.25	0.28	-0.13	-0.26	0.11	-0.31	0.24	0.00
26	-0.32	0.04	-0.10	0.51	-0.68	-0.26	-0.30	-0.26	-0.26	-0.20
Global	-0.47	-0.66	-0.78	-0.82	-0.99	-1.80	-0.58	-0.70	-0.54	-0.75

Table S1.6 As Table S1.4, but for the high tails of maximum and minimum temperature anomalies (T_{\max_H} and T_{\min_H} , respectively) of the scale parameter.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}
1	0.18	0.04	-0.11	-0.11	0.19	0.06	0.07	0.05	0.02	0.08
2	-0.02	0.09	0.07	0.09	0.11	0.10	0.19	0.15	0.08	0.05
3	0.04	0.00	-0.05	-0.05	-0.04	-0.05	-0.04	-0.03	0.02	-0.04
4	-0.09	0.06	-0.13	0.03	-0.11	-0.03	-0.02	0.03	-0.03	0.04
5	-0.05	0.06	0.00	0.04	-0.10	-0.01	0.05	0.04	0.06	0.06
6	-0.04	-0.06	-0.06	-0.06	-0.13	-0.08	0.00	0.00	-0.05	-0.03
10	0.14	0.01	0.07	0.08	0.14	0.16	0.05	0.03	0.06	0.04
11	-0.02	-0.05	0.00	-0.04	0.04	0.01	0.01	0.03	0.11	0.04
12	-0.01	-0.04	-0.04	-0.02	0.00	-0.02	0.00	0.01	0.02	-0.08
13	0.03	0.04	0.10	0.12	0.07	0.10	0.11	0.04	0.08	0.04
14	0.12	0.21	0.09	0.01	0.08	0.08	-0.02	0.01	0.06	0.02
17	0.05	-0.05	-0.04	-0.08	-0.02	-0.02	0.01	0.00	-0.03	-0.11
18	-0.03	-0.03	-0.03	-0.02	0.01	-0.01	-0.04	0.00	-0.05	-0.05
19	0.09	-0.03	0.20	-0.04	0.13	0.05	0.06	0.06	0.17	-0.03
20	0.16	0.10	0.31	0.08	0.03	0.04	0.09	0.08	0.20	-0.01
21	-0.08	-0.06	-0.03	-0.07	0.02	0.00	-0.01	-0.03	0.06	-0.06
22	0.04	0.01	0.05	0.00	0.04	0.03	0.09	-0.02	0.07	0.05
23	0.24	0.09	0.04	0.02	0.03	0.00	0.06	0.06	0.03	0.02
24	-0.01	0.04	-0.10	0.03	0.02	0.02	0.08	0.11	-0.03	0.02
25	-0.01	0.07	0.02	0.08	0.01	0.07	0.01	0.05	0.04	0.07
26	0.00	0.05	-0.06	-0.05	0.00	0.01	0.01	0.01	-0.03	0.06
Global	-0.05	-0.02	-0.04	0.01	0.05	0.05	0.02	0.03	-0.03	-0.02

Table S1.7 As Table S1.4, but for the low tails of maximum and minimum temperature anomalies (T_{\max_L} and T_{\min_L} , respectively) of the scale parameter.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}
1	-0.08	-0.24	0.01	0.01	-0.21	-0.12	-0.11	-0.04	-0.25	-0.14
2	-0.05	-0.02	0.09	0.18	-0.14	-0.09	-0.04	-0.03	-0.08	0.05
3	-0.35	-0.20	-0.43	-0.15	-0.58	-0.35	-0.33	-0.25	-0.44	-0.34
4	-0.49	-0.35	-0.46	-0.43	-0.82	-0.83	-0.37	-0.40	-0.45	-0.64
5	-0.21	-0.06	-0.19	-0.08	-0.39	-0.49	-0.18	-0.16	-0.19	-0.25
6	-0.16	-0.44	-0.02	-0.20	-0.22	-0.27	-0.18	-0.25	-0.14	-0.32
10	0.12	0.02	0.05	0.02	-0.06	0.00	0.05	0.18	0.11	0.08
11	0.08	0.01	-0.03	-0.07	-0.19	-0.20	-0.10	-0.05	0.04	-0.07
12	-0.04	-0.01	-0.09	-0.01	-0.24	-0.46	-0.04	-0.09	0.02	0.00
13	-0.01	0.13	-0.03	0.08	-0.06	-0.38	0.02	0.02	0.00	0.07
14	0.01	-0.03	0.01	0.05	-0.01	0.17	0.02	-0.02	0.07	0.00
17	0.08	-0.06	0.00	0.04	-0.01	0.12	-0.01	0.03	0.11	0.27
18	0.20	0.05	0.14	0.08	-0.05	0.07	-0.09	-0.04	0.16	0.06
19	0.23	0.44	0.23	0.51	0.20	-0.35	0.14	0.15	0.19	0.25
20	0.09	0.39	0.12	0.16	0.04	-0.04	-0.03	0.03	0.00	0.19
21	-0.03	0.06	-0.05	0.13	-0.32	-0.37	-0.09	-0.06	-0.05	0.07
22	0.01	0.00	0.02	0.06	-0.04	-0.50	0.03	-0.01	0.03	0.03
23	-0.07	0.00	-0.07	-0.31	-0.27	-0.21	-0.01	-0.09	0.04	0.09
24	0.12	-0.10	0.10	-0.01	0.08	0.05	0.00	-0.06	0.09	0.03
25	0.15	-0.02	0.15	-0.05	0.07	0.01	0.11	-0.01	0.05	0.02
26	0.04	0.03	0.09	0.02	-0.05	0.07	0.08	-0.02	0.03	0.05
Global	-0.05	-0.05	-0.05	0.02	-0.22	-0.15	-0.10	-0.11	-0.10	-0.07

Table S1.8 As Table S1.4, but for the high tails of maximum and minimum temperature anomalies (T_{\max_H} and T_{\min_H} , respectively) of the shape parameter.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}	T_{\max_H}	T_{\min_H}
1	0.04	0.00	0.02	0.02	-0.02	-0.01	0.02	-0.01	0.02	0.03
2	0.02	0.00	0.02	0.01	-0.04	0.02	-0.02	0.01	0.02	-0.02
3	0.02	0.01	0.00	0.01	0.00	0.03	0.01	0.02	0.00	0.00
4	0.02	0.03	0.03	0.02	0.03	0.00	0.00	0.02	0.04	0.01
5	0.06	-0.05	0.04	0.01	0.03	0.01	-0.02	0.00	0.04	0.01
6	-0.15	0.05	-0.07	0.03	-0.05	0.02	0.03	-0.05	-0.06	0.18
10	0.03	-0.03	0.03	0.03	0.01	0.00	0.04	0.07	0.02	-0.06
11	0.04	0.05	0.05	0.03	0.01	0.05	0.00	0.00	0.01	0.02
12	-0.04	0.04	-0.03	0.05	0.03	0.06	0.01	0.06	0.07	0.02
13	0.02	-0.02	0.03	0.02	0.00	-0.02	0.03	-0.01	0.03	0.03
14	0.11	-0.01	-0.01	0.02	-0.03	0.01	0.01	0.01	-0.04	0.01
17	0.01	0.01	0.00	-0.07	0.04	-0.02	0.00	-0.07	-0.03	-0.10
18	0.06	0.01	0.03	0.01	0.03	0.01	0.03	0.00	0.04	0.01
19	0.02	0.00	0.03	0.00	0.01	-0.03	0.04	0.02	0.01	-0.03
20	-0.03	-0.02	0.05	0.03	0.05	-0.02	0.03	0.05	0.02	0.01
21	-0.03	0.02	0.00	0.00	-0.02	0.01	-0.01	0.03	-0.03	0.02
22	0.00	-0.03	0.00	-0.03	0.00	-0.04	-0.01	-0.03	-0.01	-0.04
23	0.05	0.04	0.01	0.01	0.03	-0.01	0.04	0.00	-0.10	-0.02
24	0.11	-0.04	0.07	0.00	-0.02	0.01	0.02	0.01	0.02	-0.03
25	0.00	-0.03	-0.03	-0.03	-0.02	-0.03	-0.02	-0.01	-0.01	-0.01
26	0.03	-0.05	0.03	-0.02	0.04	-0.01	0.02	-0.06	0.01	-0.05
Global	0.04	0.00	0.02	0.01	0.02	0.01	0.01	0.02	0.02	0.01

Table S1.9 As Table S1.4, but for the low tails of maximum and minimum temperature anomalies (T_{\max_L} and T_{\min_L} , respectively) of the shape parameter.

	HadGHCND		ERA-Interim		NCEP2		JRA-55		MERRA-2	
Region	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}	T_{\max_L}	T_{\min_L}
1	-0.04	-0.01	-0.04	-0.05	0.02	0.05	-0.05	-0.06	-0.04	-0.03
2	-0.02	0.05	0.04	0.01	0.01	-0.02	0.00	-0.01	0.03	0.02
3	0.01	0.02	0.04	-0.01	0.02	0.01	-0.01	-0.04	0.01	0.02
4	-0.01	0.01	-0.01	0.03	0.03	0.05	-0.05	-0.08	0.00	0.11
5	-0.02	0.03	-0.01	-0.01	-0.02	0.05	-0.01	-0.03	0.02	0.06
6	-0.02	-0.06	-0.07	-0.06	0.02	-0.13	0.00	-0.08	0.01	-0.08
10	0.01	0.02	0.02	0.01	-0.03	-0.03	0.02	0.06	0.01	0.05
11	-0.11	-0.05	-0.06	-0.01	-0.08	0.01	-0.06	-0.07	-0.08	-0.02
12	0.02	-0.03	0.01	-0.04	0.06	0.02	0.01	0.03	-0.01	-0.01
13	-0.03	-0.04	-0.02	-0.02	-0.06	-0.07	-0.02	-0.06	-0.02	-0.01
14	-0.01	0.05	-0.01	-0.05	-0.03	-0.04	0.05	-0.02	0.08	-0.02
17	0.03	-0.12	0.02	-0.01	0.00	0.04	0.04	-0.01	0.05	-0.02
18	-0.05	-0.04	-0.01	0.01	-0.01	-0.01	-0.01	-0.01	-0.04	-0.03
19	0.03	-0.02	0.06	0.00	0.00	0.05	0.03	0.02	0.03	0.00
20	-0.02	-0.07	0.01	-0.01	0.00	0.04	-0.03	-0.08	0.02	-0.02
21	-0.06	-0.08	-0.02	-0.07	-0.01	0.04	-0.06	-0.06	-0.03	-0.02
22	-0.01	0.02	0.01	0.04	-0.04	-0.02	-0.01	0.02	-0.01	0.02
23	0.04	0.07	0.03	0.00	0.17	0.00	-0.03	0.01	0.04	0.01
24	-0.02	-0.02	-0.01	-0.02	-0.03	-0.06	0.00	-0.04	-0.04	-0.04
25	0.02	0.00	0.00	-0.04	0.01	-0.01	-0.04	0.04	-0.01	-0.03
26	0.05	0.00	0.03	-0.02	0.06	0.03	0.04	0.00	0.01	0.03
Global	-0.02	-0.01	0.02	0.01	0.00	0.01	-0.03	-0.02	0.02	0.01

S.2 SUPPLEMENTARY MATERIAL FOR CHAPTER 3

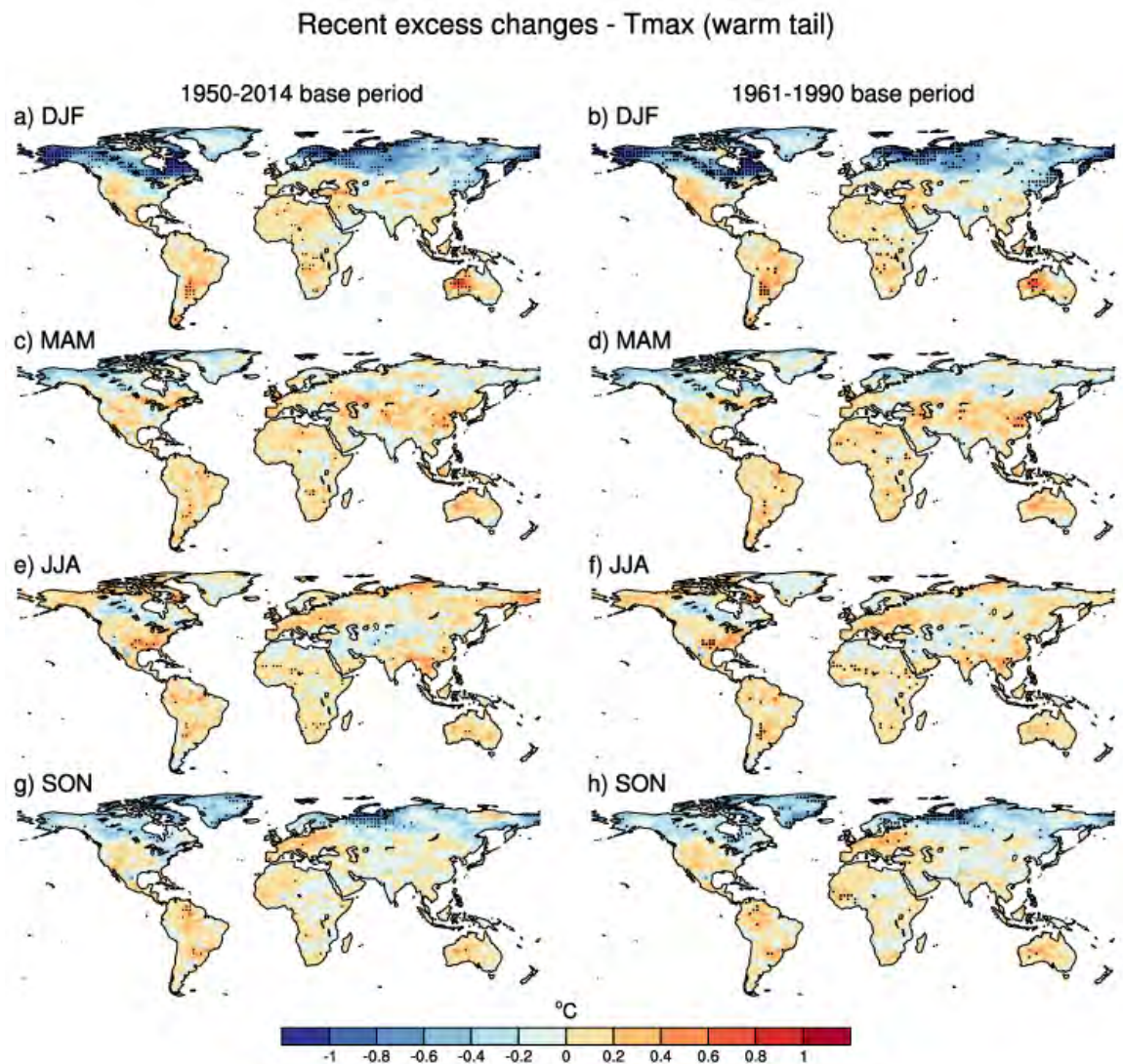


Figure S2.1 Comparison between choice of base period used for temperature anomaly calculation for recent excess changes in hot extremes (1982-2014 – 1950-1981) in the CMIP5 multi-model mean for DJF (a,b), MAM (c,d), JJA (e,f) and SON (g,h). The panels on the left show excess changes calculated from temperature anomalies relative to the entire period of analysis (1950-2014). The panels on the right show excess changes calculated from temperature anomalies relative to the 1961-1990 base period. Stippling indicates grid cells where at least 20 out of 26 models agree on the sign of excess change.

Recent excess changes - Tmin (cold tail)

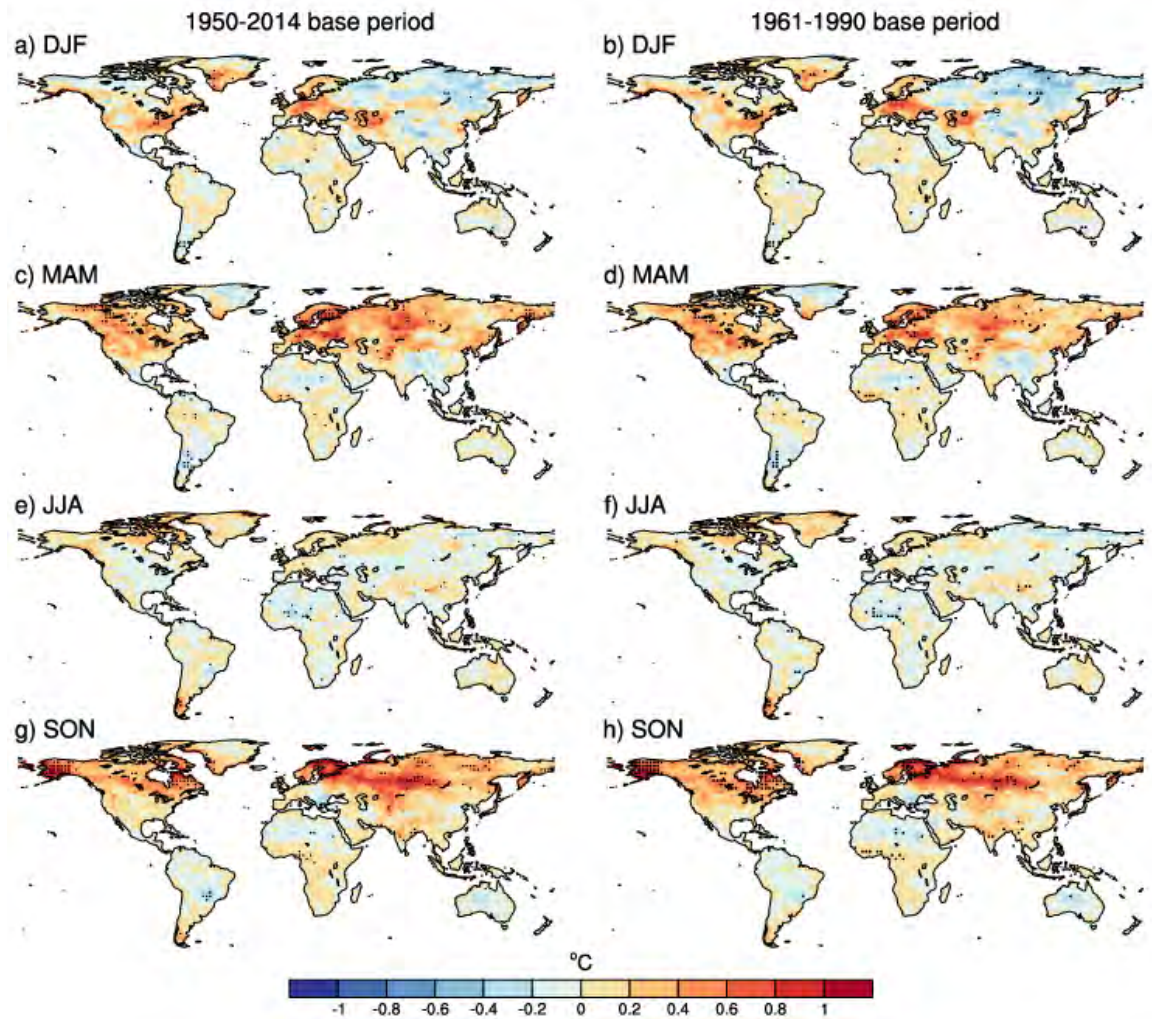


Figure S2.2 As Figure S2.1, but for extremes in the cold tails of the distribution relative to daily minimum temperatures.

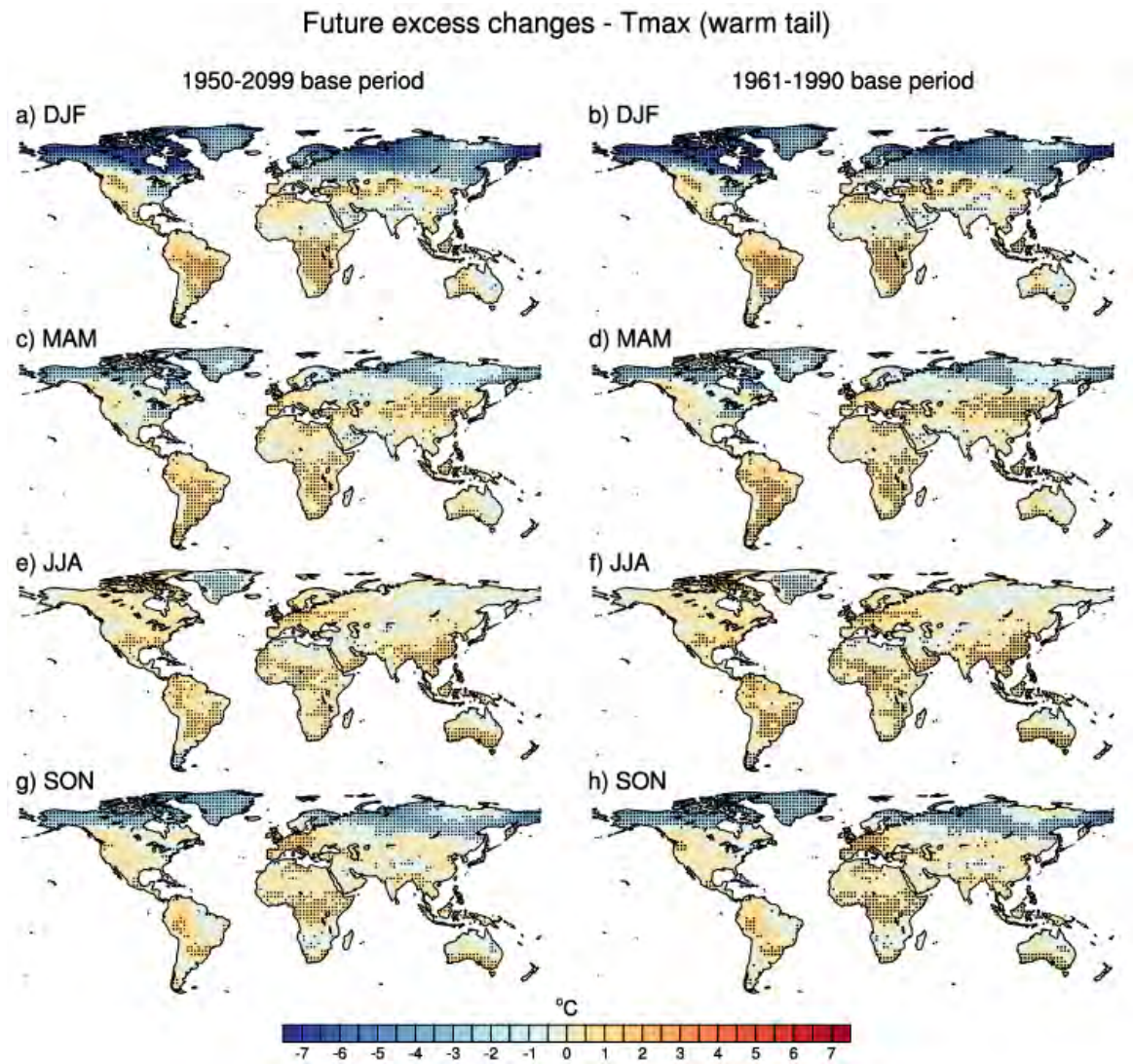


Figure S2.3 As Figure S2.1, but for future excess changes.

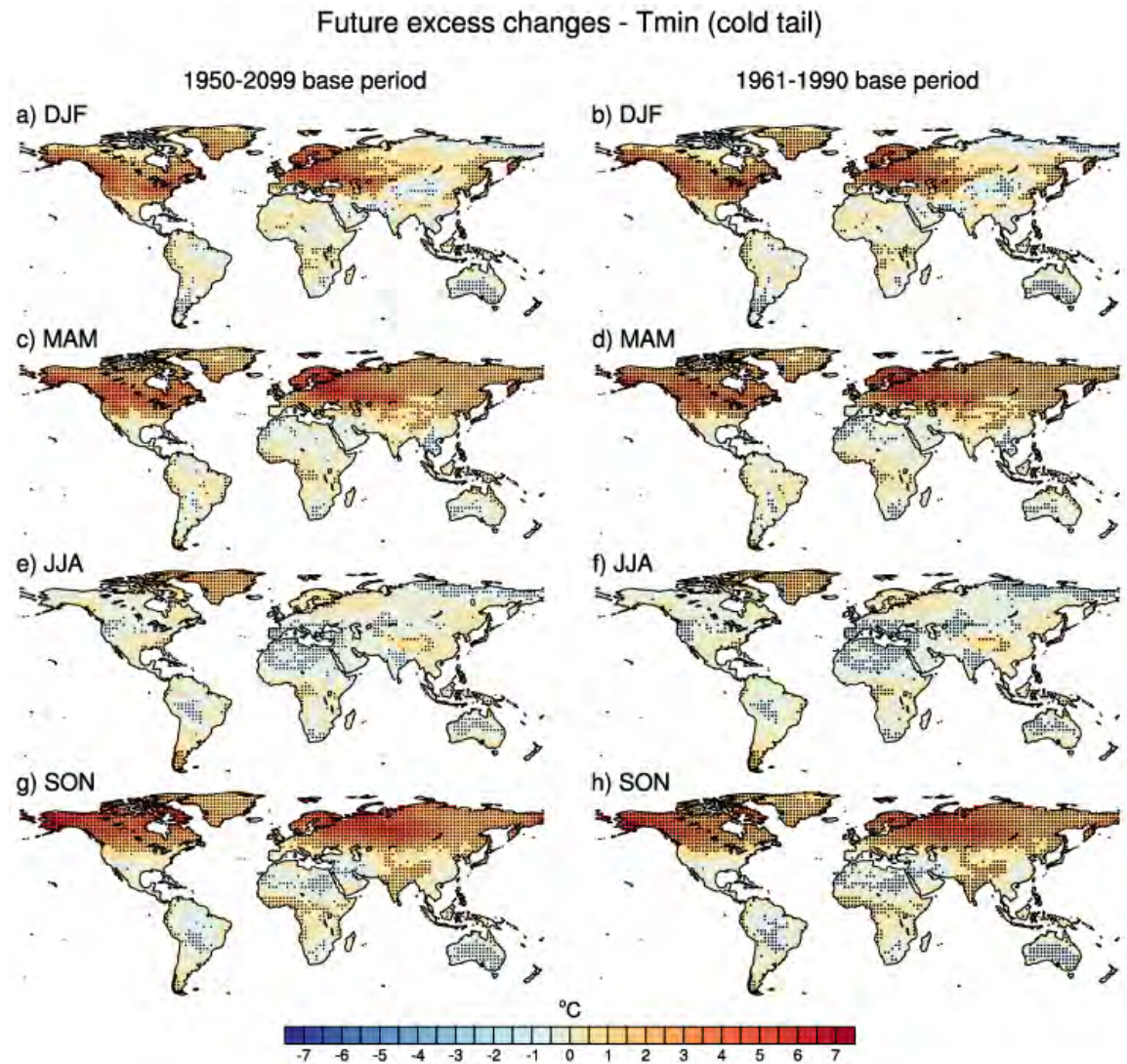


Figure S2.4 As Figure S2.3, but for extremes in the cold tails of the distribution relative to daily minimum temperatures.

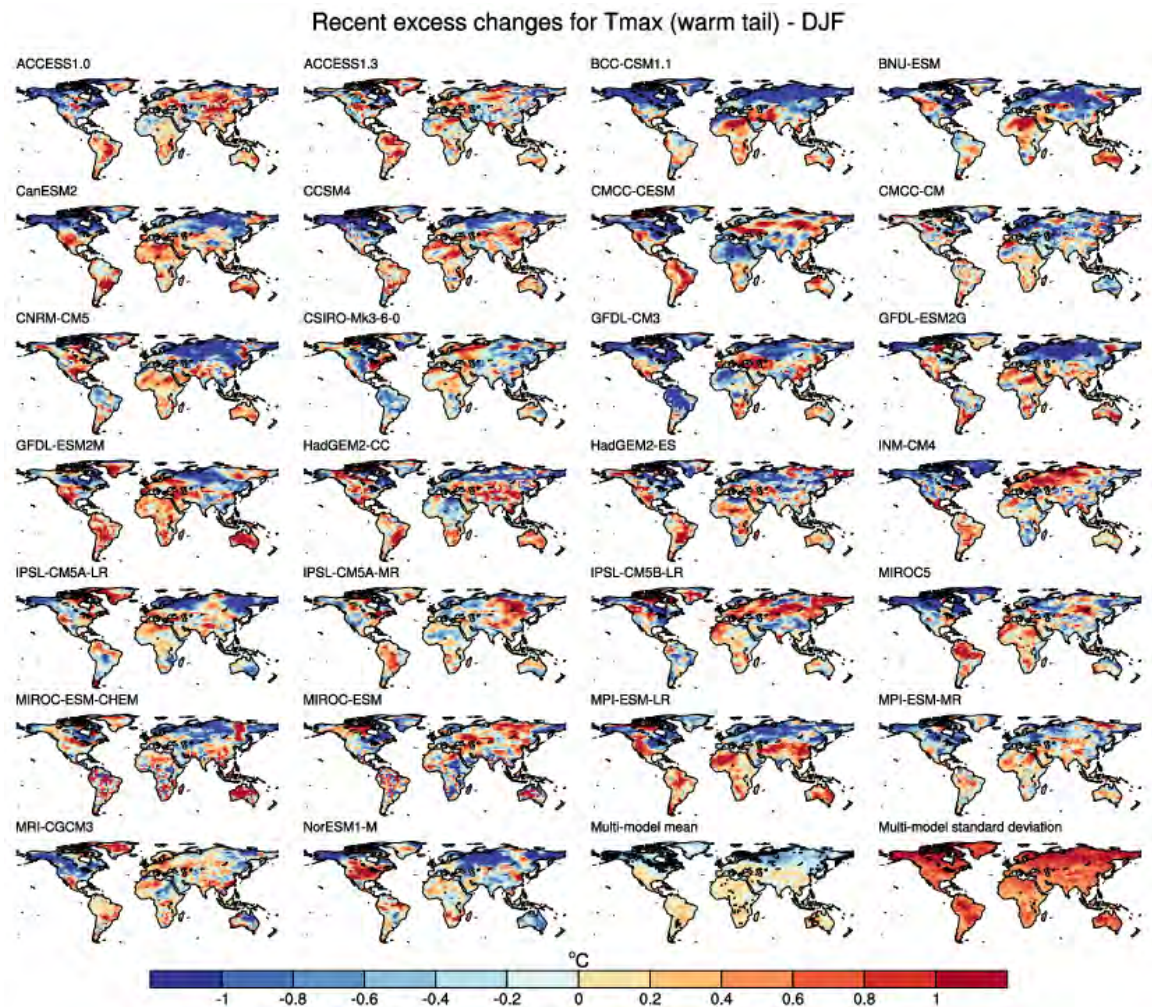


Figure S2.5 Recent excess changes in the individual CMIP5 models used to calculate the multi-model mean shown in Figures 3.2 and 3.4, for extremes in the warm tail using daily maximum temperatures for December – February.

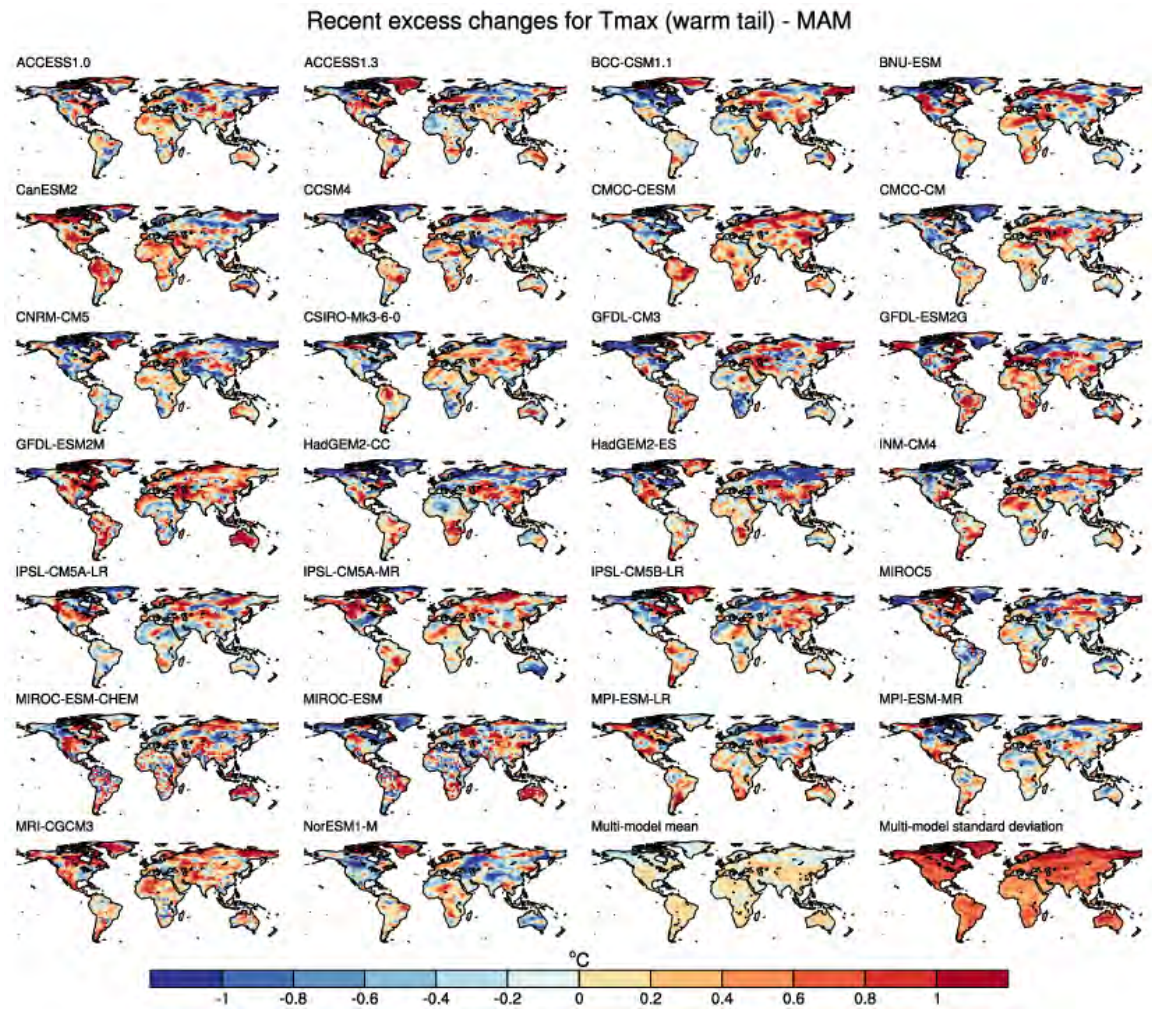


Figure S2.6 As Figure S2.5, but for March – May.

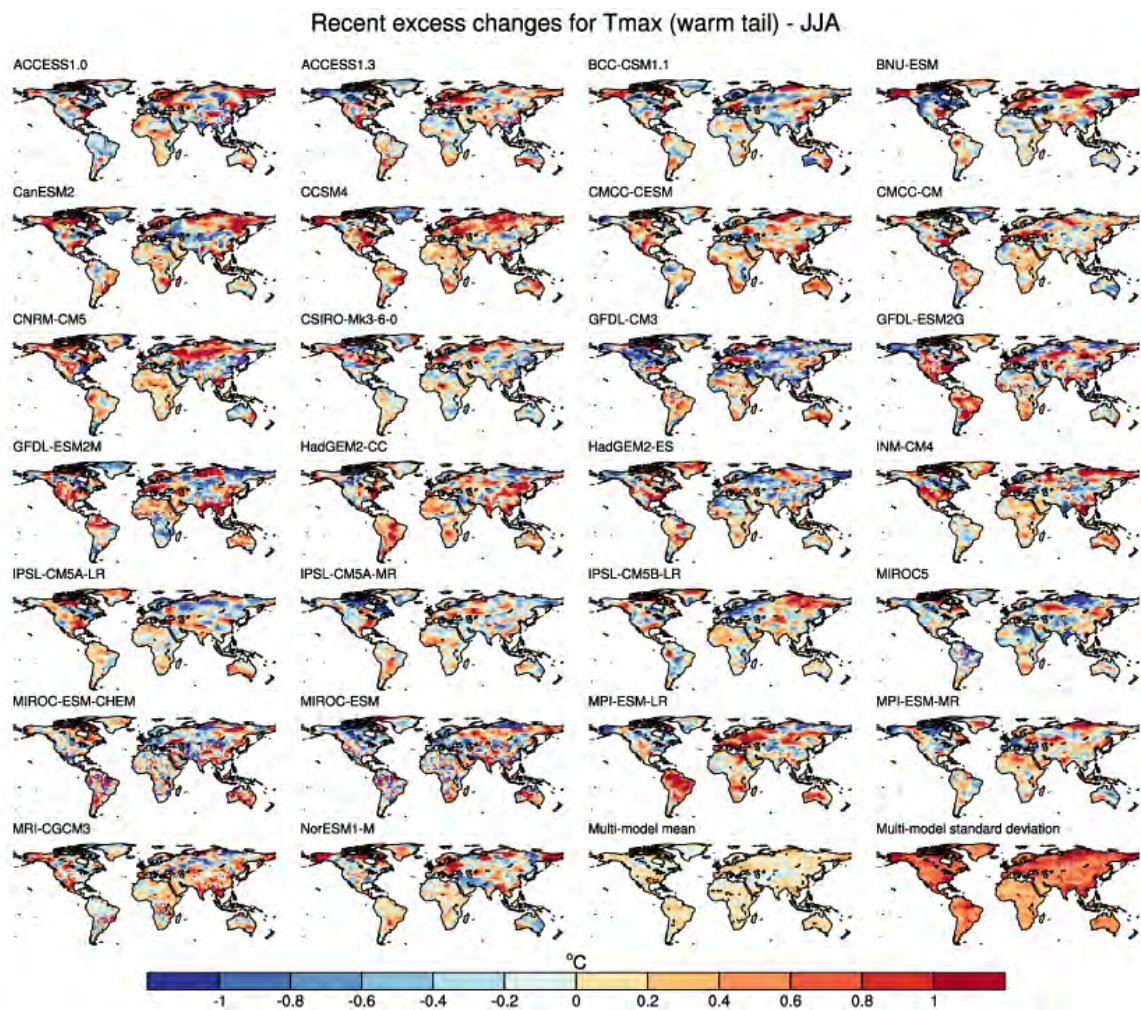


Figure S2.7 As Figure S2.5, but for June – August.

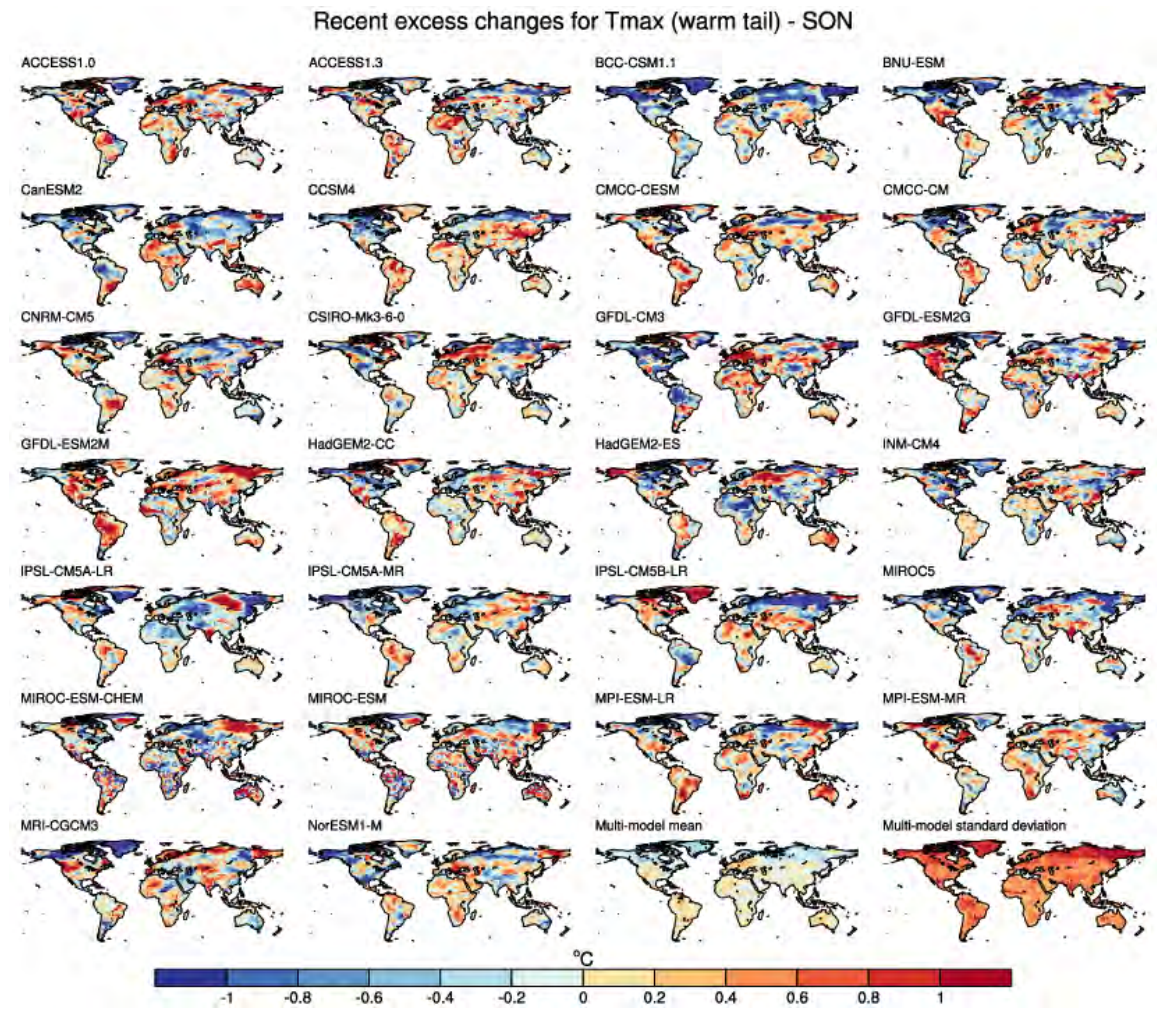


Figure S2.8 As Figure S2.5, but for September – November.

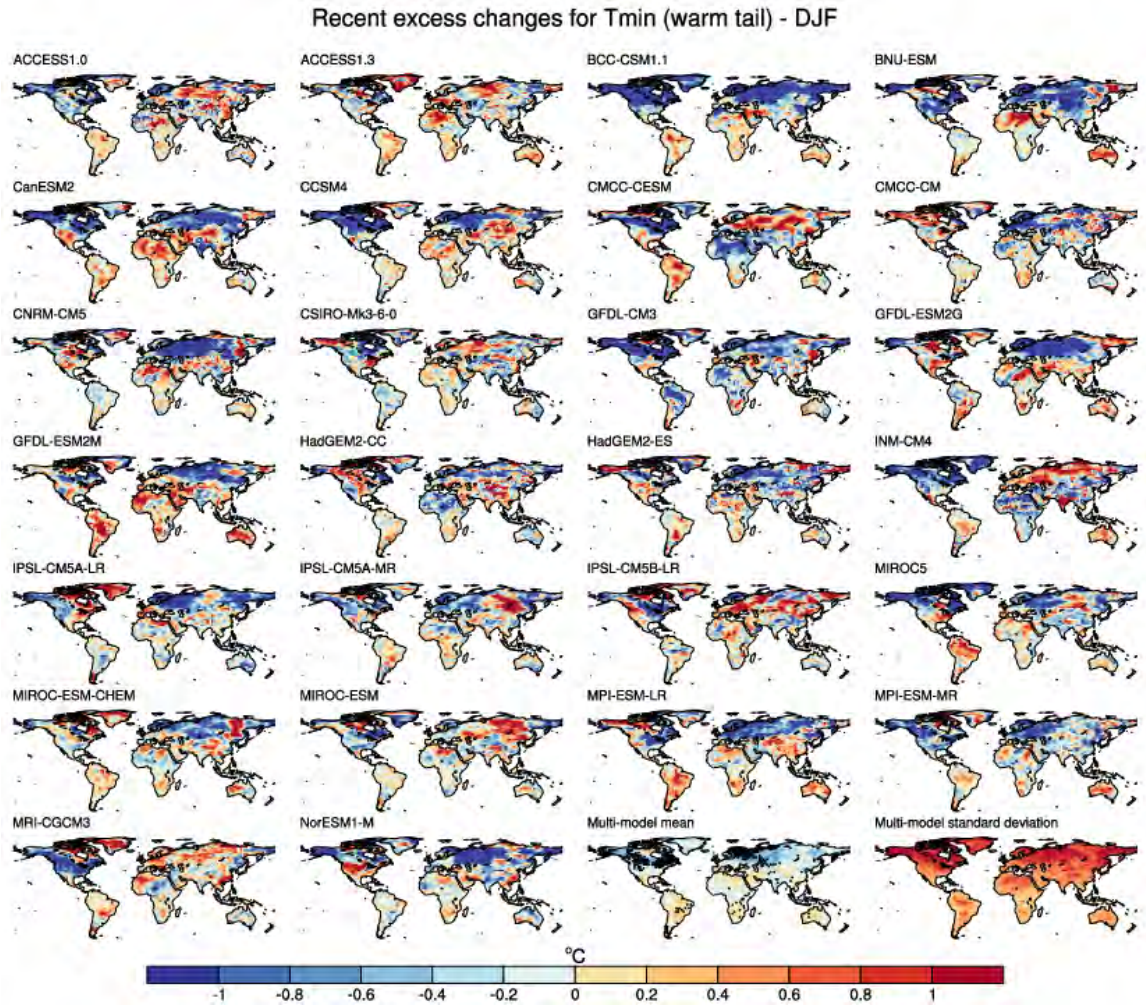


Figure S2.9 As Figure S2.5, but for the warm tails of daily minimum temperature.

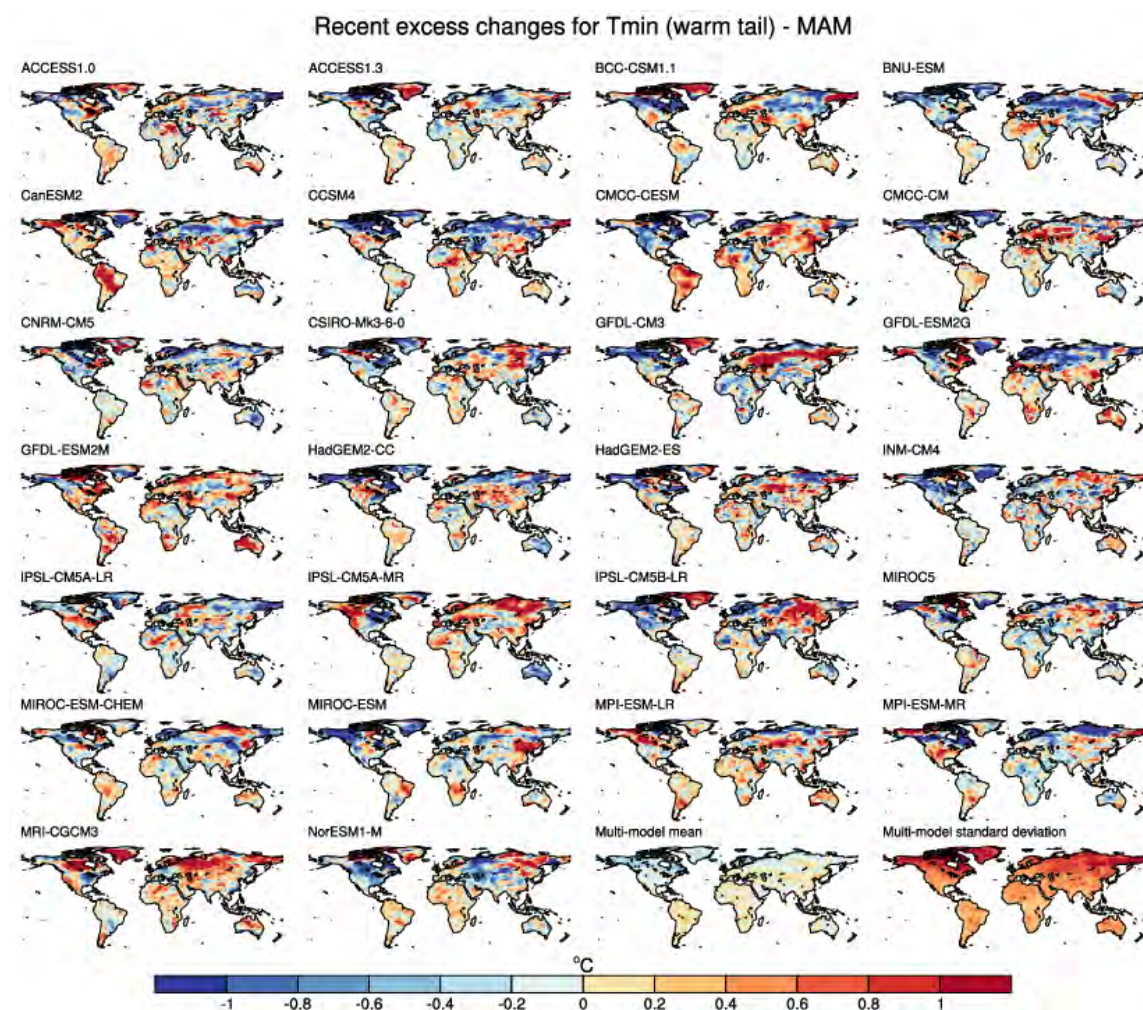


Figure S2.10 As Figure S2.9, but for March – May.

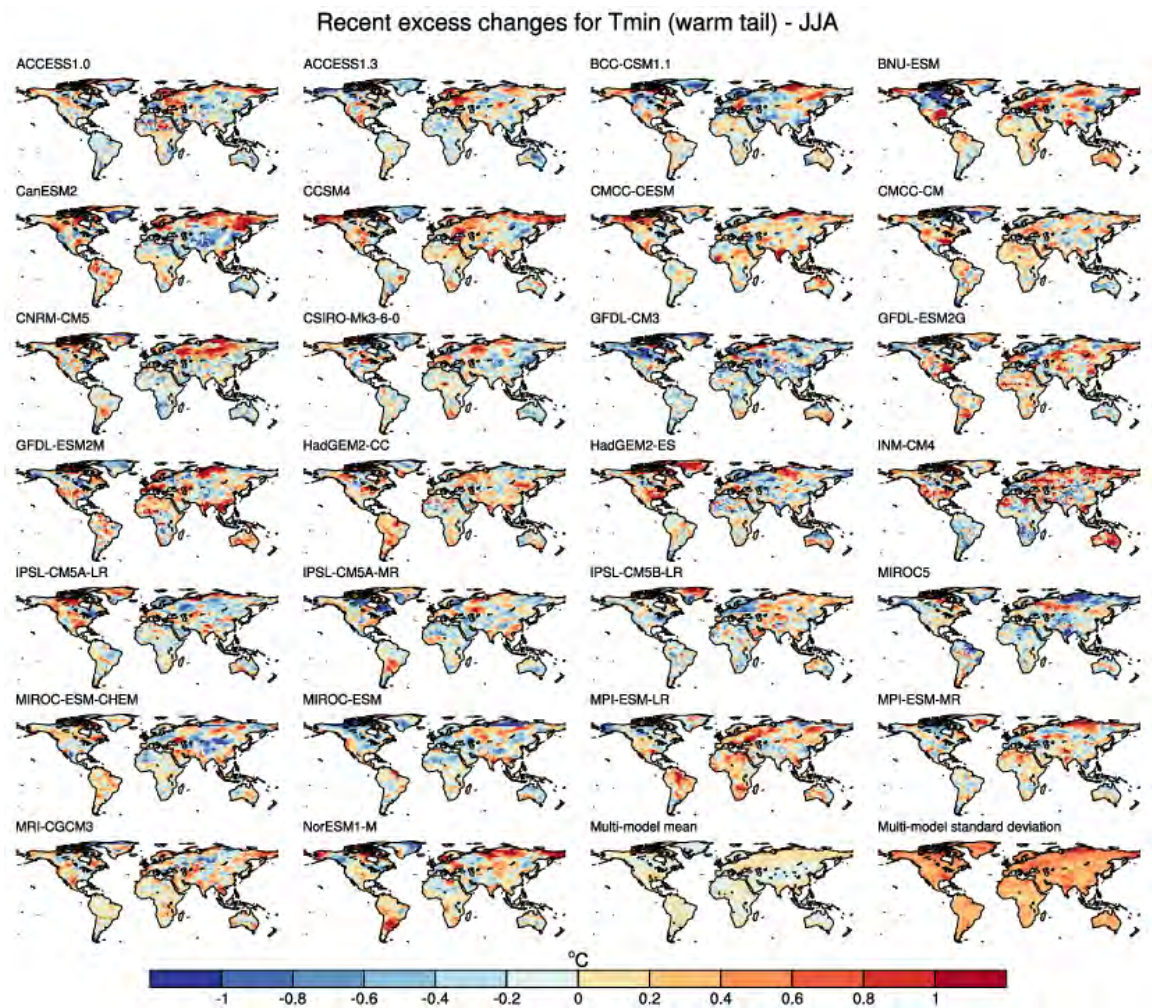


Figure S2.11 As Figure S2.9, but for June – August.

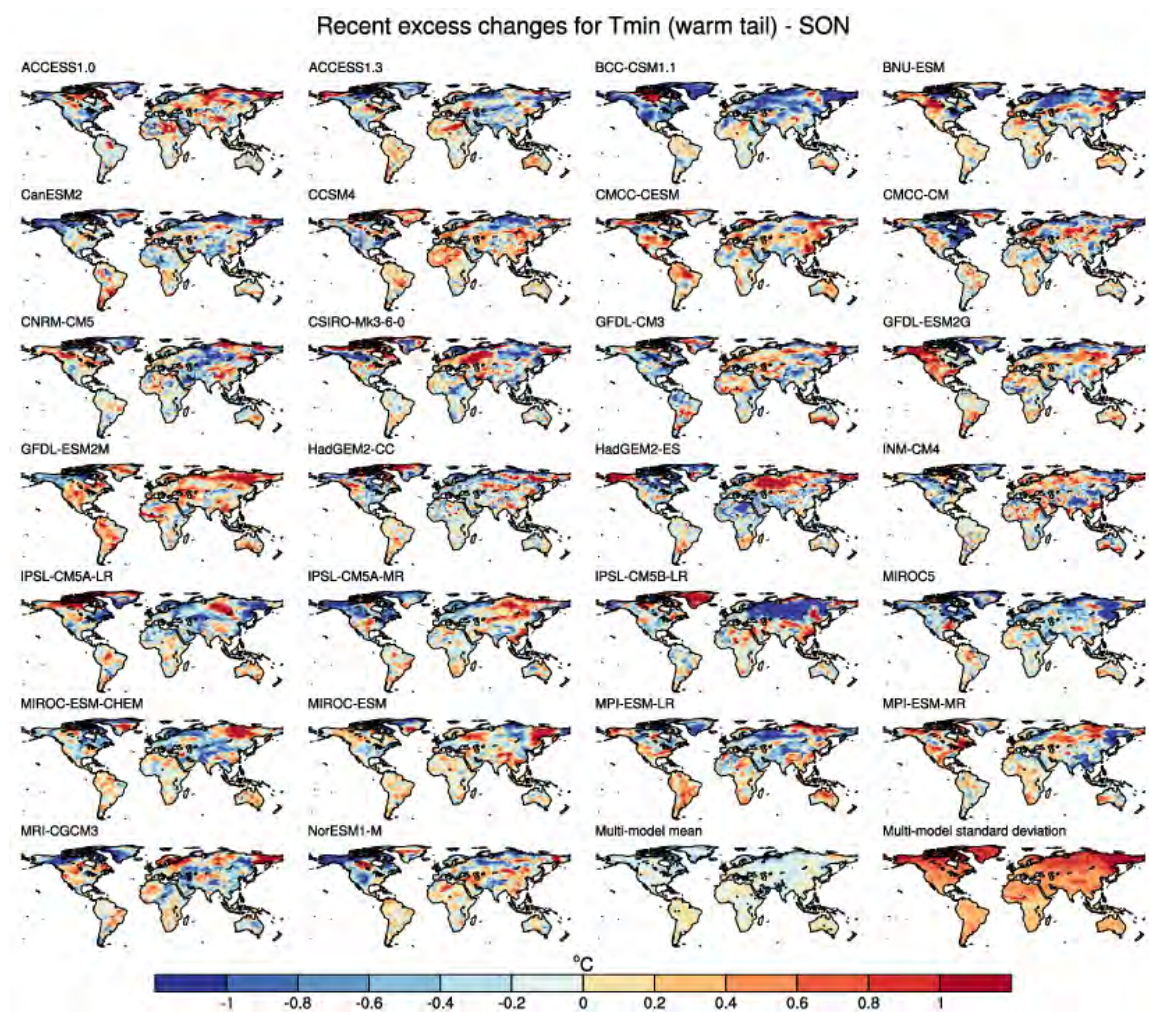


Figure S2.12 As Figure S2.9, but for September – November.

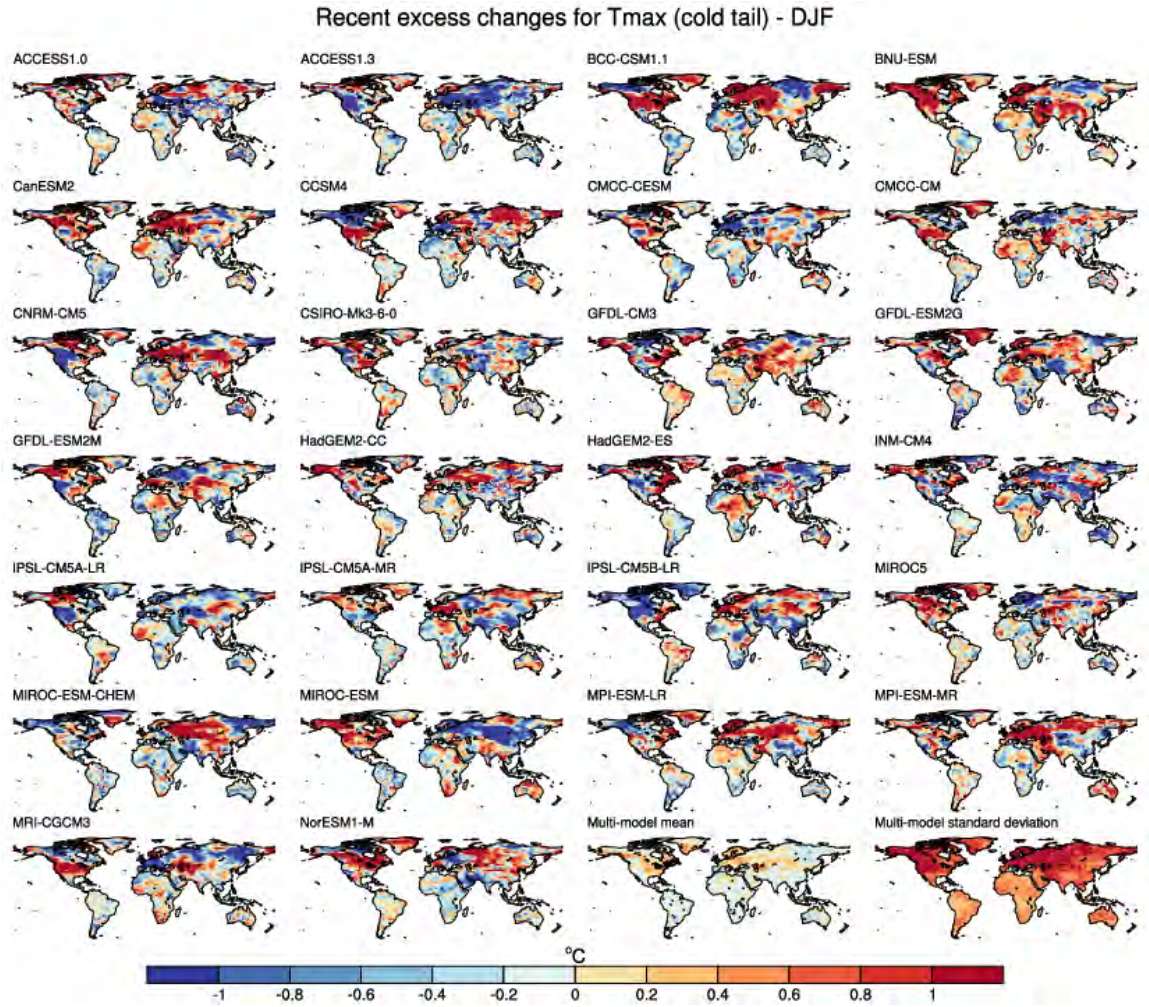


Figure S2.13 As Figure S2.5, but for the cold tails of daily maximum temperature.

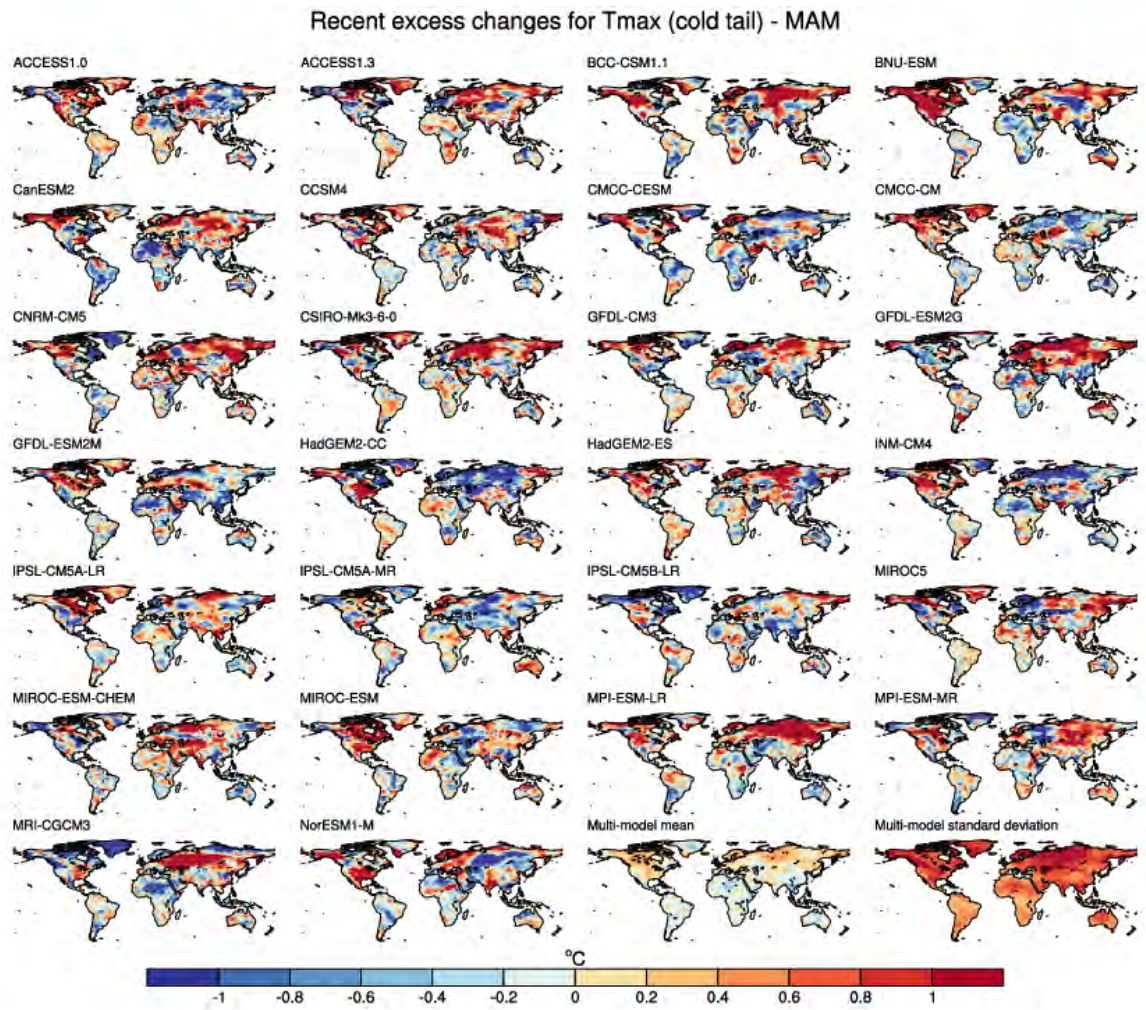


Figure S2.14 As Figure S2.13, but for March – May.

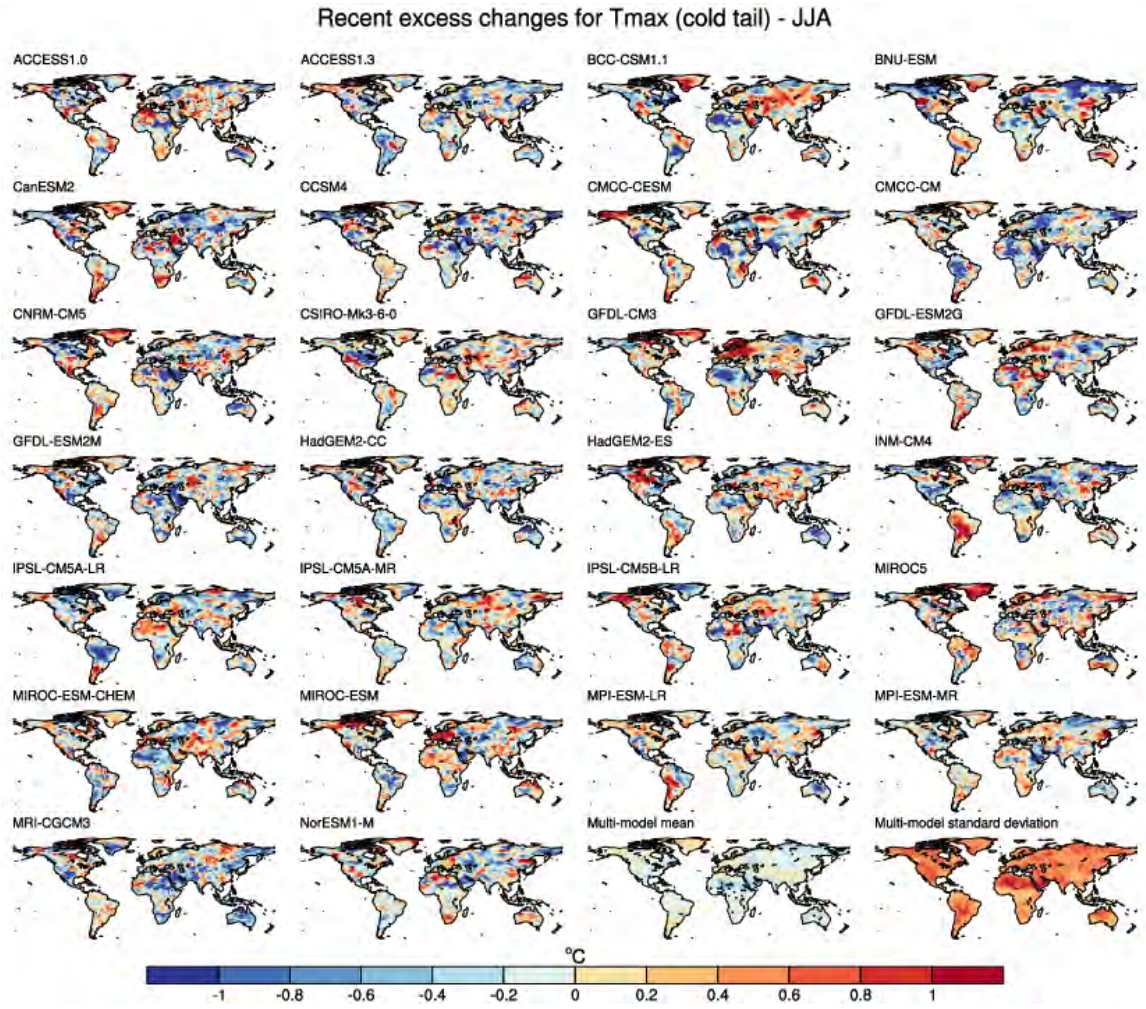


Figure S2.15 As Figure S2.13, but for June – August.

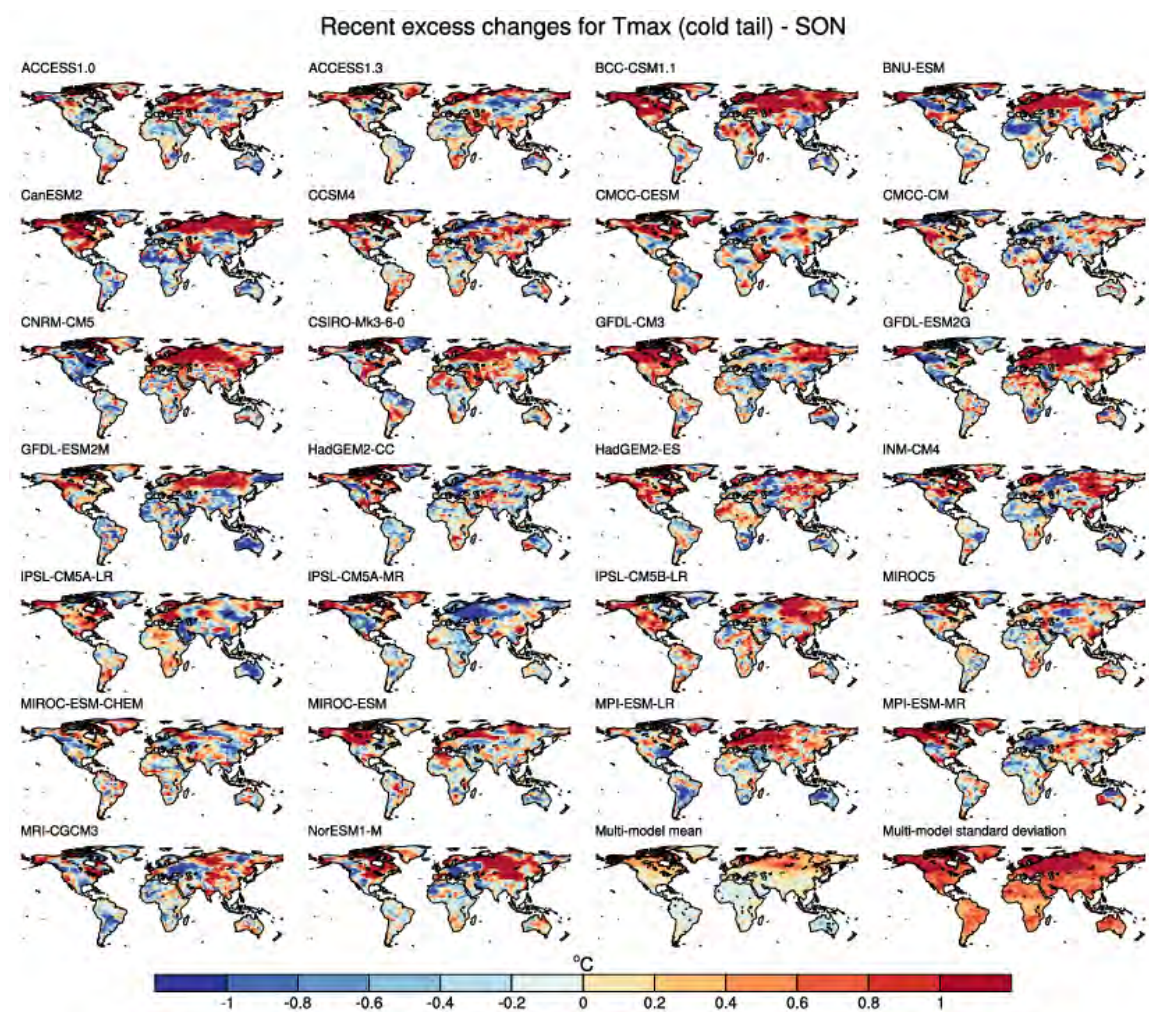


Figure S2.16 As Figure S2.13, but for September – November.

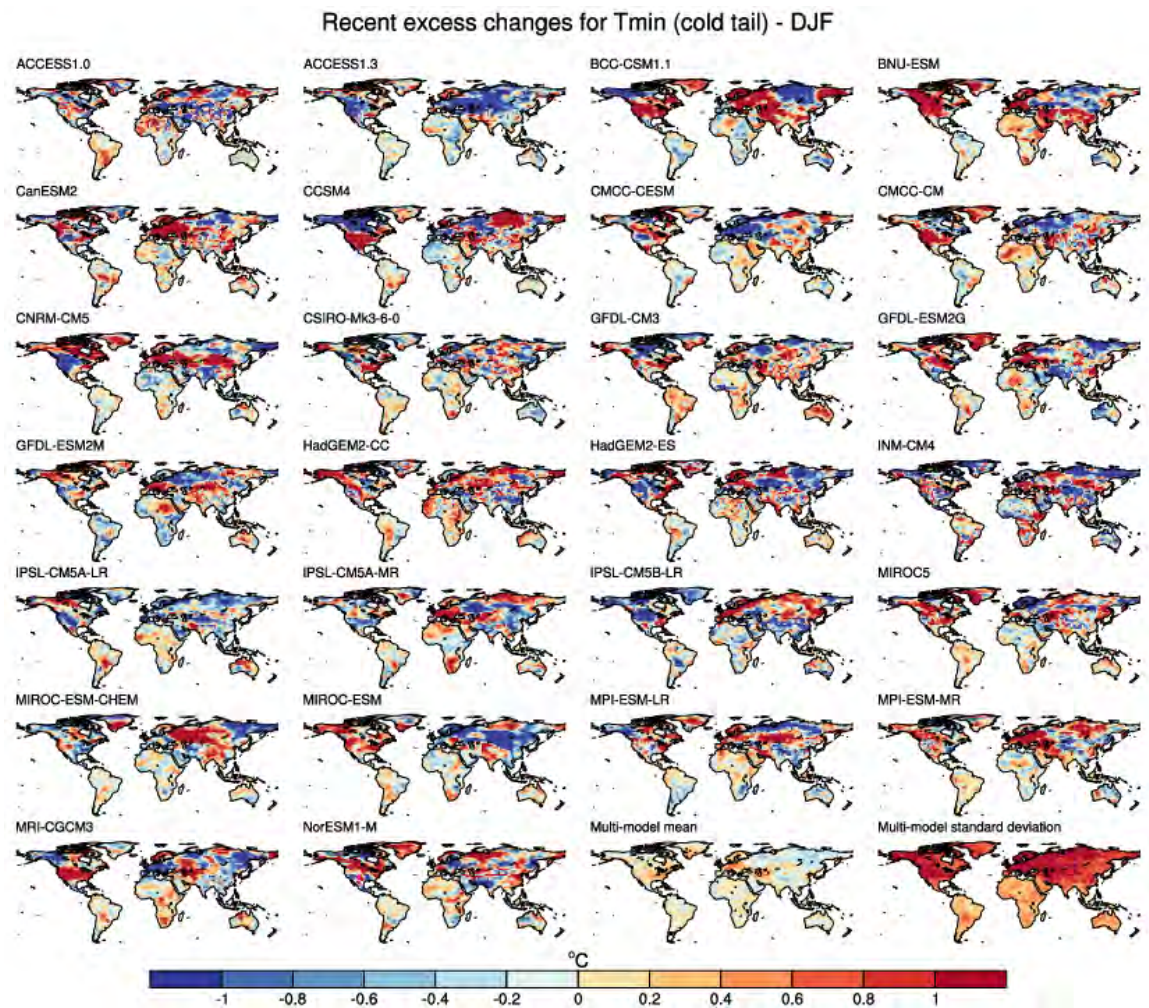


Figure S2.17 As Figure S2.5, but for the cold tails of daily minimum temperature.

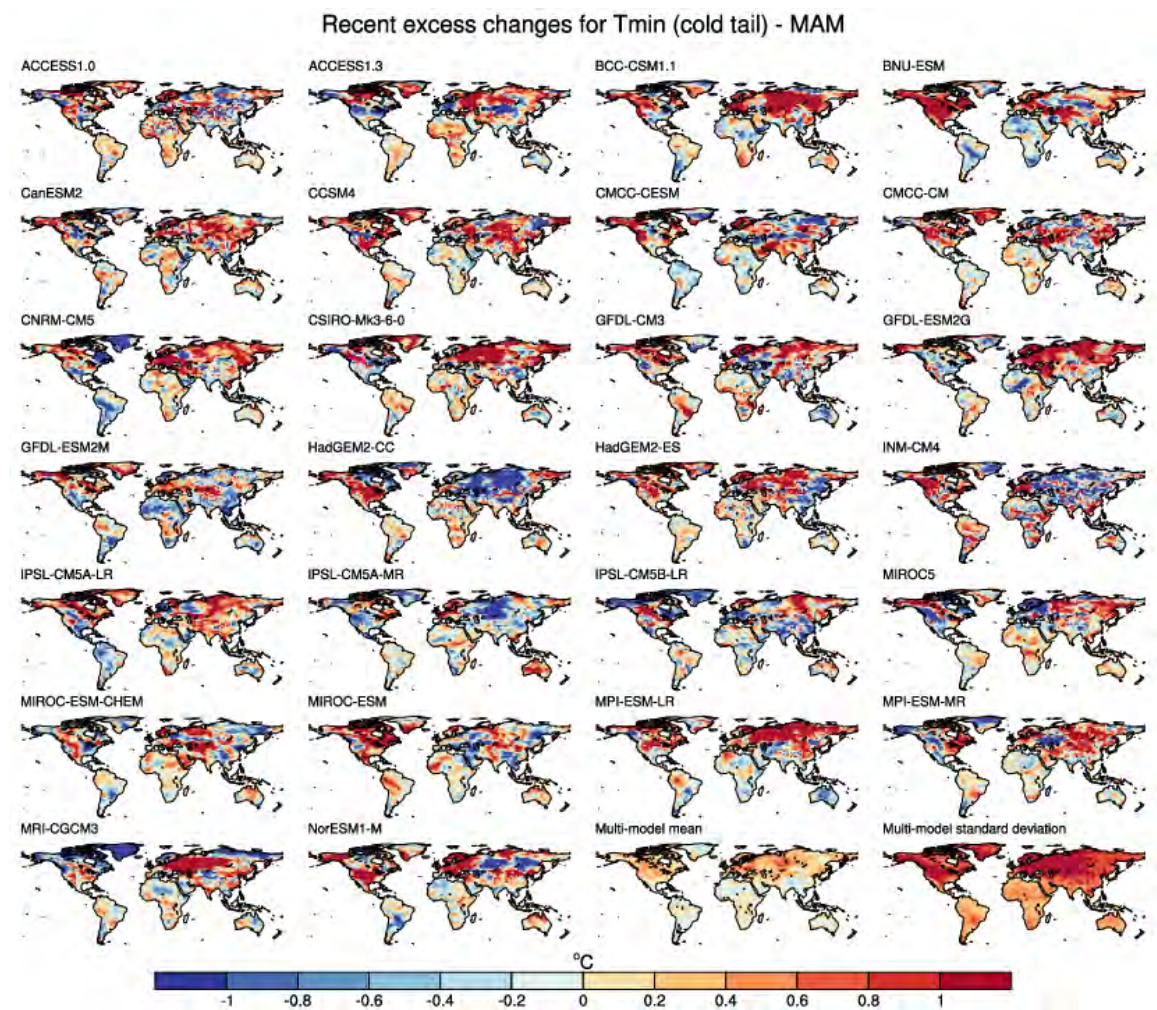


Figure S2.18 As Figure S2.17, but for March – May.

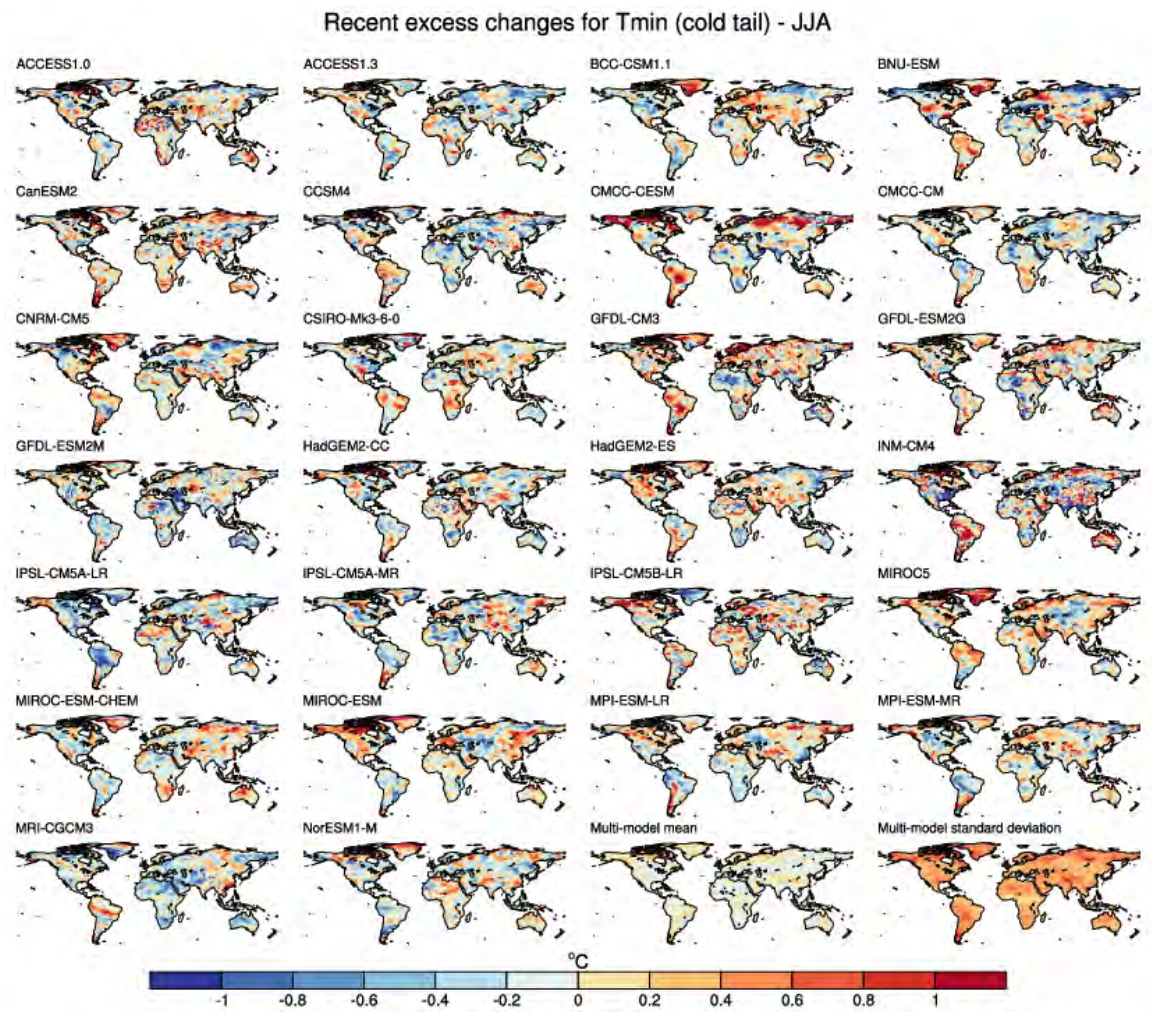


Figure S2.19 As Figure S2.17, but for June – August.

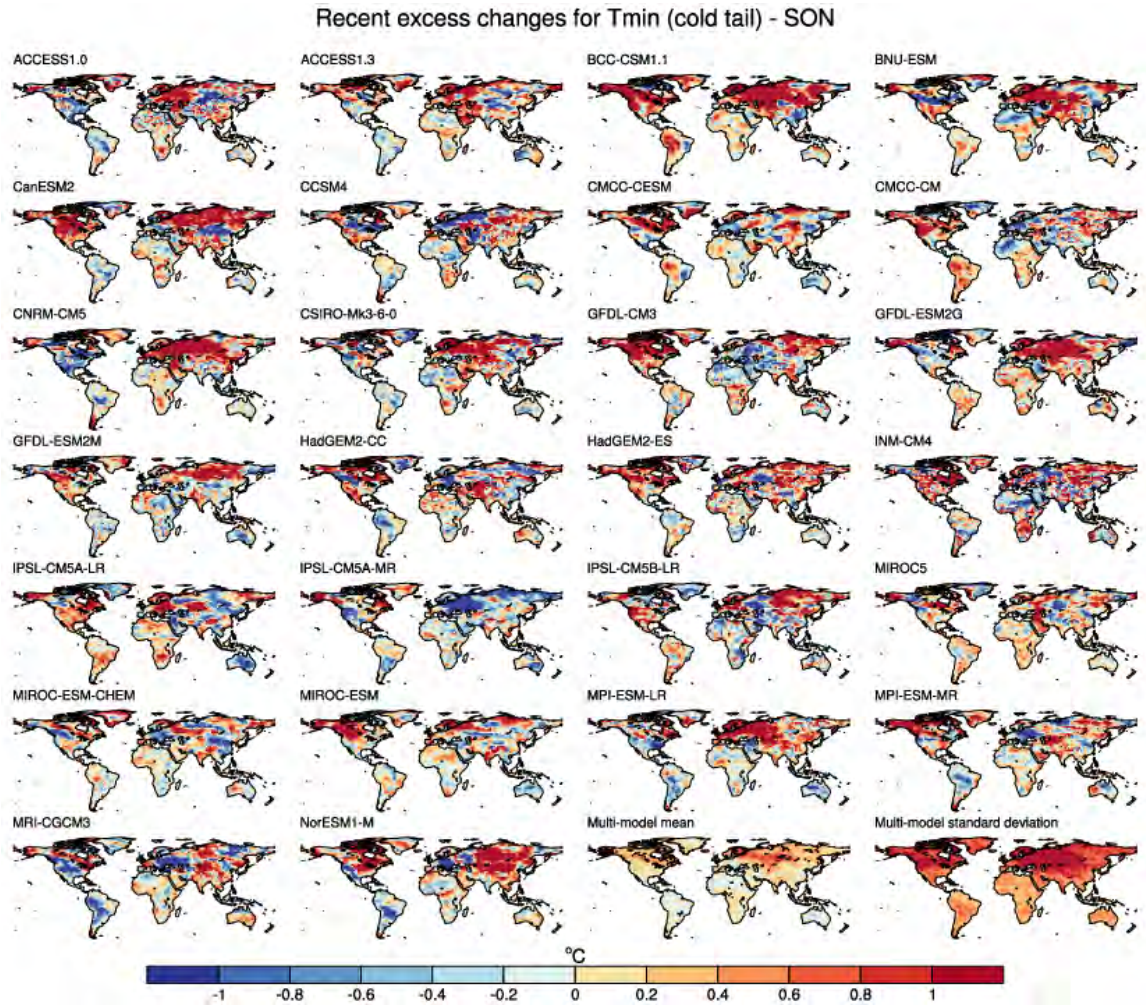


Figure S2.20 As Figure S2.17, but for September – November.

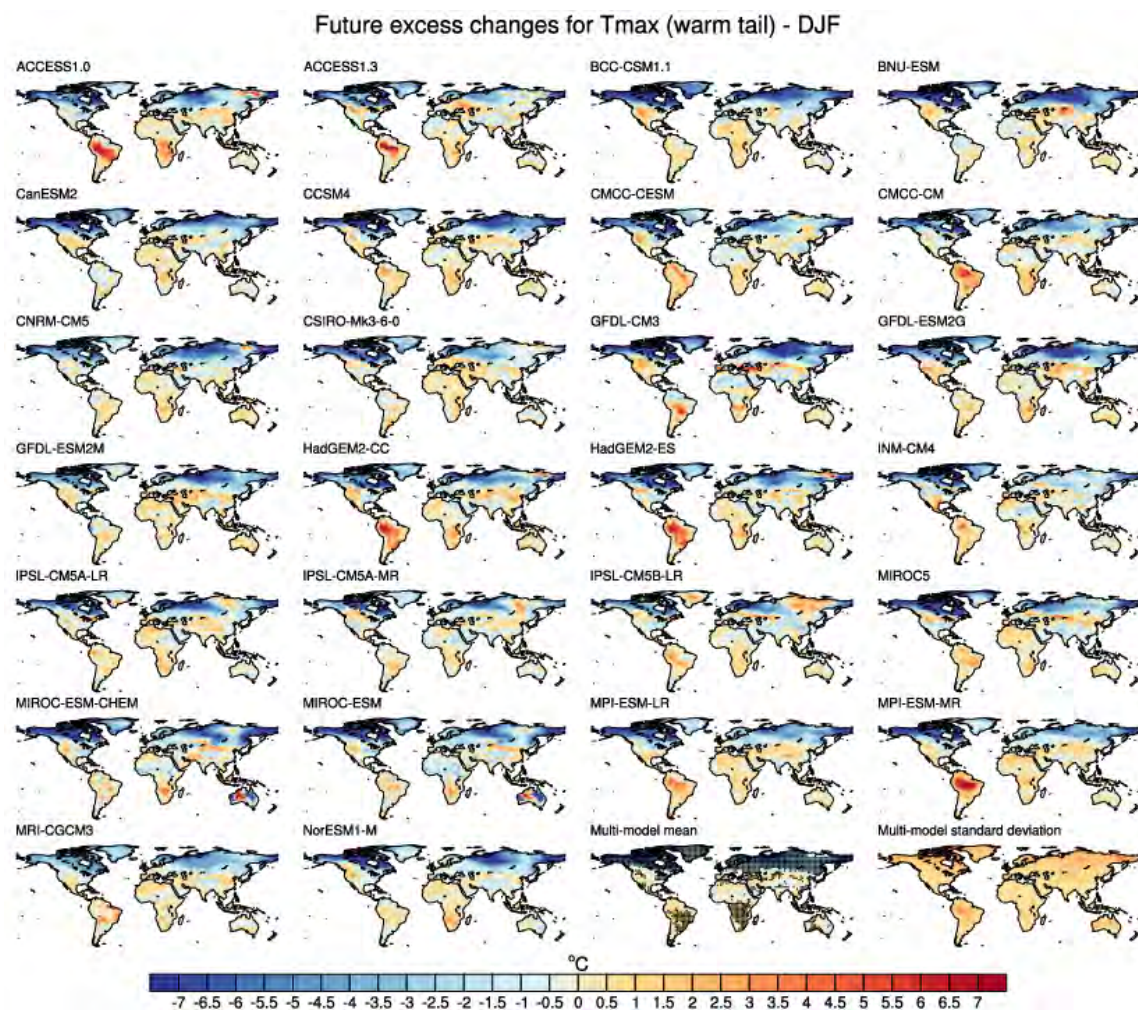


Figure S2.21 As Figure S2.5, but for future excess changes.

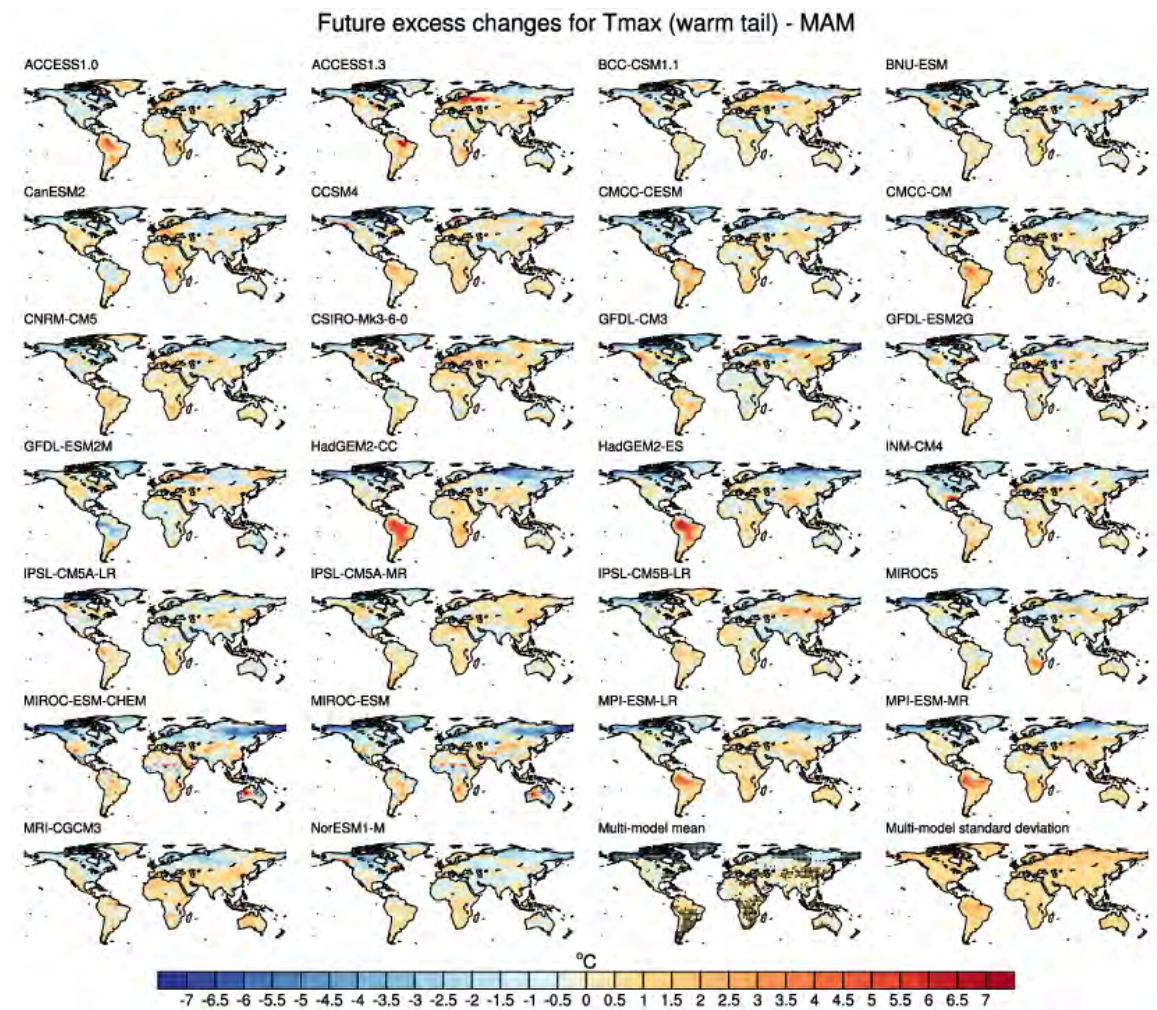


Figure S2.22 As Figure S2.21, but for March – May.

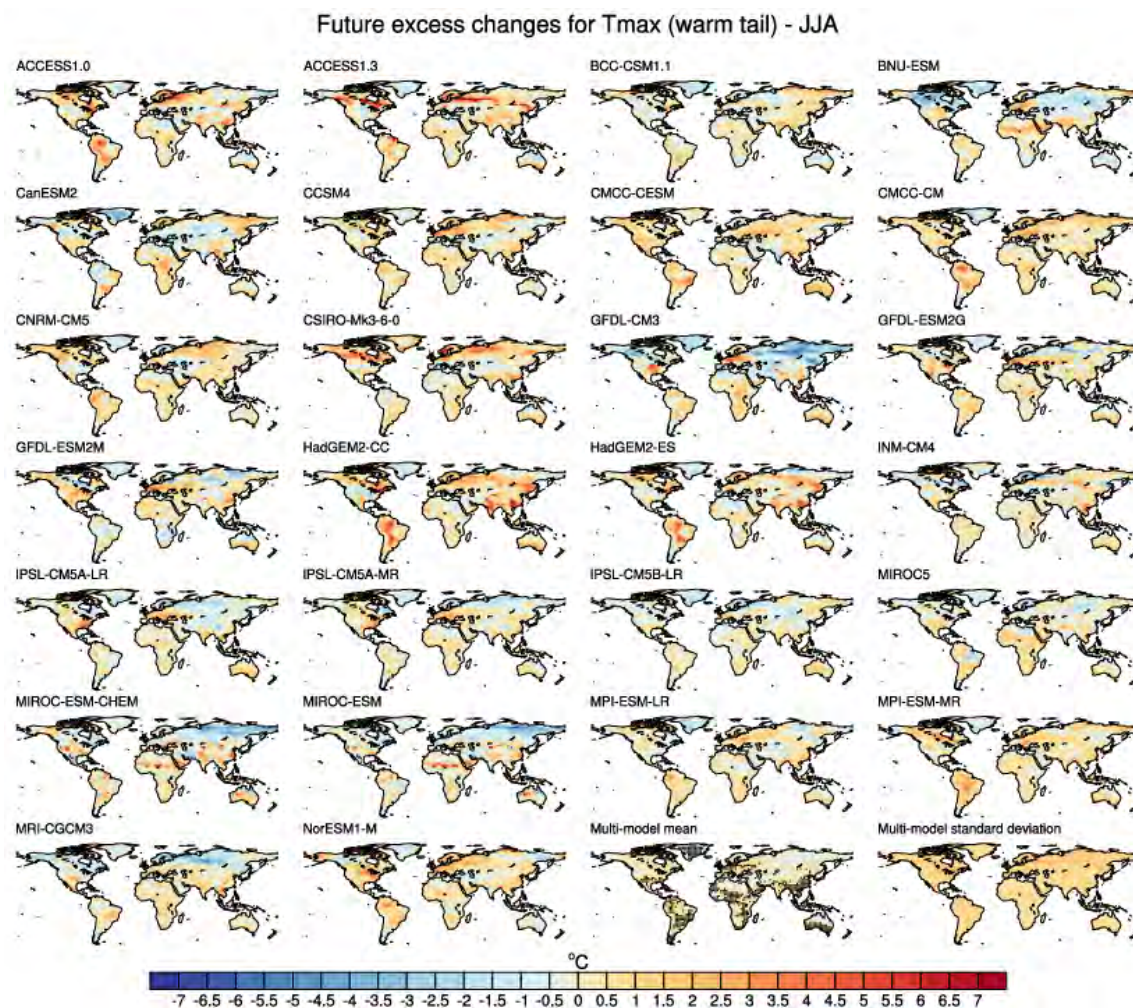


Figure S2.23 As Figure S2.21, but for June – August.

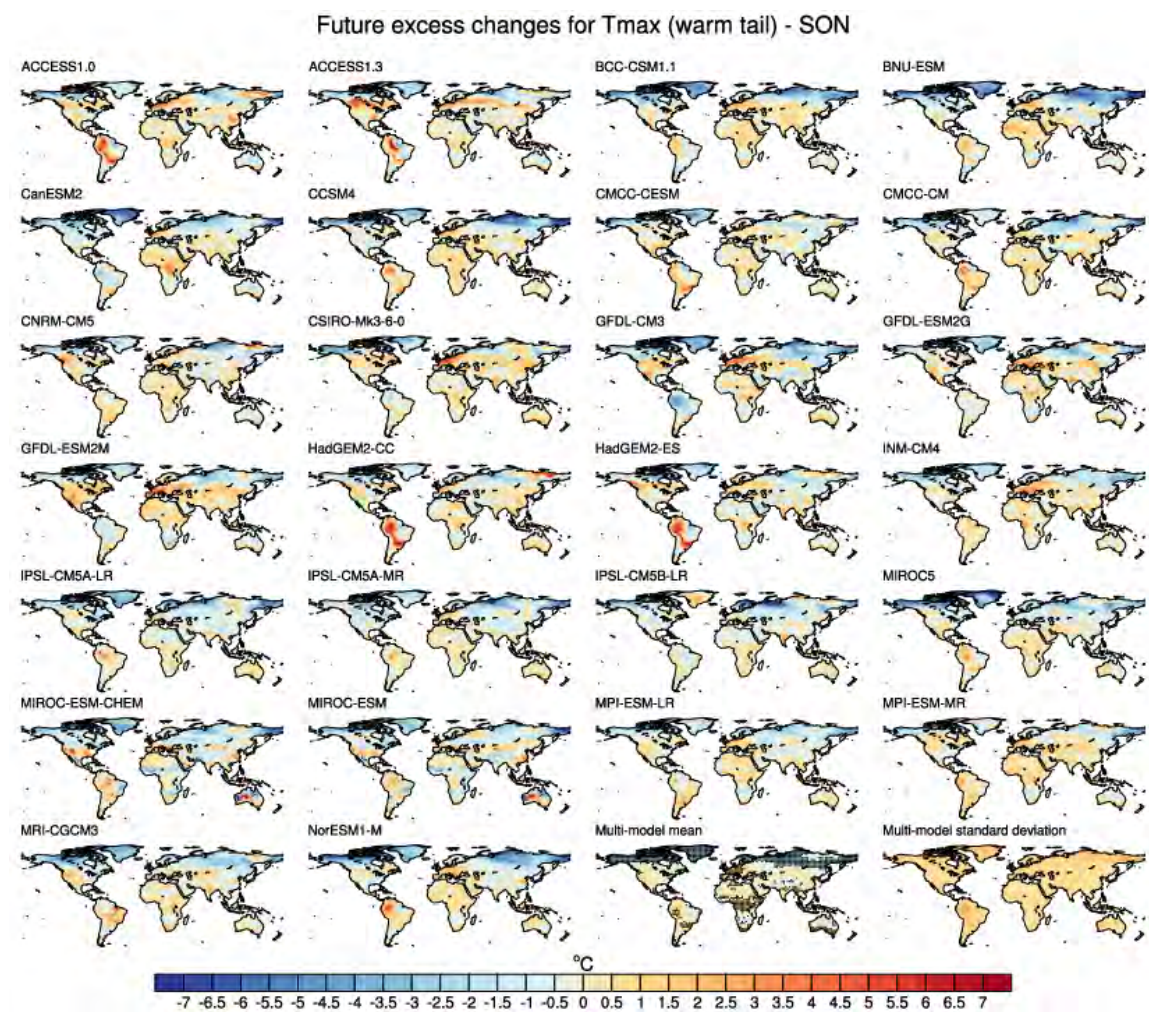


Figure S2.24 As Figure S2.21, but for September – November.

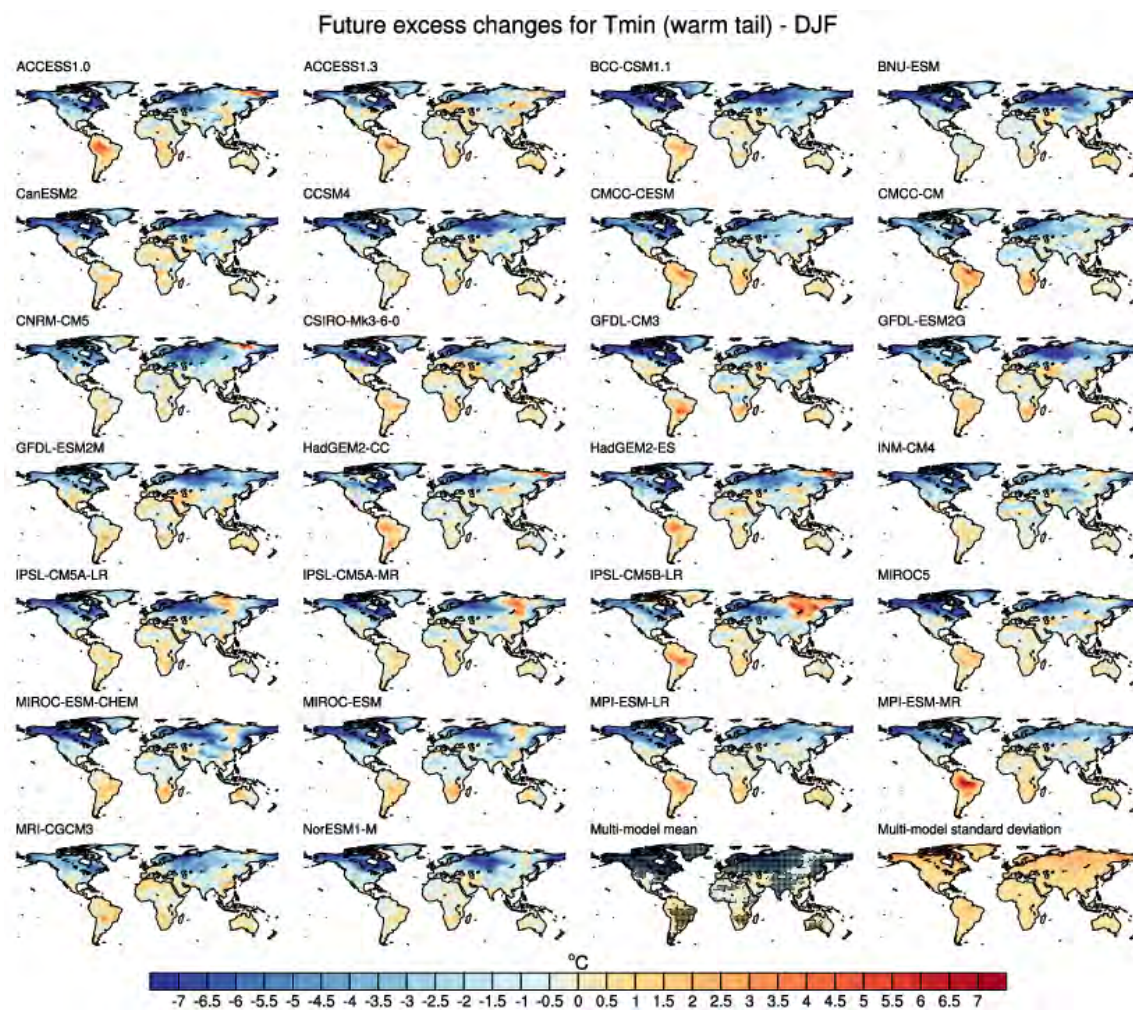


Figure S2.25 As Figure S2.21, but for the warm tails of daily minimum temperature.

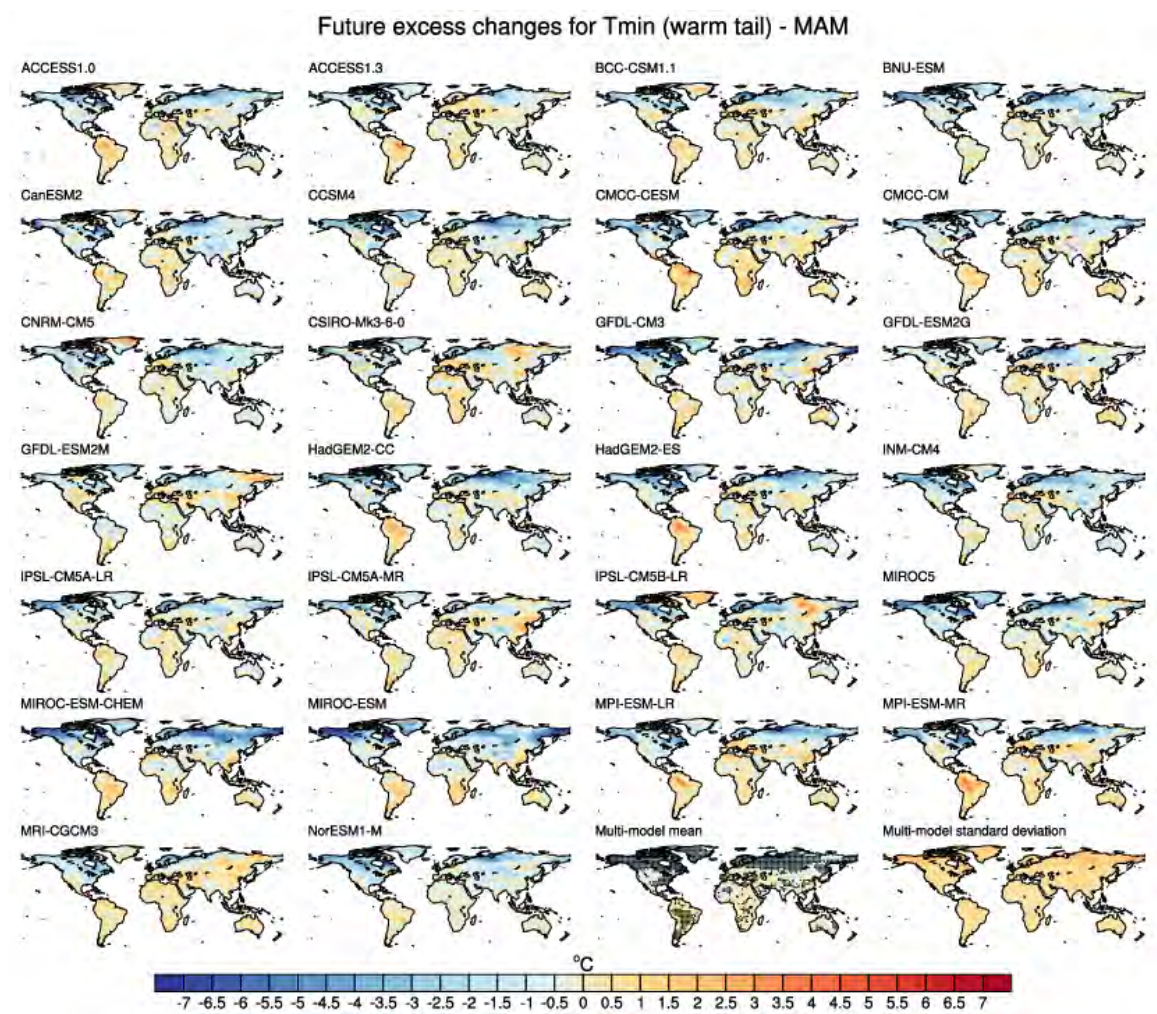


Figure 2.26 As Figure S2.25, but for March – May.

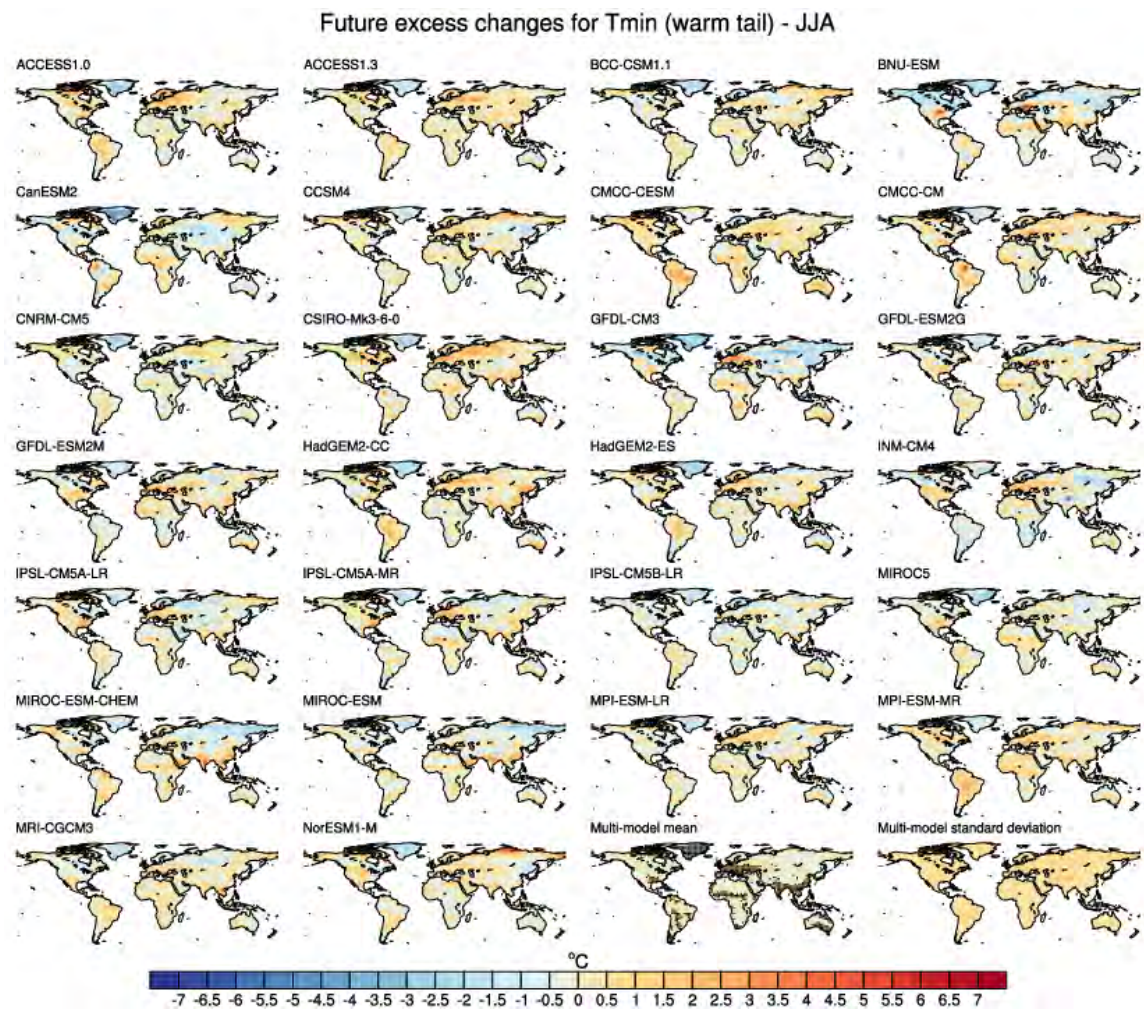


Figure S2.27 As Figure S2.25, but for June – August.

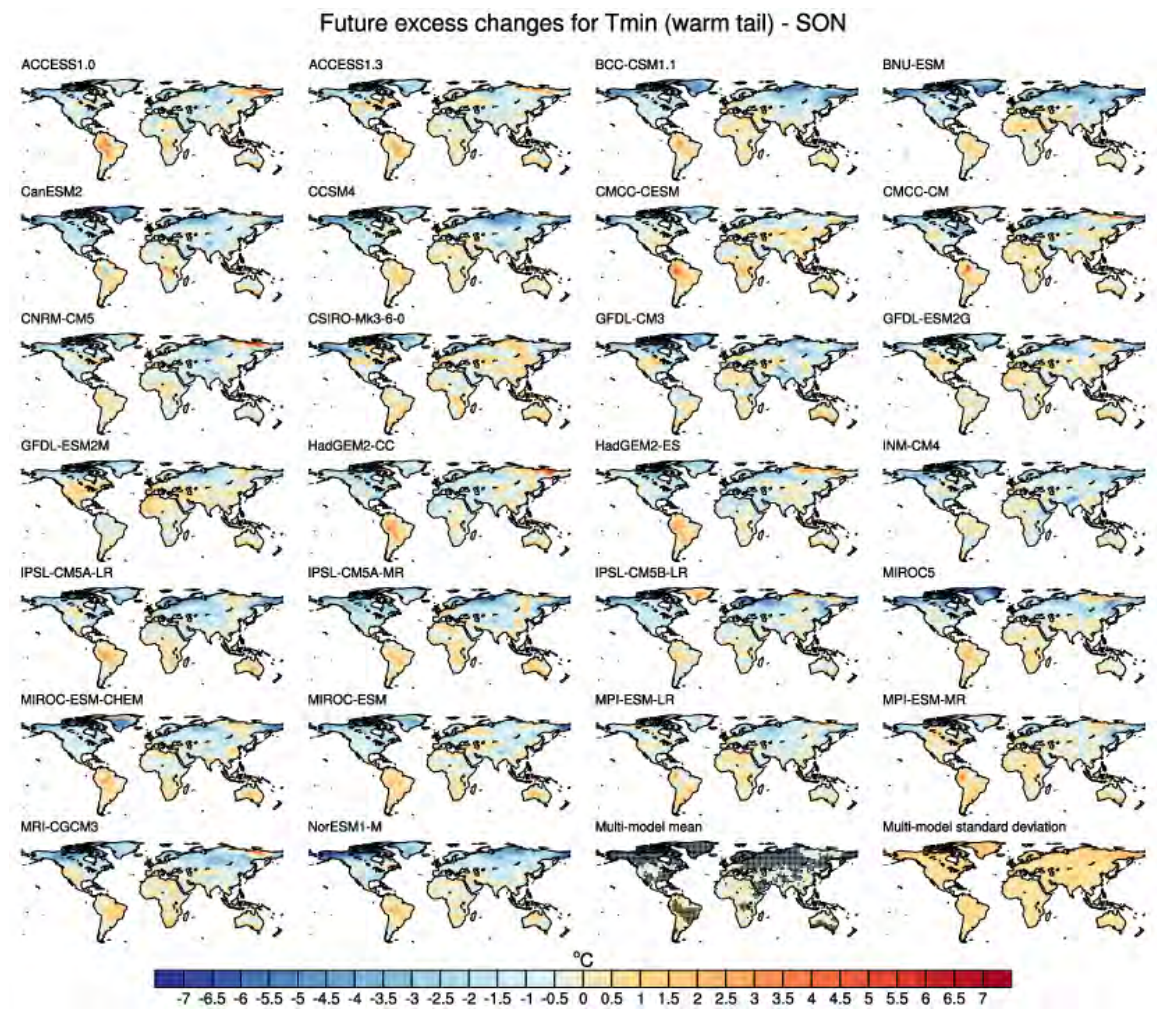


Figure S2.28 As Figure S2.25, but for September – November.

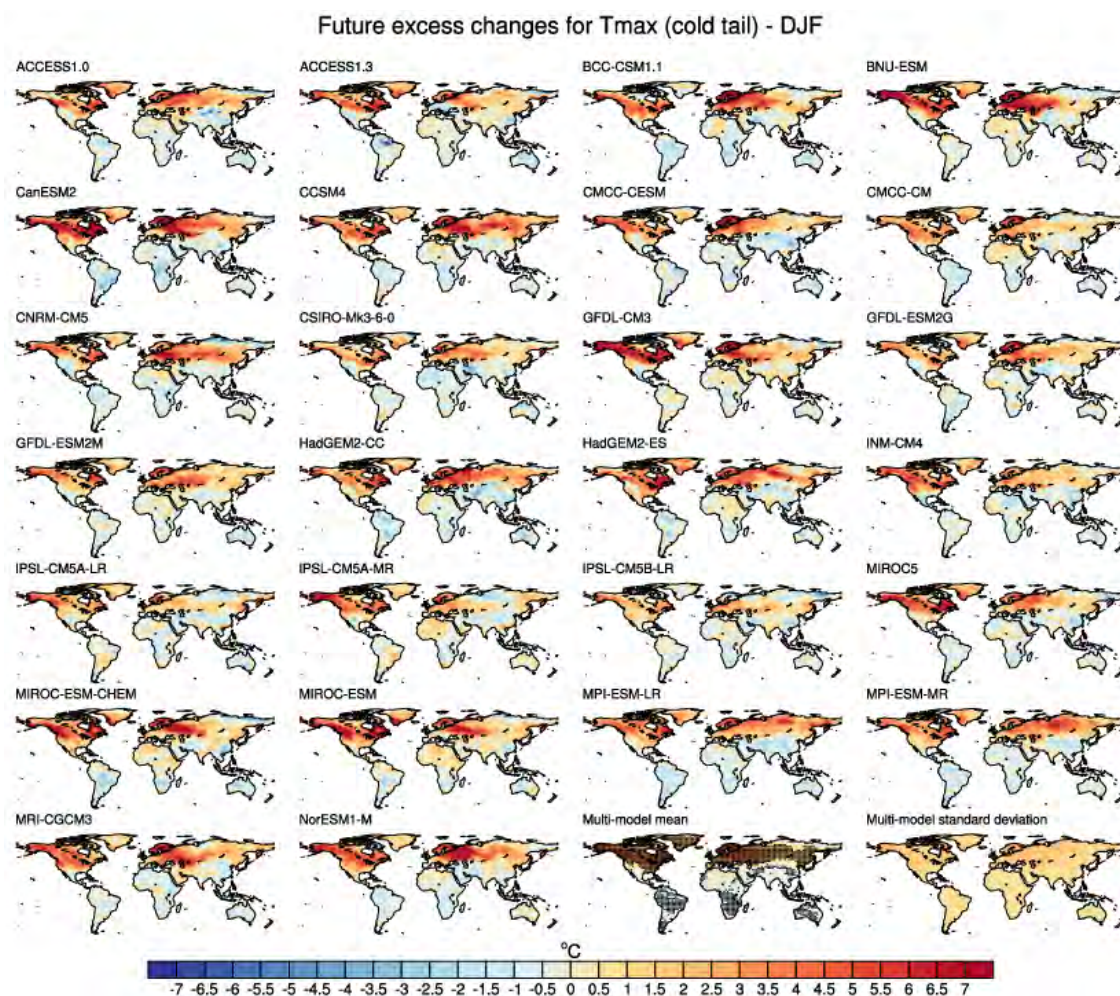


Figure S2.29 As Figure S2.21, but for the cold tails of daily maximum temperature.

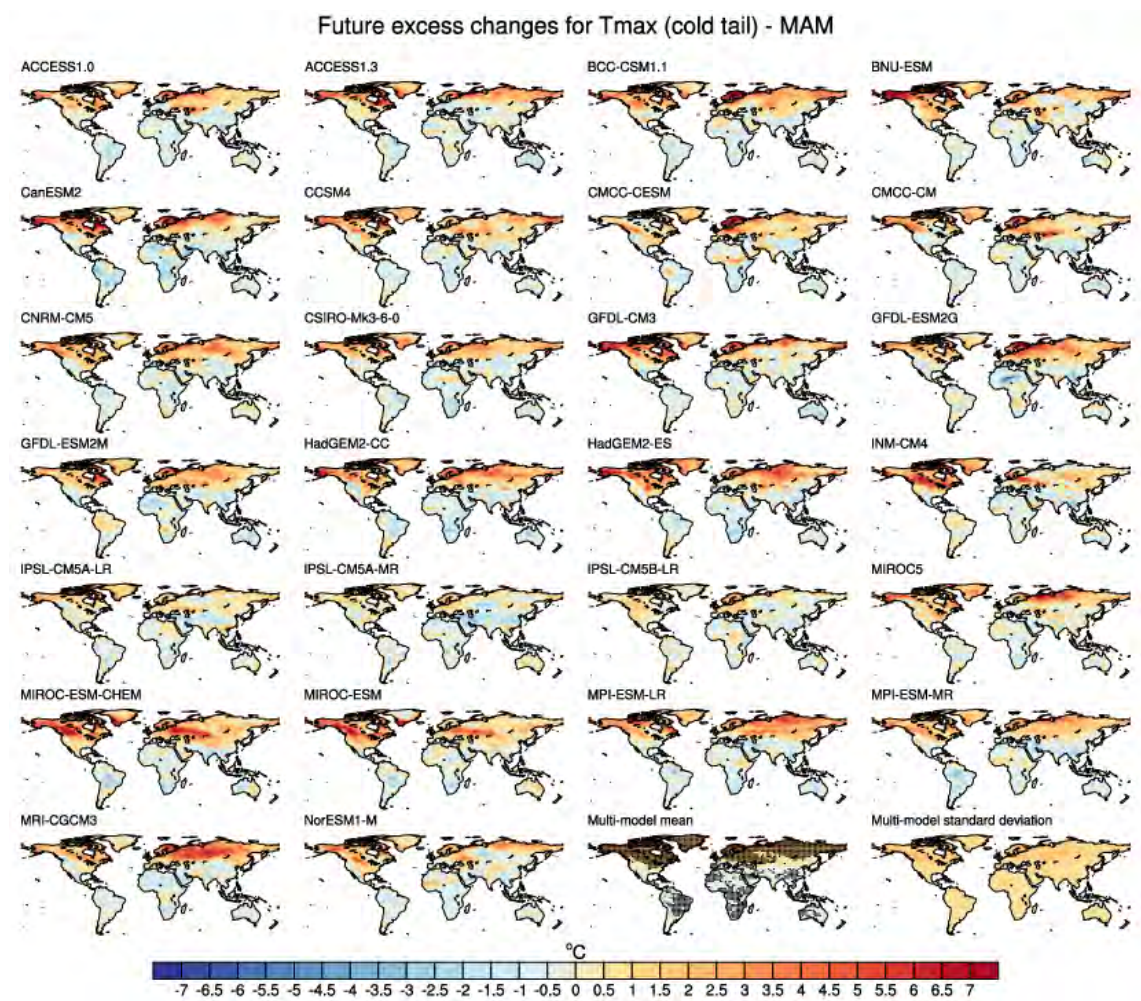


Figure S2.30 As Figure S2.29, but for March – May.

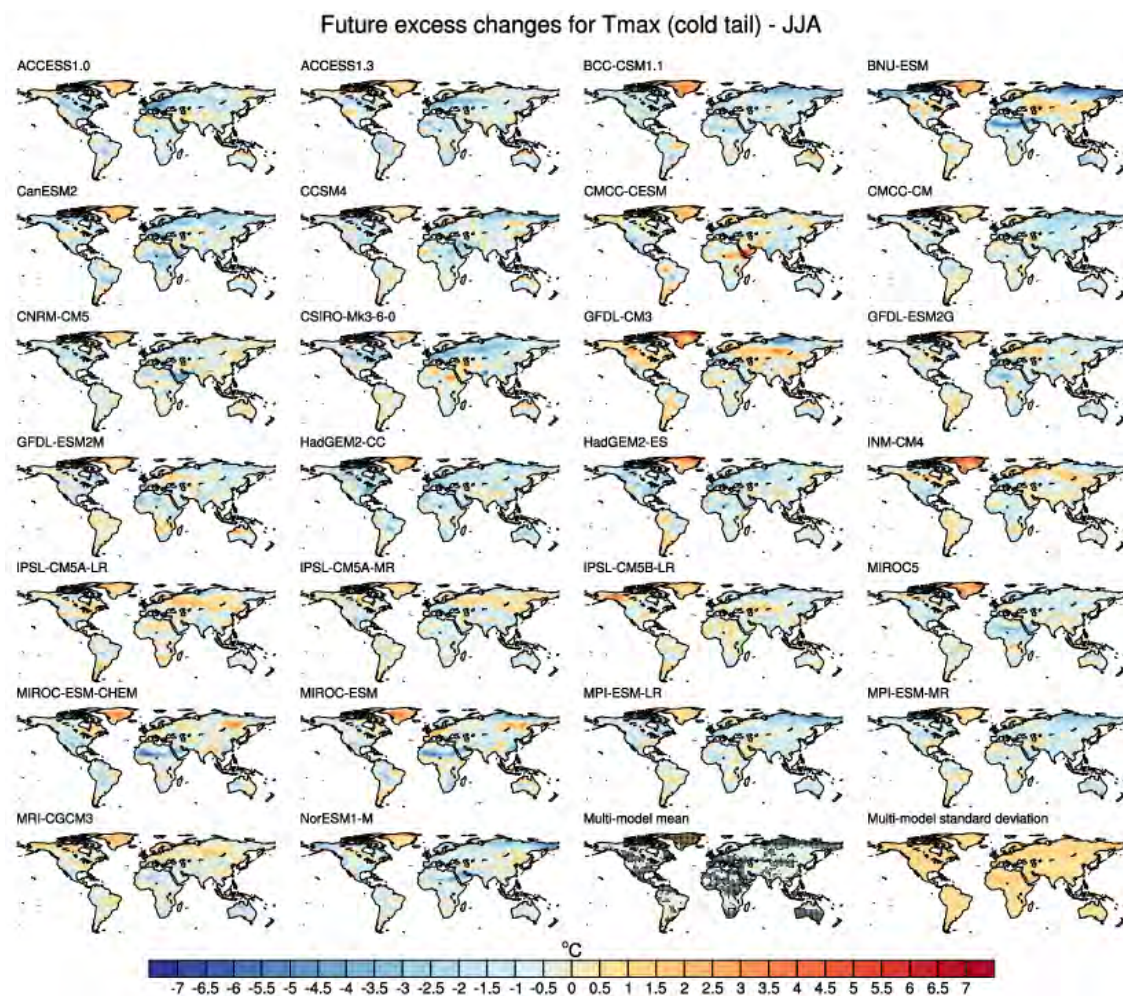


Figure S2.31 As Figure S2.29, but for June – August.

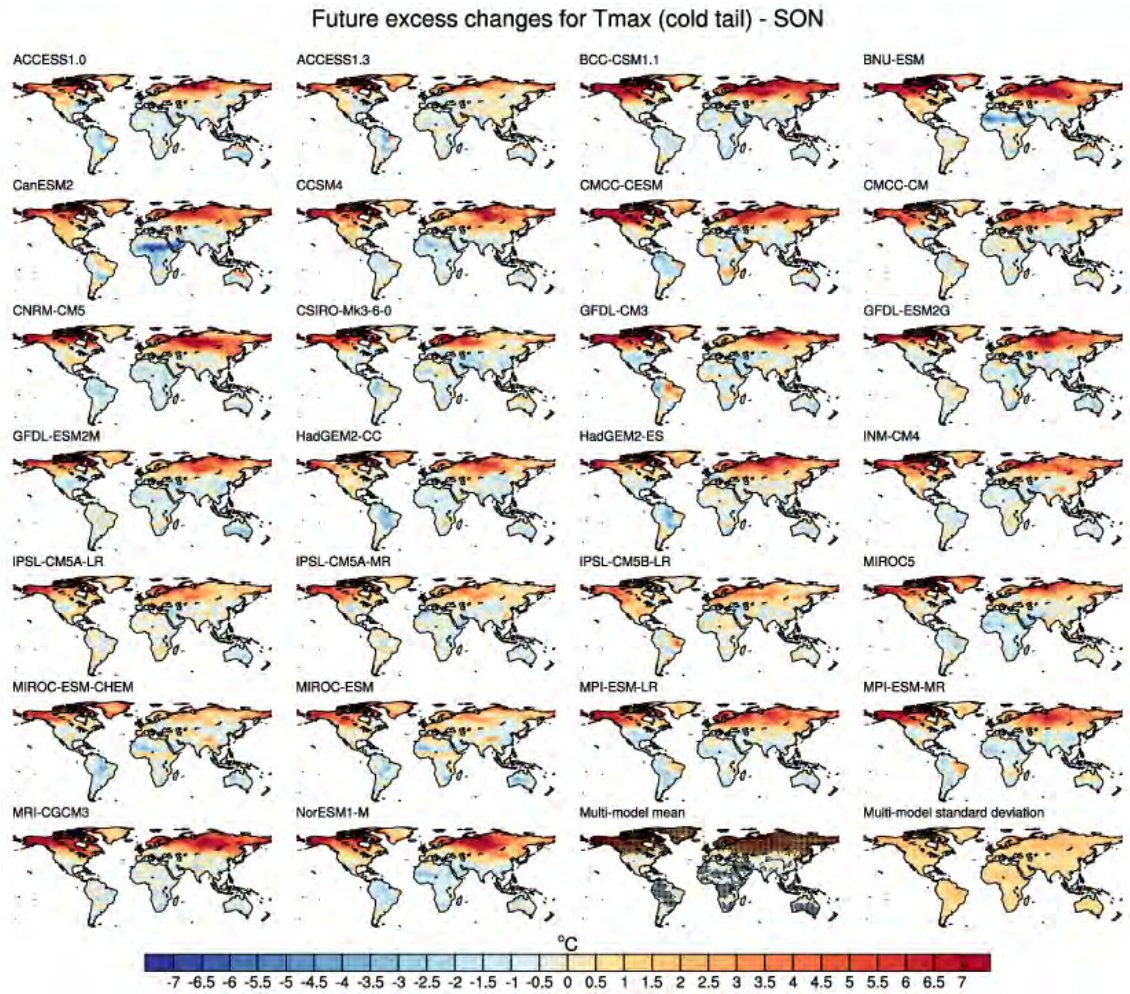


Figure S2.32 As Figure S2.29, but for September – November.

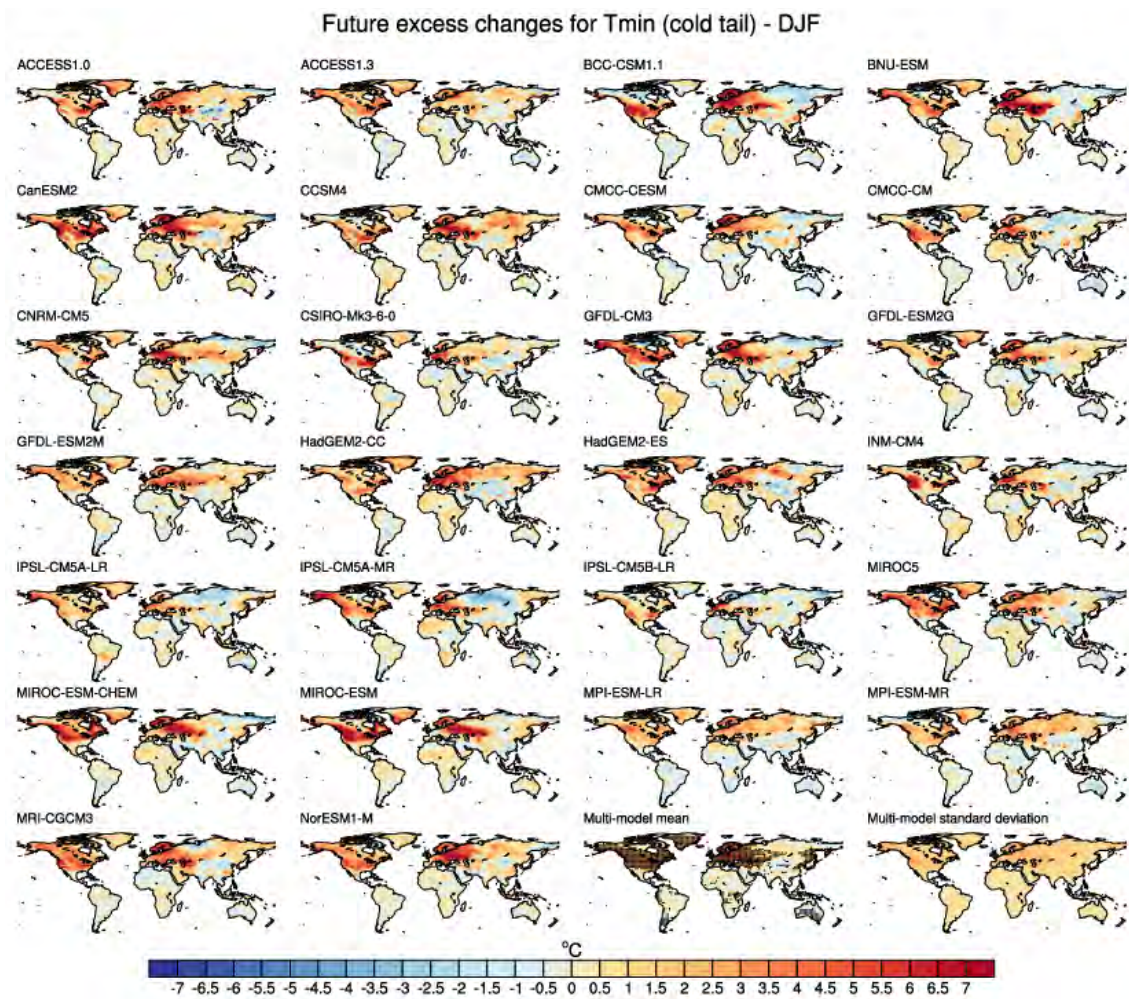


Figure S2.33 As Figure S2.21, but for the cold tails of daily minimum temperature.

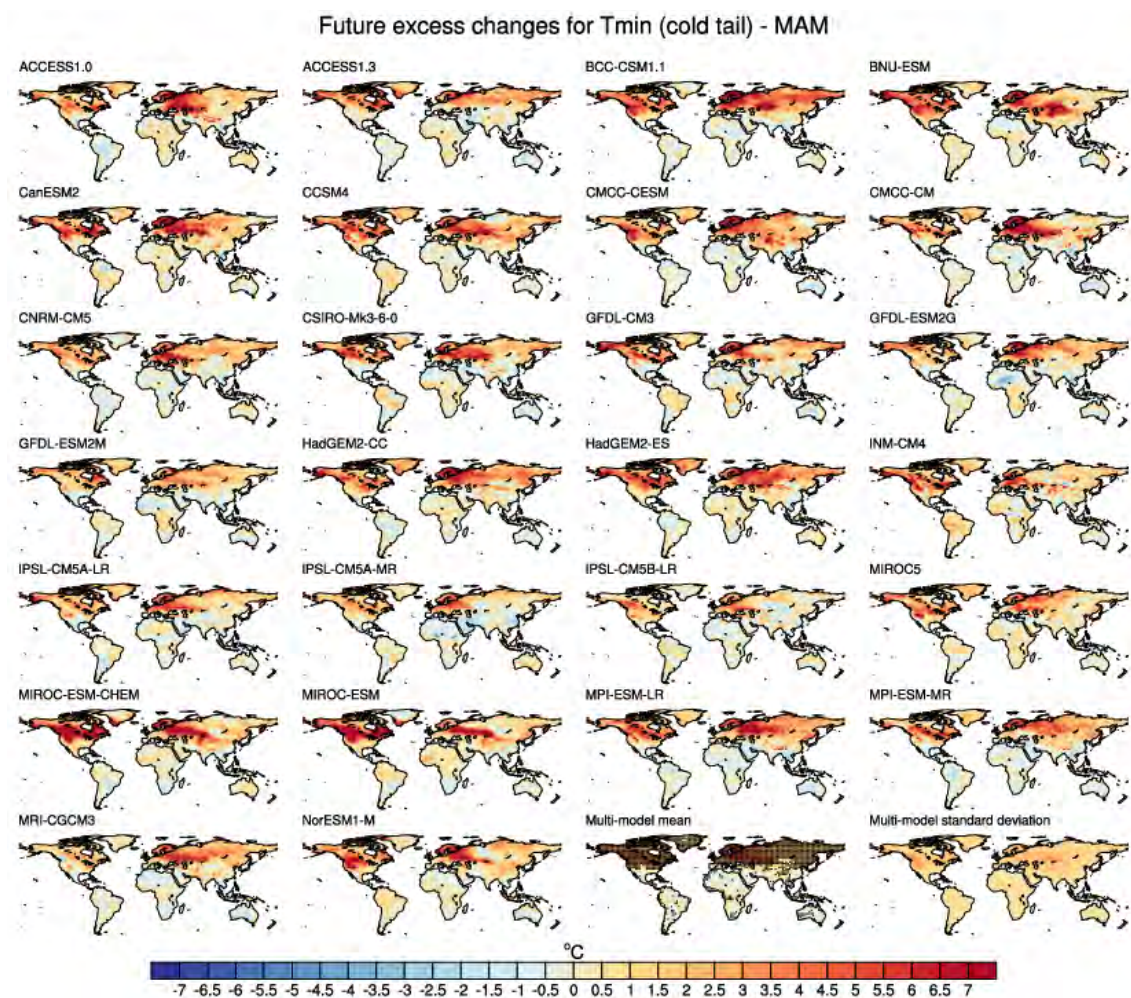


Figure S2.34 As Figure S2.33, but for March – May.

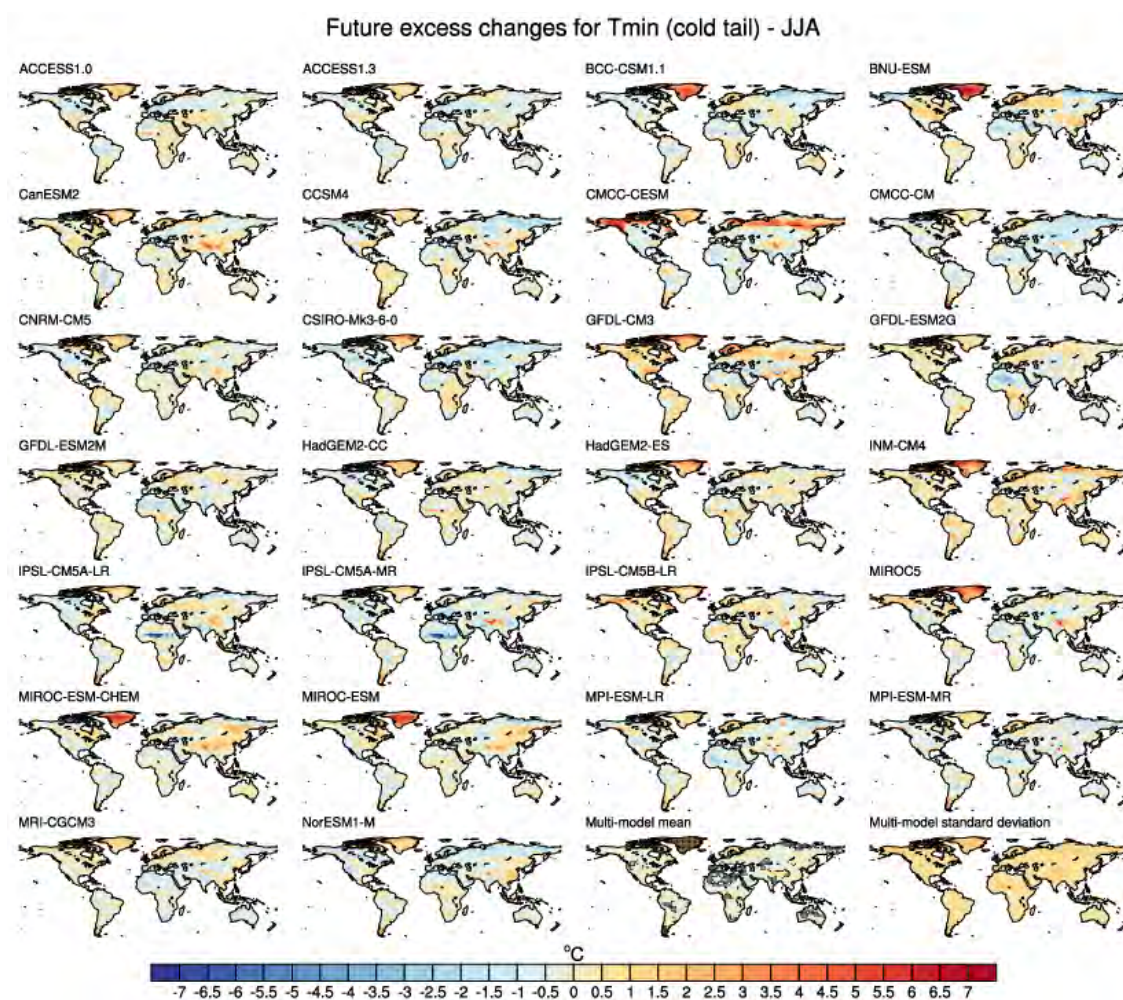


Figure S2.35 As Figure S2.33, but for June – August.

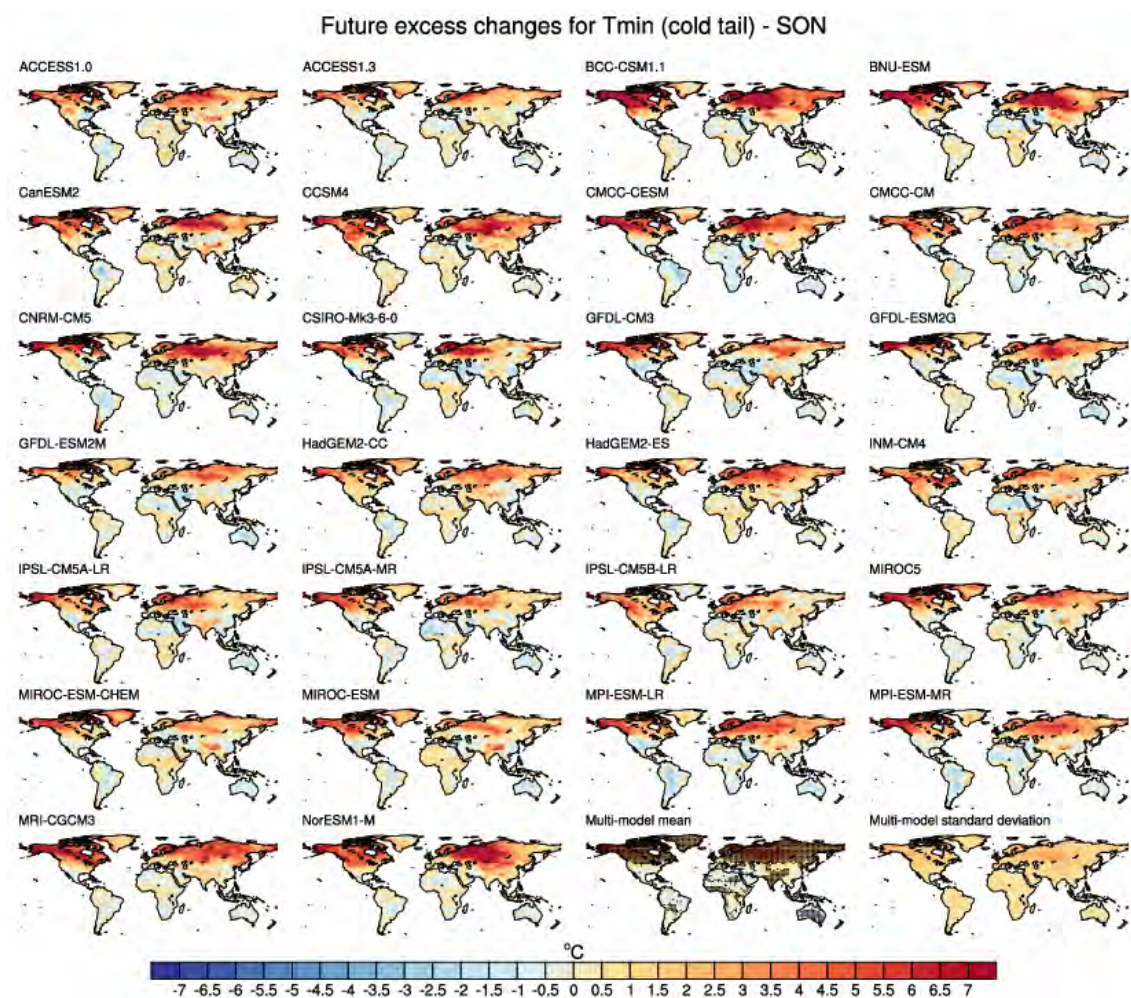


Figure S2.36 As Figure S2.33, but for September – November.

S.3 SUPPLEMENTARY MATERIAL FOR CHAPTER 4

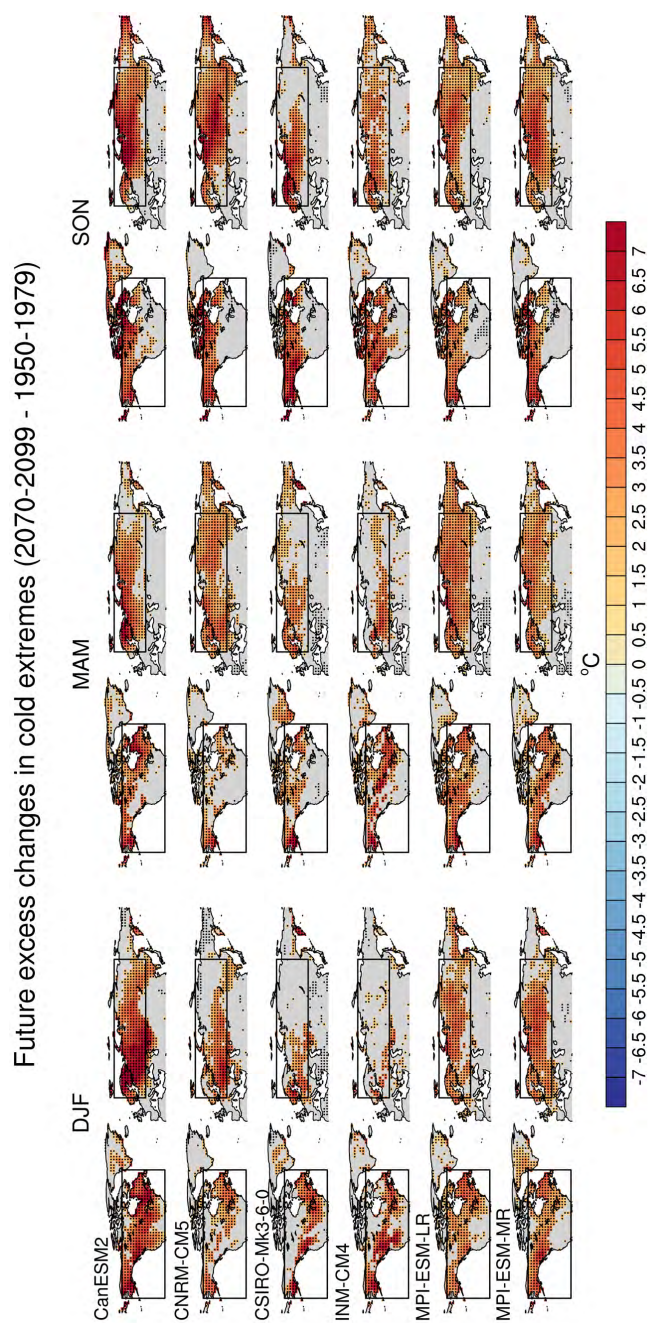


Figure S3.1 Boxed regions used for calculating area-averages for the scatter plots and correlations in Figures 4.7 – 4.9. Two regions are used: North America and northern Eurasia. The criteria specifies to only include grid cells which are both significant and that have future excess changes exceeding 1°C.

Future changes in temperature advection (3 day average prior to day of extreme)

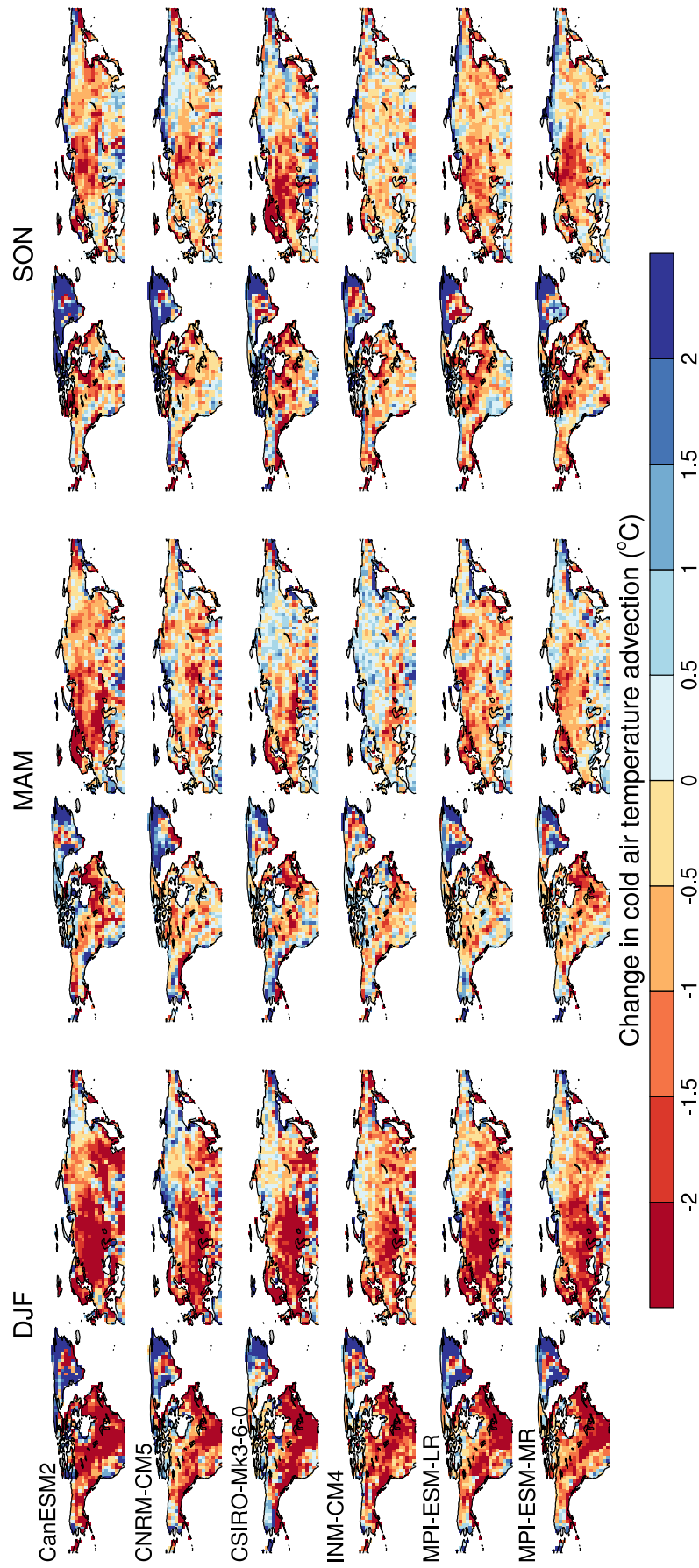


Figure S3.2 Future changes (2070-2099 – 1950-1979) in actual cold air temperature advection, calculated using the average values of advection for the three days prior to the day the annual seasonal minimum occurs.

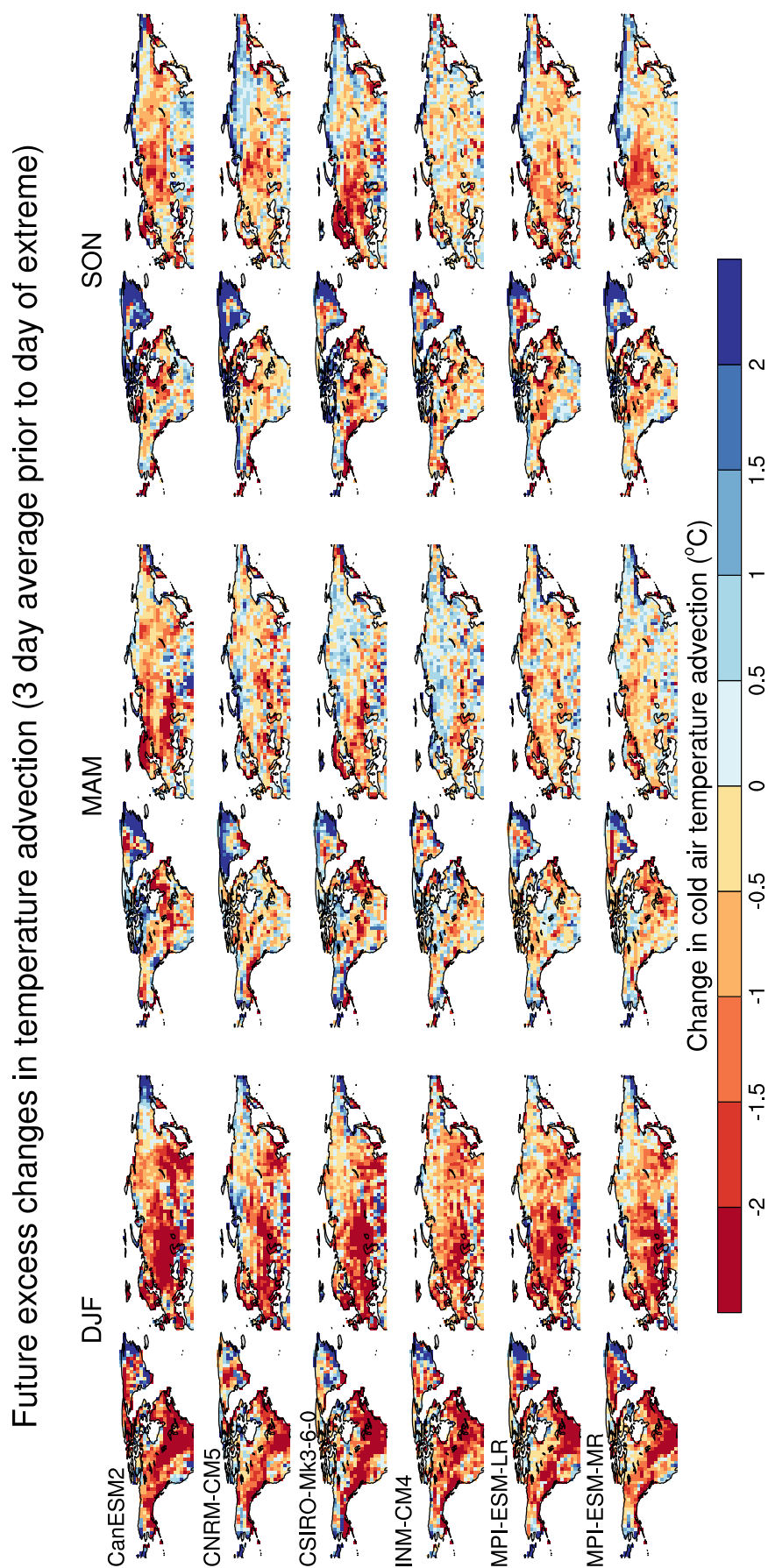


Figure S3.3 As Figure S3.2, but for excess cold air temperature advection.

Future changes in snow cover (on day of extreme)

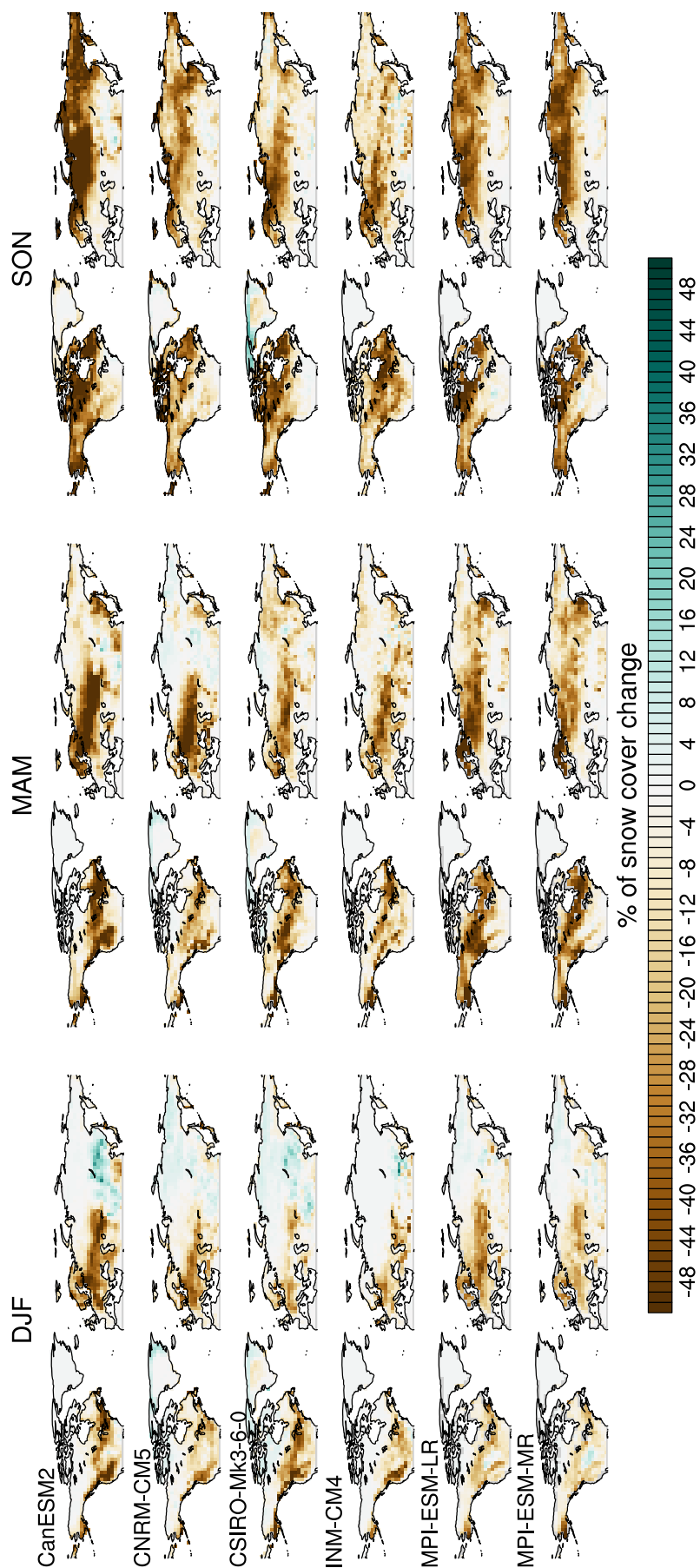


Figure S3.4 As Figure S3.2, but for future changes in actual snow cover, calculated using values of snow cover for the day the annual seasonal minimum occurs.

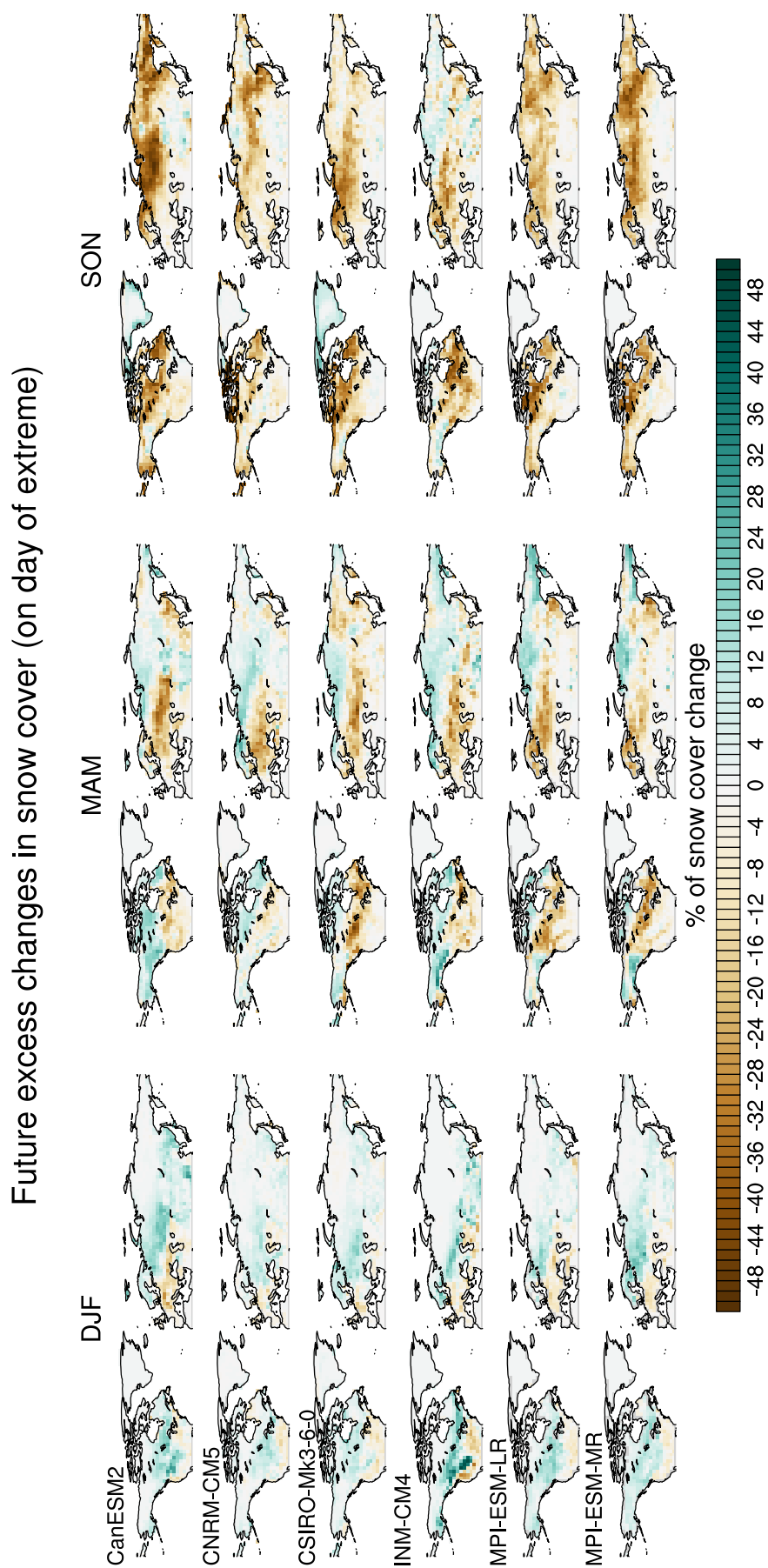


Figure S3.5 As Figure S3.4, but for excess snow cover.

Future changes in snow amount (on day of extreme)

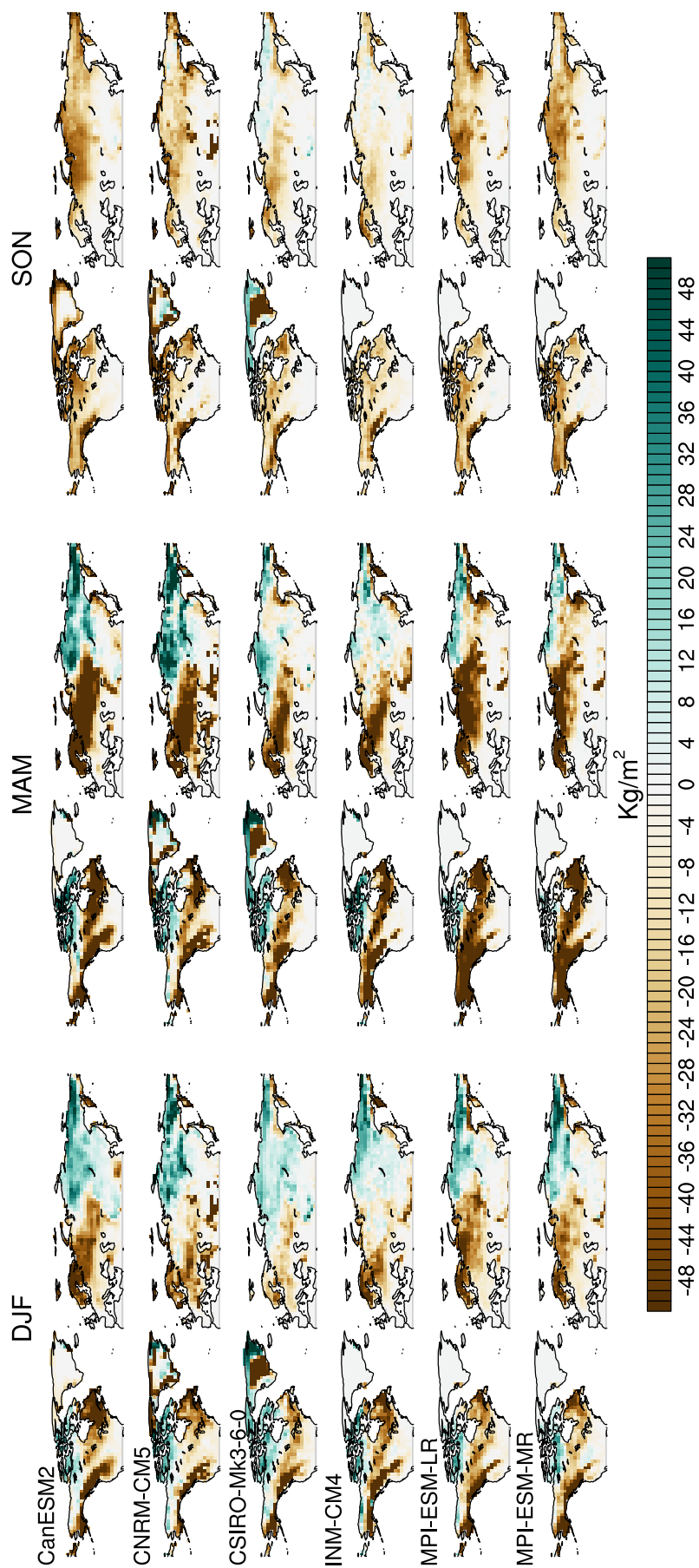


Figure 3.6 As Figure S3.4, but for actual snow amount.

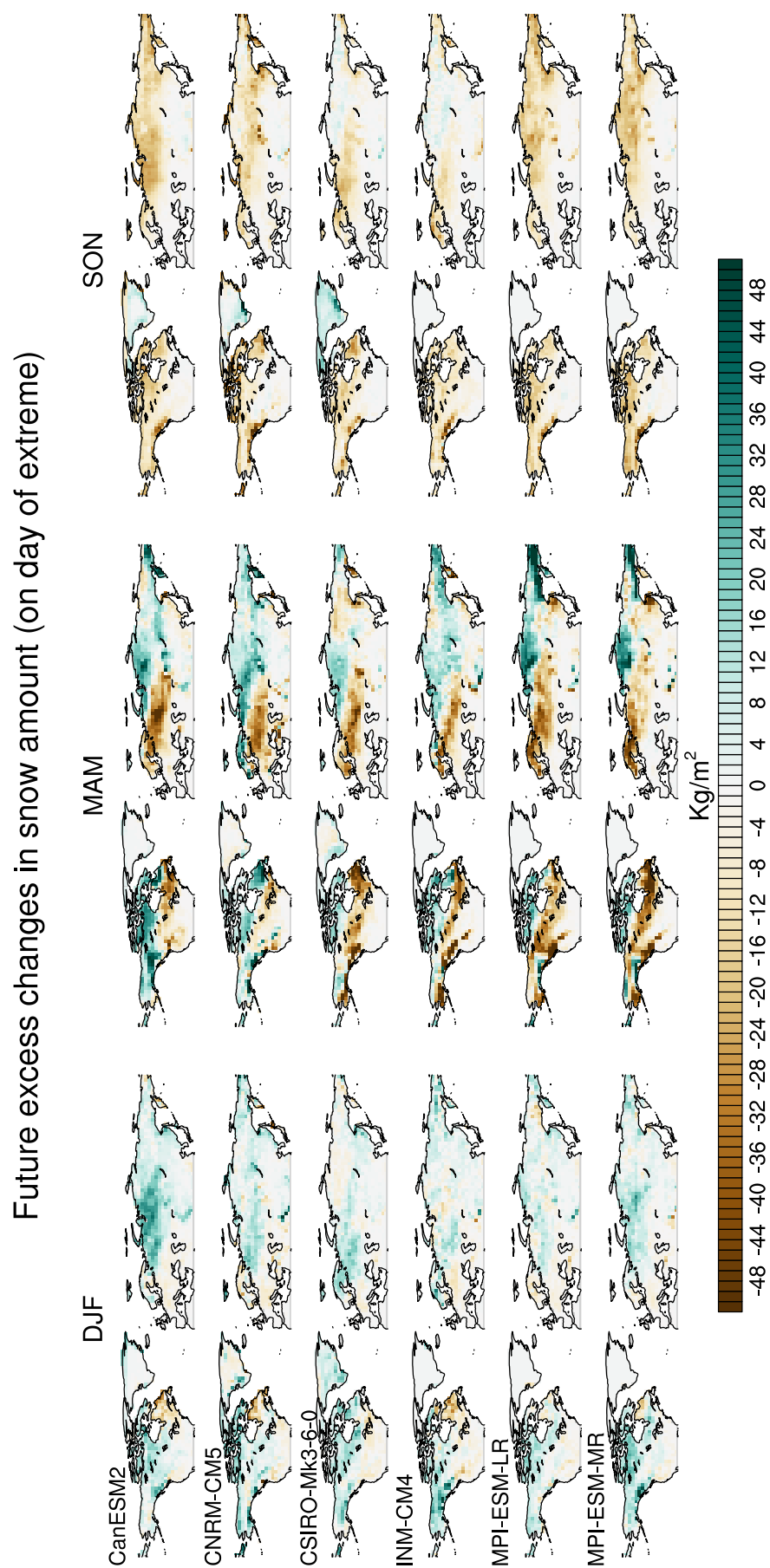


Figure S3.7 As Figure S3.4, but for excess snow amount.

Future changes in albedo (on day of extreme)

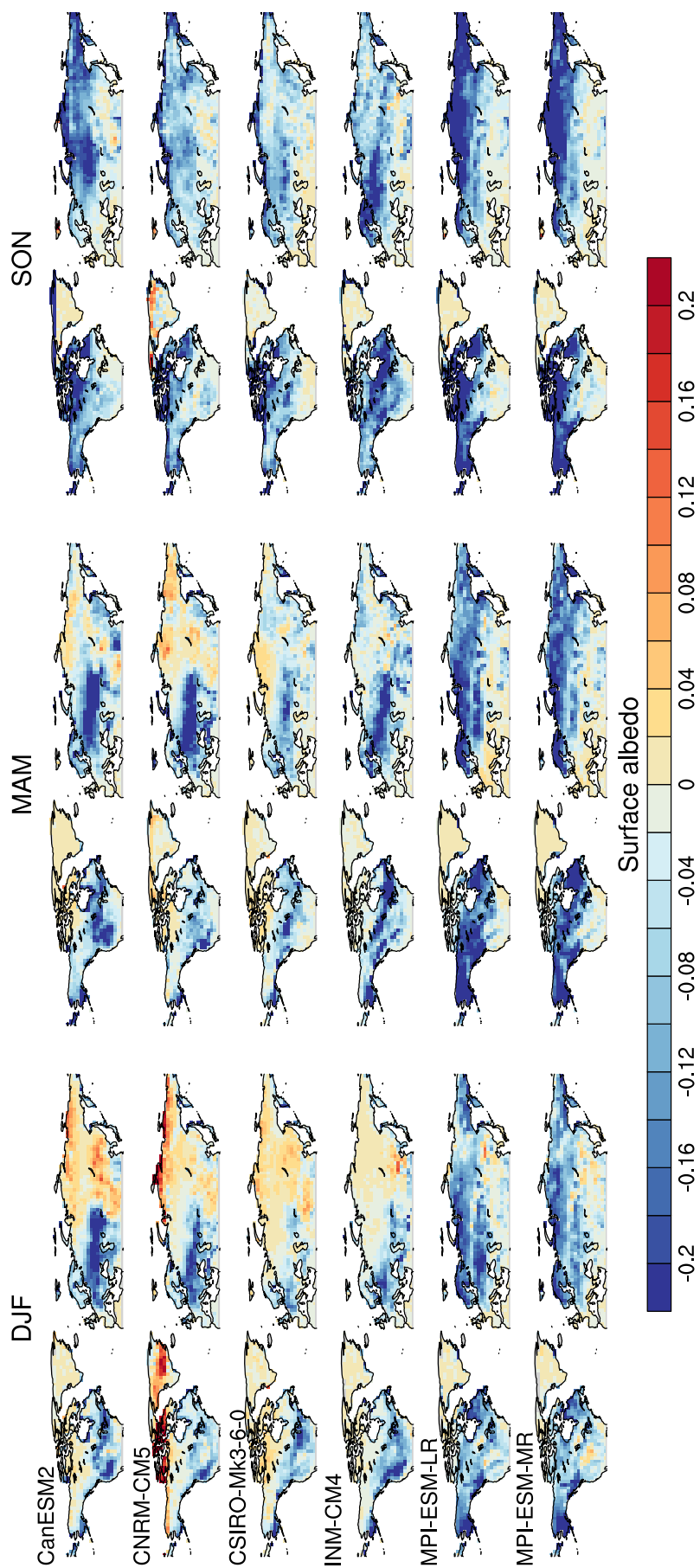


Figure S3.8 As Figure S3.4, but for actual surface albedo.

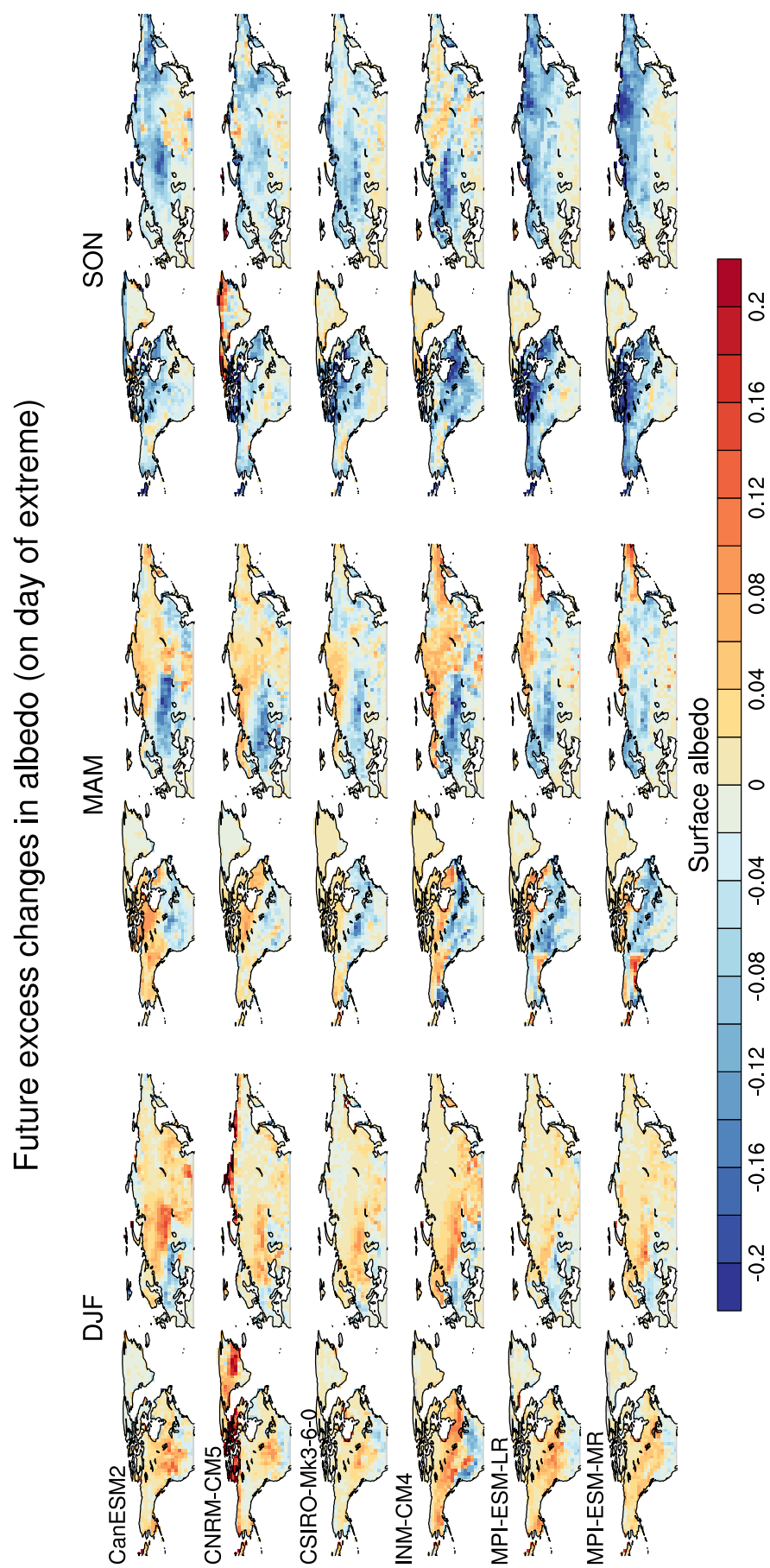


Figure S3.9 As Figure S3.4, but for excess surface albedo.

Future changes in the timing of cold extremes

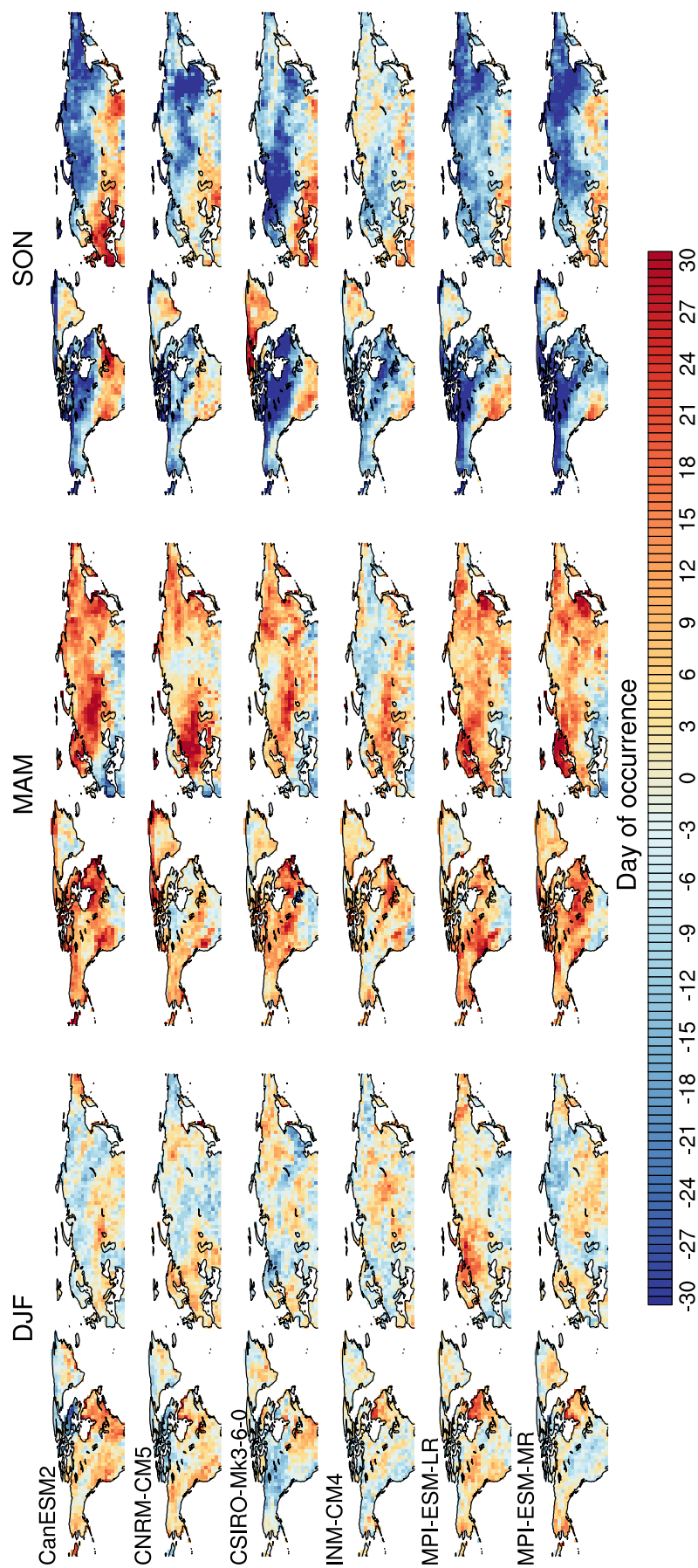


Figure S3.10 Future changes in the timing of the anomalously coldest days in the season for December – February, March – May and September – November. Negative values indicate grid cells where the cold extremes are projected to occur earlier in the season, and positive values show where cold extremes are projected to occur later in the season.