The Impact of externally forced Changes on Temperature Extremes

Simone Morak

Thesis submitted in fulfilment of
the requirements for the degree of
Doctor of Philosophy
to the
University of Edinburgh — 2013

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own.

Simone Morak
January 2013

Abstract

This thesis investigates changes in temperature extremes between 1950-2005, analysing gridded data sets of observations and climate model simulations. It focuses on changes in the frequency of extreme temperatures occurring in single days or over periods of six or more consecutive days. The study aims to quantify the significance of changes in extreme temperature events and answer the following questions. Are external or human-induced forcings together with natural forcings responsible for the observed change in temperature extremes or can these changes be explained due to natural climate variability alone? Are the observed changes consistent with those from climate model simulations? And are the changes in extremes linked only to changes in the mean climate, or only to those in climate variability or both? The analysis concentrates on changes from global to regional scale and from annual mean to seasonal scale. A detection method is applied to assess if changes are significantly different with respect to the internal climate variability. Results show that there has been a significant increase in warm daily extremes and a decrease in cold ones, both on large and small spatial scales. The increase in warm extremes has been found to be highly correlated with the increase in mean temperature. The changes in daily extremes are well represented in climate model simulations. Changes in the persistent extremes show a detectable increase in the frequency of warm and a decrease in cold events and are reproducible by models.

Acknowledgements

First I owe my gratitude to my supervisor Prof. Gabriele Hegerl, for her useful advice and guidance. I also thank my second supervisors Prof. Alexander Tudhope and Dr. Christopher J. Merchant for their encouraging words and open doors. I am indebted to my collaborators and co-authors Dr. Nikolaos Christidis and Dr. Jesse Kenyon, as well as to Prof. Francis Zwiers for all their help and advice. I have also benefited much from the steady support given by Prof. Simon Tett, Dr. Mike Mineter and the Edinburgh Compute and Data Facility (ECDF) officers.

A big thank you goes to the colleagues in my working group, Dr. Alessio Bozzo, Dr. Helen Hanlon, Dr. Anders Lindfors, Dr. Andrew Schurer, Dr. Tom Russon, Dr. Debbie Polson, Ms Carley Iles, Ms Emma Turner and Mr Darren Slevin, who have been always there for me for scientific discussions or personal encouragement. Thanks to my old and new friends, Manuela, Julia, Astrid, Gillian, Matt, Rachel, Matteo and Michele who always had some motivating words for me. A particular thank you goes to my friend Caroline who just always knows the exact right thing to say or do, to make it a good day and to Helen who made every coffee break a memorable event.

There are no words which could stand for the gratitude I owe my family and in particular Gabi, my mum, but of course also Maxi, Alessio, Marie-Luise, Roland, Nicolas and Lorenz and of course Arthos. Thanks for being there for and believing in me!

Contents

List of Tables

List of Figures

Chapter 1

Motivation and Background

1.1 Introduction

Instrumental temperature records, which go back to 1850 show clear changes and feature a long-term warming throughout the 20th century (Trenberth et al. [2007]). A strengthening of this trend occurred over the second half of the last century. Along with this trend in mean surface temperature, several indicators of temperature variability and extremes show changes as well, such as an increase in warm day and night extremes and a decrease in cold day and night extremes (Alexander et al. [2006]). The observed increase in mean temperature also lead to an increasing moisture content (Santer et al. [2007]), which further lead to changes in precipitation properties, such as amount, frequency, intensity duration and type (Trenberth et al. [2007]). However, changes in precipitation are often harder to detect, as they are underlying more complex dynamic processes and also regional effects.

The diagnostic and particularly the prediction of these extreme climate events is of great importance, as these are factors that have a strong and severe influence on society and ecosystems. Changes in extreme temperature events, for instance, can have severe impacts, for example on human health (Huynen et al. [2001], Christidis et al. [2010]), on agriculture and vegetation (Inouye [2000], Marino et al. [2011]), or energy consumption.

This chapter will give an overview on some of the previous research on changes in mean temperature (section 1.2) and extreme temperatures (section 1.3) as well as detection and attribution studies (section 1.4), followed by an introduction to the project (section 1.5) and an explanation of the data (section 1.6) and methods (section 1.8) used.

1.2 Changes in Mean Temperature

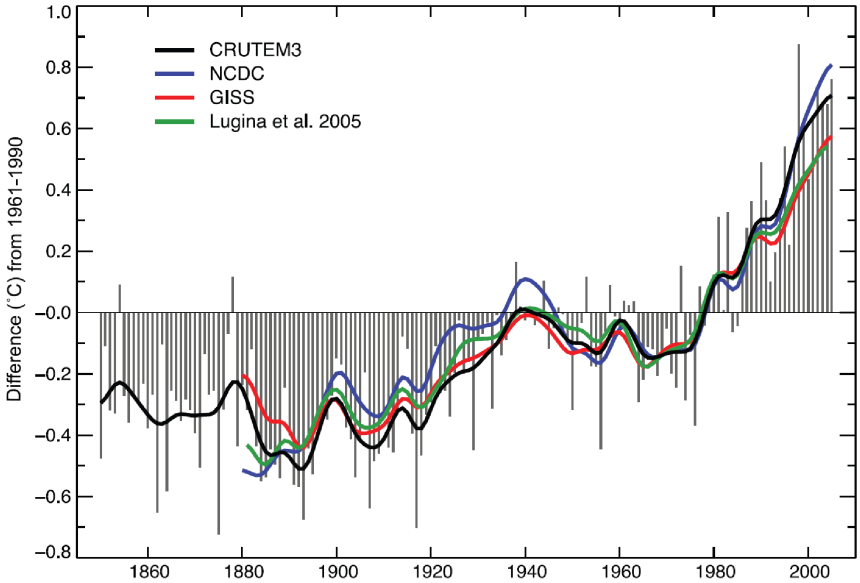


Figure 1.1: Annual anomalies of global mean land-surface air temperature from 1850-2005. Anomalies are computed relative to 1961-1990 mean of CRUTEM3. Smoothed lines show the decadal variation for CRUTEM3 (black), NCDC (blue), GISS (red) and Lugina et. al 2003 (green) (figure from Trenberth et al. [2007]).

The record of the global mean surface temperature for the period 1906-2005 shows a 100-year trend of +0.74 ±0.18 degree Celsius (Trenberth et al. [2007]). The data are based on averaging surface temperature measurements for stations and ship records in grid boxes, yielding a gridded set of temperature data, which is then averaged. Fig.1.1, which originates from the IPCC report 2007 (see Trenberth et al. [2007]), shows that all four of the processed data sets of global observed mean land-surface temperature, CRUTEM3 (Brohan et al. [2006]), NCDC (Smith and Reynolds [2005]), GISS (Hansen et al. [2001]) and Lugina et al. [2006] agree in the main features over the last century. Also the recently released data set by Berkeley Earth Surface Temperature (Rohde et al. [2012]) shows comparable results. This agreement gives confidence in the robustness of the following findings.

All data sets show an increase in surface temperature which began at least in the 1910s. This increase was interrupted by a stagnating period or even a period of decrease in the 1950s and 1960s. This stagnation is associated with a minimum in solar activity. From the 1970s on the temperature has kept on increasing and is still continuing to do so. The highest peaks within this global mean time series can be seen in the years 1998 and 2005 (see Fig.1.1). The observed linear trend of the second half of the 20th century nearly doubled in comparison to the trend of the entire century (see Fig.1.1). However, these are just the results for the global mean, regional climate can vary quite a lot due to regional circulations and atmosphere and ocean interactions, like the North Atlantic Oscillation (NAO), El Ninõ and others (see Scaife et al. [2008], Kenyon and Hegerl [2008]). Since 2000 the global mean temperature has still increased, however the magnitude of the change appears to be smaller (Hansen et al. [2010]). This slow-down in the increase of mean surface temperatures could be due to a lack in strong El Nino events, which for example contributed to the high peak in 1998.

1.3 Changes in Climate Extremes

Changes in the probability of climate extremes are generally harder to detect and understand than those in climatological mean state, due to their higher sampling variability. In order to facilitate the analysis of extremes events, several different indices have been defined and analysed. The Expert Team on Climate Change Detection and Indices (ETCCDI) for example defined 27 core indices (Zhang et al. [2005], Zhang et al. [2011]). These indices describe different aspects of changes in extremes and climate variability in temperature and precipitation. One example is the number of frost days, which are the number of days with a minimum temperature below the freezing point. Another is the growing season length which is standing for the number of days in a year from the first span of of at least 6 days with daily mean temp $> 5°C$ and first span of at least 6 days with daily mean temp $< 5°C$ after the 1^{st} July.

Whereas most of the indices are globally relevant, there are some, such as the tropical nights or icing days, which are probably climatologically just relevant for certain regions. A short description of some of the indices is given in Table 1.1.

Within the temperature indices one has to differentiate between frequency and intensity indices. Frequency indices such as TN10, TN90, TX10, TX90 (for description see table 1.1) are computed by counting the occurrences of events that exceed a defined threshold and expressing it as percent of available days. Intensity indices on the other hand determine the absolute maximum and minimum value that was recorded in a particular period. The advantage of using frequency indices rather than intensity ones is that those are less sensitive to outliers.

Many publications such as the results of Synthesis and Assessment Product 3.3 (SAP 3.3) of the Climate Change Science Program (CCSP) report by Karl et al. [2008], the paper by Alexander et al. [2006] and the work by Tebaldi et al.

Index	**Acronym**	**Definition**
Frost days	FD	annual number of days with a daily minimum temperature $< 0°C$
Summer days	SU	annual number of days with a daily maximum temperature $> 25°C$
Icing days	ID	annual number of days with a daily maximum temperature $< 0°C$
Tropical days	TR	annual number of days with a daily minimum temperature $> 20°C$
Growing seasonal length	GSL	annual amount of days from the first span of at least 6 days with daily mean temp $> 5°C$ and first span of at least 6 days with daily mean temp $<$ 5°C after the 1^{st} July
Monthly maximum values	TXx	monthly maximum values of daily maximum temperature
Monthly minimum values	TNn	monthly minimum values of daily minimum temperature
Cold nights	TN10 (= TN10p)	number of days in a year that exceed the 10^{th} percentile of the minimum temperature (given in percent)
Warm nights	TN90 (= TN90p)	number of days in a year that exceed the 90^{th} percentile of the minimum temperature (given in percent)
Cold days	TX10 (= TX10p)	number of days in a year that exceed the 10^{th} percentile of the maximum temperature (given in percent)
Warm days	TX90 (= TX90p)	number of days in a year that exceed the 90^{th} percentile of the maximum temperature (given in percent)
Warm spell duration index	WSDI	annual number of days with at least 6 days with TX $> 90^{th}$ in series
Daily temperature range range	DTR	monthly mean difference between TX and TN
Consecutive dry days	CDD	maximum number of days in series with a daily precipitation < 1 *mm*
Consecutive wet days	CWD	maximum number of days in series with a daily precipitation $>= 1$ *mm*
Annual total precip. in wet days	PRCPTOT	annual total precipitation in wet days

Table 1.1: Outline of some climate extreme indices as defined by the ETCCDI

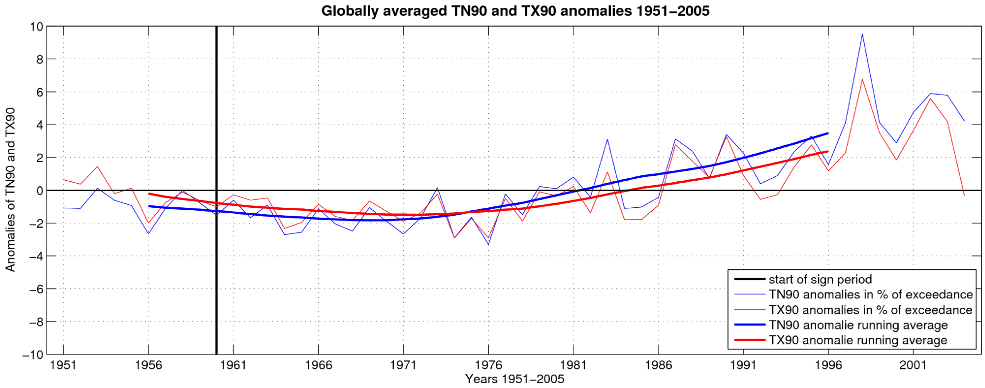


Figure 1.2: Anomalies of annual number of warm nights (thin blue line) TN90 (exceedances of the 90th percentile of the minimum temperature distribution) and annual number of warm days (thin red line) TX90 (exceedances of the 90th percentile of the maximum temperature distribution) for the period 1951-2005. The bold lines show the 11-year running average for TN90 (bold blue line) and TX90 (bold red line). The anomalies have been calculated subtracting the 1951-2005 mean. The vertical line at 1960 marks the beginning of the period where the data coverage is nearly global. (Data set by Jesse Kenyon, based on data from HadEX project (Alexander et al. [2006])

[2006], show increases in temperature extremes, led by a significant increase in the number of warm minimum temperatures (exceedances of the 90th percentile of the minimum temperature distribution), followed by an increase of a smaller magnitude in the number of warm maximum temperatures as shown in Fig.1.2. On the other hand a decrease in the cold tails of both minimum (exceedances of the 10th percentile of the minimum temperature distribution) and maximum temperature has been reported (Alexander et al. [2006], Tebaldi et al. [2006], Field et al. [2012]). From this change in the cold and warm edges of the distribution results a positive trend in growing seasonal length (GSL) and a negative trend in the daily temperature range (DTR) over parts of the late 20th century.

At this point arises the question whether the increase in mean has lead to the increase in extremes or if there are other explanations for these changes. This is one of the questions this project will concentrate on. SAP 3.3 of the CCSP report by Karl et al. [2008] focuses on the explanation that the changes in mean affect the changes in extremes, which would mean that the whole temperature distribution is being shifted to higher values as is schematically shown in Fig.1.3a. Such a development would lead to a higher mean and a disproportionally higher frequency of hot events. Extremely cold events on the other hand, would decrease in this scenario. Another possible scenario would be an increase in variability which would make the tails of the distribution larger, while keeping the mean constant (see Fig.1.3b), this scenario however does not seem adequate for the change in temperature, as an increase in mean surface temperature has already been found and verified. The last scenario suggests an increase in both mean and variability (see Fig.1.3c).

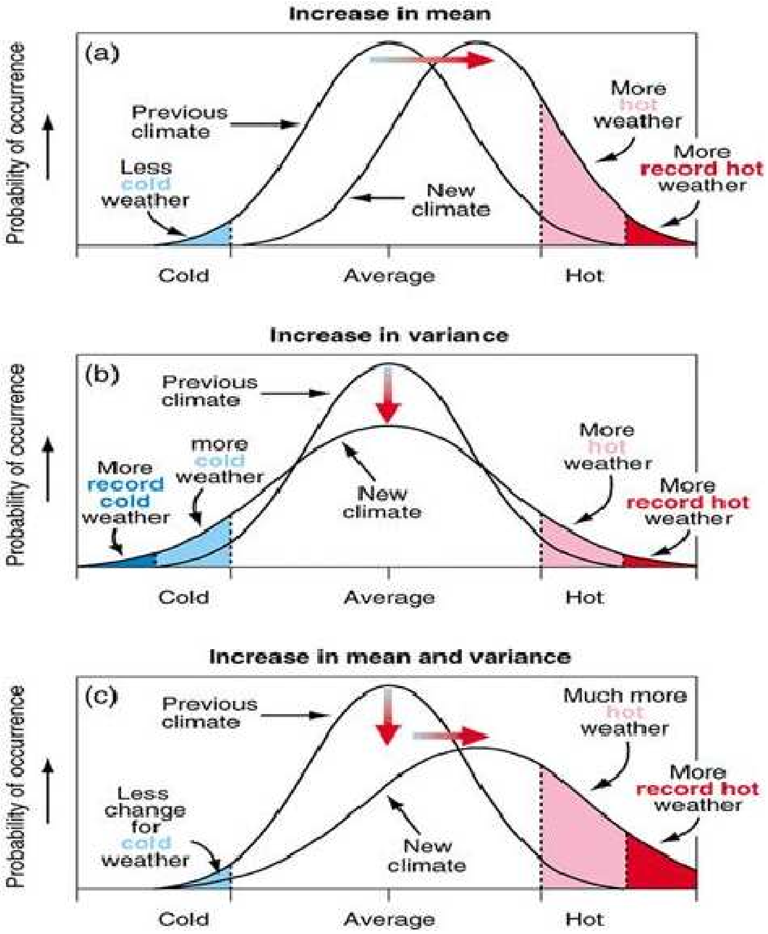


Figure 1.3: Effect of an increase in a) mean temperature, b) variance and c) mean temperature and variance on the changes in temperature extremes (IPCC (2001a, 2001b))

Together with the increase in mean and extreme temperature an increase in heat waves and a corresponding decrease in frost days has also been recorded (Tebaldi et al. [2006]).

Besides changes in extreme temperature events, changes in precipitation have also been recorded. These changes have caused already wet regions to become even wetter and dry regions to become dryer (Alexander et al. [2006], Karl et al. [2008]). Furthermore there has been a detectable increase in the intensity of heavy precipitation events (Min et al. [2011]). This change leads to a rising risk of flooding but also the risk of regional increases in frequency of droughts and dry spells, which is caused by a combination of the trend in temperature and the trend in precipitation.

1.3.1 Regional Changes in Extremes and Mechanisms

The changes in extremes described above show distinct geographical patterns. The following patterns of changes over the second half of the 20th century, as described by Tebaldi et al. [2006], are supported by being consistent across the models. Such changes are for instance a particularly strong warming and an increasing intensity of wet events in the high latitudes of the Northern hemisphere, compared to lower latitudes. Furthermore Tebaldi et al. [2006] find a significant decrease in frost days and the associated increase in growing season length in the north western region of the US and Eastern Europe. South Western North America shows a development towards higher numbers of heat waves (Tebaldi et al. [2006]). Severe changes have been taking place in Australia, where the Northern parts show the largest increase in heat waves over the continent, the South East shows the largest decrease in frost days and increase in growing season length (Tebaldi et al. [2006]).

Such changes towards warmer temperatures especially if occurring during winter can play an important role for spring vegetation, since an increase in temperature gives some plants the signal to grow. However, the wintertime dynamics often causes these false spring events to be interrupted by cold snaps, which often kill off this spring vegetation. These particular events are described by Marino et al. [2011].

Studies by Schär et al. [2004], Stott et al. [2004] and Fink et al. [2004] focus on the explanation of the European heat wave in 2003. This event showed anomalously hot temperatures exceeding the long term mean by 3 degree C (equivalent to more than 5 standard deviations). Schär et al. [2004] concentrate on the role of the rising summer temperature mean and variability in 2003 on the European summer heat waves. Stott et al. [2004] apply a method which allows them to quantify the contribution of external forcing to the risk of such events to happen. Fink et al. [2004] on the other hand explain this event by analysing the synoptic situation.

The work by Schär et al. [2004] shows the influence of increasing temperature on European summer heat waves by comparing temperature observations with the results of the control run and the greenhouse gas forced run of a regional climate model (RCM) as well as with the general climate model (GCM) run. Schär et al. [2004] state that changes in summer climate in Europe may result from an increase in variability, possibly due to a change in greenhouse gases, which would explain unusually hot extremes like the summer 2003 and would increase the likelihood of future heat waves and droughts.

Stott et al. [2004] use detection and attribution to show that the regional mean summer temperatures have likely increased due to human-induced forcing. Having found this, they use HadCM3 to estimate the change in likelihood of such events to happen, under the influence of global warming. Stott et al. [2004] find that the anthropogenic forcing has doubled the risk of events like the heat wave in 2003 to happen.

Fink et al. [2004] attribute the heat wave in 2003 to the anticyclonic situation during spring and summer 2003. This stable high pressure system led to low soil moisture and favoured a further increase in temperature and dryness. Fink et al. [2004] come to the conclusion that under the given synoptic condition an event like the 2003 heat wave would have even occurred without the current impact of anthropogenic drivers, not taking into account any possible changes in dynamics due to changes in climate.

Similar to the studies by Schär et al. [2004] (and Stott et al. [2004]) and Fink et al. [2004], also Dole et al. [2011] and Rahmstorf and Coumou [2011] (see also Otto et al. [2012]), who discuss the attribution of the Moscow heatwave 2010, show fairly opposed conclusions. Dole et al. [2011] state that this extreme event most likely resulted from atmospheric dynamical processes, causing a blocking situation, rather than from anthropogenic influences. They support this statement with their analysis of the mean surface temperature and its variability across western Russia, which does not seem to show the globally observed increase. Rahmstorf and Coumou [2011] who developed an approach to quantify the influence of long term trends on the occurrence of extreme events, found that for the case of the Moscow heatwave 2010, the general observed warming trend has led to an increased probability of such events to happen. The conclude that without climate warming this event might not had happened.

Summarising it can be said, that the occurrence of extreme temperature events, like the summer heat wave in Europe 2003 or Moscow 2010, is likely to be influenced by the observed warming trend, which has been at least partially attributed to human-induced forcing. This observed increase in mean climate state and climate variability increases the probability of hot extreme events to happen. However, for any extreme event to occur, unusual synoptic conditions, which cause a large dynamic contribution compared to a small change in mean temperatures, are needed. This is not in contradiction with substantial changes in the probability or risk of such events (Otto et al. [2012]).

Regional changes in mean temperature and temperature extremes have been found to be affected by changes in modes of variability, such as the El Niño Southern Oscillation (ENSO), the Pacific Decadal Oscillation (PDO), the North Atlantic Oscillation (NAO) or the Northern Annular Mode (NAM) (Kenyon and Hegerl [2008]). El Niño events lead to warmer than normal temperatures across the North Western US, Western Canada, Southern Alaska (Deser and Wallace [1987]) and also India and Eastern Asia (Deser and Wallace [1990]), whereas they cause cooler and wetter conditions across South-Eastern North America (Deser and Wallace [1987], Kiladis and Diaz [1989], Rasmusson and Carpenter [1982]). The Pacific Decadal Oscillation (PDO) is the leading mode of variability in the North Pacific and is characterised by the sea surface temperature anomalies in the North Pacific compared to those along the west coast of North America. Furthermore the North Atlantic Oscillation (NAO) influences the precipitation and temperature patterns over Northern and Southern Europe (Scaife et al. [2008]). Both the NAO and the related Northern Annular Mode (NAM) show a decrease in sea level pressure over the Arctic region combined with an increase over the subtropical north Atlantic.

The paper by Kenyon and Hegerl [2008] shows that these modes of climate variability are drivers for changes in temperature extremes and that in order to understand and predict future changes in temperature extremes, it is necessary to take into account changes in modes of the climate variability. Also Scaife et al. [2008] use the changes in NAO to explain aspects of European climate changes in winter between the 1960s and 1990s. They conclude that in the Northern Hemisphere, changes in the NAO index are responsible for increases in European heavy precipitation in winter as well as for a decreasing trend in unusually cold winter days. A positive NAO index causes a decrease in heavy precipitation and a slighter decrease in frost days in Southern Europe and Northern Africa. Nevertheless, according Scaife et al. [2008], there is at least one open question concerning future changes due to the NAO, which is based on the question as to whether its changes are naturally or anthropogenically influenced. If the increase in the NAO index between the 1960s and the 1990s is naturally induced, future changes in the NAO index could also show a reversal, which would lead to an increase in cold winters in Europe. But if the development in the NAO is anthropogenically forced, this could lead to a rapid change of the European winter climate towards higher temperatures and even more severe changes in heavy precipitation. The impact of this scenario is not very well represented by the models, due to the difficulties of reproducing the observed NAO increase (Scaife et al. [2008]).

Looking at regional extremes, one can also find areas which show the opposite trend to the one globally observed. The paper by Portmann et al. [2009] which concentrates on the changes in North America, shows one such case on the example of South East North American region (SENA). In this work SENA is defined as spanning from 30-40 degree N and from 100-70 degree W. The records for SENA from 1950-2006 show a negative trend in the daily maximum temperature. This changes in daily maximum temperature are highly anti-correlated with mean precipitation. The highest correlation between maximum temperature and mean precipitation occurs in May-June and the weakest or no correlation is shown in November-December. In the paper by Portmann et al. [2009] the cause of the "warming hole" (see Kunkel et al. [2006], Pan et al. [2004]), as this phenomena is referred to by the climate community, has not been fully addressed. The authors however, state a few hypotheses. One hypothesis tries to attribute the changes to land use changes in this regions, which would affect the concentration of biogenic aerosol and the hydrological cycle. Other possible explanations include an enhanced secondary aerosol effect, caused by a high aerosol concentration due to the dense population, or regional scale circulation effects.

1.4 Detection and attribution

Changes in climate can occur due to changes in the climate mean state or its variability, or both (see Fig. 1.3).

Many studies have been undertaken which aim to test if observed changes in climate are significantly different from climate variability. This process is referred to as detection. The detection analysis determines whether the observed change could have happened in response to a certain external forcing, that can be retrieved from climate model simulations or from physical understanding. A change is detectable if the likelihood of this change happening just due to internal variability alone, is evaluated to be small (Hegerl et al. [2007], Hegerl et al. [2010]). If the detection analysis fails to detect a certain signal, this could have happened due to various reasons, for example, the signal could be weak and the internal variability large in comparison, or it could be that the signal has not occurred. Detection does not attribute a cause to a detected change. Attribution is the process which tries to establish the most likely cause of the detected change at a defined confidence level.

Commonly the response to the same forcing from climate model simulations is used to characterise the expected response to a certain forcing. This response is often also referred to as fingerprint.

Detection and attribution studies have shown that the changes in mean surface temperature during the second half of the 20th century, on the global and hemispheric scale (Hegerl et al. [1997], Tett et al. [1999]) are detectable. The response is also detectable across all inhabited continents (Stott [2003], Zwiers and Zhang [2003]) and for Antarctica (Gillett et al. [2008]). Studies on the detectability of seasonal, rather than annual or decadal changes show that these can not be explained just by natural variability alone (e.g. see Jones et al. [2008]).

Jones et al. [2008], for example, analyse the changes in summer mean temperature during the last century and Zhang et al. [2006] focus on changes in seasonal temperatures.

These findings are not just limited to the global or continental scale, but have also been found in smaller scale studies, for example in the work Wu and Karoly [2007] which exhibits a detectable warming signal for regions of about 500km scale. Further Karoly and Stott [2006] found a detectable increase in temperature in the Central England temperature record and Dean and Stott [2009] found the New Zealand warming signal to be significant. However, due to the lower signal to noise ratio on regional scales, it is sometimes very difficult to attribute the changes to the possible sources of forcing.

There are also a few detection and attribution studies that deal with the changes in temperature extremes in both observations and models, which started with Kiktev et al. [2003], who showed, using atmosphere-only models, that the increase in warm extremes is more realistically simulated in a model forced with an increase in greenhouse gases. Hegerl et al. [2004] used detection and attribution in a model only study to show detectability of changes in temperature extremes and precipitation. They find that modelled changes in temperature extremes at many locations are significantly different to changes in mean temperatures and therefore a detection analysis of the changes in the mean should not be used to replace a detection analysis of extremes. Furthermore Christidis et al. [2005] show that the changes in warm and cold night time and in cold day time extreme temperatures during the second half of the last century are detectable. Changes in warm day time temperatures, however were not found to be significant at that time. A more recent paper by Christidis et al. [2011] shows that also the changes in warm day time temperature are significantly different from changes which would be expected solely due to natural climate variability. Christidis et al. [2011] and also Zwiers et al. [2011] are using generalised extreme value approach for their detection analysis. They fit a generalised extreme value (GEV) distribution to their observation, in order to account for the actual distribution of the extreme temperatures, instead of assuming that the noise on top of the signal shows a Gaussian distribution. They argue that this is a more suitable approach when dealing with extreme values, as they often show a skewed distribution. They find detectable changes in annual maxima of daily minimum and daily maximum temperatures, and in annual minima of daily minimum and daily maximum temperatures on global scale and also for some regions.

1.5 Outline of the Project

The observed changes in extremes as found by Alexander et al. [2006], Tebaldi et al. [2006], Karl et al. [2008] and Field et al. [2012] raise a number of questions. This project “The Impact of externally forced Changes on Temperature Extremes” concentrates on the following: Can external forcing explain the recent changes in climate extremes that can be seen in observational records, or did these changes arise from climate variability.

The term external forcing stands for both natural and human-induced forcing which have an impact on the climate. Natural drivers of the climate system are for instance volcanic eruptions, natural changes in tropospheric aerosols (e.g. natural concentrations of carbonates, nitrates, or soil dust) and changes in solar intensity. Human-induced or anthropogenic influences with a positive radiative forcing are for example, an increasing emission of long-lived greenhouse gases, such as carbon dioxide or methane, increases in the concentration of tropospheric ozone or changes in surface albedo, such as black carbon on snow. The term internal climate variability as used throughout this work refers to changes in climate which occur in absence of any external forcing. These values are estimated from data resulting from control simulations using climate models.

A secondary question is, if the signals are detectable, are they consistent with climate model simulations?

Furthermore, it is investigated whether changes in extremes are only driven by changes in means, which would imply that an increase in temperature extremes only results from an increase in the mean (see Fig.1.3 subfigure (a)), or whether this development is also influenced by a change in variance (see Fig.1.3 subfigure (c)).

In order to identify the cause of changes in observational records, the influence of external forcing needs to be quantified and possible alternative explanations such as internal climate variability have to be ruled out.

The work focuses on changes in temperature extremes and their variability around the world since the mid-20th century.

It consists of three sub-studies (chapter 2-4, see also Morak et al. [2011], Morak et al. [2012a], Morak et al. [2012b]). All three of them deal with changes in daily extreme temperatures and focus on answering the questions above.

The project has been designed in order to complement and extend studies such as Kiktev et al. [2003], Hegerl et al. [2010], Christidis et al. [2005], Christidis et al. [2011], Zwiers et al. [2011]. It concentrates on a detection analysis of extreme temperature indices not only on the local scale but also across many regions. This is important as these small scale changes and their impacts vary a lot from region to region and have not yet fully been investigated. Going to smaller regional scales, however, raises the problem of increasing noise and a harder to detect signal.

A further improvement to previous works is also the half-annual (chapter 3) or seasonal (chapter 4) analysis which allows a statement about the inter-annual changes of these events throughout the analysis period.

The study, presented in chapter 4, shows one of the first investigations of changes in persistent warm daytime and cold nighttime events covering the entire Northern Hemisphere and additionally focusing on selected regions.

Chapter 2 shows a comparison study of changes in the annual frequency of warm nights in observations and model simulations of different global climate models (GCMs), as well as a simple detection analysis, using a multi-model fingerprint. The analysis has not only been done globally but also for selected regions, previously used for climatological studies.
Chapter 3 also presents a comparison analysis but this time concentrating not only on the changes in the annual frequency of warm nights but on changes in the frequency of warm and cold day and nighttime extremes during both winter and summer half-year. By separating the analysis period into a winter and a summer half-year, the cold months which show a generally higher natural variability are divided from the warm months with a smaller natural variability and allow a better detection of the actual signal in the changes of these extremes. Chapter 3, however only uses ensemble data of one GCM, as this was the only data available at the time. In this study an advanced version of the detection analysis used in chapter 2 has been applied.
Chapter 4 discusses the changes in frequency of persistent warm daytime and cold nighttime extremes during winter across the Northern Hemisphere. The study is restricted to the Northern Hemisphere as only the boreal winter season (December-February) is considered. It also shows a comparison analysis between observed changes and model simulated changes, using the same model data as chapter 3, followed by a detection analysis, similar to the one used in chapter 3.

Each of the three studies has been written up as a paper and submitted for peer review. Two of them have been accepted and published by the time of final submission of this thesis and one is in preparation.

This project deliberately does not discuss any changes in extreme precipitation events, as this would be a work on its own, which has actually been partially dealt with in the Master project by Katie Noake (Noake [2011], see also Noake et al. [2012] and Polson et al. [2012]).

1.6 Data

In this project, two types of data have been used, namely observed and model data. All data sets are regular gridded ones, which cover most parts of the global land-mask, or at least of the landmasses of the Northern Hemisphere. The grid size, however, varies between the data sets, and therefore a common grid had to be defined for each study and the data sets re-gridded, in order to allow a comparable analysis. Furthermore all model data sets have been masked in space and time to match the coverage of the used observational data. The temporal resolution of the data sets spans from daily mean to annual mean. Grid-boxes with no available observational data have been excluded from the analysis ensuring that model and observed data are compared across the same domain.

1. Observations:

    - Data sets of observed daily minimum and maximum surface land temperature which originate from the HadGHCND data (Caesar et al. [2006]). The data set spans the period 1950-2007 and has a grid resolution of 3.75° × 2.5°. The observations were binned in 100 kilometre intervals applying a interpolation method which fits a two-degree polynomial to the mean correlation determined for each 100km within a 2000 kilometre distance (on the basis of the interpolation method by Shepard [1968]).
    - CRUTEM3 data set of monthly mean, minimum, mean and maximum

surface temperature from the Climatic Research Unit (CRU) (Brohan et al. [2006]). The data covers the period 1850-present and has a grid resolution of 5° × 5°.

- Observational data set of the percentile extreme temperature indices TN10, TX10, TN90 and TX90 as described in section 1.3 . The data was provided by Jesse Kenyon, is of monthly temporal resolution and spans the period 1860-2007 (see also chapter 2, section 2.2). The grid resolution is 5° × 5°. The dataset is based on averaging station based indices lying within a 5° × 5° grid-box.
- Observational data set of the percentile extreme temperature indices TN10, TX10, TN90 and TX90 provided by the Hadley Centre. This dataset originates from the HadEX project, is of annual temporal resolution and covers the period 1950-2003 (see also chapter 2, section 2.2). The grid resolution is $3.75° \times 2.5°$ and the data have been gridded using the same approach as Caesar et al. [2006] (see also Shepard [1968]).

2. Model Data: This study did not involve running any global climate model, but concentrated on the data analysis and processing.

    - Daily minimum and maximum surface temperature data sets from the HadGEM1 model simulations (see Martin et al. [2006], Stott et al. [2006]). The set of HadGEM simulations consists of two ensembles with four members each, one forced with both anthropogenic and natural forcing, and one only forced by anthropogenic drivers. All span the period 1950-2005. Additionally a 1000 year long control run from a HadGEM1 simulation is used. The grid resolution is 1.25° × 1.875°. A detailed description of the forcing can be found in chapter 3, section 3.2.
    - Model simulated data of TN90 from five different climate models

from the CMIP3 archive were provided by Claudia Tebaldi and Julie Arblaster (Tebaldi et al. [2006]). The data consists of three single runs for each of the five different global climate models, GFDL2.0, GFDL2.1, MIROC, MRI and PCM1. All these climate models were forced using the 20C3M scenario. This scenario includes anthropogenic forcing, such as the observed concentrations of green-house gases, as well as natural forcing such as the recorded changes in volcanic aerosol or changes in solar activity for the 20th century. All runs are of annual resolution and cover the period 1950-1999. The grid resolution of all data sets is $2.5° \times 2.5°$, which does not correspond to the native grid of these models but is the common grid chosen by Tebaldi and Arblaster (Tebaldi et al. [2006]). A detailed description of the models used and an explanation of the 20C3M scenario can be found in chapter 2, section 2.2.

3. Derived Data: The following data sets were produced in the course of this thesis.

    - Percentile extreme temperature indices TN10, TX10, TN90 and TX90 at monthly resolution as computed from the daily minimum and maximum surface temperatures, from the HadGEM model simulations (see section 1.8).
    - Frequency and duration of warm and cold spell events in winter as derived from the daily observed minimum and maximum surface temperature data of HadGHCND.
    - Frequency and duration of warm and cold spell events in winter as derived from the minimum and maximum surface temperatures data sets from the HadGEM model simulations using the all-forcing scenario.

More details about the data sets used can be found in chapter 2, section 2.2, chapter 3, section 3.2 and chapter 4, section 4.2.

1.7 Regions of interest

Table 1.2: Regions used in chapter 2. Names and the latitudinal and longitudinal extent.

Name of Region	Latitude degree North	Longitude degree East
Global	-85/85	-175/180
Southern Australia	-45/-30	110/155
Alaska	60/85	-170/-105
Western North America	30/60	-130/-105
Central North America	30/50	-105/-85
Eastern North America	25/50	-85/-60
Norther Europe	45/75	-10/40
Mediterranean	30/45	-10/40
Northern Asia	50/70	40/180
Western Asia	30/50	40/75
Tibet	30/50	75/100
Eastern Asia	20/50	100/145
South-East North America	30/40	-100/-75
Central Europe	45/50	0/20

Besides analysing the changes in temperature extremes on global and hemispheric scales, some smaller, mostly Northern Hemispheric regions have been chosen to focus on more regional changes. These smaller regions vary slightly from chapter 2 to chapter 4.

Chapter 2 mainly concentrates on a selection of the so-called Giorgi regions (seeGiorgi and Francisco [2000]), which are commonly used in climate studies, plus the regions covering Central Europe and South-Eastern North America (see table 1.2). However these regions are quite small and the small sampling size causes the signal-to-noise ratio to decrease. Therefore, the regions chosen in chapter 3 are slightly bigger but the overall area they are covering is roughly

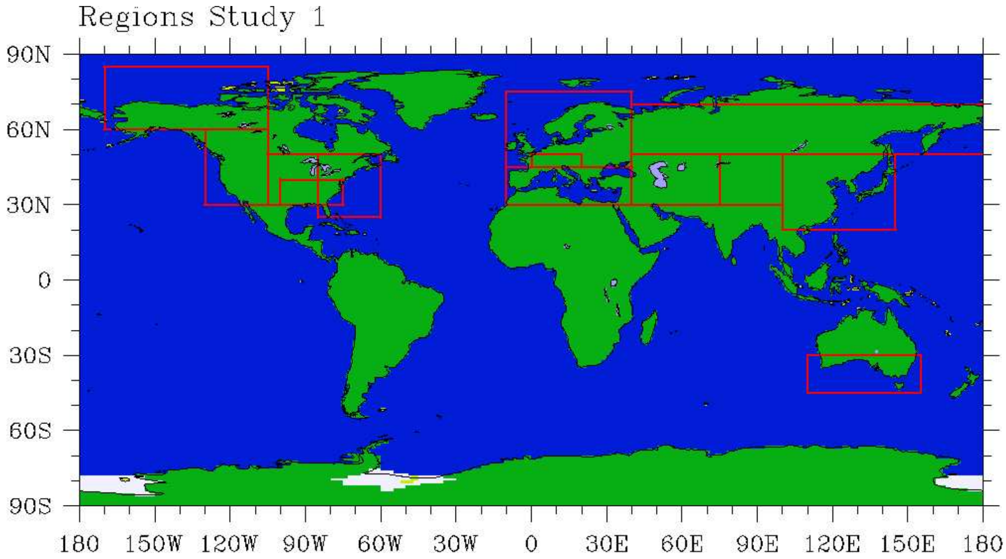


Figure 1.4: Regions as listed in Table 1.2, with the exception of the global region.

Table 1.3: Regions used in chapter 3. Names and the latitudinal and longitudinal extent.

Name of Region	Latitude degree North	Longitude degree East
Global	-90/90	-180/180
Northern Hemisphere	0/90	-180/180
Southern Hemisphere	-90/0	-180/180
Europe	35/75	-10/40
Southern Asia	10/45	40/180
Northern Asia	45/80	40/180
Australia + New Zealand	-60/-10	100/180
Western North America	25/55	-135/-100
Eastern North America	25/55	-100/-45
Northern North America	55/75	-165/-45

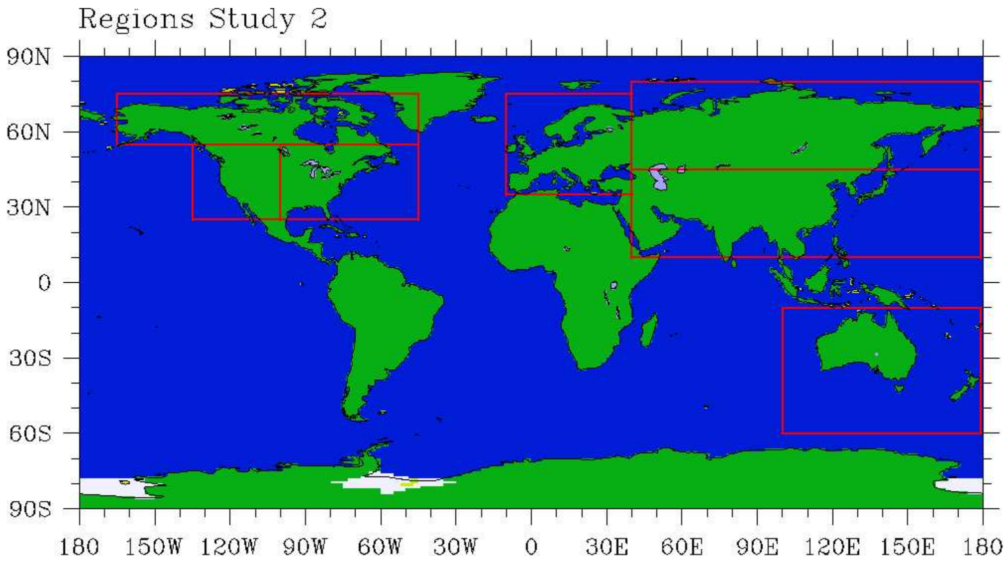


Figure 1.5: Regions as listed in Table 1.3, with the exception of the global and hemispheric regions.

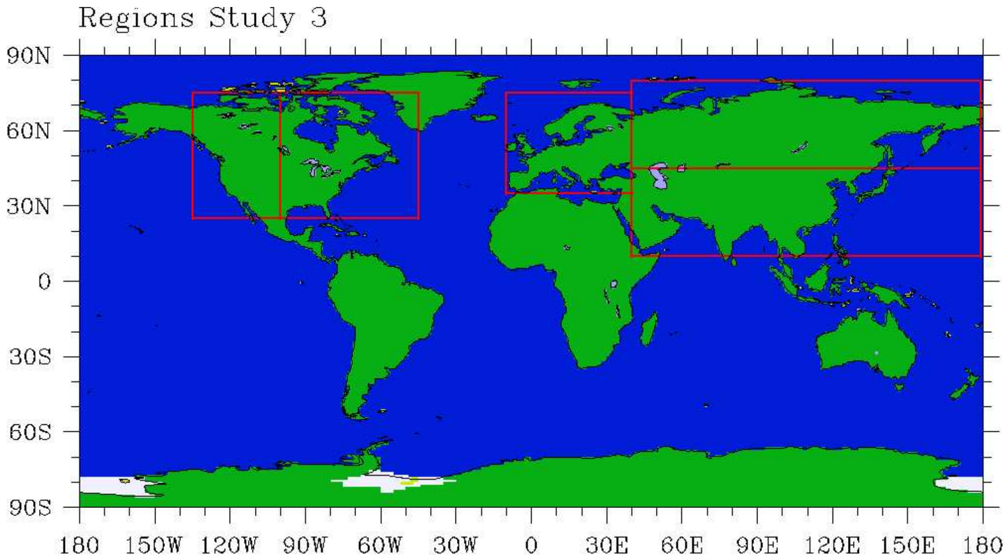


Figure 1.6: Regions as listed in Table 1.4, with the exception hemispheric region.

Table 1.4: Regions used in chapter 4. Names and the latitudinal and longitudinal extent.

Name of Region	Latitude degree North	Longitude degree East
Northern Hemisphere	0/90	-180/180
Europe	35/75	-10/40
Southern Asia	10/45	40/180
Northern Asia	45/80	40/180
Western North America	25/75	-135/-100
Eastern North America	25/75	-100/-45

the same (see table 1.3). The main changes are that the European regions from chapter 2 have been merged to one big region, Asia in chapter 3 is only split into South and North and the regions Central and South-Eastern North America from chapter 2 are not separately analysed anymore. In chapter 4 only one region really changes with respect to chapter 3, namely the region covering Northern North America. It is split into an eastern and a western part and added to Eastern and Western North America. Furthermore the Southern Hemispheric regions are excluded as this work is a Northern hemispheric study (see table 1.4).

1.8 Methods

This section describes the methods used to compute the various temperature extremes indices and to determine the detectability of the observed and modelled changes in those variables.

1.8.1 Computing the percentile temperature indices

This procedure aims to determine the monthly percentile extreme temperature indices which are computed by counting how often the daily maximum or minimum temperature exceeds a defined threshold.

To start with, the climatological base period has to be chosen, which is generally considered to be 1961-1990 or 1971-2000, depending on the analysis period. The distribution of the extreme values, for example, of the 3rd of January over the climatological period is determined by choosing all values within a 5-day window, centred on the 3rd of January, for each of the 30 years. The distribution is computed separately for each day of the year. A five consecutive day window is chosen in order to smooth out values of anomalous high variability, as the emphasise on extremely high or low values is suppressed by increasing the sample size. An alternative method, to chose the 5-day window, would be, to select 5 days, centred on the 3rd of January with a lag of five days in between each day. This way, one would possible get a sample of more independent values, as the values of days further apart from one another, are not expected to be that strongly autocorrelated.

After calculating the distributions and choosing a desired percentile threshold, the index for a certain month is given as percentage of days per month for which the daily values exceed the determined threshold.

The paper by Zhang et al. [2005] shows that when analysing the extreme value of a day within the climatological period, it is important to replace this day, or rather the 5-day window, with a randomly selected 5-day window from another year within the climatological period for determining the distribution. This way the day within the base period, is independent of the distribution it is compared to, like the days which lie outside the climatological period.

1.8.2 Computing the warm and cold spell indices

The duration of warm and cold spell events has been computed following the definition by the ETCCDI, which defines the warm spell duration index (WSDI) as 6 or more consecutive days above the 90th percentile of the daily maximum surface temperature of the considered base period. Correspondingly the cold spell duration index (CSDI) is defined as 6 or more consecutive days below the 10th percentile of the daily minimum surface temperature of the considered base period.

By counting the number of single events within the analysis period (e.g. the winter season as in chapter 4) the frequency of these events can be determined. The intensity or duration of these events within the analysis period (e.g. the winter season) can be computed by summing the number of days dedicated to these events. The base period selected for determining percentile thresholds used in the study in chapter 4 was 1961-1990.

1.8.3 Detection techniques

Detection methods aim to determine whether an observed change in climate is significantly different from changes solely due to natural internal variability.

Over the last 20-30 years various types of detection techniques have been introduced and used. One example is the optimal fingerprint method by Hasselmann [1993] (see also Hasselmann [1997], Allen and Tett [1999]) or the more recent optimal fingerprint method by Allen and Stott [2003], which is an advanced version of the one introduced by Hasselmann [1993] (described in more detail below). These methods aim to estimate the contribution of the forcing in question on the the signal.

Besides Hasselmann's approach (Hasselmann [1993]), there are also other detection methods such as the optimal weighted average by Bell [1986] whose goal it is to average the space-time results by suppressing small scale signals, to reach a better detection of large-scale changes in mean, or the optimal filtering, introduced by North et al. [1995] which aims to construct an optimal filter to determine the size of a given signal, e.g greenhouse warming within the noise contaminated data.

Throughout this work the detection method of Hasselmann [1993] (see also Hasselmann [1997], Allen and Tett [1999]) as well as that by Allen and Stott [2003] have been used. The main difference between these two methods is that Hasselmann [1993] uses an ordinary least square approach for his scaling factor calculation, whereas Allen and Stott [2003] chose a total least square one.
The prior assumption made in both this methods is that the present climate can be represented by a linear combination of the forced signal and the internal climate variability. The expected response to the considered forcing, also referred to as “fingerprint” is a known value and originates from model runs. The internal climate variability can be estimated, for example, from control run data, which is climate model data with no forcing.

The fingerprint method requires four different input data sets. Those include, one to-be-tested data set, in this study these are observations, a forced data set (the fingerprint), which is the ensemble average of a forced model simulated data set, as well as a control simulation, with two sets of realisations (one for the non-optimised case).

For the non-optimised analysis the noise in the original data has been reduced beforehand by applying a spatial or temporal smoother (depending on the kind of analysis undertaken) to the data. This avoids that the analysis uses small-scale features, which are generally less reliable. Furthermore the region size has been chosen to be big enough, to avoid a small signal-to-noise ratio, due to a small sampling size.

The ordinary least squares approach

In this method, introduced by Hasselmann [1993] (also used in Allen and Tett [1999]), the observations are expressed as

$$Y = \sum_{i=1}^{n} \alpha_i X_i + \nu_0 \tag{1.1}$$

where Y represents the observations (or the samples of noise), X_i stands for the i'th "fingerprint" in response to the respectively applied forcing and ν_0 is the observed internal variability. In a least-square estimate, the scaling factor α is expressed as

$$\alpha = \frac{X^T Y}{X^T X} \quad for \quad n = 1 \tag{1.2}$$

To estimate the uncertainty in the scaling factor the regression analysis is also applied to samples of noise, which were retrieved from the control run samples. The uncertainty range is then estimated by calculating the 5-95% range of the scaling factors arising from the individual noise samples using a t-distribution. If the confidence bar is inconsistent with zero, the observed change is detectable.

If the scaling factor is 1, it means that the fingerprint does not need to be scaled to match the observations.

All studies within this work only use a single fingerprint, so n = 1.

The total least squares approach

A total least square regression instead of an ordinary one is a more appropriate way to estimate the magnitude of fingerprints, if both X and Y are error contaminated (see Allen and Stott [2003]). The best fit in a total least square regression is therefore measured perpendicular to instead of along the axis, as done in the ordinary least squares.

$$Y = \sum_{i=1}^{n} \alpha_i (X_i - \nu_i) + \nu_0 \tag{1.3}$$

with ν_i representing the model internal variability of one forced ensemble mean and ν_0 again representing the observed variability.
In order to compute the scaling factor α the matrix Z has to be created, where

$$Z = [\frac{X_i}{fact}, Y_i] \tag{1.4}$$

where $fact = \frac{1}{\sqrt{ensemblesize}}$, serves as a scaling factor for the fingerprint, which has a smaller variance than the observations due to averaging over the ensemble members. In order to compute the regression coefficient with the best fit, we apply a singular value decomposition to Z and determine the smallest eigenvalue and its corresponding eigenvector, which give the solution that minimises the perpendicular distance from each point to the best fit.

$$Z = U \Lambda V^T \tag{1.5}$$

and after sorting V, into descending order we get the estimated n'th Eigenvector

$$\tilde{\mathbf{v}} = \mathbf{v}_{n'} \tag{1.6}$$

which is the n'th Eigenvector of $\tilde{v}$ and represents the best fit scaling parameter on the observations. This eigenvector $\mathbf{v}_{n'}$ is the slope line between the model data and the observations. To express it as scaling factor for the model data

$$\alpha_i = \frac{[\tilde{v}]_i}{[\tilde{v}]_{n'}} \tag{1.7}$$

The best fit "noise reduced" reconstruction of both observations and fingerprint can now be expressed as

$$\tilde{Z} = Z - Z\tilde{v}\tilde{v}^T \tag{1.8}$$

The uncertainty in this scaling factor(s) has been computed by adding random noise (which was provided by the control run samples) onto the best fit reconstruction of both the observations and the fingerprint using a bootstrap method. The 2.5-97.5% uncertainty range has been computed from the distribution of the computed scaling factors.

Finally the residuals of the regression analysis have been calculated in order to see whether they were lying within the 5-95% range of the model estimates of internal variability. Cases where the observed variability is greater than that simulated cannot be trusted. They have, therefore not been interpreted further.

Optimising

Optimal detection analysis, as introduced by Hasselmann [1993] (see also Hasselmann [1997], Allen and Tett [1999]), is used as it increases the signal-to-noise ratio by normalising the observations Y and the fingerprints X_i by the climate internal variability. For the optimised analysis the whole procedure starts by applying an EOF-analysis to one set of the control data. This is done to retrieve the "pre-whitening" operator P.

$$P = \Lambda^{-1/2}E \tag{1.9}$$

where Λ stands for the diagonal matrix of the eigenvalues and E is the matrix of the eigenvectors.

It is important that the chosen number of EOFs is not bigger than the number of individual samples from the control data, as this would lead to an underestimation of the higher EOFs (North et al. [1995]). In the next step the truncation level is determined. Truncation makes it possible to include only the EOFs which account for most of the explained variance and to exclude the ones that do not add any further information, or which reflect small scale variability not well reproduced in model simulations. This way the noise of the signal can be reduced. The truncation level can be determined either by checking the singular value spectrum for sharp cutoffs (see Hegerl et al. [1997]), which would indicate an appropriate truncation level, or by using a Chi-Squared test (see Allen and Tett [1999]), which determines up to which truncation level the observed variance is lying within the 5-95% range of the model variability. Fig.1.7 from the paper by Allen and Tett [1999], which focuses on detection and attribution of trends in atmospheric vertical temperature based on radiosonde measurements, illustrated how to determine an appropriate truncation level. The solid line shows the

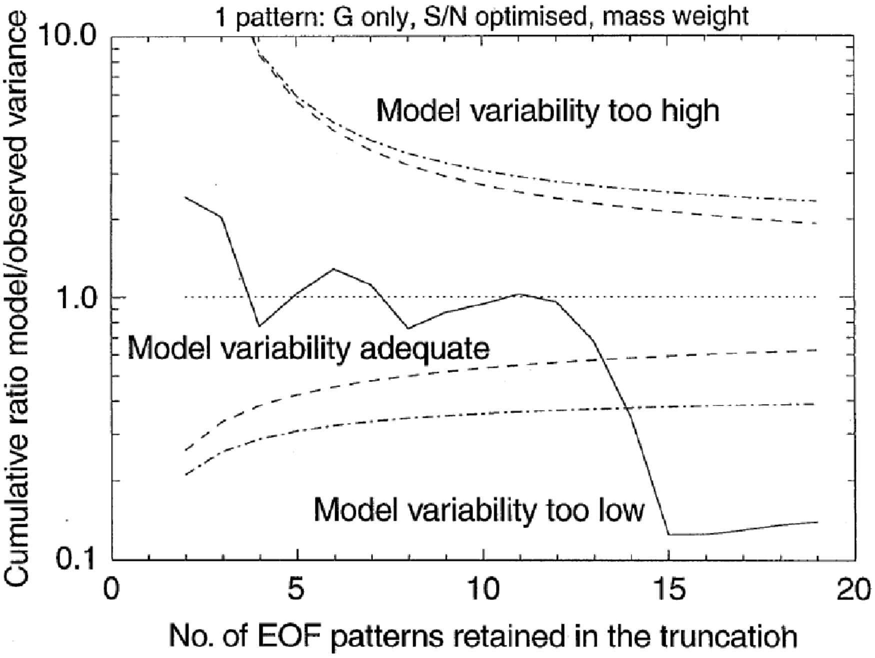


Figure 1.7: Solid line shows the change of the ratio of the cumulative model and observed residual variance with truncation. The dashed (dash-dot) lines give the 5-95% range for the Chi-Squared test (F test) (Fig. 5 from the paper by Allen and Tett [1999]).

change of the ratio of the cumulative model and observed residual variance with truncation. The 5-95% range is plotted for the Chi-Squared test (dashed) and the F test (dash-dot) approach. The truncation level is now determined by looking at the evolution of the solid line. As long as the solid line lies around one, which means that the observed and model simulated variability are about the same, the level is suitable for truncation. In this case the ideal range is 5-12. A truncation level bigger than 12 is not suitable as the residual is to large compared to the model estimate, indicating that the analysis is focused in unrealistically simulated or poorly sampled features.

After the evaluation of the truncation level, the observed and the finger print data sets are projected into the EOF space using the P operator.

$$y = PY \tag{1.10}$$

$$x = PX \tag{1.11}$$

where Y is the original observed data, y is the observed data after projection into the EOF space, X is the original fingerprint data and x is the fingerprint data after projection into the EOF space.

A practical example for the non-optimised ordinary least squares method can be found in chapter 2, section 2.3, one for the non-optimised total least squares case is shown in chapter 4, section 4.5 and an example for the optimised total least squares analysis is shown in chapter 3, section 3.4.

Chapter 2

Detectable Changes in Warm Nights

The paper equivalent to this chapter is authored by myself, Prof. Gabriele C. Hegerl (School of GeoSciences, The University of Edinburgh, Edinburgh, United Kingdom) and Dr. Jesse Kenyon (Nicholas School for the Environment and Earth Sciences, Duke University, Durham, North Carolina) and has been published in Geophysical Research Letter in 2011. During the preparation of this manuscript Prof. Gabriele C. Hegerl assisted me with her scientific advice and Dr. Jesse Kenyon provided the observed, gridded, data set of the extreme temperature index tn90 and proofread the manuscript.
The included version varies slightly from the published original to avoid unnecessary repetition.

2.1 Introduction

Studies of observational temperature records over the last 50-100 years have found evidence for an increase in both observed mean (Trenberth et al. [2007]) and extreme near-surface temperatures (Frich et al. [2002], Alexander et al. [2006], Caesar et al. [2006], Brown et al. [2008]).
Studying changes in temperature extremes, as well as attributing and predicting those changes, is of great importance as extreme temperature events can seriously affect human health, ecosystems and economy (IPCC AR4 WG1 et al. [2007], Karl et al. [2008], Christidis et al. [2010], Huynen et al. [2001]).

Several detection studies have shown (e.g., Allen and Tett [1999], Allen and Stott [2003], Hegerl et al. [1997], Hegerl et al. [2007], Tett et al. [1999], Stott [2003], Stott et al. [2010]) that observed changes in mean surface temperature since the mid-20th century, globally but also across some regions, can be largely attributed to an increase in greenhouse gases.
There are notably fewer studies that examine recent changes in temperature extremes (Kiktev et al. [2003], Christidis et al. [2005], Christidis et al. [2011] and Zwiers et al. [2011]).

This study analyses gridded observed and multi-model simulated trends in the annual number of warm nights during the second half of the 20th century.
Thirteen regions with a high density of observational data over two datasets are defined, for which the observed and simulated trends from 20th century simulations are compared. The main analysis period is 1951-1999, with a sub-period of 1970-1999. In order to investigate whether observed trends changed past 1999, the periods of 1955-2003 and 1974-2003 are also analysed.
Detection and attribution is used to identify if a detectable climate change signal is caused by external forcing.

This chapter is structured as follows: First the data sets and pre-processing methods are described, then the detection method used is introduced. Following, the results and possible causes are discussed and lastly the findings are summarised.

2.2 Data and Pre-processing

This study is based on a comparison of observed and modelled annual data of the index TN90. This index is defined as the mean number of warm nights with a minimum temperature exceeding the 90th percentile of the daily minimum temperatures of the climatological period 1961-1990 at a given location and over one year. We used five different coupled climate models, each of which consist of three individual runs, giving 15 model simulations and compare these to two observational datasets. For our analysis we selected eleven out of 32 regions introduced by Giorgi (Giorgi and Francisco [2000]), whose data coverage for TN90 is deemed sufficient to provide a credible regional average (see Table2.1, regions 1-11). We also analyse two smaller subregions, one in the South East of North America (SENA) and one in Central Europe (CEU), for a more detailed description of these two regions of interest see Table2.1 (regions 12 and 13).

Table 2.1: Table of regions . Column 1-3 show the number of the region, its acronym and its name. Column 4 and 5 list the latitudinal and longitudinal span of each region. Column 6 shows the correlation coefficient of inter-annual variation (trend subtracted) in observed TN90 with the observed annual mean Tmean. Column 7 explains how much of the observed trend in TN90 can be explained by the observed trend in Tmean. Bold numbers highlight the regions where at least 50% of the observed trend in TN90 can be explained by Tmean.

Nr	Acronym	Name of Region	Latitude degree North	Longitude degree East	$Corr$ T_{mean}	$\frac{Trend_{Tmean}}{Trend_{TN90}}$ in %
0	GLOB	Global Mean	-85/85	-175/180	0.8804	**72%**
1	SAU	Southern Australia	-45/-30	110/155	0.8254	**71%**
2	ALA	Alaska	60/85	-170/-105	0.8858	**91%**
3	WNA	Western North America	30/60	-130/-105	0.7890	45%
4	CNA	Central North America	30/50	-105/-85	0.7909	18%
5	ENA	Eastern North America	25/50	-85/-60	0.8038	13%
6	NEU	Northern Europe	45/75	-10/40	0.8630	**60%**
7	MED	Mediterranean	30/45	-10/40	0.7599	**53%**
8	NAS	Northern Asia	50/70	40/180	0.8814	**64%**
9	WAS	Western Asia	30/50	40/75	0.8419	**58%**
10	TIB	Tibet	30/50	75/100	0.7338	19%
11	EAS	Eastern Asia	20/50	100/145	0.8839	**90%**
12	SENA	South-East North America	30/40	-100/-75	0.7717	1%
13	CEU	Central Europe	45/50	0/20	0.7999	**75%**

1. Observational Data Sets: The two observational datasets are based on similar input stations (Kenyon and Hegerl [2008], Alexander et al. [2006]) but on different processing techniques to arrive at gridded data for TN90. Thus, they provide a first estimate of the role of processing uncertainty:
    (a) "Duke data set", covering 1886-2005: This data set, with a grid resolution 5° × 5° was produced by binning values of the TN90 index into grid boxes and then averaging to provide a grid box value. The station data are the same as used by Kenyon and Hegerl [2008]. The ETCCDI station data were provided from Lisa Alexander (Alexander et al. [2006]). The gridded values are based on a varying number of observations per grid and thus can be noisy if there is just a single or a few point measurements representing the value of an entire grid box. Furthermore there are quite a lot of unobserved regions in this data set, as grid boxes without stations remain blank.
    (b) HadEX Data Set (grid resolution 3.75° × 2.5°): This data set was produced by the Hadley Centre (Alexander et al. [2006]), by applying a spatial interpolation scheme to the ETCCDI index data. The HadEX data set shows a higher spatial coverage than the gridded station-based data set, but interpolates further between stations. This makes the data set much more spatially uniform, but also less anchored in close-by station data. The data set covers the years 1950-2003.
    For analysing causes of the observed change, we also used monthly mean daily minimum, mean and maximum surface temperature data from the Climatic Research Unit (CRU) (Brohan et al. [2006]).

2. Modelled Data Sets: The model simulated data sets were calculated from daily data from CMIP3 climate model simulations by Julie Arblaster and Claudia Tebaldi (Tebaldi et al. [2006]). We picked five models (listed below) that offered at least three single runs. We use three runs each in order to

give equal weight to each of the models in the multi-model mean. The data cover the years 1951-1999 and have been linearly interpolated to a grid resolution of $2.5^\circ \times 2.5^\circ$ by Tebaldi and Arblaster (Tebaldi et al. [2006]). All climate model simulations were forced with estimates of observed forcing over the 20^{th} century, including changing concentrations of greenhouse gases, tropospheric and stratospheric aerosols, changes in solar radiation and volcanoes, although forcing details and implementation vary between models. We use data of the following OAGCMs:

(a) GFDL-CM2.0 has an atmospheric resolution of the $2^\circ \times 2.5^\circ$ and L24. The oceanic resolution is $0.3^\circ - 1.0^\circ \times 1.0^\circ$.

(b) GFDL-CM2.1 has the same resolution as GFDL-CM2.0 but the atmospheric circulation is based on semi-Lagrangian transports.

(c) MIROC3.2 (hires) has an atmospheric resolution of T106 ($\sim 1.1^\circ \times 1.1^\circ$) and L56. The resolution of the ocean is $0.2^\circ \times 0.3^\circ$ and L47.

(d) MRI-CGCM2.3.2 has a atmospheric resolution of T42 ($\sim 2.8^\circ \times 2.8^\circ$) and L30, with its top at 0.4 hPa. The oceanic resolution is $0.5^\circ - 2.0^\circ \times 2.5^\circ$ and L23.

(e) PCM1 has an atmospheric resolution of T42 ($\sim 2.8^\circ \times 2.8^\circ$) and L26. Its oceanic resolution is $0.5^\circ - 0.7^\circ \times 1.1^\circ$ and L40.

For comparison with the Duke gridded indices, the model data and the HadEX data are re-gridded to a resolution of $5^\circ \times 5^\circ$ per grid box by linear 2D interpolation. Furthermore, the model data and the HadEX data are masked in space and time to reflect the sampling of the Duke data (referred to as "MASK DUKE"). The longest period that is covered by both observed and model simulated data is 1951-1999. Having 49 years of data, we first calculate the decadal trend for each $5^\circ \times 5^\circ$ grid box by fitting a slope-line with a least square fit (Fig.A1). Linear trends are only fitted to grid-boxes where at least five years are available during the first and last decade of the trend period. In order to reduce the noise of the spatial trend pattern we apply a 5-point smoother to all trend patterns, by averaging each grid-box with its nearest neighbours, or a subset thereof, based on availability (see Fig.2.1 for an example of the smoothed pattern). After computing the trend values and after the smoothing we determine the spatial average of the trend for all 13 regions as well as the global-scale trend by averaging over 11 out of 13 regions (see Fig.2.3a, b). The South-Eastern North American region (SENA) and the Central European region (CEU) are excluded from averaging as they overlap with other larger regions.

2.3 Methods

We determined if the observed trend could be explained by the fingerprint of externally driven changes plus variability. The fingerprint is derived from the average of all climate model simulations used, while the deviation from the ensemble mean, which is distinct for each ensemble member, is used to estimate variability. For the fingerprint analysis we calculate how the spatial pattern of trends from the multi model mean (fingerprint f) needs to be scaled to best match the observed trends y. This is done for individual regions as well as for a global pattern of regional mean trends. The scaling factors (α) are estimated using least squares regression (we do not use an "optimal" fingerprinting method, see Hegerl et al. [2007] for a description of fingerprint methods):

$$y = \alpha f + res \tag{2.1}$$

$$\alpha = \frac{f^T y}{f^T f} \tag{2.2}$$

f denotes the fingerprint vector $(f_1 ... f_j)$, where f_i is the decadal trend value of TN90 of the multi-model mean for one grid box, y $(y_1 ... y_j)$ represents the decadal trend value of TN90 of the observations (either Duke or HadEX) or of the samples of variability (see below), where y_i is the decadal trend value of TN90 for one grid box. To determine the uncertainty range of this scaling factor we estimated scaling factors from samples of trend patterns associated with internal climate variability.

These samples for uncertainty due to internal climate variability $(i = 1, ..., j)$ are estimated from the model simulated variability of each individual model $model_i$, around the mean change $(\overline{model})$ using

$$uncertainty_i = [model_i - \overline{model}] * correction \tag{2.3}$$

The model uncertainty $uncertainty_i$ covers the model error combined with the internal variability. By applying the fingerprint analysis to all available values of "$uncertainty_i$", we get an estimate of the internal variability. The correction against bias due to subtraction of the mean is given by:

$$correction = \sqrt{\frac{n}{n-1}}, \tag{2.4}$$

with n being the number of individual simulations, 15 in this case. The uncertainty in α is estimated by calculating the 5-95% range of scaling factors arising from these individual noise samples, using a t-test with 14 degrees of freedom (see von Storch and Zwiers [1999]). If the scaling factor α, estimated from the observations, is significantly larger than explained by noise externally forced climate change is detected. If α is consistent with 1, given its uncertainty, the multi-model mean does not need to be rescaled to match the observations.

We compare the regression residuals "res" from the observations with the samples of uncertainty "$uncertainty_i$", for both time intervals 1951-1999 and 1970-1999.

For the long period we find that, in 9 out of 13 regions, the residuals lie within the range of model uncertainty. Exceptions are Central Europe (CEU), Eastern Asia (EAS), Western Asia (WAS) and the Tibetan Regions (TIB). Over the subperiod of 1970-1999, 11 out of 13 regions are within the range, with only the residuals of EAS and TIB outside the model range.

2.4 Results

The 1951-1999 decadal global spatial trend pattern predominantly shows increases in the number of warm nights (see Fig.2.1).

All regional trend values of both observations and the model ensemble means are positive (see Fig.2.3a), however, the spread of individual model simulations do include some regional negative trends (see grey bars in Fig.2.3a) . The spread of the model trend covers the observational trends in all regions, which underlines that the models perform reasonably well in explaining the observed values, taking into account internal variability.

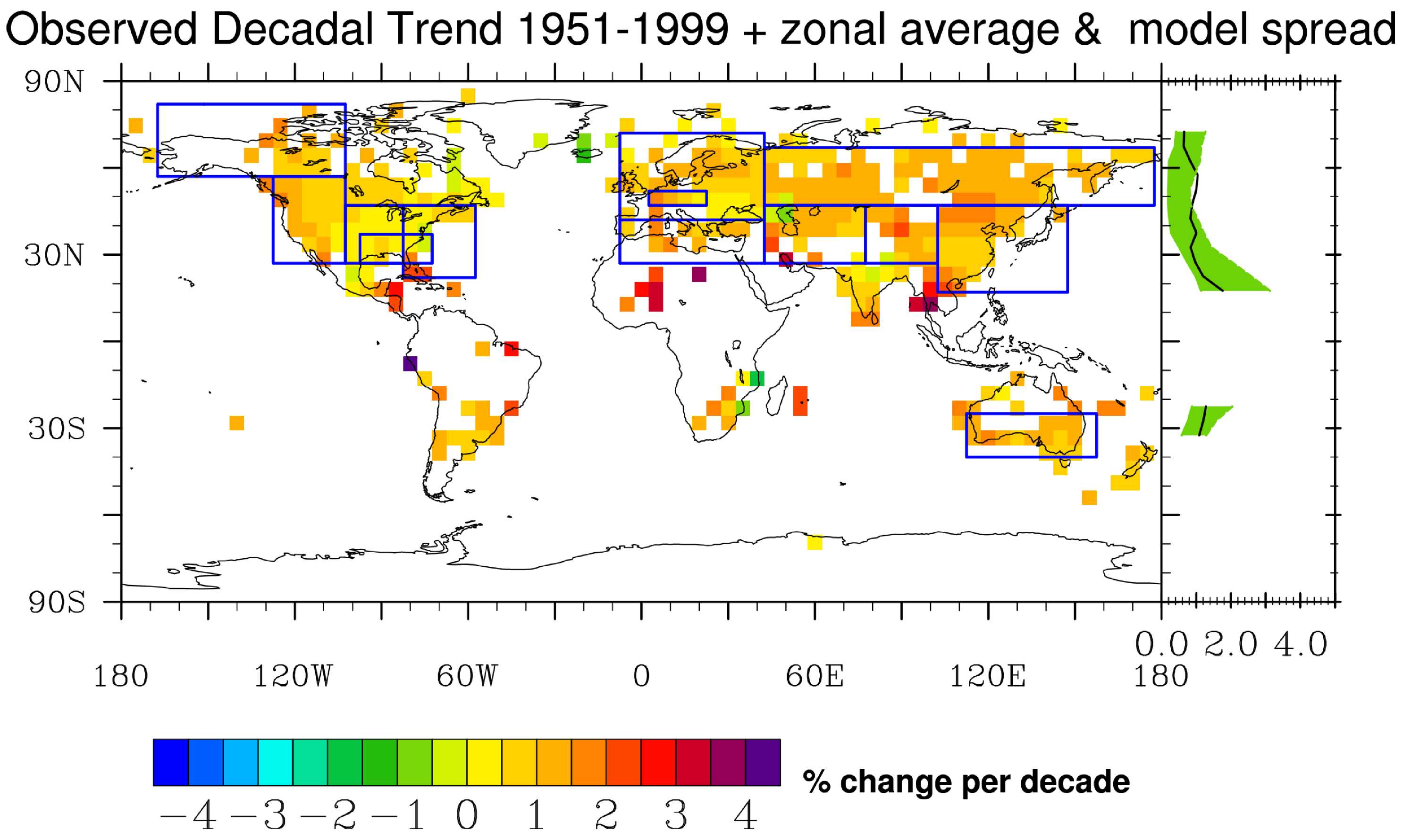


Figure 2.1: 1951-1999 observed decadal trend of TN90 (in % change per decade) derived from the Duke data set. The zonal average of the observations (black line) and the model spread (green shaded area), is shown on the side of the plot. The model spread is giving the minimum and maximum of the zonal mean values computed over all model runs.

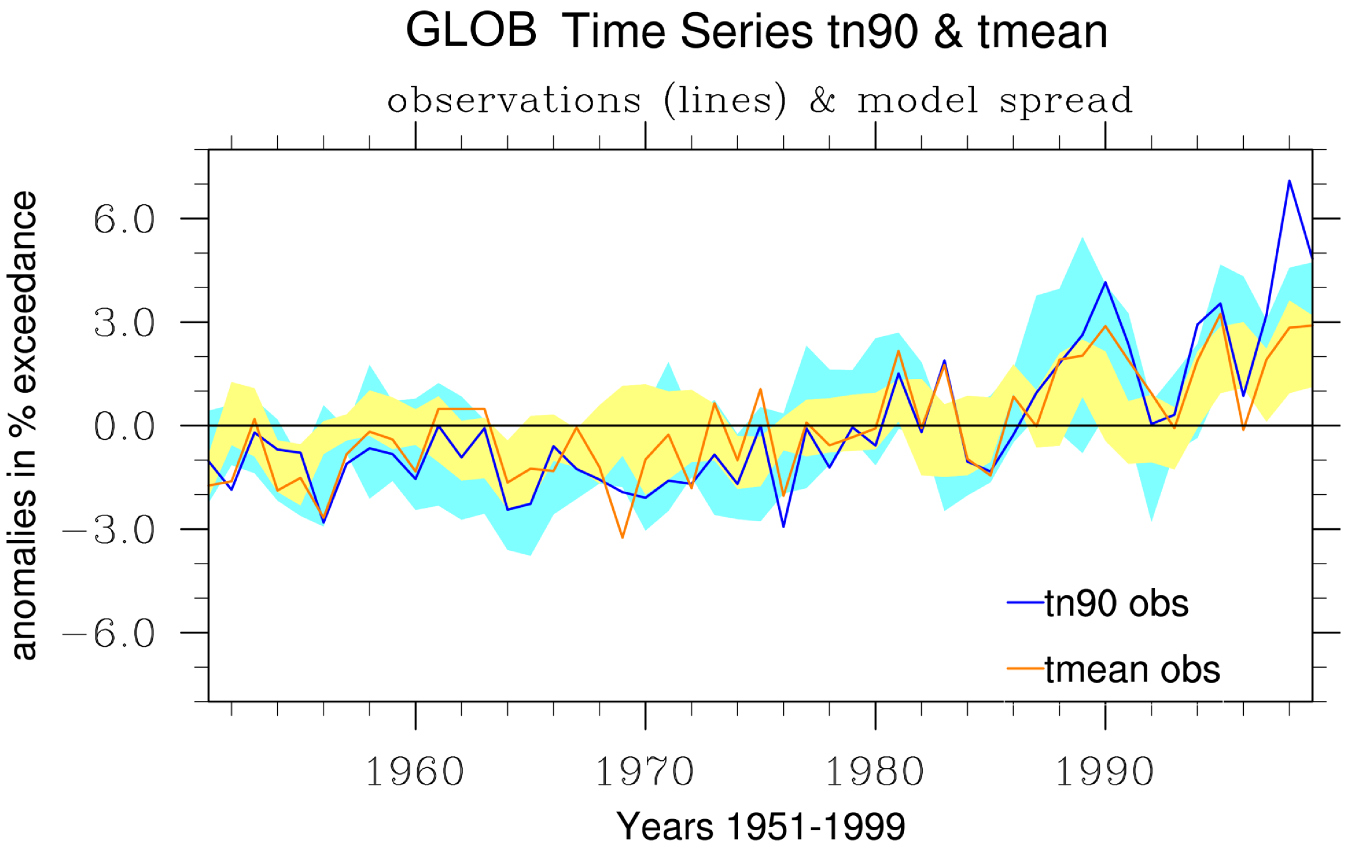


Figure 2.2: Global mean time series of observed TN90 (blue line) and regressed time series of observed Tmean (orange line) onto TN90. Shading shows the model spread of TN90 in blue and the regressed model spread of Tmean on TN90 for individual simulations in yellow.

The sub-period 1970-1999 (Fig.2.3b) shows an overall increase in magnitude of the decadal trend values compared to the period 1951-1999, which agrees well between models and observations (Fig.2.3a,b). The shorter period captures the temperature increase which followed a period of stagnation of the 1950s and 1960s. This stagnation or rather reduction in extreme values of TN90 during the 1960s can be seen in all regions, in both the observed and the modelled data (see Fig.2.2). This resembles changes for both observed and modelled data of the mean surface temperature.

As the 20^{th} century simulations only extend to 1999, but the observations extend at least up to 2003, we also computed observed trends over the periods 1955-2003 and 1974-2003. This allows us to assess whether more recent trends in the observed data show the same tendency as the earlier trends of 1951-1999 and 1970-1999 (see black markers Fig.2.3a, b). The results show an increase in trend for all regions, which is generally more pronounced for the 30 year period 1974-2003 (see black markers Fig.2.3b).

Fig.2.3c shows the results of the fingerprint analysis. 9 out of 13 regions show a fingerprint that is significantly detectable at the 5% level, since the scaling factors of the observations are significantly larger than those from noise only, and of those 5 show a residual variability within the model range. For many regions the scaling factors are larger than one, which indicates that the observations show stronger changes than the model ensemble mean, however this difference is not significant.

The results for sub-period 1970-1999 are very similar to those of the entire period (see Fig.2.3d). We lose the ability to detect a change in the Southern Australian (SAU) region, but now detect changes in Northern Europe (NEU) and most regions show variability within the model range. For assessing if changes in warm nights are detectable globally, we performed the regression on a vector of regional means for all 11 non-overlapping regions using area weighting and find a highly significant change both for the long and the short period (see Fig.2.3c and d, at

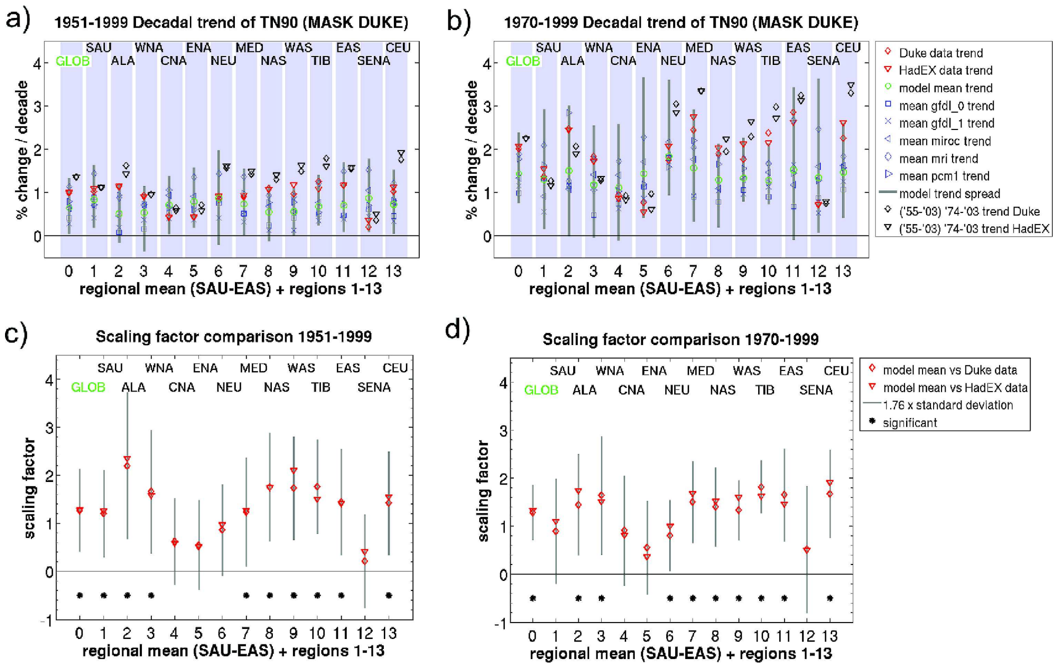


Figure 2.3: Spatially averaged trend of TN90 from observations (red markers) and model ensemble means (blue markers) as well as the spread of individual simulations (grey bar) for the selected regions (labelled on top and depicted against values 1-13 on x-axis) and the regional mean (labelled "GLOB", and plotted against '0' at x-axis) for the periods (a) 1951-1999 and (b) 1970-1999 (in % per decade). The black markers show results for the observed trend (a) 1955-2003 and (b) 1974-2003 . Figure c) and d) show the scaling factors (red markers) of observed changes onto the multi-model mean fingerprint for the period c) 1951-1999 and d) 1970-1999. Its estimated 5-95% uncertainty range is shown by the grey bar, which has been placed around the scaling factors for the Duke data. Regions with significantly detectable trend (5%) are marked by an asterisk (*)

$x = 0$, labelled "GLOB"). If the trends ending in 2003 are regressed onto the model fingerprints (for the period ending in 1999), detection results are similar.

2.5 Possible Causes

Having found a significant externally forced increase in the number of warm nights during the second half of the 20th century, a question arises, what caused the change?

Over the same period as considered here, global daily mean temperature has increased (Trenberth et al. [2007]), along with a rise in the number of warm days between 1951-1999 and a decrease in the diurnal temperature range (DTR), Vose et al. [2004].

In order to relate these changes to each other, we investigated how monthly mean, daily mean, minimum and maximum temperature are related to the number of warm nights on inter-annual to inter-decadal timescales. Table 2.1, column 6, shows a tight correlation of de-trended TN90 with annual mean daily mean temperatures for all regions and global-scale data. TN90 correlates most strongly with the observed mean surface temperature, followed by the observed annual daily mean, minimum and maximum temperature (see Fig.A3). The northern regions, ALA, NEU and NAS show a higher correlation of TN90 with the observed annual daily mean maximum than with the mean minimum temperature (see Fig.A3). In order to see what fraction of the significant changes in TN90 can be explained by changes in the mean surface temperature, we first regress the time series of annual mean daily mean temperature of each region on the one of TN90. Both time series have been de-trended before the regression analysis in order to base the connection on well-sampled inter-annual fluctuations only. The values of the regression coefficients are used to scale the raw time series of Tmean to determine which aspects of the changes in TN90 are explained by changes in Tmean. If the same physical connection between mean and extreme temperature operates on trend-timescales as on the shorter timescales, a large part of the trend in TN90 would be explained by the trend in mean temperature. This is found to be the case for global-scale data and many regions for both observations and model data (Fig.2.2, see also Fig.A4). Table 2.1 shows that for most regions more than half of the observed trend is explained by the observed trend in Tmean.
The smoothed global-scale spatial pattern of trends in TN90 (mean subtracted)

correlates, at 0.7 with that of trends in mean temperatures over the same time period (pattern not shown). This indicates that regions where mean temperatures warm stronger than the large-scale mean also tend to show stronger trends in TN90.

2.6 Discussion and Conclusions

This study shows an observed increase in the frequency of warm nights across large parts of the world since the mid-20th century. This tendency is well captured by diverse climate models, however the spatial pattern of the models does vary from the observations, which could be due to a limited resolution of the climate models and local effect that might not be resolved.

Fingerprints for the response to external forcing are detected in 9 out of 13 regions, showing that the observed trend can not be explained by internal variability alone. The residuals, unexplained variability from observations is consistent with the range of inter-model variability for most regions, which leads to a robust detection of changes globally and for SAU, ALA, WNA, MED and NAS and additionally, for the 30-year trend only, for WAS and CEU.

Based on a regression of mean temperature on the number of warm nights, most of the observed trend in the number of warm nights globally, as well as for many regions, is predicted by changes in mean surface temperature.
Studies by Hegerl et al. [2007] and Stott et al. [2010] assessed that greenhouse gases very likely play a key role in the positive trend in the global and continental mean temperature records of the 20th century. This assumption, allows the application of a multi-step attribution (see Hegerl et al. [2010]) in this particular study, which suggests that the global increase in the number of warm nights is probably in part due to anthropogenic influences. This may also be the case for

regions with detectable changes, SAU, ALA, WNA, MED and NAS. However, as regional changes can also be affected by small-scale forcings and are more difficult to attribute to causes (see Hegerl et al. [2007], Stott et al. [2010]), the multi-step attribution is only conclusive for global scale data.

Chapter 3

Detectable Changes in Temperature Extremes

The paper equivalent to this chapter is authored by myself, Prof. Gabriele C. Hegerl (School of GeoSciences, The University of Edinburgh, Edinburgh, United Kingdom) and Dr. Nikolaos Christidis (Met Office Hadley Centre for Climate Change, Exeter, United Kingdom) and was submitted for publication to Journal of Climate. During the preparation of this manuscript Prof. Gabriele C. Hegerl assisted me with her scientific advice and Dr. Nikolaos Christidis provided model data from HadGEM1 simulations for the analysis and proofread the manuscript. The included version varies slightly from the published original to avoid unnecessary repetition.

3.1 Introduction

Detection studies by Kiktev et al. [2003], Christidis et al. [2005], Christidis et al. [2011],Zwiers et al. [2011] and Morak et al. [2011] found evidence of significant changes in the intensity and frequency of temperature extremes around the world. Most of these studies, however, concentrate on changes in annual extremes and therefore loose any information on inter-annual changes. Besides an analysis of inter-annual changes of temperature extremes, also an additional regional analysis is important, as the occurrence of these extremes is often affected by regionally specific climatic features, which are neglected in global or hemispheric analyses.

This study improves and extends the work by Morak et al. [2011], by examining not only the changes in the annual frequency of warm nights, but changes in both the warm and cold tails of the daily minimum and maximum temperature distribution during boreal cold (ONDJFM) and warm (AMJJAS) season separately. The work concentrates on changes from global and hemispheric scales to regional scales. It shows a comparison analysis of observed and model simulated changes in these events, as well as a detection analysis and examines the benefit of using an optimised approach rather than a simple total-least squares one.

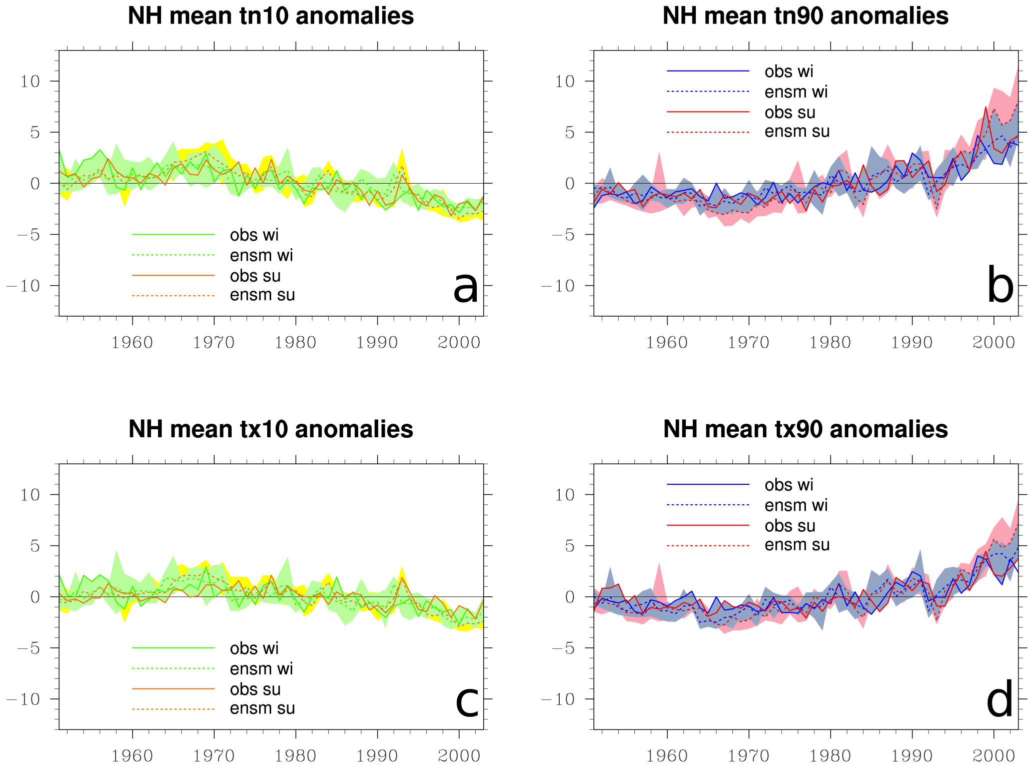


Figure 3.1: Northern hemisphere mean a) change in the frequency of cold nights (tn10), b) warm nights (tn90), c) cold days (tx10) and d) warm days (tx90). Changes are expressed as anomalies relative to the average over the period 1951 to 2003 and are expressed in % change in the frequency of days. Solid lines show observed anomalies, dashed lines represent the ensemble mean anomalies, and the shaded areas show the ensemble spread (max minus min). Changes in cold extremes (tn10 and tx10) during the boreal cold season (ONDJFM) are shown in green (light-green shading) and those during the boreal warm season (AMJJAS) in orange (yellow shading). Changes in warm extremes (tn90 and tx90) during the boreal cold season are displayed in blue (grey-blue shading) and those during the boreal warm season in red (light-pink shading).

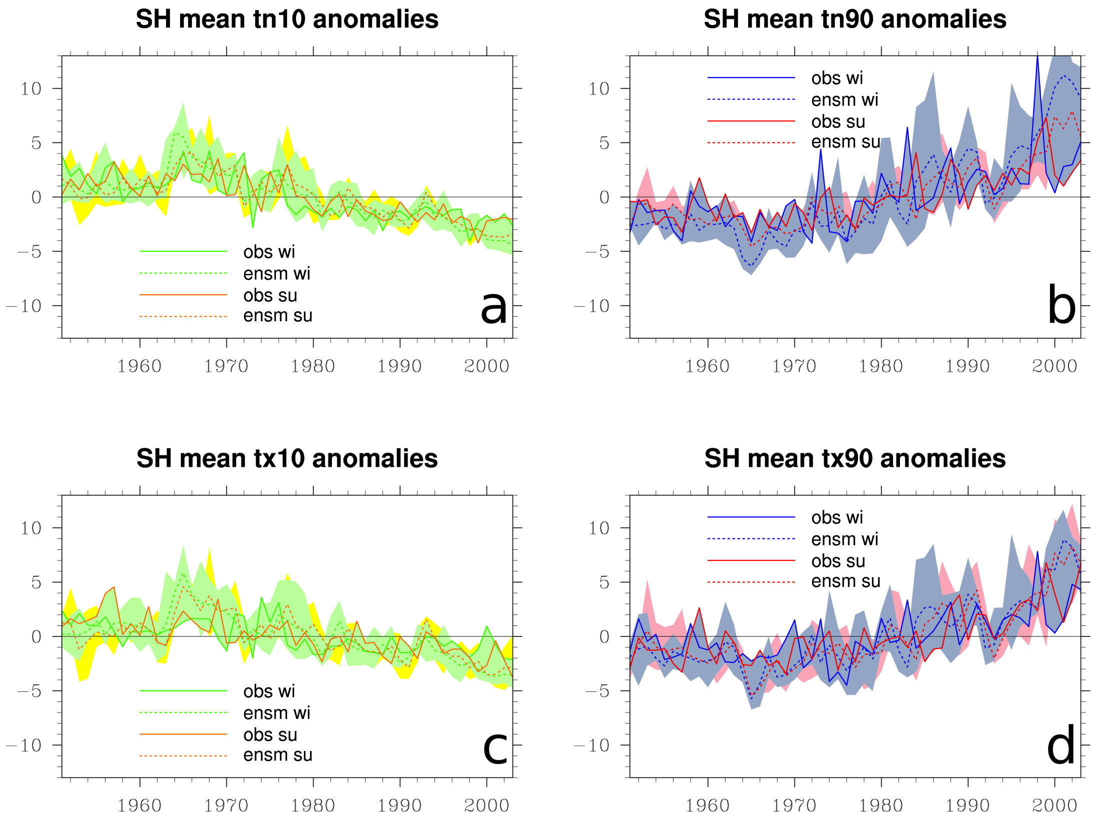


Figure 3.2: As figure 1, but for Southern hemisphere mean a) tn10, b) tn90, c) tx10 and d) tx90 anomalies.

The work is structured as follows: First, the data and processing methods are described (section 3.2). Then the observed and model simulated changes in the frequency of cold and warm extremes on global and regional scales are compared (section 3.3). The detection analysis is described in section 3.4. Results are presented in section 3.5, followed by a summary of the study and conclusions (section 3.6).

3.2 Data and Processing

In this work, we compare observed and climate model simulated trends in mean values of temperature extreme indices splitting the year into the dynamically active boreal cold (ONDJFM) and warm (AMJJAS) season (as used in studies by Meehl et al. [2004], Cook et al. [2011]). We define our indices relative to the 10th and 90th percentiles of the daily maximum and minimum temperatures during the base period 1961-1990, which we use to compute threshold values. Index tn10 (frequency of cold nights) is defined as the percentage of days per month where the daily minimum temperature does not reach the 10th percentile of the base period, and tn90 (frequency of warm nights) represents the percentage of days where the 90th percentile is exceeded. Similarly, tx10 (frequency of cold days) is defined as the percentage of days per month where the daily maximum temperature does not reach the 10th percentile of the base period and tx90 (frequency of warm days) is the percentage of days where the 90th percentile is exceeded. The indices have been computed separately for each grid box, using the grid box specific climatology. The advantage of using percentile indices, rather than absolute values, is that outliers do not have a strong impact on the index values, as an outlier only counts for one exceedance and does not introduce a bias due to its actual magnitude or anomaly. The use of percentile indices also allows for comparison of changes across climatologically different regions (see Alexander et al. [2006]).

This study uses one set of observed gridded indices as well as indices computed from daily minimum and maximum temperatures from climate model simulations that have been driven with both natural and anthropogenic forcings. To determine variations in extremes due to internal climate variability we also use data from a 1000 year long control run without forcings apart from the seasonal cycle (see below for more details).

1. Observational Data: The observed gridded data sets of tn10, tn90, tx10 and tx90, which cover the period 1886-2005, were produced by binning the monthly values of the indices, computed from station data, into $5° \times 5°$ grid boxes (see Morak et al. [2011]). The index station data are the same as used by Kenyon and Hegerl [2008] and were produced as part of the HadEX project (Alexander et al. [2006]). The values of each grid-box are based on a varying number of observations and can therefore be noisy, especially if the value from a grid-box represents just a small number of measurements. Unobserved regions remain blank, i.e. no interpolation routine has been applied to fill them. A previous study (Morak et al. [2011]) used this data set and the HadEX data set and found that the results were not sensitive to the observational data used. For this study, only data to 2003 are used since the spatial and temporal coverage drops substantially after 2003.
2. Model Data: We use modelled daily minimum and maximum surface temperature data derived from simulations with the Hadley Centres ocean-coupled global climate model HadGEM1, which has a grid resolution of $1.25° \times 1.875°$. HadGEM1 is the first of the HadGEM family of models, which have a non-hydrostatic dynamical core and employ a semi-implicit, semi-Lagrangian time integration scheme (Davies et al. [2005]).

> Model simulations with HadGEM1 have been used in several detection and attribution analyses and also feature in the 4th Assessment Report of the IPCC (Christidis et al. [2012]; Hegerl et al. [2007]). The equilibrium climate sensitivity of HadGEM1 is 4.4 K (Stott et al. [2006]).These model data come from an ensemble of four 20th century simulations and a 1000 year long control run without any forcings (Martin et al. [2006], Stott et al. [2006]). :
>
> (a) The 20th century simulations include both natural and anthropogenic drivers, such as time varying volcanic aerosol and solar forcing, as well as changes in greenhouse gas concentrations, land use, anthropogenic aerosols and black carbon. We use data from the experiment for the period 1950-2005.
>
> (b) The control simulation does not include any forcing apart from the seasonal cycle. From the 1000 years of available daily data, we extracted thirty-two overlapping segments, each which are 53 years long. When performing the optimised analysis, these are split into one set of 16 chunks used to pre-whiten the data, and another to determine the uncertainty due to internal variability.

The percentile indices are computed using the daily maximum and minimum temperature data using the fortran code provided by ETCCDI (Klein Tank and Können [2003], Zhang et al. [2005], Alexander et al. [2006]). This computation produces monthly values of tn10, tx10, tn90 and tx90 indices from each of the four runs with all forcings, as well as the control experiment. A bootstrap method (Zhang et al. [2005]) is employed to avoid inhomogeneities between the climatological base period and the subsequent and precedent time.

After the index computation, the model data sets are re-gridded onto a 5° × 5° grid in order to compare to observations. All the model data sets, including the control simulation segments, are masked in space and time to reflect the data coverage of the observations.

In all figures, results based on the boreal cold season are labeled as “wi” and results for the boreal warm season are labelled as “su”.

In order to investigate the temporal change in the frequency of extremes, the regional mean time series of all indices, estimated with both observations and data from the model runs, and averaged over the Northern and Southern hemispheres, are plotted in Fig.3.1 and Fig.3.2, and are discussed below.

For the next part of the analysis, spatial trend patterns expressed as trend per decade for each index are computed separately for the boreal cold (winter half-years 1950/1951-2002/2003) and boreal warm season (summer half-years 1951-2003), from both observations and climate model simulations.
The linear trend for each $5^\circ \times 5^\circ$ grid-box is computed by fitting a slope line using a least squares fit and is expressed as percent per decade. The linear trend is only computed for grid-boxes where at least five years of data are available during the first and last decade of the analysis period. These linear trend patterns are shown in Fig. 3.3 to 3.6 for cold and warm extremes in winter and summer. Linear least-square trend patterns are often used in detection and attribution analyses that focus on determining causes of large-scale changes (e.g., Hegerl et al. [1997], Hegerl et al. [2004], Zhang et al. [2007]) and generally capture recent changes well. Although Fig. 3.1 and 3.2 suggest the trend would be stronger when focusing on a more recent period, the longer analysis period improves the signal-to-noise ratio by being less affected by internal climate variability (Morak et al. [2011] for tn90).

In order to reduce the amount of spatial noise and to focus on spatial scales larger than the grid-point scale, a 5-point smoother is applied to all data sets

Table 3.1: Table of regions used in this study. Column 1-3 numbers, lists the acronyms and gives the name of the ten regions. Column 4 and Column 5 give the latitudinal and longitudinal extent of each region.

Nr	Acronym	Name of Region	Latitude	Longitude
1	GLOB	Global	-90/90	-180/180
2	NH	Northern Hemisphere	0/90	-180/180
3	SH	Southern Hemisphere	-90/0	-180/180
4	EU	Europe	35/75	-10/40
5	SAS	Southern Asia	10/45	40/180
6	NAS	Northern Asia	45/80	10/180
7	AUS	Australia + New Zealand	-60/-10	100/180
8	WNA	Western North America	25/55	-135/-100
9	ENA	Eastern North America	25/55	-100/-45
10	NNAM	Northern North America	55/75	-165/-45

by computing the average of each grid-box and its adjacent four grid-boxes prior to plotting and the detection analysis. In order to illustrate whether the trend patterns of the individual ensemble members are represented well by the ensemble mean, grid- boxes where all four runs have the same sign are stippled in Fig.3.3b - Fig.3.6b. This is a simple nonparametric test of where the simulated trend pattern is robust relative to internal climate variability.

The detection analysis encompasses a range of spatial scales (global, hemispheric, regional) (see Table 3.1). Regions that are bigger than the more frequently used Giorgi regions (Giorgi and Francisco [2000]) are used, in order to increase the signal-to-noise ratio and adjust the analysis to data availability. However, our regions encompass many of the smaller Giorgi regions.

3.3 Observed and model simulated changes in frequency of extremes

The time series of Northern and Southern Hemispheric land average frequency in cold days and nights (see Fig. 3.1a, c and Fig. 3.2a, c) shows a decrease for both observations and model simulations which started in the 1970s and lasted until the end of the analysis period. The observed changes expressed in % change in the frequency of days, corresponds to a decrease of approximately 5-6 days over the analysis period. In the Northern Hemisphere, the decrease in the number of cold nights (tn10; Fig. 3.1a) is slightly larger than the decrease in the number of cold days (tx10; Fig. 3.1c) for both the boreal cold (green lines and light-green shading) and boreal warm season (orange lines and yellow shading). The changes in tx10 during the boreal warm season are smaller than the ones during the boreal cold season (Fig. 3.1c). The Southern Hemispheric mean time series (Fig. 3.2a and c) displays higher variability than the Northern Hemisphere equivalent, probably due to fewer land areas and poorer data coverage in the Southern Hemisphere.

Hemispherically averaged time series of the number of warm days and nights (tn90 and tx90) (Fig. 3.1b, d) show an increase since the 1970s. The observed changes expressed in % change in the frequency of days, corresponds to an increase of approximately 7-9 days over the analysis period. In the Northern Hemisphere, the warm nights index (Fig. 3.1b) shows a larger change than the number of warm days (Fig. 3.1d) during the boreal warm season. Changes in Northern Hemispheric tn90 and tx90 during the boreal cold season are smaller in magnitude than those during the boreal warm season (see blue lines and grey-blue shading in Fig. 3.1b, d).

tn90 and tx90 display much larger fluctuations in the Southern Hemisphere for both model and data (Fig. 3.2b, d). Observed changes in Southern Hemisphere tn90 during the austral warm season (boreal cold season) and austral cold season (boreal warm season) have a similar magnitude to changes in the Northern Hemisphere (compare Fig. 3.2b with Fig. 3.1b). However, observed changes in tx90 during the austral warm and cold season (boreal cold and warm season) across the Southern Hemisphere are larger than those of the Northern Hemisphere (compare Fig. 3.2d with Fig. 3.1d).

We now turn to an analysis of the spatial pattern of change in extremes.
The observed and simulated spatial trend pattern in the number of cold nights (tn10) during the boreal cold season (Fig.3.3A) shows an overall decrease that is evident in most parts of the world. The strongest decrease is observed across Central and Southeast Asia, and parts of Northern Asia (see Fig.3.3A, subfigure a). On the other hand, there are some grid-boxes across the Northeast Coast of the United States and South America where a weak increase is observed (see Fig.3.3A, subfigure a), even in regions where the model shows consistent (stippled) decreases. The model simulated ensemble mean trend pattern shows a large-scale decrease of cold spells that are similar to those observed in many regions (Fig.3.3A, subfigure b) . However, the model does not reproduce the very strong observed decrease over parts of Asia. (see Fig.3.3A).
The observed tn10 trend during the boreal warm season (Fig.3.3B, subfigure a) shows smaller changes than during the boreal cold season (Fig.3.3A, subfigure a). The largest observed decreases are found across most parts of Asia, while very little decrease is observed across Eastern Europe (Fig.3.3B, subfigure a). The simulated trends also indicate a large scale decrease in the frequency of cold nights.

The observed trend pattern for the frequency of cold days (tx10) in boreal winter (Fig.3.4A) also shows a general decrease in the number of cold days, with some pronounced regional exceptions, such as in Eastern North America, where a slight increase in the number of cold days is observed. Similar to the winter changes in tn10, the strongest observed decrease occurs over parts of Asia. The model simulated trend pattern shows an overall decrease, which is robust across ensemble members (stippled) even in some regions where the observations show a weak increase. Changes in both observed and model simulated tx10 (see Fig.3.4A) are generally of lower magnitude than those of tn10 (see Fig.3.3A).

The observed trend pattern for the number of cold days (tx10) during the boreal warm season shows a small decrease across most parts of the world (Fig.3.4B). As for the cold season, there are a few regions showing changes of the opposite sign, such as large parts of Eastern North America (where in some points the model, in contrast, shows robust decreases across all 4 ensemble members), South-Eastern Asia, and parts of Eastern Europe. The simulated ensemble mean pattern shows larger trends than observations over many regions and does not feature any increases (see Fig.3.4B, subfigure b).

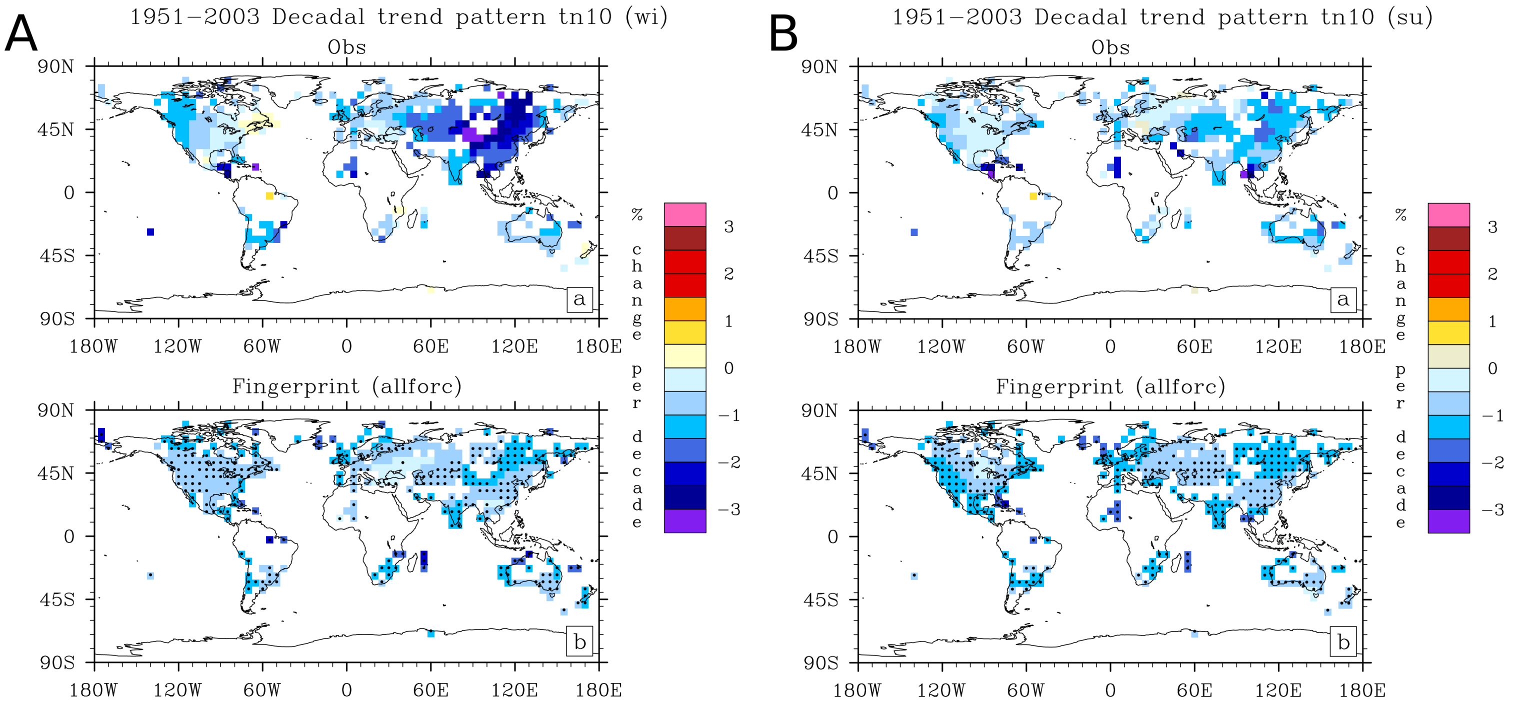


Figure 3.3: Spatial trend pattern [in % per decade] over the period 1951-2003 in the frequency in cold nights (tn10) during boreal A) cold and B) warm season for the observations (upper panel) and the ensemble mean (lower panel). The stipples superimposed on the ensemble mean trend pattern mark the grid-boxes where all four ensemble members agree in the sign of their trend. Spatial trend patterns have been smoothed by a 5-point spatial smoother (see body of paper) in order to focus on large-scale features.

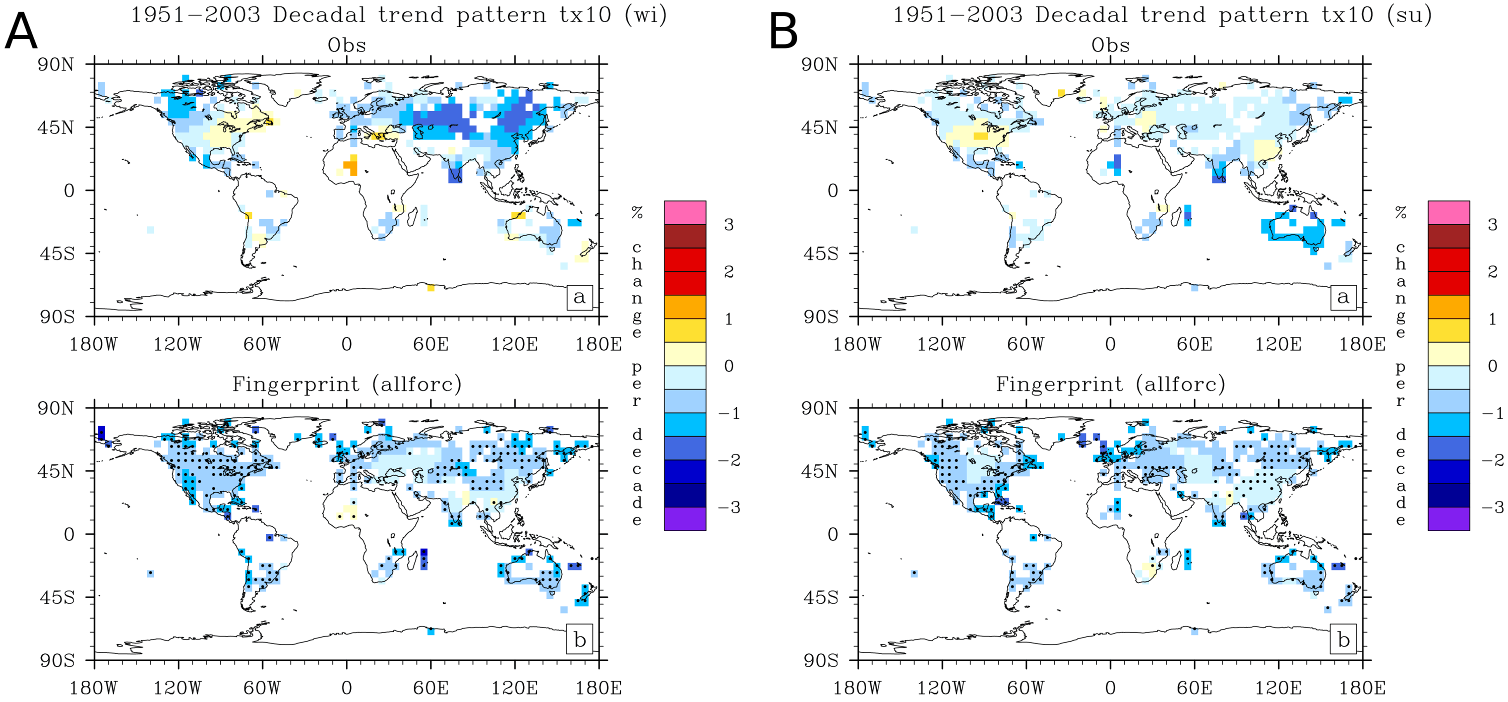


Figure 3.4: As figure 3, but showing the change in the frequency in cold days (tx10) during boreal A) cold and B) warm season for observations (top panel) and model ensemble mean (bottom panel).

The frequency of warm nights and warm days shows a general increase in models and observations (Fig. 3.5 and Fig. 3.6). The observed trend pattern of number of warm nights (tn90) during the boreal cold season shows an overall increase across most parts of the world (Fig.3.5A), that is strongest in Central Asia, Northern Europe, and over many tropical grid points. This enhancement of trends in the tropics was also observed in Morak et al. [2011] and is probably due to smaller climate variability in the tropics, leading to a tn90 threshold that is easier to exceed with warming in both models and observations. Smaller changes, and even a small area of slight decrease is observed in many parts of Eastern North America (see Fig.3.5A, subfigure a). The model manages to reproduce the overall increase in tn90, with changes that are robust across ensemble members over most grid points. However, the model does not reproduce the strong increase over Asia in winter.

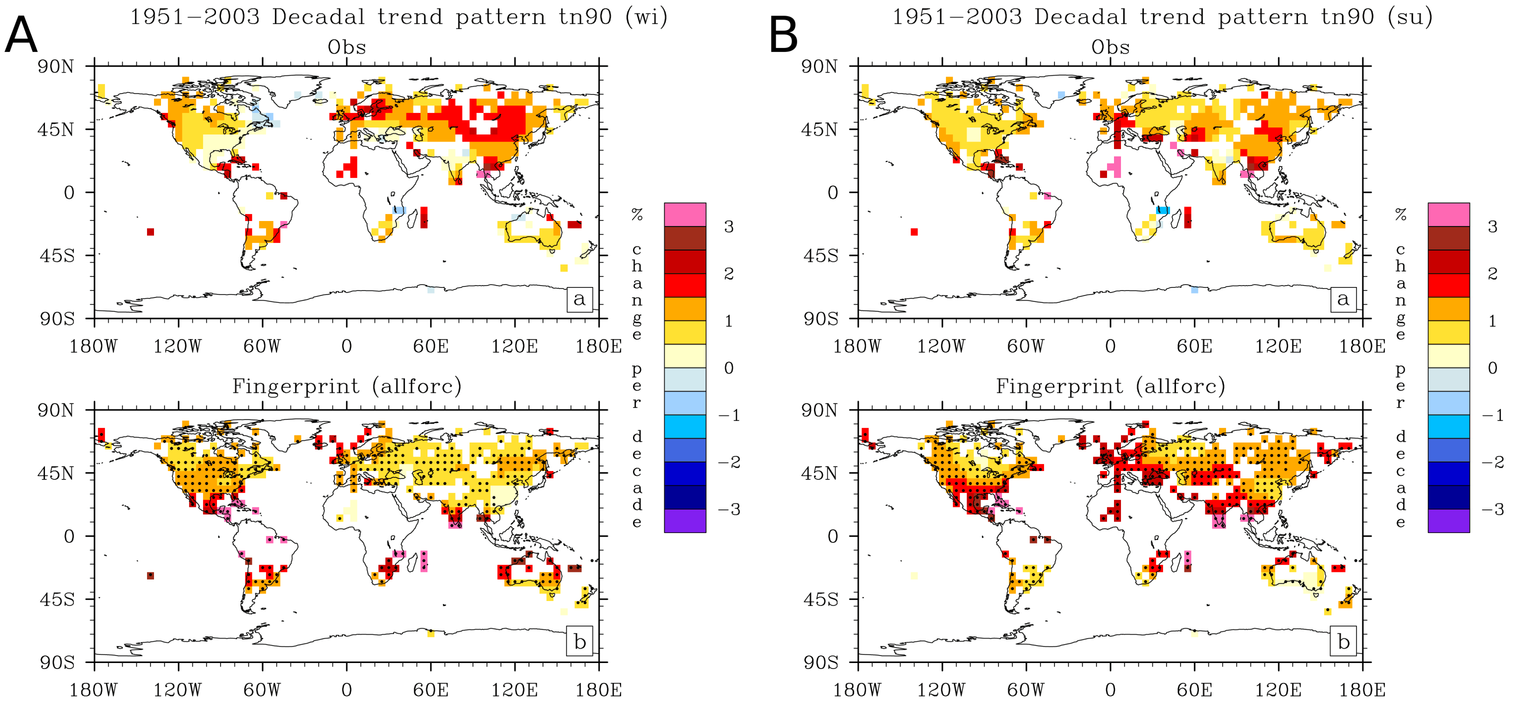


Figure 3.5: As figure 3, but showing the change in the frequency in warm nights (tn90) during boreal A) cold and B) warm season for observations (top panel) and model ensemble mean (bottom panel).

Fig.3.5B shows the observed and simulated trend pattern of tn90 during the boreal warm season. In some areas where the observed trend (Fig.3.5B, subfigure a) is smaller than during the boreal cold season (Fig.3.5A, subfigure a), the model indicates larger trends (compare Fig.3.5B, subfigure b with Fig.3.5A, subfigure b).

The trend pattern of the frequency of warm days (tx90) during the boreal cold season also indicates a widespread increase (see Fig.3.6A). However, this increase is smaller than the increase in tn90 (see Fig.3.5A, subfigure a). Trend values of small magnitude are observed in the United States and South East Asia, some of which are reproduced by the model. The simulated changes in the frequency of warm days are larger at low latitudes (see Fig.3.6A, subfigure b), and larger than the observed ones in parts of North America. Observed trends in parts of western Australia in the austral summer are much smaller and are even regionally negative in the observations.

The observed trend (see Fig.3.6B, subfigure a) in the tx90 index during the boreal warm season is smaller than that observed during the boreal cold season (see Fig.3.6A, subfigure a), with significant areas of decrease in the frequency of hot days, including parts of North America and Asia. The ensemble mean from the model does not reproduce these regional decreases, and instead shows robust increases over some of the same regions (see Fig.3.6B, subfigure b).

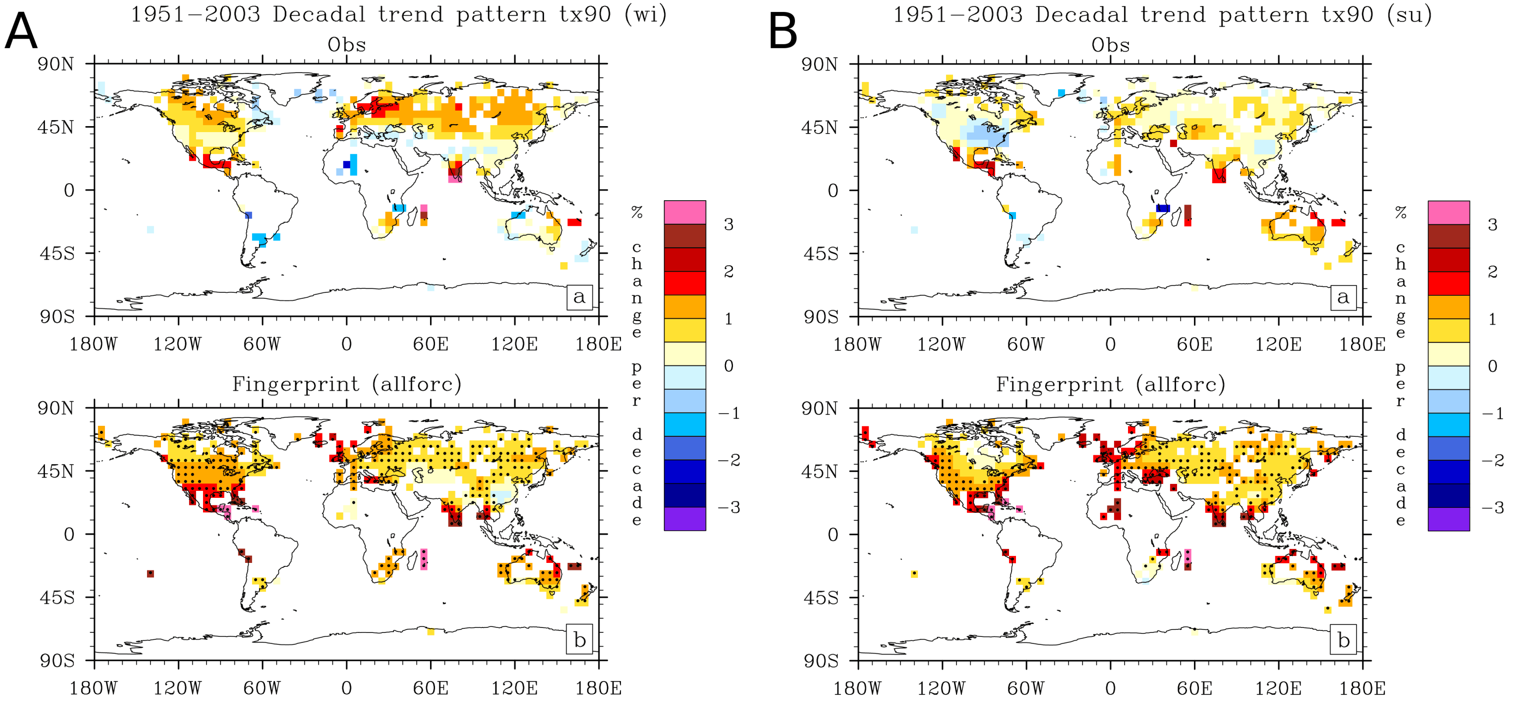


Figure 3.6: As figure 3, but showing the change in the frequency in warm days (tx90) during boreal A) cold and B) warm season for observations (top panel) and model ensemble mean (bottom panel).

Figures 3.7 and 3.8 illustrate the regionally averaged trends of the changes in the frequency of extremes (taking into account the varying grid-box size) and their variability across grid-boxes from observations (red symbols) and simulations (blue) for boreal winter. The figures also show the spread (maximum minus minimum) of the regional mean trend (grey bar) of the individual ensemble members in order to illustrate to what extent the regionally averaged trends vary across individual simulations. The figures show that the spatial variability of tn10 and tx10 (Fig.3.7) is smaller compared to tn90 and tx90 (Fig.3.8). We also find that in Western and Eastern North America, Europe and Northern Asia, there is a large variation of the regional mean trend values within the ensemble. The observed area mean trend is within the model range for many, but not all regions.

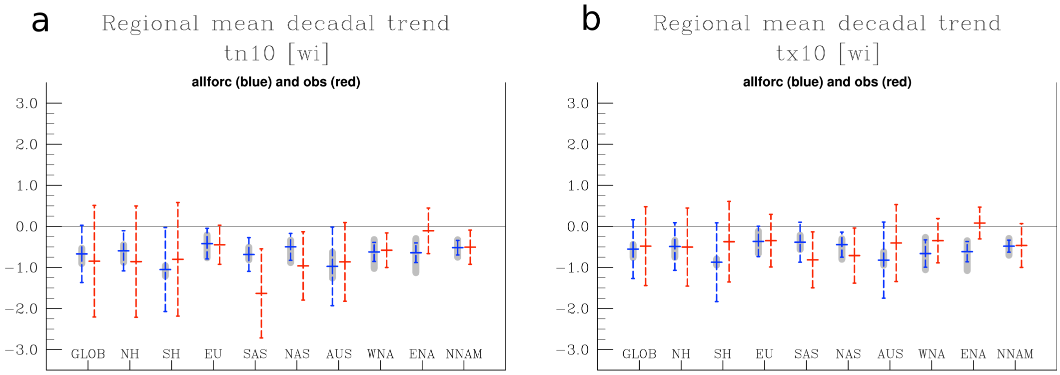


Figure 3.7: a) tn10 and b) tx10 regional mean decadal trend for the boreal cold season (horizontal line marker) and spread of two standard deviations computed across all grid-boxes in the region, representing the spatial variability (vertical line) [in % per decade]. Red symbols represent the observed values and blue ones those of the all-forced ensemble. The grey shading in the background of the blue symbols represents the range of the regional mean trend values spanned by the four ensemble members.

3.4 Detection Analysis

In order to determine whether the observed trend is significantly larger than expected from variability generated within the climate system, and whether it

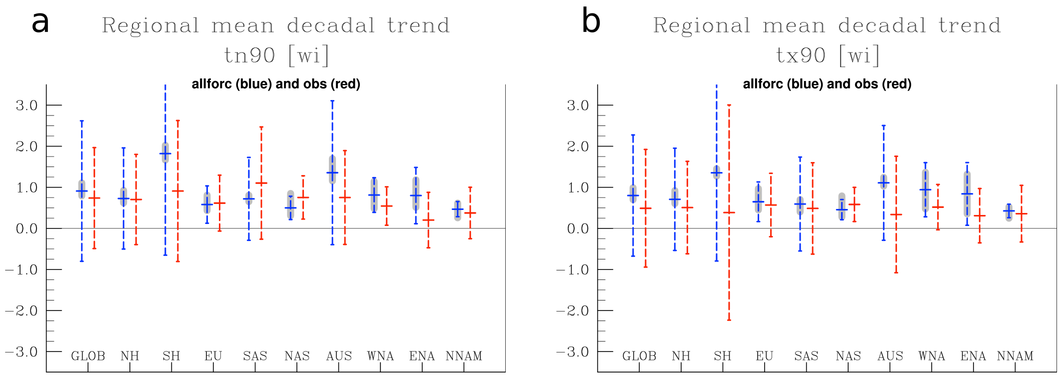


Figure 3.8: As Figure 7, but for a) tn90 and b) tx90 regional mean decadal trend for the boreal cold season.

is consistent with the fingerprint of forced changes plus variability, an optimal detection analysis has been applied. The fingerprints are derived from the ensemble mean of simulations with all forcings. Two sets of sixteen segments extracted from the control run are used to provide estimates of the internal climate variability and to carry out the pre-whitening process (see below).

The fingerprint analysis is based on a total least squares regression and aims to explain the observed changes as a linear combination of changes due to anthropogenic and natural forcings as well as changes due to internal variability (see Allen and Stott [2003] for a detailed mathematical description of this analysis).

$$Y = (X - \nu)\alpha + res \tag{3.1}$$

Y represents the vector of the observations $(Y_1, ..., Y_i)$, where Y_i denotes the observed grid-box trend values of the index for the boreal cold or warm season. X is the model fingerprint vector $(X_1, ..., X_i)$, comprising model estimates of the grid-box trend values computed from the ensemble mean of the model simulations with all forcings. ν stands for the model internal variability and is estimated from the control segments (the fingerprint is averaged from 4 simulations and the noise variance is hence reduced by a factor of 4). α is the scaling factor that determines the magnitude of the fingerprint in observations. res is the regression residual and is assumed to be due to internal climate variability.

The scaling factor α is the factor by which the fingerprint has to be scaled in order to best match the observations. The uncertainty in α has been computed by adding noise onto both the fingerprint and the observations and repeating the scaling factor calculations (Allen and Stott [2003]).

In the optimised analysis, all data are first pre-whitened using the inverse covariance matrix of climate variability (see Allen and Tett [1999]). In order to be able to invert the covariance matrix, all data are projected onto the space of the first few empirical orthogonal functions (EOFs) of control run variability, thus reducing the spatial dimension of the covariance matrix (Hegerl et al. [1997], Allen and Tett [1999]). The regression equation is the same as equation 3.1, although the regression is now computed on pre-whitened data expressed in the space of the truncated eigenvectors rather than grid points (Allen and Tett [1999]).

An important step is to determine the level of truncation that sufficiently represents the spatial fingerprint trend patterns, excluding EOFs that do not add any further information, or which reflect small scale variability that is not well reproduced in model simulations. The truncation level is determined by analysing the regression residual res in a Chi-Square test (Allen and Tett [1999]), ensuring that the observed variance lies within the 5-95% range of the model variability. In the best case, the ratio of the modelled and observed variance is close to unity of the chosen truncation level. Figure 3.9 illustrates the results of the Chi-squared test for the global trend in tx90 during the boreal warm season and indicates that a truncation level of about 12 is adequate. (An additional example is given in the supplementary material, see Fig.B1 and B2). Different truncation levels are chosen for each spatial region, and where no truncation level yields residuals consistent with model variability, this indicates that either model variability in the chosen region is inappropriate, or the observed variability contribution is unusually high, or the fingerprint does not describe the observed trend pattern (e.g.due to model deficiencies or missing forcings).

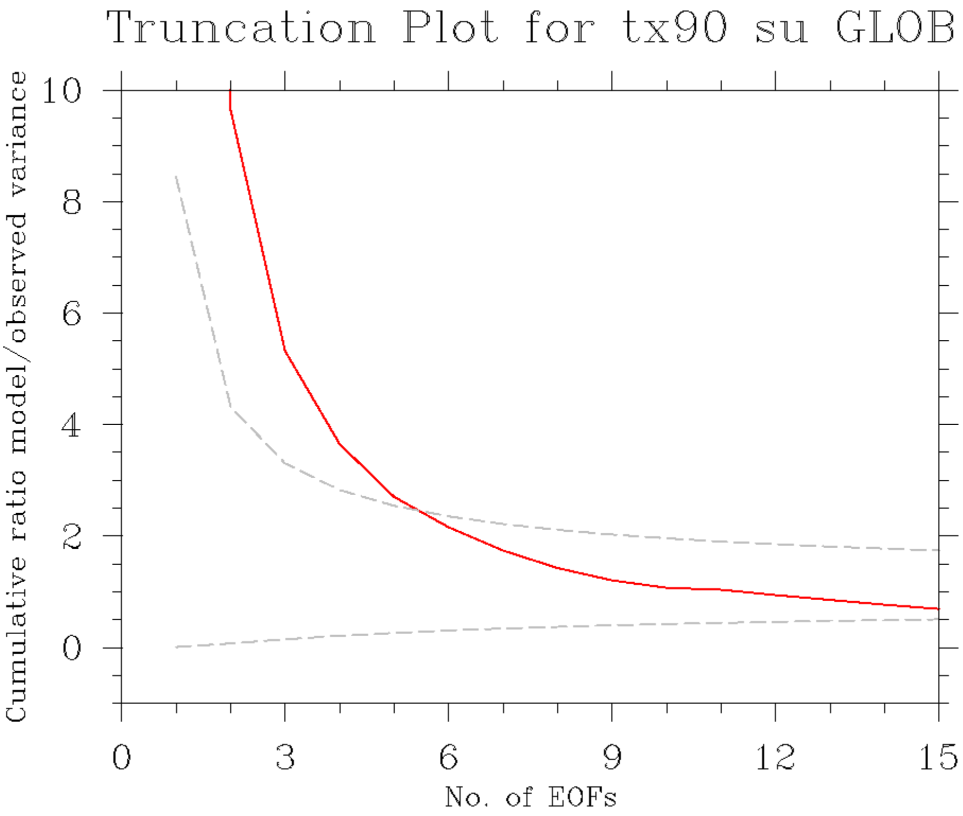


Figure 3.9: Solid red line shows the change of the ratio of the cumulative model and observed residual variance with truncation for the global decadal trend in tx90 during boreal warm season for each truncation level between 2 and 15 in the optimal fingerprint analysis. The dashed grey lines give the upper and lower limit of the 5-95% range for the Chi-Squared test.

The scaling factors α are computed for all indices, all regions and both half-year trends using both the non-optimised and optimised approach. Following this, the residuals of the regression are calculated and are tested to determine if they are consistent with the model estimates of internal variability. Regions where the variability is significantly larger than simulated are excluded from the analysis. If the fifth percentile of the scaling factor is greater than zero, the observed climate change is said to be detectable. We consider a change detected if it is detectable at least for the optimal analysis, as that is expected to be more powerful. In cases where the regression residual is only consistent with model variability in the non-optimised case, this may be due to a strong loss of spatial fingerprint information due to truncation. Hence we consider the few cases where this occurs, such as for warm austral winter days in the Australian region (see below), as successful despite being based on the non-optimal analysis.

3.5 Results of the Detection Analysis

In this section we present results of the detection analysis for all indices and investigate potential benefits from the optimised approach. Changes are detectable when they are significantly different from those driven by internal variability alone. A scaling factor of 1 implies that the ensemble mean trend perfectly represents the observed trend and, therefore, it does not need to be scaled, and a scaling factor consistent with 1 given uncertainty indicates that the model is statistically consistent with the observations.

3.5.1 Detectability of trends in Cold Extremes

Fig. 3.10a and 3.10b show results from the optimised (solid) and non-optimised (dashed) fingerprint analyses of the trend in tn10 during the boreal cold and warm seasons, respectively. The results for the boreal cold season (see Fig.3.10a) show a high detectability with the exception of the SAS region for which the regression residual is not consistent with noise using either of the analysis methods. We also find that optimising leads to a larger number of detectable regions, while in the non-optimised spatial trend analysis several regions show regression residual variability that is not consistent with that in the model control simulation. This may be due to the larger spatial variability across data points than resolved in the model, a problem that is avoided by truncating to a small number of spatial EOFs in the optimised analysis.

During the boreal warm season (see Fig.3.10b), all ten regions show a trend which is found to be significantly different to changes solely due to internal variability, when using an non-optimised analysis. Eight out of ten regions show detectable changes when a optimised analysis is used. Changes in tn10 during the boreal warm season show the largest detectability, with all regions containing significant changes that are generally consistent with the model simulated changes (scaling factor ranges encompassing 1). In the boreal winter analysis only Northern Asia (which has very large observed trends) yields regression residuals that are not consistent with climate model variability.

Changes in tx10 during the boreal cold season (see Fig.3.10e) are detectable in a larger number of regions in the optimised analysis (nine out of ten) compared to the non-optimised one (six out of ten). The only region which does not show significant changes using either approach is Europe.

Changes in the tx10 index during the boreal warm season (see Fig.3.10f) are found to be detectable in seven out of ten regions using an optimised analysis and in six out of ten for the non-optimised. No detectable changes are found in Europe and Eastern North America. Optimisation leads to a slight improvement of the number of regions with detectable signals (see Fig.3.10f).

Overall, the model-simulated trend pattern expected in response to external forcing is detected in the observed change in the global frequency of cold daytime and nighttime extremes. Changes are also detected in the Northern and Southern hemisphere and across most regions (over Europe for the frequency of cold winter and summer nights; for Southern Asia for all indices except the frequency of cold winter nights; for Eastern North America for all but the frequency of cold summer days; and for Northern Asia , Australia and New Zealand, Western North America and Northern North America for all indices in the cold tail).

Using an optimised analysis improves the number of detectable regions for all cold extreme indices, except the changes in the frequency of cold summer nights.

3.5.2 Detectability of trends in Warm Extremes

Fig. 3.10c and 3.10d show results from the optimised and non-optimised detection analyses of the trend in tn90 during the boreal cold season and the boreal warm season, respectively. We find detectable changes during the boreal cold season, in eight out of ten regions for with optimisation and in five out of ten regions without optimising. The Southern Hemisphere is the only region where no detectable change is found using either approach, since the regression residual is again not consistent with noise.

The fingerprint analysis of the trend in frequency of warm nights in the boreal warm season suggests a slightly lower number of detectable regions for the optimised analysis compared to the boreal cold season. The optimised analysis indicates detection in seven out of ten regions and the non-optimised analysis shows detectable changes in six out of ten regions. Global and Southern Hemispheric changes are detectable only when using the optimised method, whereas significant changes and consistent residuals across Northern North America are only found in the non-optimised analysis. Thus, overall, all regions, except of the Northern Hemispheric and the Southern Asian regions, show detectable changes, with the regression residual being problematic for the Northern Hemisphere as a whole and Southern Asia, even when optimising, and for several regions in the non-optimised analysis.

The detection analysis of the trend in tx90 during the boreal cold season (see

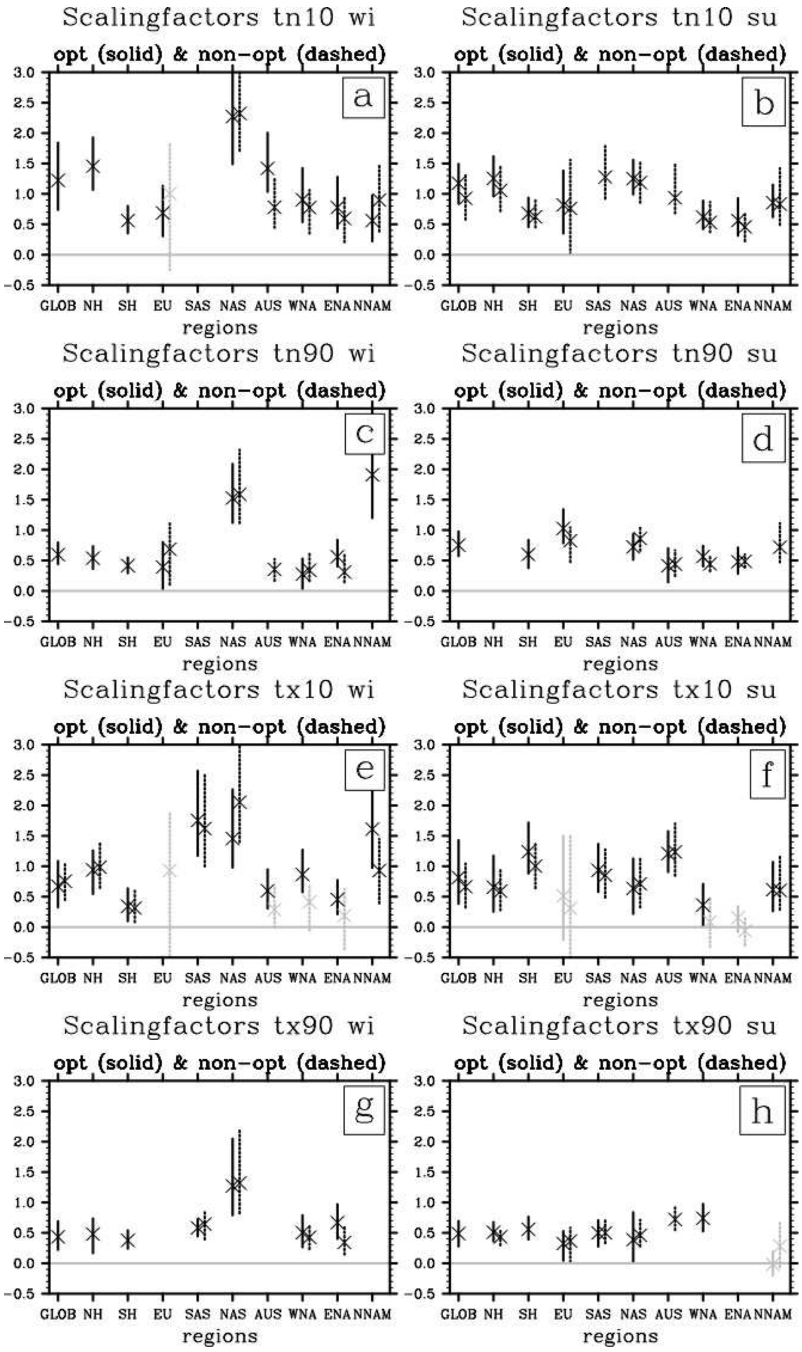


Figure 3.10: Scaling-factors (eqn. 1) by which the model mean fingerprint has to be scaled in order to reproduce the observed trend, plus its 5-95% uncertainty range for a) tn10 during boreal cold season, b) tn10 during boreal warm season, c) tn90 during boreal cold season, d) tn90 during boreal warm season, e) tx10 during boreal cold season, f) tx10 during boreal warm season, g) tx90 during boreal cold season and h) tx90 during boreal warm season. Black lines and symbols denote detectable regions, grey lines and symbols stand for those not detectable. Missing symbols indicate regions for which the regression residual is statistically inconsistent (at the 5% level) with the model estimate of internal variability.

Fig.3.10g) shows detectability in seven out of ten regions for the optimised analysis and in four out of ten regions for the non-optimised analysis. Europe is the only region not showing significant changes in either analysis, due to the residual being inconsistent with model variability.
Fig.3.10h shows the results of the tx90 analysis during the boreal warm season and indicates detectability in seven out of ten regions for the optimal and in five out of ten regions for the non-optimised approach. Global and Southern Hemispheric changes as well as changes in Western North America are only detectable using an optimal approach. Changes in Northern North America are not outside the range of internal climate variability.

In summary, the model-simulated trend pattern expected in response to external forcing is detected in the observed change in the global, northern hemispheric (in all indices except for warm summer nights) and southern hemispheric frequency of warm day and nighttime extremes. Detectable changes have also been observed for all indices in Europe (except warm winter days), Southern Asia (only for changes in warm winter and summer days), Northern Asia (where winter changes are significantly larger than in the model), Australia and New Zealand, Eastern North America (for all indices except changes in warm summer days), Western North America and Northern North America (only for changes in warm winter and summer nights). Changes in the frequency of hot summer days and warm winter days are significantly larger in the model than observed in almost all regions, particularly so in the summer.

Optimising improves the results of the detection analysis for the changes in the frequency warm extremes.

3.6 Discussion and Conclusions

This study shows the observed decrease in the number of cold extremes (tn10 and tx10) and the increase in the number of warm extremes (tn90 and tx90) during boreal cold and warm seasons are generally well reproduced by model simulations with HadGEM1 forced using both natural and anthropogenic drivers. The largest number of regions with detectable changes correspond to the tn10 index and its change during the boreal warm season.

The model significantly underestimates changes in some regions, particularly in winter across large parts of Asia (scaling factors not consistent with 1), and has a tendency to overestimate changes in the frequency of hot days in both the winter and summer seasons over most regions, and in the global and hemispheric mean. It also overestimates changes in the frequency of warm winter days on larger scales, which may be due either to too large a forcing or model response over that period. The model reproduces changes in cold extremes within the uncertainty range on large scales, but does underestimate them in some regions. Despite the observed overall decrease in tn10 and tx10 and increase in tn90 and tx90, there are some regions with trends of the opposite sign, such as changes in tx90 and tx10 during the boreal warm season, tx90 during the austral warm season in western Australia, and tx10 during the boreal cold season across large parts of Eastern North America.

The particular regional feature across Eastern North America is not evident in the simulated trend pattern and is often addressed as the "warming hole" (Kunkel et al. [2006], Pan et al. [2004], Portmann et al. [2009], Meehl et al. [2012]). Portmann et al. [2009] speculate that the "warming hole" could be related to changes in land use in the region, which would affect the concentration of biogenic aerosol and the hydrological cycle. Other possible explanations include regional scale circulation effects, like cold air advection during winter and moisture convergence at low levels during summer (Meehl et al. [2012]).

In contrast, the observed changes indicate stronger warming (decrease in cold extremes, increase in warm extremes) over large parts of Eurasia in the boreal cold season. The change is most pronounced in the number of cold nights, where it extends far into Eastern Asia, but also occurs in the frequency of warm nights, and least in the frequency of warm days. The strong changes in Eurasia may be caused by changes in circulation, particularly, the upward trend of the Northern Annular Mode over much of the late 20th century (Thompson et al. [2000]), which will lead to changes in all temperature extremes indices over Asia in the cold season (Kenyon and Hegerl [2008]). This is consistent with Fig. 3.3-3.6 and probably leads to very large scaling factors found for Northern Asia (Fig.3.10). In contrast, the strong trends in Southern Asia in cold winter night anomalies prevent detection, as the regression residual is inconsistent with model estimates. This suggests that the change there is spatially more complex than can be addressed by inflating the model simulated change, pointing at regional circulation changes or forcings.

The detection analysis indicates a clear benefit from using an optimised approach, with the exception of the changes in the frequency of cold summer nights. Optimising particularly improves the detectability for global and hemispheric changes in the number of cold nights, and warm days and nights during boreal cold season and in warm days and nights during boreal warm season. In some regions, particularly those with changes that are different from expectation, the analysis yields inconsistent residuals and with it no detection. Examples are Southern Asia for winter minima, and Eastern North America for summer maxima.

In conclusion, we find that there is a significant increase in the trend in warm temperature extremes and a decrease in cold extremes during both boreal cold and warm season over the second half of 20th century, which is detectable on the 5% confidence level globally and over many regions. These findings are in agreement with studies by Christidis et al. [2005], Christidis et al. [2011], Zwiers et al.

[2011] and add the information on inter-annual changes, as the previous studies only concentrated on annual changes. We have not attempted to attribute these changes to a combination of forcings by applying a multi-fingerprint approach (Hegerl et al. [1997]; Tett et al. [1999]). Morak et al. [2011] showed that change in the frequency of warm nights corresponds well to changes in annual mean temperatures, even on temporal scales not affected by the trend, suggesting that detectable changes in warm nights are at least partially due to greenhouse gas increases. The changes detected in the frequency of warm nights and days, and hot days are also consistent with the expectation from the observed change in mean temperature. This suggests that greenhouse gas forcing has contributed to the observed changes. However, only an attribution analysis can determine the magnitude of the greenhouse gas contribution to the changes detected in the frequency of temperature extremes.

Chapter 4

Changes in the Frequency of Warm and Cold Spells in Winter

This paper is authored by myself, Prof. Gabriele C. Hegerl (School of GeoSciences, The University of Edinburgh, Edinburgh, United Kingdom) and Dr. Nikolaos Christidis (Met Office Hadley Centre for Climate Change, Exeter, United Kingdom). During the preparation of this manuscript Prof. Gabriele C. Hegerl assisted me with her scientific advice and Dr. Nikolaos Christidis provided model data from HadGEM1 simulations for the analysis and proofread the manuscript.
The included version varies slightly from its paper version to avoid unnecessary repetition.

4.1 Introduction

Recent detection studies found evidence of changes in extreme surface temperatures (Kiktev et al. [2003], Christidis et al. [2005], Christidis et al. [2011], Hegerl et al. [2004], Zwiers et al. [2011]). These changes are of particular interest, as their impact on society and nature is much more severe than the impact of an increase in mean temperature, as it is much harder for humans, flora and fauna to adapt to changing extremes. If extreme events occur over consecutive days, like it is the case during cold or warm spells, the overall impact can be even more severe. For example, cold spells in winter can affect society by increasing the mortality due to severe cold (Huynen et al. [2001],Christidis et al. [2010]). They can also have a strong impact on agriculture, vegetation (Frich et al. [2002]) and energy consumption. Conversely, warm spells in winter could positively affect economy and society, as they lead to a decrease in energy consumption and less cold related deaths. However, warm spells can also have negative effects, for example on snow accumulation, and thus on the water storage and on winter tourism. Warm spells can even have an negative impact on vegetation, as a warming in winter can initiate the growing season of plants, which are subsequently killed as cooler temperatures return (Inouye [2000], Marino et al. [2011]).

This study focuses on observed changes in the frequency of warm and cold spell events during winter since the second half of the 20th century. It shows a comparison analysis of observed and model simulated events as well as a detection analysis using a non-optimised fingerprint method. The analysis focuses on the changes in these events across the Northern Hemisphere during the winter months December, January and February. Besides a hemispheric analysis of the changes in these persistent winter extremes, a regional analysis has also been undertaken.

The chapter is structured as follows: First the data sets are explained, then the methodology which has been applied is introduced. Following the results of the spatial and the time series analysis are shown and discussed. Next, the procedure and results of the significance testing are shown and in the end the findings are summarised in a concluding section.

4.2 Data

This study uses observed and model simulated data sets of gridded daily mean minimum and maximum near surface temperature. The observational data originate from the Hadley Centre gridded observed data set HadGHCND, which covers the period 1951-2007 (Caesar et al. [2006]), and has a grid resolution of 2.5° × 3.75°.

The model data consist of two output fields, daily minimum and maximum near surface temperature, of four members of one ensemble using the Hadley Centre, ocean-coupled global climate model HadGEM1. HadGEM1 has a grid resolution of 1.25° × 1.875°. The model simulations were run using both natural and anthropogenic forcings, including drivers such as volcanic, aerosol, solar forcing, changes in greenhouse gases, land use changes, sulphate aerosol and black carbon. We use daily data for the period 1950-2005 (Stott et al. [2006]). Indices of the number of cold and warm spells have been computed for individual simulations and averaged for the ensemble mean change.

Additionally ten 55-year long segments extracted from the HadGEM1 control simulation of daily minimum and maximum daily surface temperature are used (Stott et al. [2006]), for estimating the internal climate variability in the detection analysis.

Before starting the analysis the model data have been re-gridded to match the grid resolution of the observations. Furthermore the model data have been masked in space and time to the same data coverage as available from the observations.

4.3 Methodology

As defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) (Karl et al. [1999], Alexander et al. [2006]) a warm spell consists of six or more consecutive days with a daily maximum surface temperature above the 90th percentile of the daily climatological temperature distribution. Correspondingly, a cold spell is defined as six or more consecutive days with a daily minimum surface temperature below the 10th percentile of the daily climatological temperature distribution. The period 1961-1999 has been chosen to calculate the 10th and 90th percentile threshold. The threshold has been computed separately for the observations and for each simulation.

By counting the number of occurrences of warm spells (WSP) or cold spells (CSP) within the three month period and for each grid point, we retrieved the number of events.

Early on, a second variable was used, namely the total duration of warm (WSPD) or cold spells (CSPD), which gives the total number of days in winter (DJF) that are dedicated to one of these events. However, early results showed that the duration indices behave very similarly to the frequency indices, featuring an increase in the duration of warm spells and a decrease of the duration of cold spell events over the analysis period. Therefore it has been decided to concentrate on the frequency indices in all further analysis.

The work focuses on the changes of these events across the Northern Hemisphere on both hemispheric and regional scale.

The indices have been computed for both observations and ensemble data. To facilitate a fingerprint analysis we also applied this analysis to the data of the control simulation.

As these events are quite rare and do not occur in the same place each year, the year-to-year spatial patterns of number of cold and warm spells are very patchy. Therefore, instead of focusing on a pattern analysis, 6 regions have been defined for which the regional mean time series has been computed. The calculation of the area average included a latitudinal weighting function to account for latitudinal-dependent grid-box area. More details on the selected regions as well as their longitudinal and latitudinal extend can be found in table 4.1. Regional mean time series have been computed for the observations, the ensemble data and the control run data for both frequency of cold and warm spell events. Following the time series computation anomalies relative to the average over the entire period (1951/'52-2004/'05) have been calculated and a 5-year running average has been applied to all data in order to reduce the noise. Due to the computation of the running average the resulting time series spans over the winters 1953/'54 to 2002/'03.

4.4 Results

This section discusses the results of the spatial analysis, the analysis in time and space and the regional mean time series.

4.4.1 Spatial analysis

Fig.4.1 shows the spatial pattern of the cumulative number of cold spells on the Northern Hemisphere in (a) observations, (b) ensemble mean. The observed and the ensemble mean patterns are quite similar, indicating that spatial variations in the model captures the frequency in cold spells. Patterns vary between individual ensemble members (c-f). Furthermore there are some regional differences between all ensemble members and the observations, like for instance across Eastern North America and the Iberian Peninsula.

The spatial patterns of the cumulative number of warm spells across the northern hemisphere as shown Fig.4.3 shows that there have been less of these events during the analysis period compared to the total number of cold spells. Although the patterns of observations and ensemble mean look similar, the magnitude of the model pattern is somewhat smaller than the one observed over a large part of the Northern Hemisphere. Single ensemble members show a quite noisy pattern, which however preserves the model tendency to underestimate (see Fig.4.3c-f).

In order to demonstrate the changes in the observed and ensemble mean frequency of cold spells in both time and space, the percentage of cold spell events occurring relative to the cumulative value over the entire period has been computed and plotted (see Fig.4.2) for nine year chunks from the 50s through to 2005. More than 65% of all observed winter cold spells across most parts of the Northern Hemisphere are shown to have happened between 1951/'52 and 1977/'78, with a general decrease in cold spells since the 50s. Europe and Eastern Asia show a relative decline already during the 1970s. The changes in the ensemble mean show a somewhat weaker decrease.

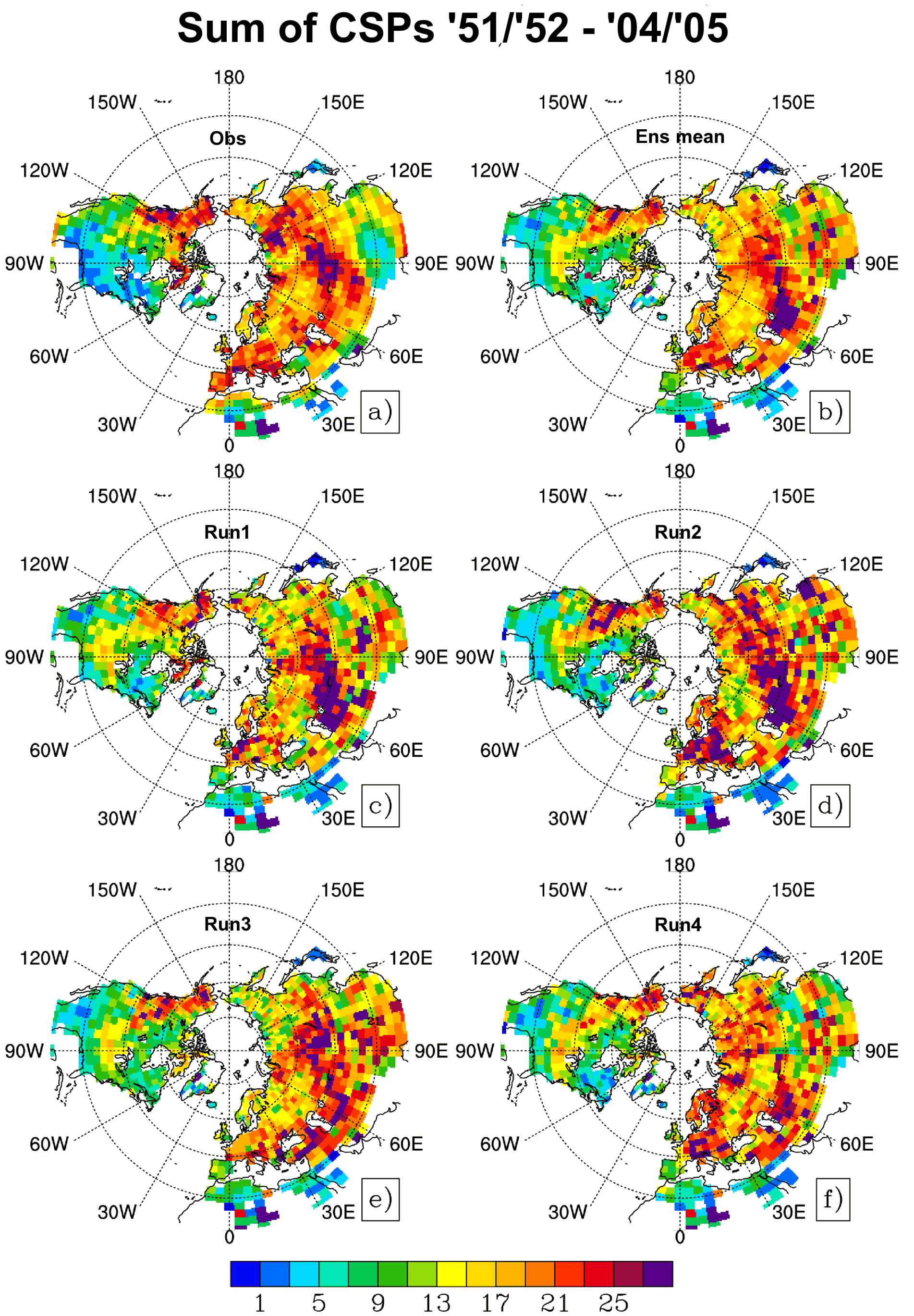


Figure 4.1: Cumulative number of cold spells during the winters 1951/'52-2004/'05 for a) observations, b) ensemble mean, c) ensemble run 1, d) ensemble run 2, e) ensemble run 3 and f) ensemble run 4.

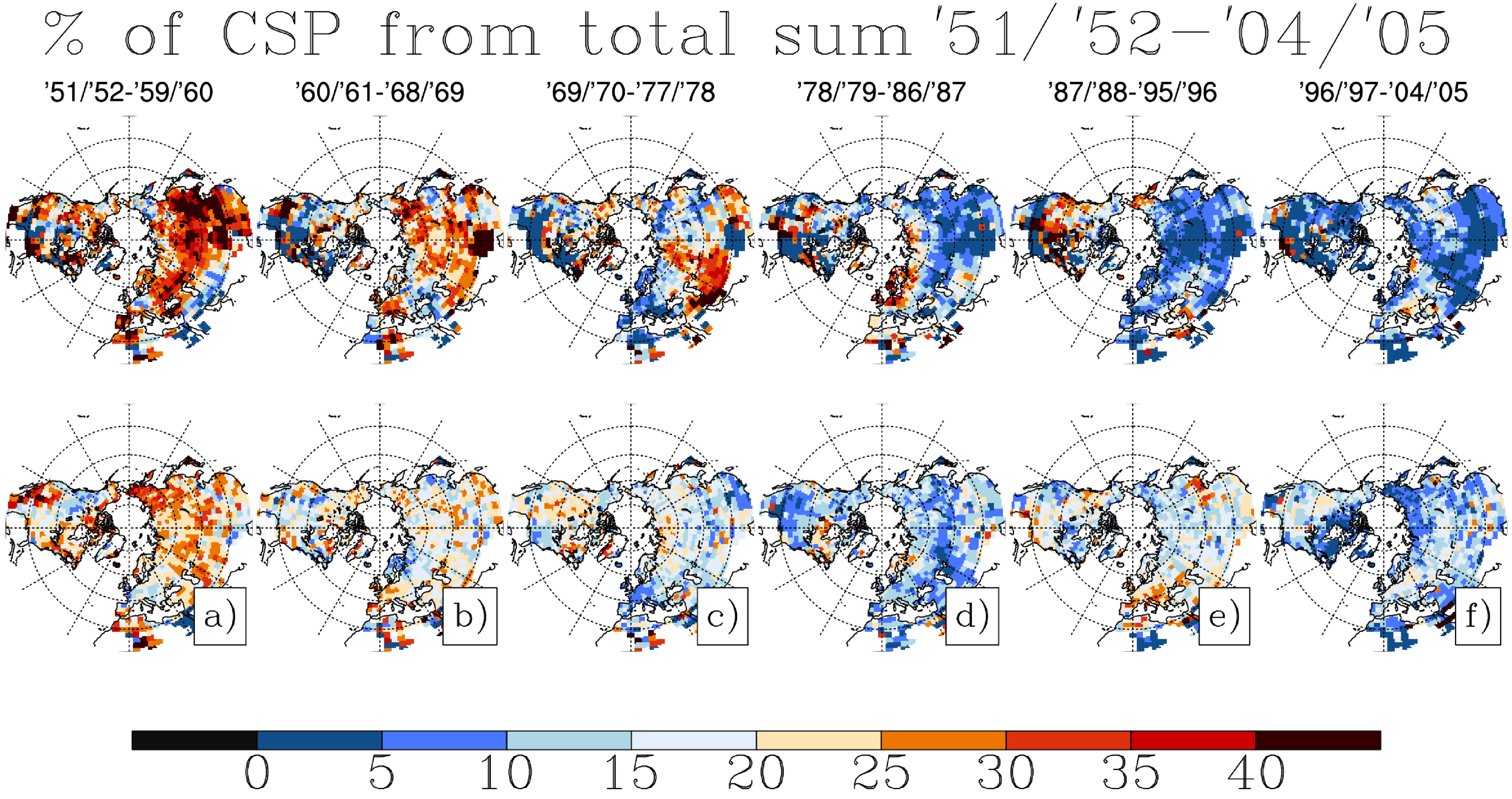


Figure 4.2: Percentage of total number of cold spells during the 1951/52'-2004/05' for observations (top) and ensemble mean(bottom) and periods a) '51/'52-'58/'59, b) '59/'60-'68/'69, c) '69/'70-'78/'79, d) '79/'80-'88/'89, e) '89/'90-'98/'99 and f) '99/'00-'04/'05.

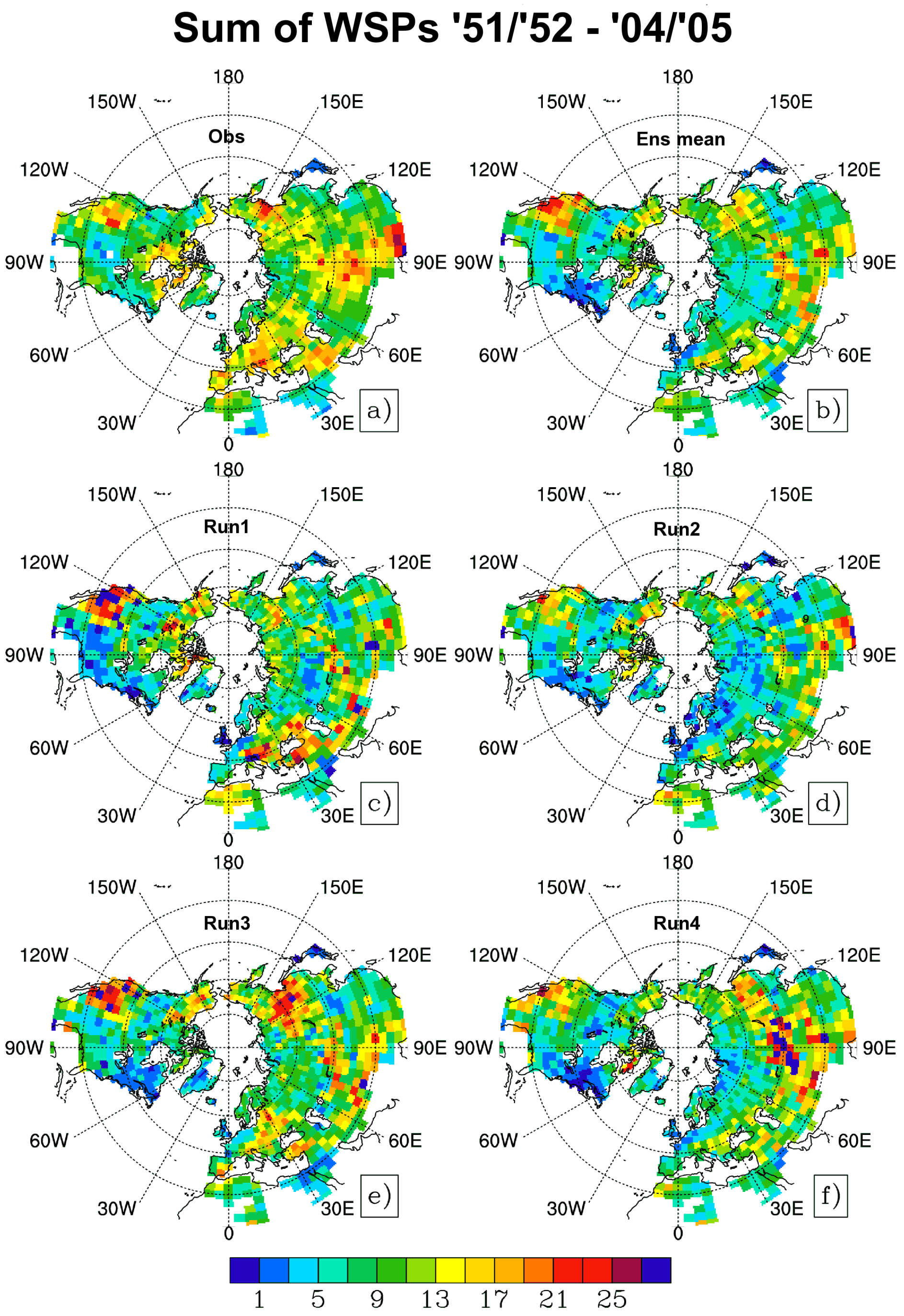


Figure 4.3: Cumulative number of warm spells during the winters 1951/'52-2004/'05 for a) observations, b) ensemble mean, c) ensemble run 1, d) ensemble run 2, e) ensemble run 3 and f) ensemble run 4.

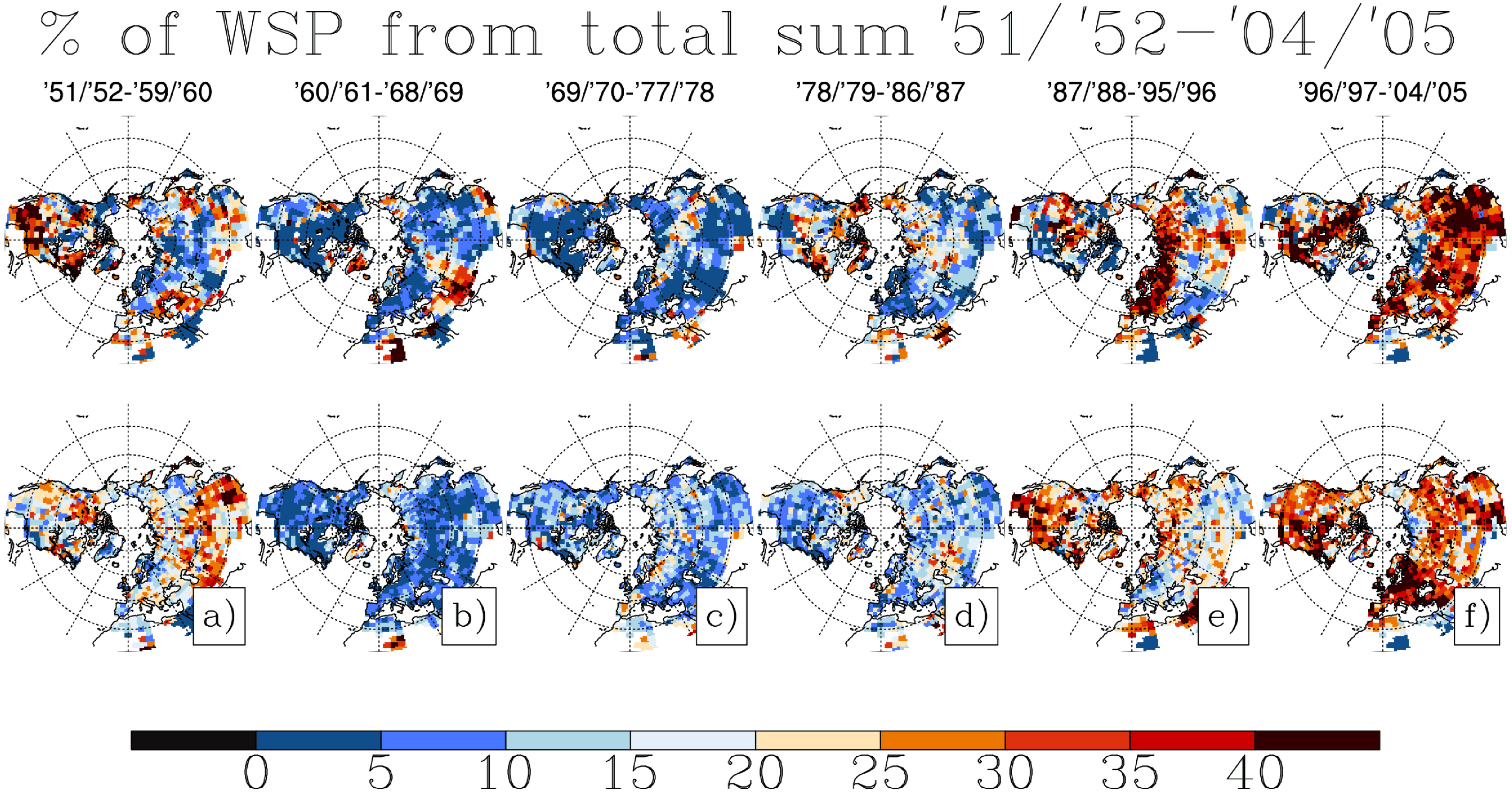


Figure 4.4: Percentage of total number of warm spells during the 1951/52'-2004/05' for observations (top) and ensemble mean(bottom) and periods a) '51/'52-'58/'59, b) '59/'60-'68/'69, c) '69/'70-'78/'79, d) '79/'80-'88/'89, e) '89/'90-'98/'99 and f) '99/'00-'04/'05.

The top panel of Fig.4.4 shows the contribution of 9 year chunks of winter periods to the observed total number of warm spells within the analysis period. Fig.4.4e and Fig.4.4f (top) show that across most parts of the Northern Hemisphere sub-period '87/'88-'95/'96 and '96/'97-'04/'05 together account for about 50% of the entire number of warm spells between 1951/'52 and 2004/'05. Furthermore, more warm spells occurred over the 50s, mainly across the US (see Fig.4.4a, top) in both model and observation. Overall the ensemble mean reproduces the general observed evolution of warm spells during the sub-periods, although the spatial pattern is somewhat different.

4.4.2 Time series analysis

Fig.4.5 a-f shows the smoothed time series of anomalies in the frequency of cold spells for selected regions which, despite smoothing, exhibit large inter-decadal variability, particularly for the observations (black bold solid line). The ensemble mean shows smaller magnitudes as it results from an average over 4 ensemble members. Single ensemble members show fluctuations with a magnitude comparable to the one observed (see red thin dotted lines). The observations show a negative trend in most of the regions, which is largest for the time series of Southern Asia and smallest for Eastern North America. This observed trend remains pretty much constant comparing the 50-year trend (dark grey solid line) with the 34-year (1969/'70-2002/'03) trend (light grey solid line), with the exception of the European region. The ensemble mean trend (dark grey dashed line) also shows a general decrease in these events, however the magnitude is smaller than the one observed (dark grey solid line). In contrast to the observations (light grey solid line) the ensemble mean trend (light grey dashed line) decreases when looking at the sub-period 1969/'70-2002/'03.

The smoothed time series of the anomalies in the regional frequency in warm

Table 4.1: Column 1-3 is giving the number, the acronym and the name of the ten regions. Column 4 and Column 5 state the latitudinal and longitudinal extent of each region.

Nr	Acronym	Name of Region	Latitude	Longitude
1	NH	Northern Hemisphere	0/90	-180/180
2	EU	Europe	35/75	-10/40
3	SAS	Southern Asia	10/45	40/180
4	NAS	Northern Asia	45/80	10/180
5	WNA	Western North America	25/75	-135/-100
6	ENA	Eastern North America	25/75	-100/-45

spells, shown in Fig.4.6 display some similarities in the changes of observed (black bold solid line) and ensemble mean values (red bold dashed line). Even in terms of magnitude observations and models are similar, with the exception of Europe and Southern Asia. For all regions an increase in trend occurs since 1969/'70 (light grey solid line) in comparison to the entire period 1953/'54-2002/'03 (dark grey solid line). The ensemble mean trend lines (dark grey and light grey dashed lines) are shallower than those of the observations, but with a similar increase in the period '1969/'70-2002/'03 (light grey dashed line) relative to the entire period (dark grey dashed line).

4.5 Significance Testing

In order to check whether these changes in the regional frequency of cold and warm spells are significantly different from those expected solely due to internal variability a detection analysis has been applied. This detection analysis is based on a total least square regression where the observations are expressed as a linear combination of the modelled response to anthropogenic and natural forcing as

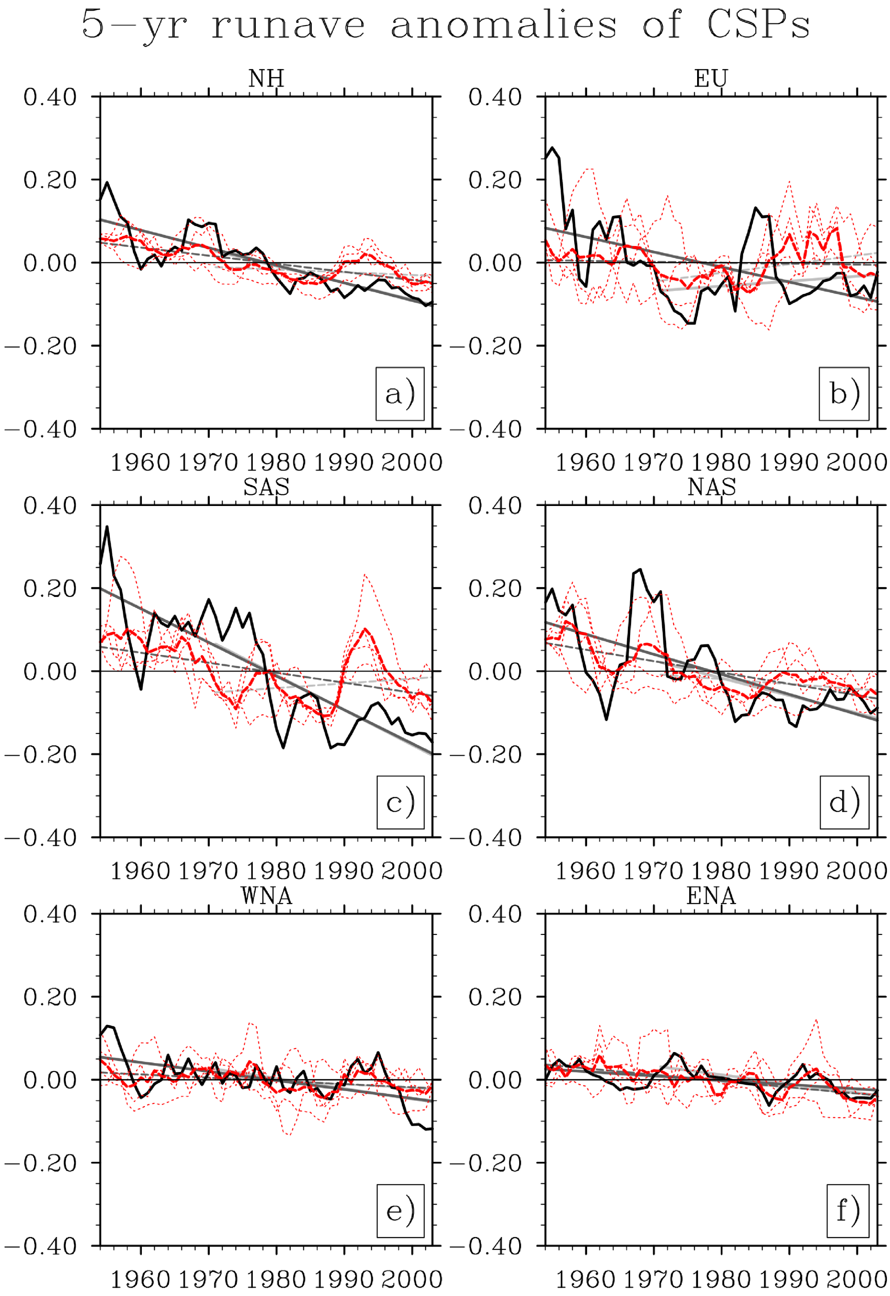


Figure 4.5: Time series of anomalies of 5-year running average of number of cold spells of observations (black bold solid line), ensemble mean (red bold dashed line) and ensemble members (red thin dotted lines). The dark grey bold solid line displays the observed linear trend for period 1953/'54-2002/'03, whereas the dark grey dashed line shows the linear trend of the ensemble mean for this period. The light grey bold solid line shows the observed linear trend for the sub-period 1969/'70-2002/'03, whereas the light grey dashed line represents the linear trend of the ensemble mean over this sub-period.

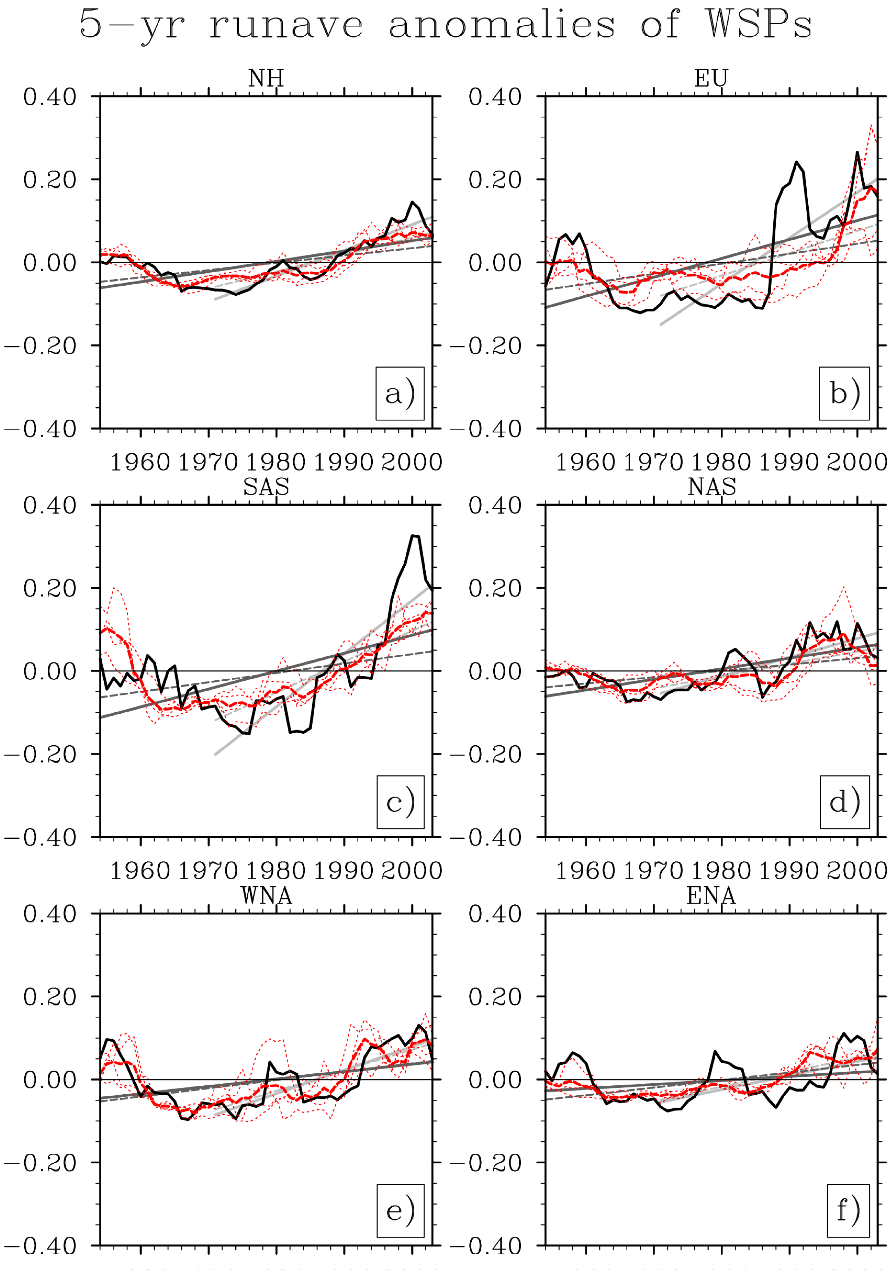


Figure 4.6: Time series of anomalies of 5-year running average of number of warm spells of observations (black bold solid line), ensemble mean (red bold dashed line) and ensemble members (red thin dotted lines). The dark grey bold solid line displays the observed linear trend for period 1953/'54-2002/'03, whereas the dark grey dashed line shows the linear trend of the ensemble mean for this period. The light grey bold solid line shows the observed linear trend for the sub-period 1969/'70-2002/'03, whereas the light grey dashed line represents the linear trend of the ensemble mean over this sub-period.

well as internal variability, v_0 and v_1 (see Allen and Stott [2003]).

$$y = (x - v_1)\beta + v_0 \tag{4.1}$$

y stands for the observed smoothed time series, x represents the ensemble mean smoothed time series, β is the scaling factor with which the ensemble mean has to be scaled in order to best match the observations y, v_0 stands for the internal variability in the observations and v_1 represents the variability in the ensemble mean that was not removed by averaging. The noise samples originate from the time series produced from the control run data.
The uncertainty range has been computed by repeating the computation of β for all combinations of the noise samples using a bootstrap method. The 2.5-97.5% range has been determined by choosing the relevant percentiles. Where the 5th percentile of the scaling factor is bigger than zero and the observed internal variability is within the range of modelled internal variability (see Allen and Tett [1999]), the changes observed are said to be detectable.

The results of the detection analysis of the changes in the frequency of cold spells (see Fig.4.7) shows that in five out of six regions the scaling factor is bigger than one, which means that the changes in these events as simulated by the model are smaller than the one observed. Furthermore the changes on the hemispheric scale, but also those across Northern Asia and Western and Eastern North America, are found to be detectable. Eastern North America however shows significantly less variability in the residual than in the model control simulation. The results of the detection analysis for the Southern Asian region seem to be detectable at first sight, however as the observed variability is larger than the model internal variability these changes can not be claimed to be significant (therefore marked by dotted lines).

Fig.4.8, which shows the results of the detection analysis of the changes in the

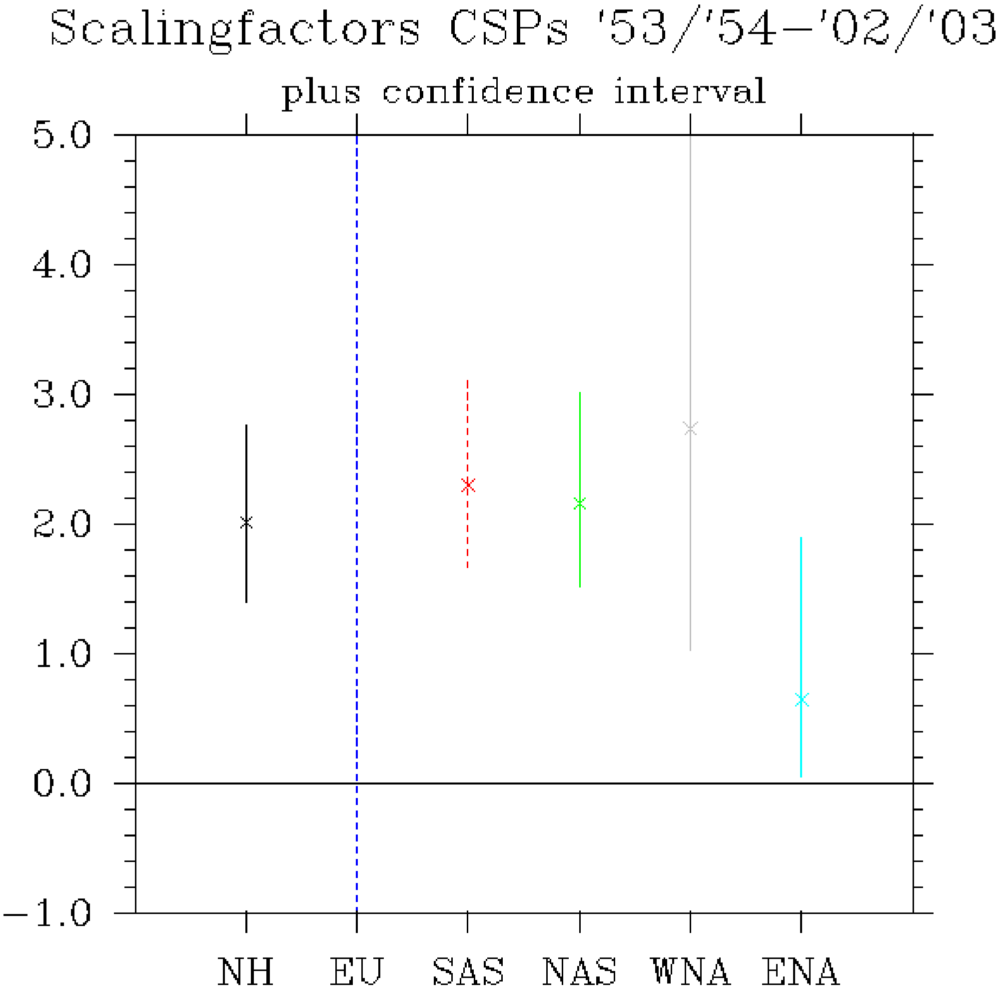


Figure 4.7: Scaling factors of changes in frequency of cold spells (symbols) and uncertainty range (vertical line) plotted for Northern Hemisphere (black), Europe (blue), Southern Asia (red), Northern Asia (green), Western North America (grey) and Eastern North America (cyan). Solid lines denote regions with a observed internal variability within the range of model internal variability. Dashed lines mark those regions with have an observed internal variability higher than the one simulated by the model.

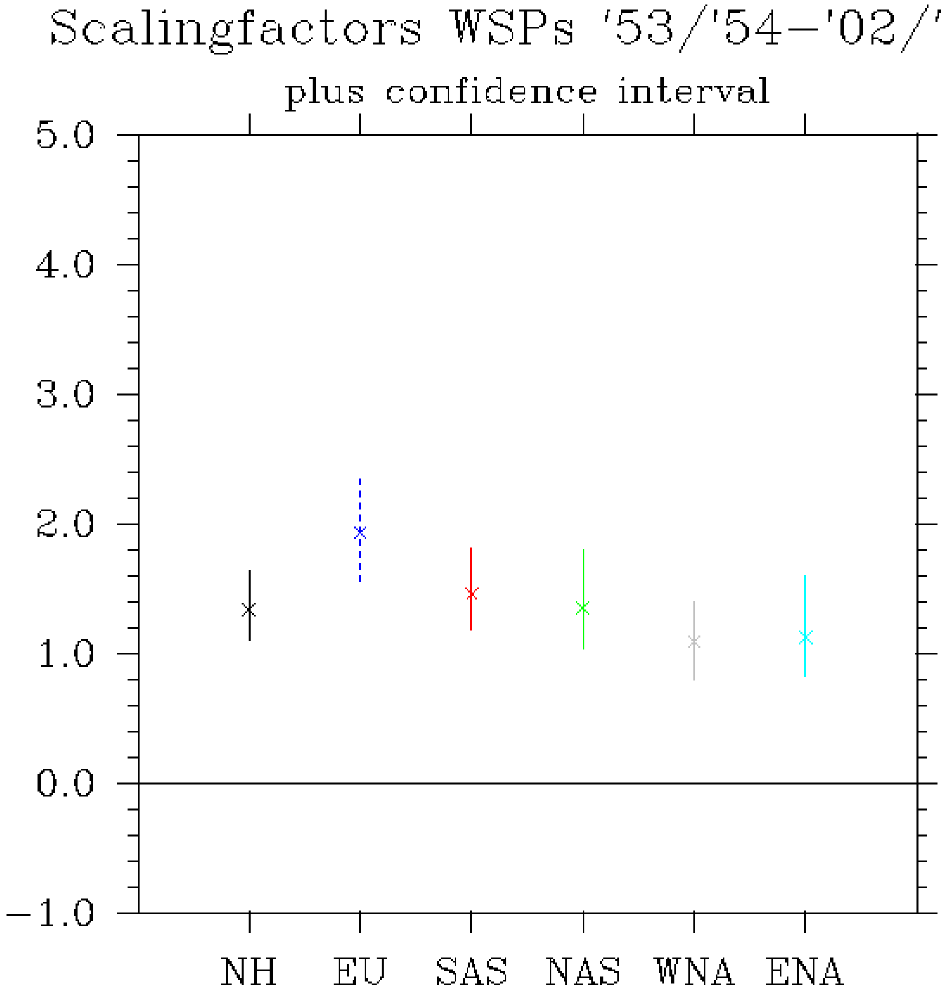


Figure 4.8: Scaling factors of changes in frequency of warm spells (symbols) and uncertainty range (vertical line) plotted for Northern Hemisphere (black), Europe (blue), Southern Asia (red), Northern Asia (green), Western North America (grey) and Eastern North America (cyan). Solid lines denote regions with a observed internal variability within the range of model internal variability. Dashed lines mark those regions with have an observed internal variability higher than the one simulated by the model.

frequency of warm spells, points out that, like for the changes in the frequency of cold spells, the model also underestimates the changes in these events, however the scaling factors are not as high as the one seen for the cold spells. It further shows that these changes are detectable at the 5% confidence level on the hemispheric scale and also across Southern and Northern Asia and Western and Eastern North America. The computed observed variability of Europe, is larger than the model internal variability, therefore the results of the analysis cannot be trusted.

The results of the analysis also show that the duration of these events changed in a similar way as shown in the results of the frequency analysis. So warm spell events didn't only become more frequent but also longer over time and cold spells became less frequent and shorter in duration.

4.6 Summary

This study shows the observed decrease in the frequency of winter cold spells as well as an increase in the frequency of winter warm spells since the 1950s. The total number of cold spells during winter over the considered period is higher than the number of warm spells in this season. This could be explained by the influence of the winter atmospheric circulation, which supports the occurrence of cold extremes.

Frequent cold spell events can be found at high altitudes, such as the Alaska range, the Northern part of the Rocky Mountains, or the high plateaus of Central Asia. The highest number of observed cold spell events between 1950 and 2005 can be found across Central Asia and parts of Siberia. The lowest number of persisting cold extremes is featured across Eastern North America and parts of Southern Asia. A possible explanation of the rare occurrence of cold spell events across Eastern North America could be the influence of ENSO, which causes warm and

dry conditions across the Eastern US during La Nina years.
The spatial pattern of the cumulative number of warm spells shows its the most frequent occurrence across Western North America and South-Eastern Asia. The lowest number of warm spell events occurred across the Central United States.

The analysis of the modelled changes in these events as simulated using HadGEM1 generally agrees with the observations, featuring a decrease in the frequency of cold spell events and an increase in warm spell events. The model, however, seems to underestimate the frequency of these events, which might be due to the models limited ability to simulate persisting extremes like these. Expectedly, the spatial pattern of the observations and the model also vary from each other, as these events do not occur very frequently and not necessarily in the same place. This is also the reason why this study refrains from undertaking a more detailed spatial analysis.

The observed changes in the frequency of cold spell events shows a steady decline over the analysis period across most parts of the Northern Hemisphere. A similar trend can be seen for the modelled changes, though this trend weakens towards the end of the period. The analysis of the observed frequency of warm spells on the other hand shows a distinct peak during the period 1987/'88-2002/'03, which accounts for about 50% of all warm spells during the entire period across nearly the entire Northern Hemisphere. These changes can also be found in the model analysis.

The detection analysis exhibit that the changes in the frequency of cold spells, on hemispheric scale but also across Northern Asia and Western and Eastern North America are significantly different to those expected due to natural variability alone. Changes in the frequency of warm spells were found to be detectable for the Northern Hemispheric regions as well as for Southern and Northern Asia and Western and Eastern North America. It stands out that all scaling factors are bigger than one, which means that the modelled changes have to be scaled up in order to match the observed. The fact that the uncertainty bar often does not encompass 1 proves that the model significantly underestimates the frequency of these persisting extreme events.

Chapter 5

Discussion and Conclusions

This chapter is going to show how these three studies fit together to make it one research project, it will present and discuss the answers of the set research question and it will clarify how this work contributes to the state of the art. These sections will also discuss possible weaknesses and improvements.

The chapter is structured as follows: First an overview over the research problems and the key findings is given, then the overall performance of the climate models used is discussed, following specific results are highlighted and some of the possible uncertainties are explored and the chapter is closed by an outlook on possible improvements and future work.

5.1 Reflection

The main aim of this work is to investigate whether the observed changes in the considered temperature extremes are caused by changes in anthropogenic and natural forcings or whether they can be explained by fluctuation due to internal climate variability alone.

The study focuses on the observed changes in the frequency of abnormally warm days and nights as well as abnormally cold days and nights. It also analyses changes in the frequency of periods of six or more consecutive days with abnormally high daytime or low nighttime temperatures, referred to as warm and cold spells. All the analysis variables are percentile based indices.

Using percentile indices rather than absolute indices has the advantage that outliers do not have such a great impact on the values of the indices, as one outlier only counts for one exceedance, rather than for the magnitude of the anomaly. Also the sample size of these so-called moderate-to-extreme extremes is larger than those of absolute ones, as they occur on a more frequent basis. This makes an analysis easier. The downside of using percentile extremes is, that they do not really account for changes in extreme events on the very tails of the distribution, such as the heat waves of Europe 2003 or Russia 2010.

The changes are analysed for different spatial (global, hemispheric and regional) and temporal scales (annual, half-annual and seasonal).

Another aim of this work is, to investigate, whether global climate model simulations reflect the observed changes. For this comparison analysis, model ensemble data from diverse global climate models are used. The study also presents a multi-model analysis.

This work furthermore aims to determine whether observed changes in extremes are linked only to changes in the mean state of the climate, or only to those in climate variability or both? This is done by investigating the relationship between changes in mean and extreme temperatures.

A fingerprint method, has been applied to determine whether the observed changes occurred solely due to internal climate variability or whether they are influenced by external.
The three studies show the evolution of this fingerprint method, from an ordinary least squares (OLS) regression, to a total least squares (TLS) approach and finally an optimised TLS method (see section 1.4).

It is to note, that

- the OLS method neglects any possible error contamination of the model simulated signal and therefore tends to underestimate the magnitude of the scaling factor.
- the TLS method takes into account the error in both observations and model simulations.
- the optimised TLS method aims to increase the signal-to-noise ratio, by truncating parts of the signal, which do not add valuable information to it. This can often lead to a higher detectability.

This work adds important knowledge and understanding to the research on changes in temperature extremes, by not only investigating global changes, but also regional changes. This research is important, especially, as the changes on regional scale are not yet fully understood, but are those that with the largest impact on both society and nature. Going to smaller spatial scales, however, raises the problem of increasing noise and a harder to detect signal. Therefore, in order to obtain robust results, it is advisable not to choose regions that are too small.
A further advancement to previous works, such as Christidis et al. [2011] or Zwiers et al. [2011], is also the half-annual (chapter 3) or seasonal (chapter 4) analysis which allows a statement about the inter-annual changes of these events

throughout the regarded period.
Another novelty of this work, is the study, presented in chapter 4, which shows one of the first investigations of changes in persistent warm daytime and cold nighttime events covering the entire Northern Hemisphere and additionally focusing on selected regions.

5.1.1 Key Findings

The main result of this study shows significant changes in the frequency of temperature extreme events since the mid-20th century, featuring an increase in warm extremes and a decrease in cold extremes. These changes are determined to be significant not only on the global scale, but also on the hemispheric scale and for many regions. Changes in extremes can be detected on annual, half-annual and even seasonal scale.

Although most regions show a general increase in warm extremes and decrease in cold extremes, there are also some regions which display an observed change of the opposite sign. These changes, however, are down to regional effects and do not contradict with the overall findings.

The study also shows that the detected increase in observed warm nighttime extremes is found to be strongly correlated to the observed increase in mean surface temperature. This relationship leads to the assumption that changes in mean surface temperature influence the occurrence of these extremes and would support the theory of a shift in the distribution towards a higher mean and an increase in the warm tail. The increase in mean surface temperature, which has already been attributed to anthropogenic forcing, together with the high correlation found with the increase of warm nighttime extremes, allows a multi-step attribution (Hegerl et al. [2010]), which concludes that the global increase in warm extremes is also likely to be influenced by human-induced forcing.

A multi-step attribution is used in cases of climate extremes and rare events, where often there is not enough data available, to get reliable estimates of observed changes. This method estimates the change in likelihood of a certain event to happen, by investigating the changes in an event, which relationship to the considered event is known and quantified (see Hegerl et al. [2010]).

5.2 Model Performance

State of the art global coupled climate models include all the relevant atmospheric physics to simulate weather and climate, however there are still limitations to the model performance, due to limited computing capacities, which do not allow long simulations with high spatial resolution yet, or because of poorly parameterised processes. This work uses in total, model data from 6 different climate models. Whereas five of them are used in study (chapter 2) one and the remaining one is used in both study two (chapter 3) and three (chapter 4).

The models generally do a good job in reproducing the general tendency of a decrease in cold extremes and an increase in warm extremes, for both the single day and the 6+ day extremes. They, however, struggle to resolve local features such as the local decrease in warm extremes across Eastern North America. This might be due to their coarse spatial resolution, which can result in a poor representation of the topography and local circulations.

The limited ability to reproduce local patterns of changes in day and nighttime extremes makes it even more important to choose decent sized regions for a

comparison or detection analysis.
Also the use of a multi-model mean helps to achieve a more realistic estimate of the error, whilst avoiding putting too much emphasis on single realisations. A multi-model ensemble also covers a wider range of possible states of the climate in comparison to one single model.

5.3 Specific Results

The following section gives a summary on the global changes, as well as changes in both hemispheres and the four considered continents, North America, Europe, Asia and Australia.

Due to the high agreement between the results of the analysis on changes in warm nights shown in chapter 2 and chapter 3, they are not going to be discussed separately in this chapter.

5.3.1 Global and Hemispheric Scale Changes

When talking about global-scale changes in this work, mostly Northern hemispheric changes are addressed, as the observational coverage in the Southern hemisphere is fairly poor. It is mainly limited to the South-Eastern tip of Africa, Southern South America and the regions covering Oceania.

The records of global scale changes show an increase in warm day and night time extremes and a decrease in cold day and night time extremes, which were found to be significant, relative to changes solely due to internal climate variability. These findings are valid for both annual mean changes and boreal winter and summer half-year changes.

Changes in warm and cold day and night time extremes across the Northern Hemisphere have also been found to be significant on both annual and half-annual time scales, with the exception of changes in warm night time extremes during summer.
Besides changes in the frequency of single day extremes, changes in the frequency of persistent winter time extremes have also been detected over the Northern Hemisphere. These changes feature a decrease in cold spell events and the corresponding increase in winter warm spells.

Despite the sparse data coverage across the Southern Hemisphere, changes in warm day and night time as well as cold day and night time extremes have been found to be detectable for both the boreal winter and boreal summer half-year analysis.

5.3.2 Changes across North America

Changes across Northern America can be generally split into changes in the West and changes in the East, as well as those in the North.

Across Western North America significant changes in warm day and night time extremes, as well as for cold day and night time extremes in both boreal winter and summer half-year have been detected.
Western North America also shows a detectable decrease in winter cold spell events and in increase in winter warm spell events.

Eastern North America is an exceptional region in this study. It is the only region that shows wide areas featuring negative trends in warm day time extremes (and positive trends in cold day time extremes) over the second half of the 20th century. This phenomenon is also referred to as the ”warming hole” (Kunkel et al. [2006],

Pan et al. [2004], Portmann et al. [2009]). Its effect has been found to be strongest during summer (Kunkel et al. [2006], Pan et al. [2004], Portmann et al. [2009]). Despite this phenomenon, detectable changes can be found for warm and cold night time extremes during both boreal winter and summer half-year and for warm and cold day time extremes during boreal winter half-year.
Winter warm and cold spells across Eastern North America show a small but detectable signal featuring a decrease in persistent cold and an increase in persistent warm extremes.

Across Northern North America a significant increase in the frequency of warm night time and a decrease in the frequency of cold day and night time extremes during both boreal winter and summer half-year as well as for the annual mean, has been detected. Changes in warm day time extremes are not found to be significant.

5.3.3 Changes across Europe

Changes across Europe often show a very noisy pattern, especially if the observations are coarsely gridded. The high noise level most likely occurs due to the fact that Europe is a relatively small region with strong ocean-land interactions in the West and a strong continental influence in the East. So, unless the observations are of high spatial resolution, like for example in the E-OBS temperature data set (Hofstra et al. [2009]), the use of even smaller subregions is not recommended as the signal-to-noise ratio will suffer due to the small sample size.

Despite its high variability, Europe shows detectable increases in the frequency of warm day time extremes in the boreal summer half-year as well as in the frequency of warm night time extremes during both boreal winter and summer

half-year. A significant decrease in the frequency of cold night time extremes has been detected for both the boreal winter and summer half-year analysis.
Changes in warm or cold spell events in winter are not found to be significant, this is most likely due to the large variability in the changes in these persisting temperature extreme events during boreal winter.

5.3.4 Changes across Asia

In this study Asia is usually split into the North and the South. Only in the paper on changes in warm nights (see chapter 2), which uses the so-called Giorgi regions, Asia is split into Western, Northern, South-Eastern Asia and Tibet.

A significant increase in the frequency of warm days and nights, as well as a significant decrease in the frequency of cold days and nights during both boreal summer and winter half-year have been found for Northern Asia.
Across Southern Asia changes in the frequency of warm and cold day time extremes during both boreal winter and summer half-year are detectable. Furthermore detectable changes in cold and warm night time extremes during the boreal summer half-year have been found.
The analysis of persistent warm winter extremes show a detectable increase across both Northern and Southern Asia. Changes in cold spell events are only found to be significant across Northern Asia.

5.3.5 Changes across Australia and New Zealand

In the Southern Hemisphere the region ”Australia and New Zealand” has been selected due to its sufficient data coverage. In chapter 2 only Southern Australia has been used for analysis. In chapter 3 the region has been extended to cover entire Australia and New Zealand in order to obtain a larger sample size.

The results show a significant increase in the frequency of warm day and night time extremes, with the exception of warm day time extremes during austral summer and a significant decrease in cold day and night time extremes during both, austral summer and winter half-year.

5.4 Possible Impacts

What are the consequences of the observed changes in temperature extremes? Both, the increase in warm extremes, as well as the decrease in cold extremes, have an impact on society and nature.

For example an increase in warm extremes during summer can lead to an increase in human mortality (Christidis et al. [2010], Huynen et al. [2001], especially if these events last for several days. Hot events in summer can also affect crop farming and can lead to a loss in harvest. A loss in harvest leads to a increase crop price and consequently to an increase in price of the products in the supermarket shelves.

A decrease in cold extremes during winter on the other hand, has a positive effect on human mortality. It also positively influences energy consumption. However, the influence of a decrease in cold extremes during winter, or rather a increase in warm events during this period, on the flora can be crucial. Warmer than normal temperatures during winter can initiate the growing season of several plants, which then often get killed off, if the temperatures drop to normal values. Furthermore, an increase in warm events during winter can have a negative effect on winter tourism.

5.5 Uncertainties and Weaknesses

This section outlines possible sources of uncertainty in this study and some discussion on how to avoid those.

5.5.1 Data Quality

Besides uncertainties in model simulated data sets, due to limited spatial resolution, possible missing processes or poor parametrisation, there are also uncertainties in the observations that have to be taken into account. Outliers for example can lead to a wrong representation of changes in temperature extremes. In this study, the uncertainties due to outliers, are reduced as all data have been quality controlled (Caesar et al. [2006]). Also, the choice of the percentile threshold indices as the main analysis variable avoids an emphasise on particularly high or low values, which could be outliers (Zhang et al. [2005]). However, there are other possible uncertainties such as high spatial variability due to changes in topography or uncertainties due to fluctuations in data coverage. To prevent uncertainties due to changes in data coverage, the start of the analysis period was set to 1950. Prior to this date, many areas, in particular across Asia showed very poor data coverage.

5.5.2 Processing

There are also processing steps that can influence the quality of the results, such as the choice of re-gridding routine. For this work the nearest-neighbour interpolation routine (see remapnn, in the CDO package) has been used. This procedure selects the nearest point to the regarded grid point. There are more sophisticated interpolation routines (e.g. remapbil), however, if only land data is available costal regions are neglected in the interpolation process, as they are yielding missing values, in the "box" of grid-points used for interpolation. This would reduce the available spatial data coverage immensely.

Also the choice of analysis variables is crucial. This study mainly concentrates on spatial pattern analysis of trend values. This type of analysis makes the prior assumption that observations could be compared to model simulated data on the grid-point scale, which is not always the case. If I would repeat this study I would possibly stick to a time-series analysis of regional mean values.

5.5.3 Scaling factors

Scaling factors vary between indices and from region to region. This suggests that the observed changes in the investigated indices are influenced by regional effects and supports the choice of a spatial analysis on regional scale.

A scaling factor of 1 indicates that the model does not need to be scaled to fit the observation and a uncertainty range which includes 1 means that the model is consistent with the observations.

Chapter 4 shows many regions where the model significantly underestimates the occurrence of the warm and cold spell events during winter. This could be due to the model's limited ability to simulate this kind of persistent weather events, because of its spatial resolution or problems in getting local circulations right. This significant underestimation can also be found for day and night time extremes across Northern Asia during the winter half year (see chapter 3). This may be due to limited representation of the changes in the NAM across Eurasia during the late 20th century.

Chapter 3 also shows a significant overestimation of warm day and night time extremes. A reason for this could be the high climate sensitivity of HadGEM1, which causes it to warm the planet by a higher rate than actually observed.

5.5.4 Residual Discrepancy

Despite the fact that many regions showed robust scaling factors, their are some regions where the scaling factors had to be neglected as the observed residual was found to be larger than the model variability, but why? One possible reason for this is the high variability in the observational trend pattern as shown in chapter 3, due to regional features, which are not resolved in the models, or small sampling size. The model also fails to reproduce the high variability across high latitudes, but rather simulates the maximum variability across the equator. In order to reduce some of the spatial noise, the data sets have been smoothed, using a five-point smoother.

For the optimised detection analysis also the truncation level has an influence on the residual, a value chosen too small or large can increase the respective residual values.

5.5.5 Multi-Step Attribution Assumption

In chapter 2 a multi-step attribution is used to indirectly link the observed changes in the frequency of warm nights to anthropogenic forcing, taking into account the strong relationship with the observed changes in mean surface temperature. It has been clearly stated that this assumption is valid only for global scale changes, as a human-induced increase in mean surface temperature across the investigated regions has not been confirmed yet. This assumption does not take into account that the increase in mean surface temperature and night time extremes could be caused by different physical processes. Changes in night time temperatures for example could be down to changes in the indirect radiative effect due to changes in aerosols concentration, whereas mean surface temperatures could be rather influenced by latent and sensible heat fluxes.

5.6 Possible Improvement and Future Work

Research on detection and attribution of climate extremes is a field of steady progress, therefore it is not surprising that there are some aspects of this particular project which could be improved or focused on in more detail.

For example the study could be repeated using an extended period, ideally back to mid-nineteenth century, before the start of the strong increase in green house gases. This would allow for a better estimate of the human-induced change and variability. However, the data coverage on a high temporal resolution particularly prior to 1900 is fairly poor, which reduces studies of daily or multi-day extremes from large, global or hemispheric, scale down to station scale and restricts them to regions with long time series of surface temperature measurements.

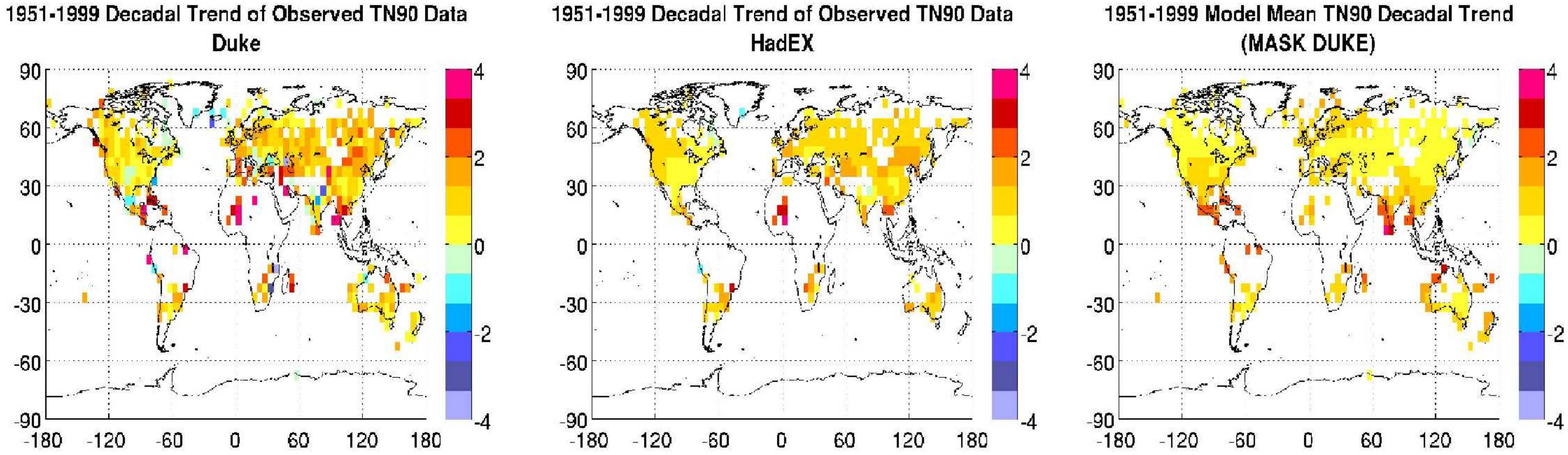
Another extension of this project could be a multi-fingerprint analysis of a multi-model ensemble. Such an analysis would allow an estimate of the impact of individual forcings, such as solar, volcanic, greenhouse gases, changes in land use or urban aerosol on the observed changes in extremes. This goal is achievable provided that further runs of differently forced experiments of daily resolution will be uploaded onto the CMIP5 archive in the near future.

As extreme temperature events are mostly regional phenomena, a regional study using high resolution regional models would be desirable. This would also allow a direct comparison with station data. The crux of a study like this is the choice of the appropriate model for the simulation of the considered phenomena and the quality of its performance. A detection analysis using a regional high resolution model, which manages to reproduce the observed change in temperature extremes on regional scales, would certainly contribute to our understanding of these events.

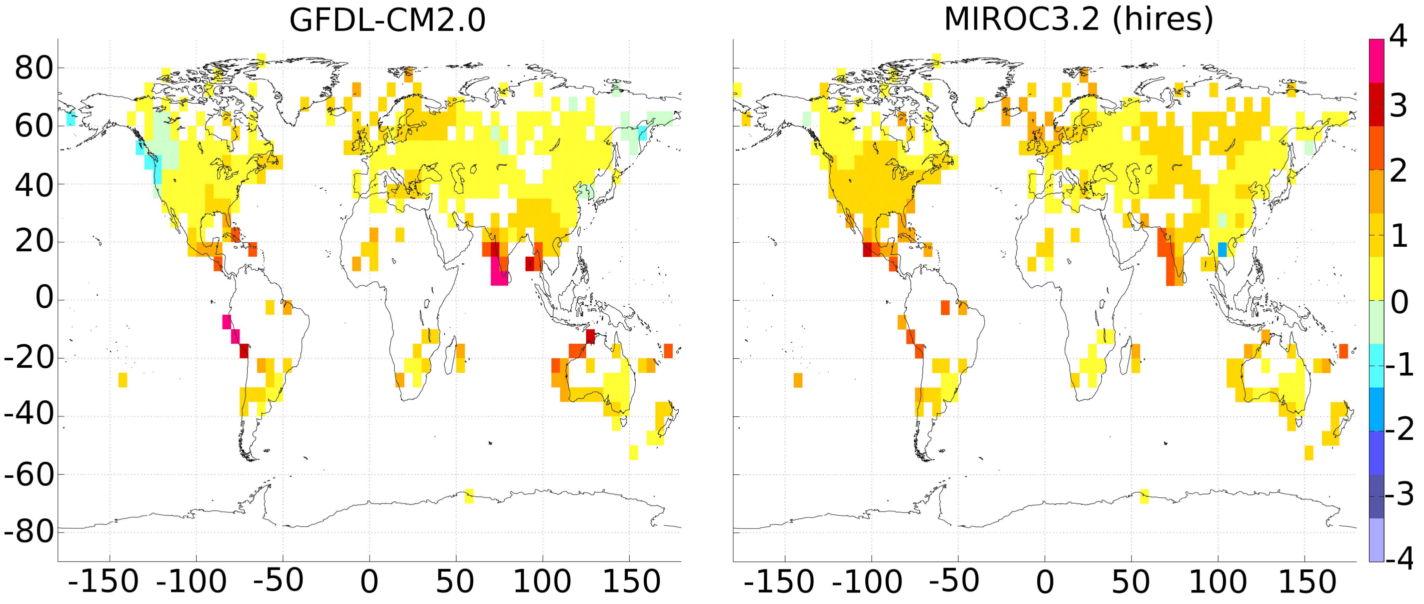
Another possible extension of this project would be the attribution of the observed and detected changes in temperature extremes, for example to changes in circulation patterns and persistent synoptic situations using reanalysis data.

Supplemental Material A

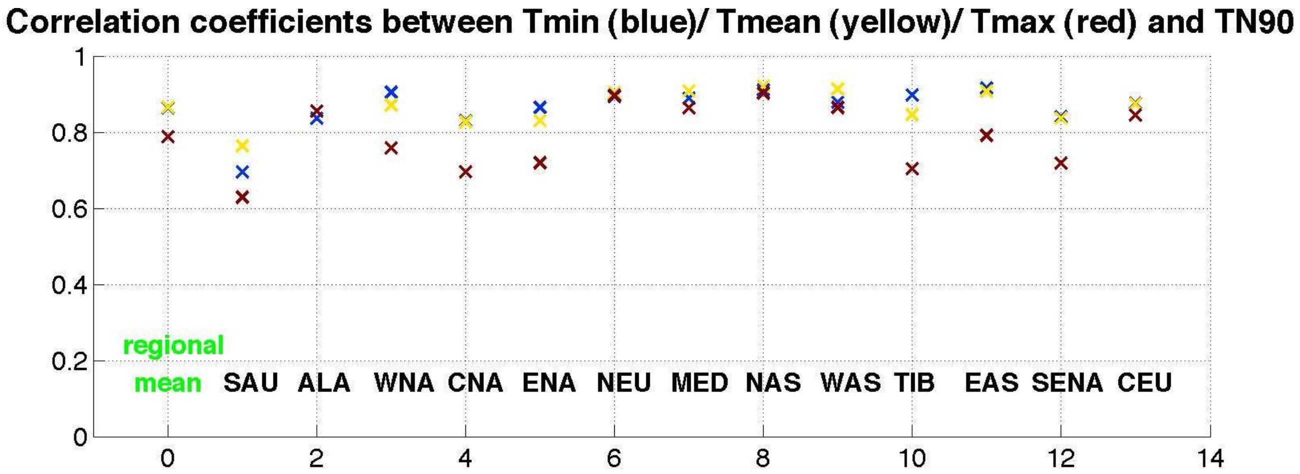
This section shows the supplementary material of chapter 2, which includes the spatial maps of the observations and the model mean (Fig.A1), as well as an example of model patterns of two selected models (FigA2). This section also includes a graph showing the global mean anomalies of regressed mean surface minimum, mean and maximum temperature onto TN90 and anomalies in TN90 (Fig.A4) and a plot presenting the correlation coefficients between the mean, minimum and maximum temperature and the frequency in warm nights (Fig.A3).



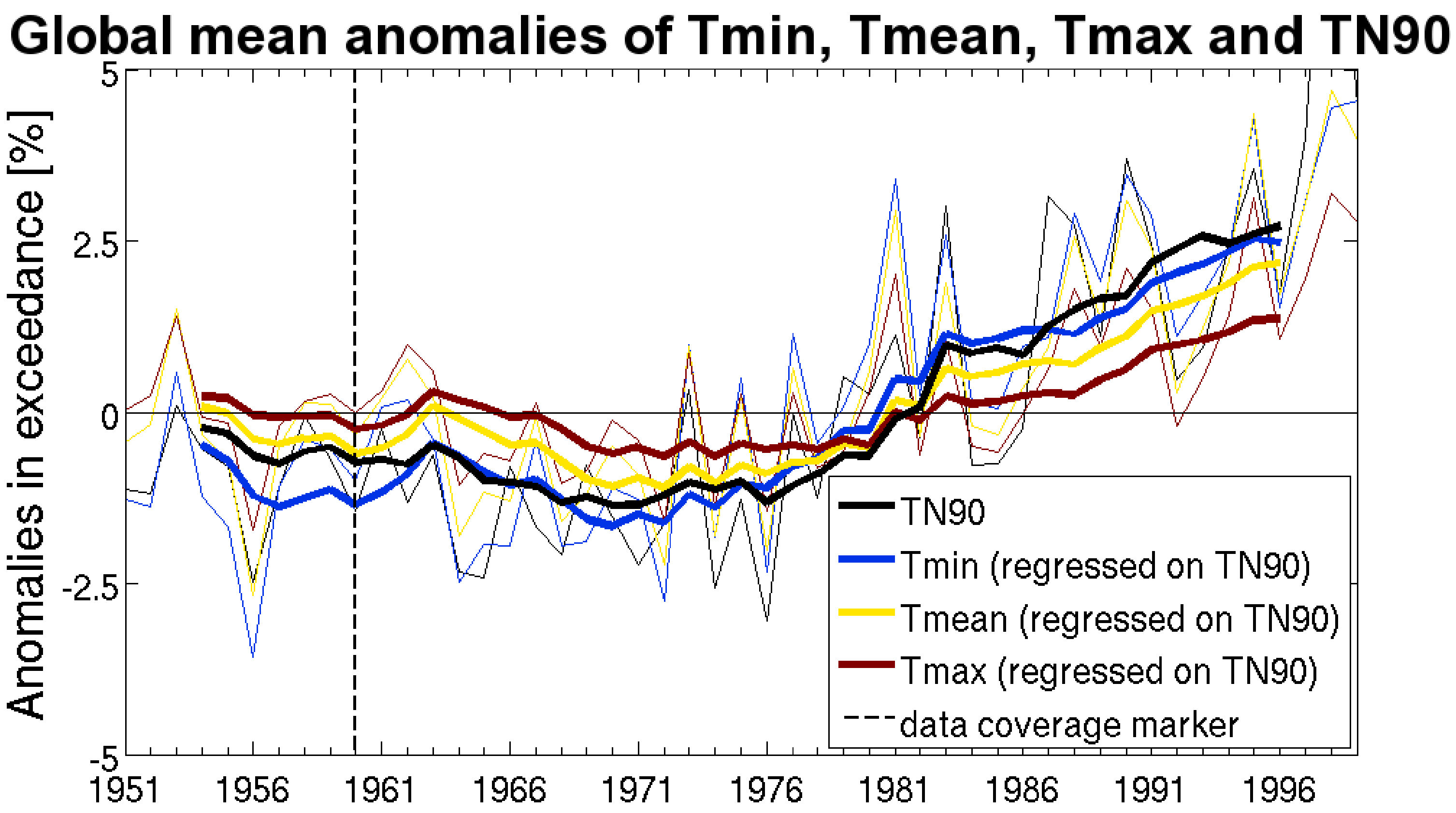
Supplemental Material, Figure A1: Non-smoothed 1951-1999 observed decadal trend pattern of TN90 (in % change per decade) derived from the Duke data set (left), HadEX data set (middle) and the multi-model mean (right). The spatial coverage of all data sets has been reduced onto the one of the Duke data set.



Supplemental Material, Figure A2: 1951-1999 decadal trend pattern in tn90 as simulated by GFDL-CM2.0 (left) and MIROC3.2 (hires) (right). This figures illustrate nicely how much the spatial pattern varies from model to model. In comparison to MIROC3.2 (hires) but also in comparison to the multi-model mean as presented in Fig.A1 (on the right), GFDL-CM2.0 is showing less of an increase and even regions with a change of the opposite sign. MIROC3.2 (hires) on the other hand shows a stronger trend, in some regions even stronger than the multi-model mean, indicating that it is one of the warmer models.



Supplemental Material, Figure A3: Correlation coefficient of the the time series of mean surface minimum (blue), mean (yellow) and maximum (red) temperature with respect to the time series of TN90 for each region.

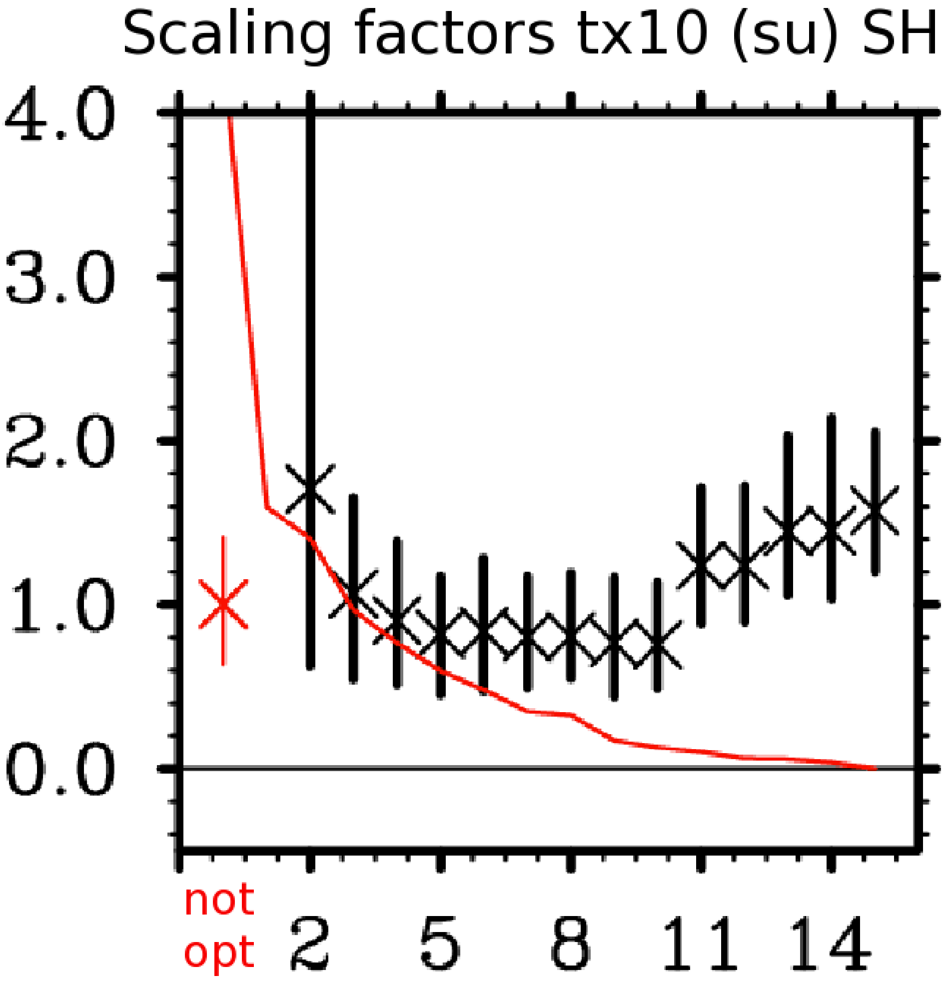


Supplemental Material, Figure A4: Global mean anomalies of regressed mean surface minimum (blue), mean (yellow) and maximum (red) temperature onto TN90 and anomalies in TN90 (black). Dashed line marks the start of the period with data coverage across most parts of at least the Northern Hemisphere. Bold lines show the running average and the fine lines display the year to year fluctuations.

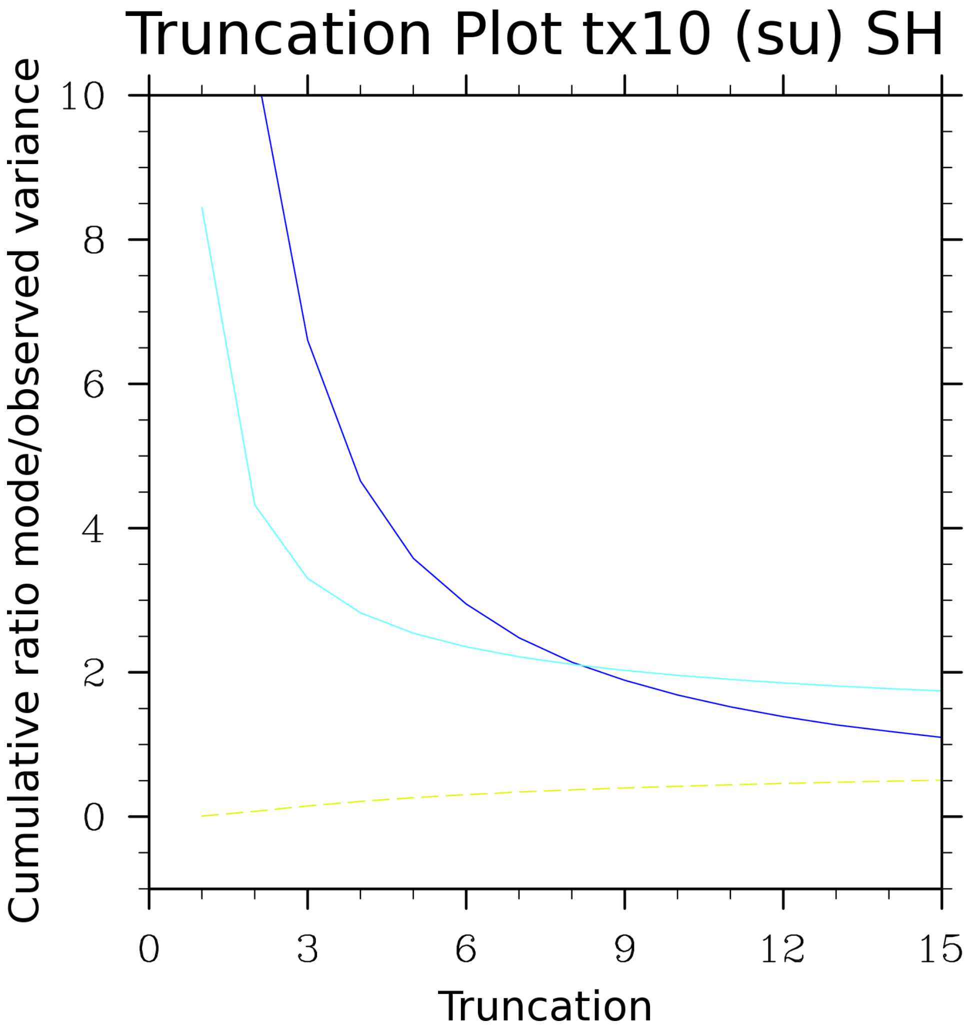
Supplemental Material B

This section shows the supplementary material of chapter 3.

Fig.B1 illustrates on the example of changes in cold summer days across the Southern Hemisphere, how the scaling factor and uncertainty range vary with truncation level. Especially at very low truncation levels (2 and 3) the uncertainty range is large, due to the lag in information at these few levels. Fig.B2, which shows the Chi-squared plot, points out that the range of possible truncation levels lies between 10 and 15. Choosing a lower truncation level than 10 would exclude to much information to describe the signal. It is, however, advisable to chose a truncation level at the lower end of the range in order to avoid the inclusion of too much unnecessary information and to reduce degrees of freedom for the tls regression. I therefore chose 11 as my truncation level, as it lies nicely within the 5-95% confidence area and includes most of the explained variance.



Supplemental Material, Figure B1: Scaling factors plus 5-95% uncertainty range for changes in cold summer days across the Southern Hemisphere, for different truncation levels (black symbols) and non-optimised fingerprint method (red symbol). The descending red line stands for the explained variance.



Supplemental Material, Figure B2: Dark blue line shows the change of the ratio of the cumulative model and observed residual variance with truncation for the Southern Hemispheric decadal trend in tx10 during boreal warm season for each truncation level between 2 and 15 for the optimal fingerprint analysis. The yellow dashed and the light blue solid line lines give the lower and upper limit of the 5- 95% range for the Chi-Squared test.

Bibliography

L. V. Alexander, X. Zhang, T. C. Peterson, J. Caesar, B. Gleason, A. M. G. Klein Tank, M. Haylock, D. Collins, B. Trewin, F. Rahimzadeh, A. Tagipour, K. Rupa Kumar, J. Revadekar, G. Griffiths, L. Vincent, D. B. Stephenson, J. Burn, E. Aguilar, M. Brunet, M. Taylor, M. New, P. Zhai, M. Rusticucci, and J. L. Vazquez-Aguirre. Global observed changes in daily climate extremes of temperature and precipitation. *J. Geophys. Res.*, 111(D5), 2006. doi: 10.1029/2005JD006290. URL `http://dx.doi.org/10.1029/2005JD006290`.

M. R. Allen and P. A. Stott. Estimating signal amplitudes in optimal fingerprinting, part i: theory. *Climate Dynamics*, 21 (5):477–491, 2003. doi: 10.1007/s00382-003-0313-9. URL `http://dx.doi.org/10.1007/s00382-003-0313-9`.

M. R. Allen and S. F. B. Tett. Checking for model consistency in optimal fingerprinting. *Climate Dynamics*, 15(6):419–434, 1999. doi: 10.1007/s003820050291. URL `http://dx.doi.org/10.1007/s003820050291`.

Thomas L. Bell. Theory of optimal weighting of data to detect climatic change. *Journal of the Atmospheric Sciences*, 43(16):1694–1710, 1986. doi: 10.1175/1520-0469(1986)043<1694:TOOWOD>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0469(1986)043<1694:TOOWOD>2.0.CO;2`.

P. Brohan, J. J. Kennedy, I. Harris, S. F. B. Tett, and P. D. Jones. Uncertainty estimates in regional and global observed temperature changes: A new data set from 1850. *J. Geophys. Res.*, 111(D12), 2006. doi: 10.1029/2005JD006548. URL `http://dx.doi.org/10.1029/2005JD006548`.

S. J. Brown, J. Caesar, and C. A. T. Ferro. Global changes in extreme daily temperature since 1950. *J. Geophys. Res.*, 113(D5), 2008. doi: 10.1029/2006JD008091. URL `http://dx.doi.org/10.1029/2006JD008091`.

John Caesar, Lisa Alexander, and Russell Vose. Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded data set. *J. Geophys. Res.*, 111(D5), 2006. doi: 10.1029/2005JD006280. URL `http://dx.doi.org/10.1029/2005JD006280`.

Nikolaos Christidis, Peter A. Stott, Simon Brown, Gabriele C. Hegerl, and John Caesar. Detection of changes in temperature extremes during the second half of the 20th century. *Geophys. Res. Lett.*, 32(20), 2005. doi: 10.1029/2005GL023885. URL `http://dx.doi.org/10.1029/2005GL023885`.

Nikolaos Christidis, Gavin Donaldson, and Peter Stott. Causes for the recent changes in cold- and heat-related mortality in england and wales. *Climatic Change*, 102(3):539–553, 2010. doi: 10.1007/s10584-009-9774-0. URL `http://dx.doi.org/10.1007/s10584-009-9774-0`.

Nikolaos Christidis, Peter A. Stott, and Simon J. Brown. The role of human activity in the recent warming of extremely warm daytime temperatures. *Journal of Climate*, 24(7):1922–1930, 2011. doi: 10.1175/2011JCLI4150.1. URL `http://dx.doi.org/10.1175/2011JCLI4150.1`.

Nikolaos Christidis, Peter Stott, Francis Zwiers, Hideo Shiogama, and Toru Nozawa. The contribution of anthropogenic forcings to regional changes in temperature during the last decade. *Climate Dynamics*, pages 1–16, 2012. doi: 10.1007/s00382-011-1184-0. URL `http://dx.doi.org/10.1007/s00382-011-1184-0`.

Benjamin Cook, Richard Seager, and Ron Miller. Atmospheric circulation anomalies during two persistent north american droughts: 1932–1939 and 1948–1957. *Climate Dynamics*, 36(11):2339–2355, 2011. doi: 10.1007/s00382-010-0807-1. URL `http://dx.doi.org/10.1007/s00382-010-0807-1`.

T. Davies, M. J. P. Cullen, A. J. Malcolm, M. H. Mawson, A. Staniforth, A. A. White, and N. Wood. A new dynamical core for the met office's global and regional modelling of the atmosphere. *Quarterly Journal of the Royal Meteorological Society*, 131(608):1759–1782, 2005. URL `http://dx.doi.org/10.1256/qj.04.101`.

S. M. Dean and P. A. Stott. The effect of local circulation variability on the detection and attribution of new zealand temperature trends. *Journal of Climate*, 22(23):6217–6229, 2009. doi: 10.1175/2009JCLI2715.1. URL `http://dx.doi.org/10.1175/2009JCLI2715.1`.

C. Deser and J. M. Wallace. El niño events and their relationship to the southern oscillation. *Journal of Geophysical Research*, 1987.

Clara Deser and John M. Wallace. Large-scale atmospheric circulation features of warm and cold episodes in the tropical pacific. *Journal of Climate*, 3(11):1254–1281, 1990. doi: 10.1175/1520-0442(1990)003¡1254:LSACFO¿2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(1990)003<1254:LSACFO>2.0.CO;2`.

Randall Dole, Martin Hoerling, Judith Perlwitz, Jon Eischeid, Philip Pegion, Tao Zhang, Xiao-Wei Quan, Taiyi Xu, and Donald Murray. Was there a basis for anticipating the 2010 russian heat wave? *Geophys. Res. Lett.*, 38(6), 2011. doi: 10.1029/2010GL046582. URL `http://dx.doi.org/10.1029/2010GL046582`.

C.B. Field, V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley, editors. *IPCC, 2012: Summary for Policymakers. In: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation.* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2012.

A. H. Fink, T. Brücher, A. Krüger, G. C. Leckebusch, J. G. Pinto, and U. Ulbrich. The 2003 european summer heatwaves and drought - synoptic diagnosis and impacts. *Weather*, 59(8):209–216, 2004.

P. Frich, L. V. Alexander, P. Della-Marta, B. Gleason, M. Haylock, A. M. G. Klein Tank, and T. Peterson. Observed coherent changes in climatic extremes during the second half of the twentieth century. *Climate Research*, 19(3):193–212, 2002. doi: 10.3354/cr019193. URL `http://www.int-res.com/abstracts/cr/v19/n3/p193-212/`.

Nathan P. Gillett, Daithi A. Stone, Peter A. Stott, Toru Nozawa, Alexey Yu. Karpechko, Gabriele C. Hegerl, Michael F. Wehner, and Philip D. Jones. Attribution of polar warming to human influence. *Nature Geosci*, 1(11):750–754, 2008. URL `http://dx.doi.org/10.1038/ngeo338`.

Filippo Giorgi and Raquel Francisco. Evaluating uncertainties in the prediction of regional climate change. *Geophys. Res. Lett.*, 27(9):1295–1298, 2000. doi: 10.1029/1999GL011016. URL `http://dx.doi.org/10.1029/1999GL011016`.

J. Hansen, R. Ruedy, M. Sato, M. Imhoff, W. Lawrence, D. Easterling, T. Peterson, and T. Karl. A closer look at united states and global surface temperature change. *J. Geophys. Res.*, 106(D20):23947–23963, 2001. doi: 10.1029/2001JD000354. URL `http://dx.doi.org/10.1029/2001JD000354`.

J. Hansen, R. Ruedy, M. Sato, and K. Lo. Global surface temperature change. *Rev. Geophys.*, 48(4), 2010. doi: 10.1029/2010RG000345. URL `http://dx.doi.org/10.1029/2010RG000345`.

K. Hasselmann. Optimal fingerprints for the detection of time-dependent climate change. *Journal of Climate*, 6(10):1957–1971, 1993. doi: 10.1175/1520-0442(1993)006<1957:OFFTDO>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(1993)006<1957:OFFTDO>2.0.CO;2`.

K. Hasselmann. Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dynamics*, 13(9):601–611, 1997. doi: 10.1007/s003820050185. URL `http://dx.doi.org/10.1007/s003820050185`.

G. Hegerl, F. W. Zwiers, P. A. Stott, and V. V. Kharin. Detectablility of anthropogenic changes in annual temperature and precipitation extremes. *Journal of Climate*, 17, 2004.

G. C. Hegerl, K. Hasselmann, U. Cubasch, J. F. B. Mitchell, E. Roeckner, R. Voss, and J. Waszkewitz. Multi-fingerprint detection and attribution analysis of greenhouse gas, greenhouse gas-plus-aerosol and solar forced climate change. *Climate Dynamics*, 13(9):613–634, 1997. doi: 10.1007/s003820050186. URL `http://dx.doi.org/10.1007/s003820050186`.

G. C. Hegerl, O. Hoegh-Guldber, G. Casassa, M.P. Hoerling, R.S. Kovats, C. Parmesan, D.W. Pierce, and Stott P. A. *Good practice guidance paper on detection and attribution related to anthropogenic climate change*. IPCC Work, Group I Tech. Support Unit, Univ. of Bern, Bern, 2010.

G.C. Hegerl, F. W. Zwiers, P. Braconnot, N.P. Gillett, Y. Luo, J.A. Marengo Orsini, N. Nicholls, J.E. Penner, and P.A. Stott. *Understanding and Attributing Climate Change. In: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Nynke Hofstra, Malcolm Haylock, Mark New, and Phil D. Jones. Testing e-obs european high-resolution gridded data set of daily precipitation and surface temperature. *J. Geophys. Res.*, 114(D21), 2009. doi: 10.1029/2009JD011799. URL `http://dx.doi.org/10.1029/2009JD011799`.

Maud M. T. E. Huynen, Pim Martens, Dieneke Schram, Matty P. Weijenberg, and Anton E. Kunst. The impact of heat waves and cold spells on mortality rates in the dutch population. *Environ Health Perspect*, 109(5), 05 2001. URL `http://dx.doi.org/10.1289%2Fehp.01109463`.

D. W. Inouye. The ecological and evolutionary significance of frost in the context of climate change. *Ecology Letters*, 3(5):457–463, 2000. URL `http://dx.doi.org/10.1046/j.1461-0248.2000.00165.x`.

IPCC AR4 WG1, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, and H.L Miller, editors. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. ISBN 978-0-521-88009-1 (pb: 978-0-521-70596-7). Cambridge, 2007.

Gareth S. Jones, Peter A. Stott, and Nikolaos Christidis. Human contribution to rapidly increasing frequency of very warm northern hemisphere summers. *J. Geophys. Res.*, 113(D2), 2008. doi: 10.1029/2007JD008914. URL `http://dx.doi.org/10.1029/2007JD008914`.

Thomas R. Karl, Neville Nicholls, and Anver Ghazi. Clivar/gcos/wmo workshop on indices and indicators for climate extremes workshop summary. *Climatic Change*, 42(1):3–7, 1999. doi: 10.1023/A:1005491526870. URL `http://dx.doi.org/10.1023/A:1005491526870`.

T.R. Karl, G.A. Meehl, T.C. Peterson, K.E. Kunkel, W.J. Gutowski Jr., and D.R. Easterling. Executive summary in weather and climate extremes in a changing climate. regions of focus: North america, hawaii, caribbean, and u.s. pacific islands. t.r. karl, g.a. meehl, c.d. miller, s.j. hassol, a.m. waple, and w.l. murray (eds.). a report by the u.s. Technical report, Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC, 2008.

David J. Karoly and Peter A. Stott. Anthropogenic warming of central england temperature. *Atmospheric Science Letters*, 7(4):81–85, 2006. URL `http://dx.doi.org/10.1002/asl.136`.

J. Kenyon and G. Hegerl. Influence of Modes of Climate Variability on Global Temperature Extremes. *Journal of Climate*, 21:3872–3889, 2008.

Dmitry Kiktev, David M. H. Sexton, Lisa Alexander, and Chris K. Folland. Comparison of modeled and observed trends in indices of daily climate extremes. *Journal of Climate*, 16(22):3560–3571, 2003. doi: 10.1175/1520-0442(2003)016<3560:COMAOT>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(2003)016<3560:COMAOT>2.0.CO;2`.

George N. Kiladis and Henry F. Diaz. Global climatic anomalies associated with extremes in the southern oscillation. *Journal of Climate*, 2(9):1069–1090, 1989. doi: 10.1175/1520-0442(1989)002<1069:GCAAWE>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(1989)002<1069:GCAAWE>2.0.CO;2`.

A. M. G. Klein Tank and G. P. Können. Trends in indices of daily temperature and precipitation extremes in europe, 1946–99. *Journal of Climate*, 16(22):3665–3680, 2003. doi: 10.1175/1520-0442(2003)016<3665:TIIODT>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(2003)016<3665:TIIODT>2.0.CO;2`.

Kenneth E. Kunkel, Xin-Zhong Liang, Jinhong Zhu, and Yiruo Lin. Can cgcms simulate the twentieth-century "warming hole"in the central united states? *Journal of Climate*, 19(17):4137–4153, 2006. doi: 10.1175/JCLI3848.1. URL `http://dx.doi.org/10.1175/JCLI3848.1`.

K.M. Lugina, P.Ya. Groisman, K.Ya. Vinnikov, V.V. Koknaeva, , and N.A. Speranskaya. *Monthly surface air temperature time series area-averaged over the 30-degree latitudinal belts of the globe, 1881-2005.* Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A, 2006.

Garrett P Marino, Dale P Kaiser, Lianhong Gu, and Daniel M Ricciuto. Reconstruction of false spring occurrences over the southeastern united states, 1901–2007: an increasing risk of spring freeze damage? *Environmental Research Letters*, 6(2), 2011. URL `http://stacks.iop.org/1748-9326/6/i=2/a=024015`.

G. M. Martin, M. A. Ringer, V. D. Pope, A. Jones, C. Dearden, and T. J. Hinton. The physical properties of the atmosphere in the new hadley centre global environmental model (hadgem1). part i: Model description and global climatology. *Journal of Climate*, 19(7):1274–1301, 2006. doi: 10.1175/JCLI3636.1. URL `http://dx.doi.org/10.1175/JCLI3636.1`.

Gerald A. Meehl, Warren M. Washington, Caspar M. Ammann, Julie M. Arblaster, T. M. L. Wigley, and Claudia Tebaldi. Combinations of natural and anthropogenic forcings in twentieth-century climate. *J. Climate*, 17(19):3721–3727, October 2004. ISSN 0894-8755. URL `http://dx.doi.org/10.1175/1520-0442(2004)017<3721:CONAAF>2.0.CO;2`.

Gerald A. Meehl, Julia M. Arblaster, and Grant Branstator. Mechanisms contributing to the warming hole and the consequent u.s. east-west differential of heat extremes. *Journal of Climate*, 25:6394–6408, 2012.

Seung-Ki Min, Xuebin Zhang, Francis W. Zwiers, and Gabriele C. Hegerl. Human contribution to more-intense precipitation extremes. *Nature*, 470(7334):378–381, 2011. URL `http://dx.doi.org/10.1038/nature09763`.

S. Morak, G. C. Hegerl, and J. Kenyon. Detectable regional changes in the number of warm nights. *Geophys. Res. Lett.*, 38(17), 09 2011. doi: 10.1029/2011GL048531. URL `http://dx.doi.org/10.1029/2011GL048531`.

S. Morak, G. Hegerl, and N. Christidis. Detectable changes in temperature extremes. *Journal of Climate*, 2012a. in press.

S. Morak, G. Hegerl, and N. Christidis. Changes in the frequency of warm and cold spells during winter on the northern hemisphere. in preparation, 2012b.

K. Noake. Anthropogenic changes in precipitation characteristics. Master's thesis, University of Edinburgh, School of Geosciences, 2011.

K. Noake, D. Polson, G. Hegerl, and X. Zhang. Changes in seasonal land precipitation during the latter twentieth-century. *Geophys. Res. Lett.*, 39(3):L03706, 2012. doi: 10.1029/2011GL050405. URL `http://dx.doi.org/10.1029/2011GL050405`.

Gerald R. North, Kwang-Y. Kim, Samuel S. P. Shen, and James W. Hardin. Detection of forced climate signals. part 1: Filter theory. *Journal of Climate*, 8(3):401–408, 2011/11/25 1995.

doi: 10.1175/1520-0442(1995)008<0401:DOFCSP>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(1995)008<0401:DOFCSP>2.0.CO;2`.

F. E. L. Otto, N. Massey, G. J. van Oldenborgh, R. G. Jones, and M. R. Allen. Reconciling two approaches to attribution of the 2010 russian heat wave. *Geophys. Res. Lett.*, 39(4), 2012. doi: 10.1029/2011GL050422. URL `http://dx.doi.org/10.1029/2011GL050422`.

Zaitao Pan, Raymond W. Arritt, Eugene S. Takle, Jr. Gutowski, William J., Christopher J. Anderson, and Moti Segal. Altered hydrologic feedback in a warming climate introduces a warming hole. *Geophys. Res. Lett.*, 31(17), 2004. doi: 10.1029/2004GL020528. URL `http://dx.doi.org/10.1029/2004GL020528`.

D. Polson, G. Hegerl, and X. Zhang. Causes of robust seasonal land precipitation changes. *J. Climate*, 2012. under revision.

Robert W. Portmann, Susan Solomon, and Gabriele C. Hegerl. Spatial and seasonal patterns in climate change, temperatures, and precipitation across the united states. *Proceedings of the National Academy of Sciences*, 106(18): 7324–7329, 2009. URL `http://www.pnas.org/content/106/18/7324`.

Stefan Rahmstorf and Dim Coumou. Increase of extreme events in a warming world. *Proceedings of the National Academy of Sciences*, 2011. URL `http://www.pnas.org/content/early/2011/10/18/1101766108.abstract`.

E. M. Rasmusson and T. H. Carpenter. Variations in tropical sea surface temperature and surface wind fields associated with the southern oscillation / el. nino. *Mon. Wea. Rev.*, 110:354–384, 1982.

R. Rohde, J. Curry, D. Groom, R. Jacobsen, R. A. Muller, S. Perlmutter, A. Rosenfeld, C. Wickham, and J. Wurtele. Berkeley earth temperature averaging process. Technical report, Novim Group, Berkeley Earth Surface Temperature Project, 2012.

B. D. Santer, C. Mears, F. J. Wentz, K. E. Taylor, P. J. Gleckler, T. M. L. Wigley, T. P. Barnett, J. S. Boyle, W. Brüggemann, N. P. Gillett, S. A. Klein, G. A. Meehl, T. Nozawa, D. W. Pierce, P. A. Stott, W. M. Washington, and M. F. Wehner. Identification of human-induced changes in atmospheric moisture content. *Proceedings of the National Academy of Sciences*, 104(39): 15248–15253, 2007. URL `http://www.pnas.org/content/104/39/15248`.

Adam A. Scaife, Chris K. Folland, Lisa V. Alexander, Anders Moberg, and Jeff R. Knight. European climate extremes and the north atlantic oscillation. *Journal of Climate*, 21(1):72–83, 2008. doi: 10.1175/2007JCLI1631.1. URL `http://journals.ametsoc.org/doi/abs/10.1175/2007JCLI1631.1`.

C. Schär, P. L. Vidale, D. Lüthi, C. Frei, C. Häberli, M. Liniger, and C. Appenzeller. The role of increasing temperature variability for european summer heat waves. *Nature*, 2004.

D. Shepard. A two-dimensional interpolation function for irregularly spaced data. In *paper presented at 23rd National Conference*, Mach., New York, 1968. Assoc. for Comput.

Thomas M. Smith and Richard W. Reynolds. A global merged land–air–sea surface temperature reconstruction based on historical observations (1880–1997). *Journal of Climate*, 18(12):2021–2036, 2005. doi: 10.1175/JCLI3362.1. URL `http://dx.doi.org/10.1175/JCLI3362.1`.

P. A. Stott, D. A. Stone, and M. R. Allen. Human contribution to the european heatwave of 2003. *Nature*, 432(2):610–613, 2004.

Peter A. Stott. Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, 30(14), 2003. doi: 10.1029/2003GL017324. URL `http://dx.doi.org/10.1029/2003GL017324`.

Peter A. Stott, Gareth S. Jones, Jason A. Lowe, Peter Thorne, Chris Durman, Timothy C. Johns, and Jean-Claude Thelen. Transient climate simulations with the hadgem1 climate model: Causes of past warming and future climate change. *Journal of Climate*, 19(12):2763–2782, 2006. doi: 10.1175/JCLI3731.1. URL `http://dx.doi.org/10.1175/JCLI3731.1`.

Peter A. Stott, Nathan P. Gillett, Gabriele C. Hegerl, David J. Karoly, DáithíA. Stone, Xuebin Zhang, and Francis Zwiers. Detection and attribution of climate change: a regional perspective. *Wiley Interdisciplinary Reviews: Climate Change*, 1(2):192–211, 2010. URL `http://dx.doi.org/10.1002/wcc.34`.

Tebaldi et al. Going to the extremes: An intercomparison of model-simulated historical and future changes. *Climate Change*, 2006.

Simon F. B. Tett, Peter A. Stott, Myles R. Allen, William J. Ingram, and John F. B. Mitchell. Causes of twentieth-century temperature change near the earth's surface. *Nature*, 399(6736):569–572, 1999. URL `http://dx.doi.org/10.1038/21164`.

David W. J. Thompson, John M. Wallace, and Gabriele C. Hegerl. Annular modes in the extratropical circulation. part ii: Trends. *Journal of Climate*, 13(5):1018–1036, 2000. doi: 10.1175/1520-0442(2000)013<1018:AMITEC>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(2000)013<1018:AMITEC>2.0.CO;2`.

K.E. Trenberth, P.D. Jones, P. Ambenje, R. Bojariu, D. Easterling, A. Klein Tank, D. Parker, F. Rahimzadeh, J.A. Renwick, M. Rusticucci, B. Soden, and P. Zhai. *Observations: Surface and Atmospheric Climate Change. In: Climate*

Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 2007.

Hans von Storch and Francis W. Zwiers. *Statistical Analysis in Climate Research*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1999.

R. Vose et al. Maximum and Minimum Temperature Trends for the Globe: an Update Through 2004. *Geophysical Research Letters*, 32:L23822.1–L23822.5, 2004.

Qigang Wu and David J. Karoly. Implications of changes in the atmospheric circulation on the detection of regional surface air temperature trends. *Geophys. Res. Lett.*, 34(8), 04 2007. doi: 10.1029/2006GL028502. URL `http://dx.doi.org/10.1029/2006GL028502`.

X. Zhang, G. Hegerl, F.W. Zwiers, and J. Kenyon. Avoiding Inhomogenity in Percentile-Based Indices of Temperature Extremes. *Journal of Climate*, 18: 1641–1651, June 2005.

Xuebin Zhang, Francis W. Zwiers, and P. A. Stott. Multimodel multisignal climate change detection at regional scale. *Journal of Climate*, 19(17):4294–4307, 2006. doi: 10.1175/JCLI3851.1. URL `http://dx.doi.org/10.1175/JCLI3851.1`.

Xuebin Zhang, Francis W. Zwiers, Gabriele C. Hegerl, F. Hugo Lambert, Nathan P. Gillett, Susan Solomon, Peter A. Stott, and Toru Nozawa. Detection of human influence on twentieth-century precipitation trends. *Nature*, 448 (7152):461–465, 07 2007. URL `http://dx.doi.org/10.1038/nature06025`.

Xuebin Zhang, Lisa Alexander, Gabriele C. Hegerl, Philip Jones, Albert Klein Tank, Thomas C. Peterson, Blair Trewin, and Francis W. Zwiers. Indices for monitoring changes in extremes based on daily temperature and precipitation data. *Wiley Interdisciplinary Reviews: Climate Change*, 2(6):851–870, 2011. URL `http://dx.doi.org/10.1002/wcc.147`.

Francis W. Zwiers and Xuebin Zhang. Toward regional-scale climate change detection. *Journal of Climate*, 16(5):793–797, 2003. doi: 10.1175/1520-0442(2003)016<0793:TRSCCD>2.0.CO;2. URL `http://dx.doi.org/10.1175/1520-0442(2003)016<0793:TRSCCD>2.0.CO;2`.

Francis W. Zwiers, Xuebin Zhang, and Yang Feng. Anthropogenic influence on long return period daily temperature extremes at regional scales. *Journal of Climate*, 24(3):881–892, 2011 2011. doi: 10.1175/2010JCLI3908.1. URL `http://dx.doi.org/10.1175/2010JCLI3908.1`.