Novel approaches to quantify the emergence of anthropogenic climate change

Luke James Harrington
New Zealand Climate Change Research Institute
School of Geography, Environment and Earth Sciences

A thesis submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Doctor of Philosophy

Victoria University of Wellington
(2017)

Understanding how the climate system will respond to ongoing human interference is a question of profound societal importance. A significant barrier to quantifying the effect of human activity on the climate system is interpreting how the signal of anthropogenic change can be isolated and distinguished from the background noise of internal variability. An emerging framework in the scientific community is now to investigate signal-to-noise ratios as a more effective measure of the impact of human influence on the climate.

As the cumulative amount of greenhouse gases emitted by the global population continues to increase, emerging (human-induced) signal-to-noise ratios in temperatures are starting to become clear for highly aggregated spatial and temporal scales. However, some other climate phenomena, such as meteorological drought, exhibit a more complex response to anthropogenic forcings. Identifying how further warming will change the characteristics of such phenomena is therefore more difficult, despite the significant policy implications for both climate adaptation and mitigation.

In this thesis, I investigate novel approaches towards separating the relative signal of anthropogenic climate warming from internal variability for these cases of low signal-to-noise ratios. By more effectively understanding the drivers of emergent changes to the climate system, these results help to quantify, and thus communicate, the increasingly damaging effects of human interference on the climate system.

Acknowledgements

I would first like to acknowledge my supervisors, without whom this thesis would not have been possible. I'd like to thank Prof Dave Frame for instilling in me his ambition to produce world-leading science, and providing a first-hand example of how to make a definitive impact in multiple fields of research, despite the relative isolation of climate science in New Zealand. In enabling the pursuit of my own scientific curiosities without constraint or cynicism, while actively encouraging participation in the wider research community, he has provided me with a unique perspective on the benefits of research diversification, and this is something I will be eternally grateful for. Many thanks also goes to Dr Sam Dean, whose patience, guidance and articulate understanding of so many different aspects of climate science, particularly in the field of extreme event attribution, has allowed me to develop as both a student and scientist over the last three years.

I would also like to thank Suzanne Rosier and James Renwick, who have both helped immensely in my understanding of technical aspects of this thesis at crucial times, and provided general guidance and counsel over many cups of coffee. Thanks also to my academic collaborators near and afar – including (but not limited to) Andrew King, Manoj Joshi, Ed Hawkins, Pete Gibson, Erich Fischer, Fraser Lott, Fredi Otto, Dann Mitchell and Sebastian Sippel – their willingness to engage in research collaborations and consider my work as an academic peer has been invaluable and enabled my growth as a scientist.

I would like to thank the many friends who have put up with me over the last seven years of my time at Victoria University of Wellington, and made life bearable during moments when the work threatened to swallow me whole. Special mentions go to Shane, Cabbey and Juliet for making undergraduate and postgraduate life a memorable experience, and to Max, Mike, John, Nick, Jack, Jackson, Cyril, Briana, Fin and Annie for making those non-work hours that much more enjoyable.

Finally, and most importantly, I need to thank my parents for their unwavering support in helping me to get this far. Their greatest gift has been to encourage Michael, Whitney and I to passionately pursue our own paths, no matter the direction. Words will never convey the debt of gratitude that I have for them, and this thesis simply would not exist without them.

Contents

Glossary of terms

20CR – 20th Century Reanalysis Project

CEPRE – Cumulative emissions of probability ratio exceedance

CMIP5 – Coupled Model Intercomparison Project Phase 5

EEI – Emergence emissions index

ESM – Earth system model

FAR – Fraction of attributable risk

GCM – General circulation model

GHG – Greenhouse gases

GMST – Global mean surface temperature

IPCC AR5 – Fifth assessment report of the Intergovernmental Panel on Climate Change

MAE – Mean absolute error

MSLP – Mean sea level pressure

PDF – Probability density function

PED – Potential evapotranspiration deficit

PR – Probability ratio

RCP – Representative Concentration Pathway

SOM – Self-organizing map

S/N – Signal-to-noise ratio

ToE – Time of emergence

UNFCCC – United Nations Framework Convention on Climate Change

WG1 – Working Group 1

Chapter 1

Introduction

1.1 Understanding human-induced climate change

> *"Warming of the climate system is unequivocal, and since the 1950s, many*
> *of the observed changes are unprecedented over decades to millennia."*
> *- IPCC AR5, Synthesis Report Summary for Policymakers.*

There is now overwhelming evidence to support the statement that the climate system is warming, and at an increasingly unprecedented rate (Stocker et al. 2013; Hartmann et al. 2013). The above quote was one of the main 'summary' statements to be released with the fifth and most recent assessment report (AR5) by the Intergovernmental Panel on Climate Change (IPCC), representing a consensus of leading international climate scientists operating under the auspices of the United Nations Framework Convention on Climate Change (UNFCCC). An arguably more important conclusion came from the Working Group 1 (WG1) Summary for Policymakers (SPM), which stated that it was "extremely likely [>95% likelihood] that more than half the observed increase in global average surface temperature from 1951 to 2010 was caused by the *anthropogenic* increase in greenhouse gas [GHG] concentrations and other anthropogenic forcings together" (WG1 SPM Page 17).

The physical mechanisms which relate increasing GHG emissions with subsequent warming of the atmosphere, otherwise known as the 'greenhouse effect', have been understood since the late nineteenth century (Arrhenius 1896; Anderson et al. 2016). More recently, a simple framework has emerged to quantify the warming influence of long-lived GHGs like carbon dioxide (CO_2). Multiple lines of evidence (Allen et al. 2009; Meinshausen et al. 2009; Matthews et al. 2009) have demonstrated the existence of a robust, near-linear, relationship between the cumulative amount of CO_2 emitted into the atmosphere and subsequent (global mean) warming (Figure 1.1, Ciais et al. 2013; Stocker et al. 2013; Frame et al. 2014). This can be easily understood due to the fact that CO_2 accumulates and stays in the atmosphere, whilst other climate pollutants like methane have much shorter atmospheric lifetimes, and thus contribute to net climate warming for only a finite period of time (Allen et al. 2016). Such a framework is especially important when considering that CO_2 emissions by far

represent the largest contribution to the enhanced greenhouse effect in recent decades (Myhre et al. 2013).
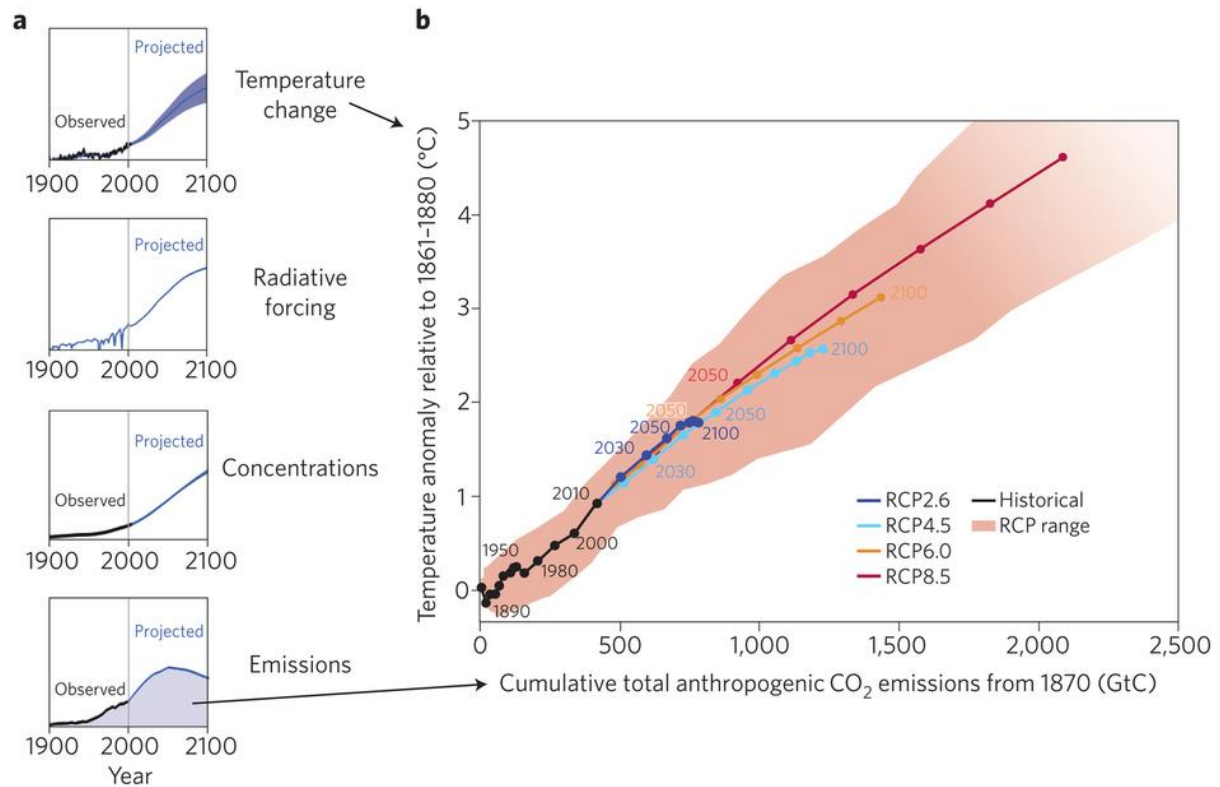
Figure 1.1: (a) In the 2007 IPCC report from Working Group I 'The Physical Science Basis', temperature change, radiative forcing, atmospheric CO_2 concentrations and CO_2 emissions were depicted separately, each in their evolution over time **(b)** In the 2013 report (Stocker et al. 2013) temperature change is shown against cumulative emissions of carbon dioxide (equivalent to the light shaded area in the bottom panel of a), with time progressing along the lines of the various emissions scenarios. Figure reproduced from (Frame et al. 2014).

While the underlying physical mechanisms which govern anthropogenic climate change are well understood, both the IPCC summary statements and cumulative emissions framework highlighted above refer only to changes in global mean temperatures, rather than being regionally-specific. Interpreting how continuing GHG emissions will in fact translate to changes in regional climate, and specifically quantifying when robust deviations from historical observations occur, is a much more difficult task. Further, it is important to recognise that members of the global population do not experience changes to the mean climate, but rather perceive the changing frequency or intensity of local extreme weather events as the physical manifestation of a changing climate (Weber 2010). Therefore, it is of crucial importance to understand how emergent changes in the mean climate translate to changes in the frequency and intensity of such societally-relevant climate extremes, both to inform future adaptation measures and motivate near-term mitigation efforts for individual nations (Hallegatte et al. 2016; Schleussner et al. 2016).

1.2 Tools to investigate anthropogenic climate change

Understanding the complexities of the global climate system requires the use of modelling techniques (McGuffie and Henderson-Sellers 2001). In other fields of study, a hypothesis can be tested by isolating a particular variable, repeating an experiment in controlled conditions and analysing the results. However, this approach cannot be implemented when considering the physical climate response to anthropogenic forcing mechanisms: there is only one planet which is too large and too complex, so the prospect of attempting large-scale controlled experiments to evaluate a particular hypothesis is impossible (Randall and Wielicki 1997). Climate models are therefore the primary tools available for investigating the response of the global climate system to anthropogenic greenhouse gas emissions, and particularly for making projections of future climate over the next century and beyond (Flato et al. 2013).
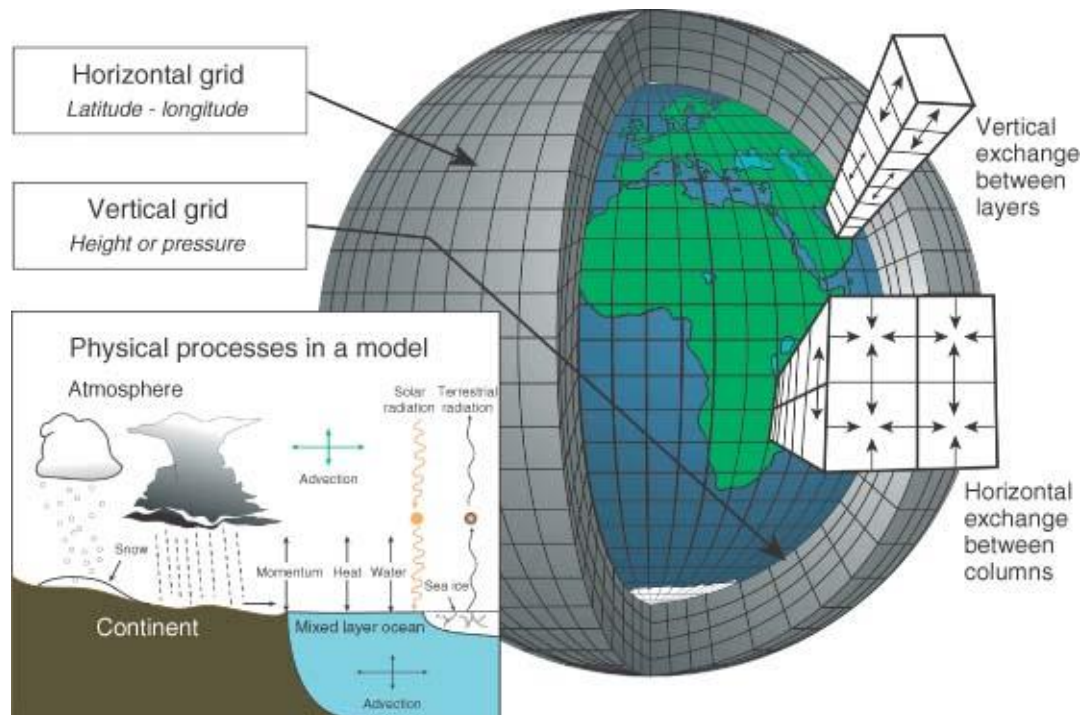
Figure 1.2: The structure of a typical general circulation model (GCM), denoting the 3-dimensional grid on which the large-scale variables are resolved. The sub-panel shows a representation of processes that occur on sub-gridscales and are therefore parametrised within the model. Reproduced from Edwards [2011], Figure 2.

Climate models are discretised, numerical representations of the Earth's climate system: a series of initial conditions are defined, and computational algorithms generate future responses based on established physical laws which govern the interacting processes (Edwards 2011). A schematic example of a general circulation model (GCM) is presented in Figure 1.2: from an initial state, this model evolves the climate system through time, and

outputs the future model state at the specified grid resolution. The spatial resolution of any climate model is limited by computational cost, and those processes which operate at sub-grid scales consequently have to be parameterised, rather than being simulated explicitly.

Any attempt to simplify the climate system into a representation of its fundamental physical laws will result in uncertainty (Knutti et al. 2008; Knutti and Sedláček 2013): first, uncertainty derives from the fact that observed processes of the climate system are not fully understood (Allen and Frame 2007; Hawkins and Sutton 2009; Flato et al. 2013); second, the complex array of processes which interact on a variety of both spatial and temporal scales result in feedback mechanisms, so any solution of the governing equations at a *sufficient* resolution will bring with it a high computational cost. These issues collectively introduce the issue of model hierarchy, and determining how complex a model needs to be for the purpose of its investigation.

1.2.1 Coupled Model Intercomparison Project Phase 5

In climate system modelling, there exists a spectrum of models with differing levels of complexity, each of which are suited to answering specific questions about the climate system (Meinshausen et al. 2011; Collins et al. 2013). Some of the most comprehensive models available for the investigation of centennial-scale anthropogenic climate change are fully coupled Earth System Models (ESMs): these include all known components of the climate system and therefore give the best representation of global climate processes and dynamics.

A sustained effort by modelling centres around the world to provide a co-ordinated and comparable suite of such model simulations resulted in the advent of the Model Intercomparison Projects; the most up-to-date iteration for coupled climate modelling is the Coupled Model Intercomparison Project Phase 5, or CMIP5 (Taylor et al. 2012). While many different experiments were performed as part of the CMIP5 project, the most important components for the purpose of this thesis were: (1) the 'Historical' experiment, which has boundary conditions most representative of the evolution of the global climate system over the period 1860-2005; and (2) a sequence of possible future climate projections which have been simulated for the period 2005-2100 and beyond, called the Representative Concentration Pathways (RCPs, Moss et al. 2010). Four RCP scenarios were developed (RCP2.6, RCP4.5, RCP6.0 and RCP8.5), each of which represent a different possible

socioeconomic pathway and hence representation of future GHG emissions. As Figure 1.3 shows, the four numbers associated with each scenario denote the corresponding effective radiative forcing expected by 2100. For the purposes of the analyses presented hereafter, RCP8.5 can be thought of as an acceleration of present-day GHG emission rates (Riahi et al. 2011); RCP6.0 is the closest analogue to a 'business-as-usual' pathway (van Vuuren et al. 2011a); RCP4.5 is consistent with a scenario where some mitigation actions have been implemented globally (Thomson et al. 2011); while RCP2.6 is a low emissions pathway where robust mitigation policies have been enacted (van Vuuren et al. 2011b).
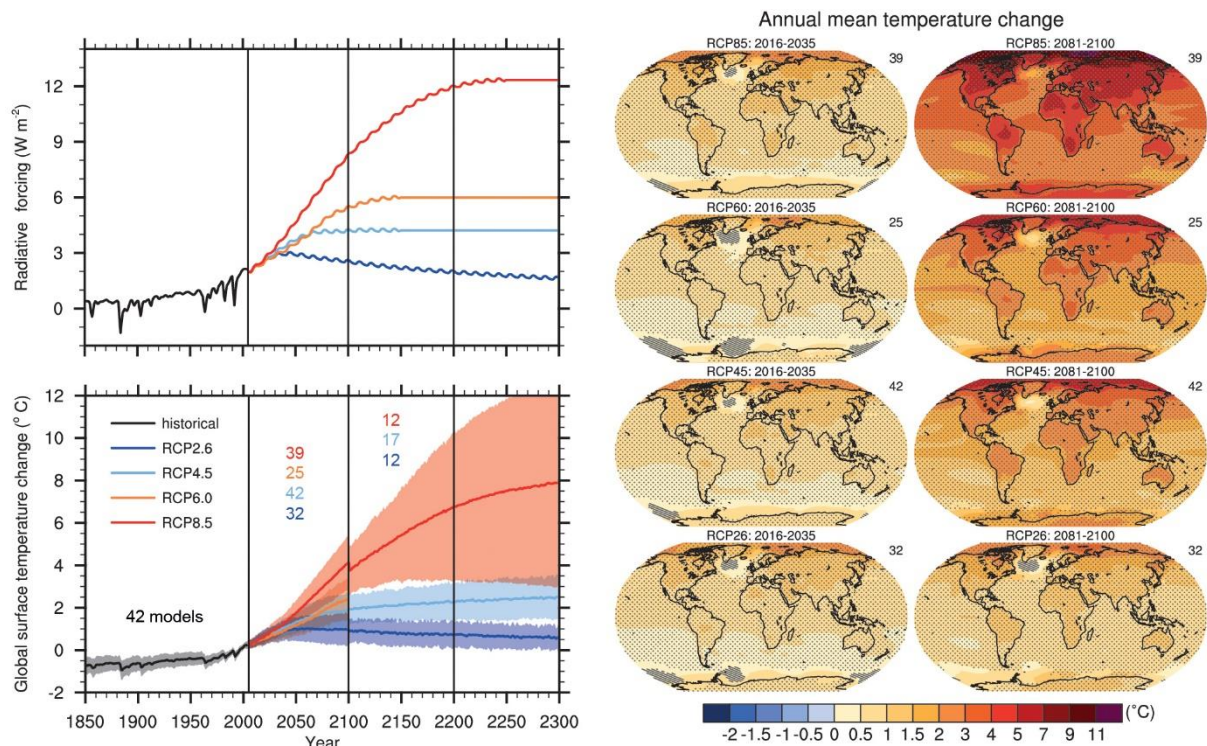
Figure 1.3: (Top left) Effective global mean radiative forcing for the four RCP scenarios based on the MAGICC energy balance model. **(Bottom left)** Time series of global annual mean surface air temperature (GMST) anomalies, relative to 1986-2005, from the key CMIP5 centennial experiments. Projections are for each RCP for the multi-model mean (thick line) and 5-95% confidence interval using all individual models. Numbers in the same colours as the lines indicate the number of different models contribution to each of the different time periods. **(Maps)** Multi-model ensemble average of GMST anomalies for 2016-2035 and 2081-2100, for each of the four RCP scenarios. Figure reproduced from the IPCC WG1 AR5 Technical Summary, Figure TS. 15 (Stocker et al. 2013).

1.3 Patterns of future climate change

Many figures in the IPCC's fifth assessment report present changes expected under future warming scenarios relative to a fixed baseline period (commonly 1986-2005). Figure 1.4 illustrates such an example, while also highlighting some of the key patterns of regional climate change which are expected with future warming.

Several robust patterns emerge when looking at regional temperature increases under future emission scenarios (Figure 1.4a). Among them, (1) land regions tend to warm faster than nearby ocean regions, owing to a variety of well-understood physical mechanisms (see Joshi et al. (2008) for further details); and (2) polar regions warm faster than the global average, owing to (among other reasons) the ice-albedo feedback process, whereby melting ice exposes darker ocean or land surfaces which can reflect less incoming radiation and hence leads to accelerated warming (Hartmann et al. 2013).
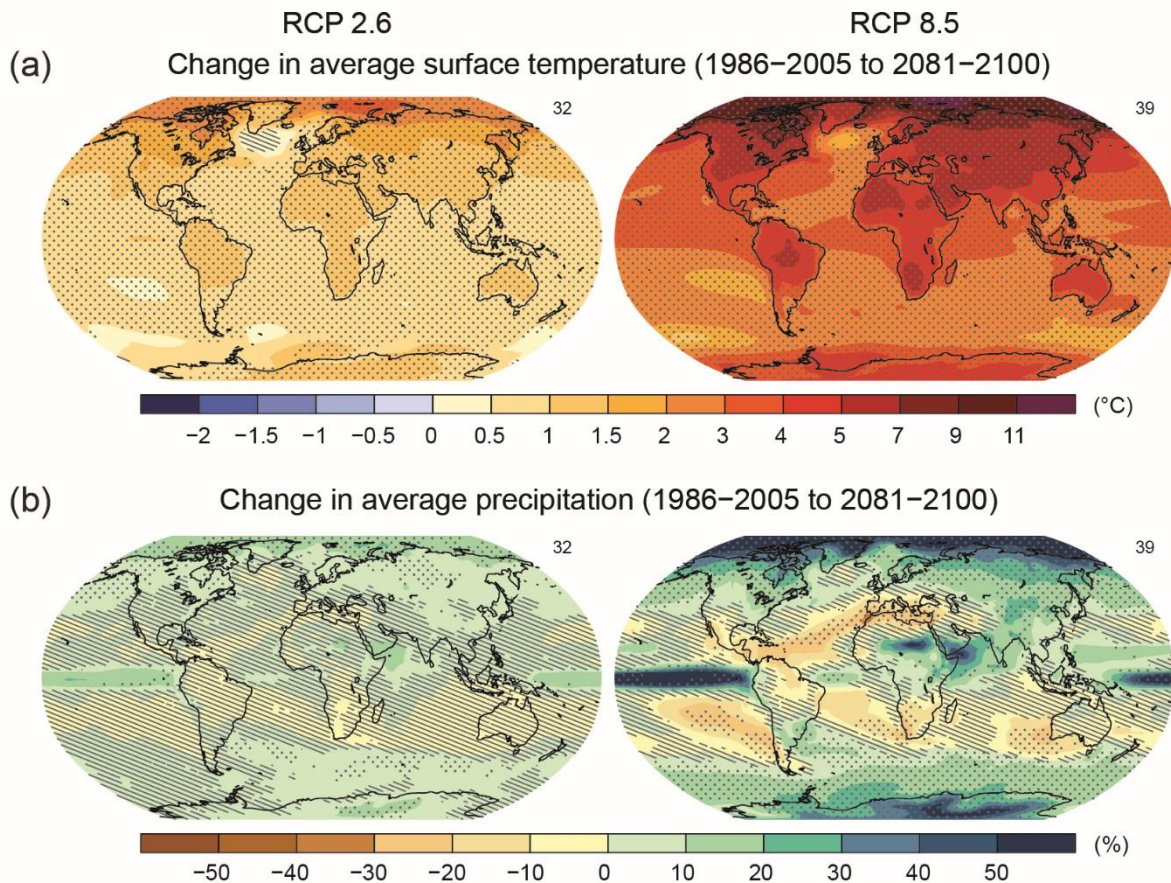
Figure 1.4: Maps of multi-model mean changes to **(a)** annual mean surface air temperature and **(b)** annual mean precipitation under a low emissions (RCP2.6, left-hand panels) and high emissions (RCP8.5, right-hand panels) scenario. Changes are presented for the period 2081-2100 relative to the period 1986-2005. Inset numbers denote number of models used for each map. Figure adapted from Figure SPM 8 in the IPCC AR5 WG1 Summary for Policymakers (2013).

When considering future changes to regional precipitation, the patterns of mean change can be broadly characterised, to first order, as following the concept of 'wet get wetter, dry get drier' (Held and Soden 2006; Allan 2014; Allan et al. 2010). Indeed, Figure 1.4b reveals a projected intensification in precipitation over equatorial and high latitude regions where climatological rainfall totals are high. Equally, a decrease in average precipitation is found over those subtropical arid regions which are associated with downwelling of the tropical

Hadley Cell (Seidel et al. 2008; Huang et al. 2016; Heffernan 2016). It is important to note that caveats to this simplistic concept do exist, especially when looking at extreme precipitation (Donat et al. 2016) and over certain land regions (Greve et al. 2014).

Interpreting how internal variability interacts with long-term warming signals has been crucial to addressing several important questions in climate science over the last decade, including, for example, the recent slowdown in the rate of global mean warming (Fyfe et al. 2013; Huber and Knutti 2014; Maher et al. 2014). Consequently, there have been increasing efforts to frame climate projections in a more policy-relevant framework: instead of quantifying what the global climate will 'look' like at future time periods (such as 2050 or 2100), analysis is focused on *when* the signal of human influence on a climate variable will emerge from the noise of natural climate variability. This concept is referred to as the 'time of emergence', or 'ToE' (Giorgi and Bi 2009; Joshi et al. 2011; Hawkins and Sutton 2012). When calculated for low- and high-emission climate scenarios, measures of ToE can provide a clear and policy-relevant demonstration of the time saved by lowering future emissions. It also represents a potentially powerful tool for adaptation planning at both global and regional scales (Hawkins et al. 2014).
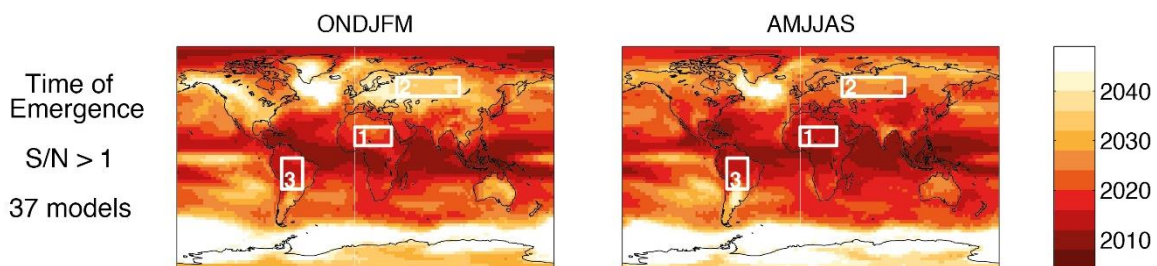
Figure 1.5: Time of Emergence (ToE) of significant local warming derived from 37 CMIP5 models under the RCP4.5 scenario. Warming is quantified as the half-year mean temperature anomaly relative to 1986–2005, and the noise as the standard deviation of half-year mean temperature derived from a control simulation of the relevant model. Central panels show the median time at which the signal-to-noise ratio exceeds a threshold value of 1 for **(left)** the October to March half year and **(right)** the April to September half year, using a spatial resolution of 2.5° × 2.5°. Reproduced from IPCC AR5 WG1 Chapter 11 *(Kirtman et al. 2014)*.

The ToE concept can be applied to a range of different statistical thresholds, depending on the research question: for example, when a specified temperature target (such as 2°C above pre-industrial levels) is exceeded (Joshi et al. 2011), or perhaps when the GHG-forced temperature signal-to-noise ratio exceeds one (Hawkins and Sutton 2012, Figure 1.5) – this is equivalent to the magnitude of the forcing signal exceeding one standard deviation for a given distribution. More recently, the ToE concept has been expanded to other climatic variables, such as sea level rise (Lyu et al. 2014) as well as certain components of the global

water cycle (Sedláček and Knutti 2014). Further discussion of variations on the ToE concept can be found in Sections 5.1 and 7.1.

In this thesis, we consider *absolute* measures of climate change as those examples where the difference is calculated between two time periods for a given variable, while measures of the *emergence* of climate change are hereafter defined as anomalies over time which have then been *normalised by local climate variability*.

1.4 The science of extreme event attribution

Understanding how the signal of climate change emerges from the background noise of internal variability is clearly an important question for the climate science community. However, while this information is useful for policy makers who are making decisions about the future, it provides little value to people in demonstrating that the impacts of climate change are in fact already being experienced today. Consequently, increased research efforts have been placed on quantifying how the likelihood of extreme weather events from the present-day are changing in response to anthropogenic influences on the climate system. These studies primarily concern events which have led to devastating societal impacts, and thus represent tangible examples of unfamiliar climatic conditions that different communities around the world can relate to. The fundamental difficulty with the concept of extreme event attribution is that these events are, by definition, extremely rare. Only in the last five-to-ten years have concerted modelling efforts enabled a proper assessment of the change in probability of these specific events occurring.

One approach to extreme event attribution is to employ a very large (often multi-thousand member) ensemble of climate model simulations which are driven with the anthropogenic forcing mechanisms of the present day (Stott et al. 2016). Depending on the specific question which is being asked, these simulations may be further conditioned with the observed atmospheric or oceanic state (such as observed sea surface temperatures), to obtain a more precise understanding of anthropogenic influences on specific mechanisms contributing to the event (Shepherd 2016). These model simulations are then compared with identical model runs, except this time with all anthropogenic components from the relevant forcing mechanisms removed. This way, one can consider the number of times the rare event in question happens in each model ensemble, and quantify a probabilistic change in the likelihood of that event occurring as a result of those anthropogenic forcing components

which have been removed (Allen 2003). Two metrics are commonly used to quantify this probabilistic change: the 'probability ratio' (also called the risk ratio) represents how many more times that event occurs in the model ensemble which includes human influences; similarly, the 'fraction of attributable risk' (or FAR) denotes what *additional* fraction of event occurrences are found in the all-forcings ensemble, relative to the ensemble with human influences removed. Further explanation of these concepts can be found in Sections 2.4, 3.1 and 4.2.

Because the field of extreme event attribution is still so new, there is no universally accepted methodology for evaluating the role of human influence for a given event - the preceding paragraph represents one of the more common approaches. It is widely accepted, however, that several independent methods of analysis should be employed, and only when each of these approaches reach a similar conclusion can further confidence be placed on the final answer. Good examples of multi-method analyses include publications by King et al. (2015b) and Uhe et al. (2016), and further discussion of alternative methodologies can be found in Section 3.1.

1.4.1 Modelling approaches for event attribution

Alongside CMIP5 models, there are a multitude of alternative modelling frameworks which have been developed in recent years, particularly in the context of addressing the requirements for probabilistic event attribution. The methodologies highlighted below should be considered only a snapshot of the range of tools available, rather than a complete and exhaustive record.

Climateprediction.net was a revolutionary project which utilised the spare computing power of home computers from thousands of volunteers to run very large climate modelling experiments (Frame et al. 2009). This technique of distributed computing has successfully generated multi-thousand member climate model ensembles, which enables a quantitative method of assessing model uncertainty for future climate simulations (Rowlands et al. 2012). A more recent type of experiment, called 'Weather@Home' (Massey et al. 2015), focused on developing initial condition (IC) ensembles for specific regions around the world. IC simulations are performed by using the same model with identical atmospheric physics, and slightly perturbing the start date from which each simulation is run. When a comparison is made between many IC model runs using a best-guess estimate of the observed climate, with

an otherwise equivalent ensemble but with anthropogenic forcing signals removed, a probabilistic estimate of changes to certain aspects of the climate, including extreme events, can be found. More recently, these techniques have been further refined by using seasonal SST forecasts to provide estimates of probabilistic changes to the likelihood of extreme events in near real-time (Haustein et al. 2016).

Variations on this approach of large initial condition ensembles, often using atmosphere-only regional climate models with observed SST anomalies as a boundary condition, have also been explored by the Weather Risk Attribution Forecast project (WRAF, see http://web.csag.uct.ac.za/~daithi/forecast/), and several publications have made use of this data (Angélil et al. 2014, 2016).

Other modelling approaches for probabilistic event attribution have included using multiple ensemble runs for fully coupled ESMs (Kay et al. 2015), as well as using seasonal forecasting models with boundary conditions adjusted to (1) present-day and (2) historical levels of greenhouse gas forcings, with changes to the extreme event of interest identified as a result (eg, Hope et al. (2016)).

1.5 Ongoing challenges and research priorities

The confidence with which we can ascribe statements about the change in likelihood of witnessing a certain type of extreme event vary greatly, depending on the type of climate extreme being considered. The synthesis of the state of event attribution science by the US National Academy of Sciences in 2016 helped to summarise this point in Figure 1.6: here it can be seen that those types of extreme events which are direct, first-order manifestations of a warming world, like extreme heat and cold events, are well-understood in terms of both the expected changes with climate change, as well as our ability to simulate these changes. Conversely, drought is an example of an extreme climate event where there is less confidence in the capability to make reliable attribution statements, due to both an incomplete understanding of how the complex mechanisms which lead to severe drought will change in a warming world, and lower confidence in the ability of models to simulate the complex coupling mechanisms which result in severe drought.
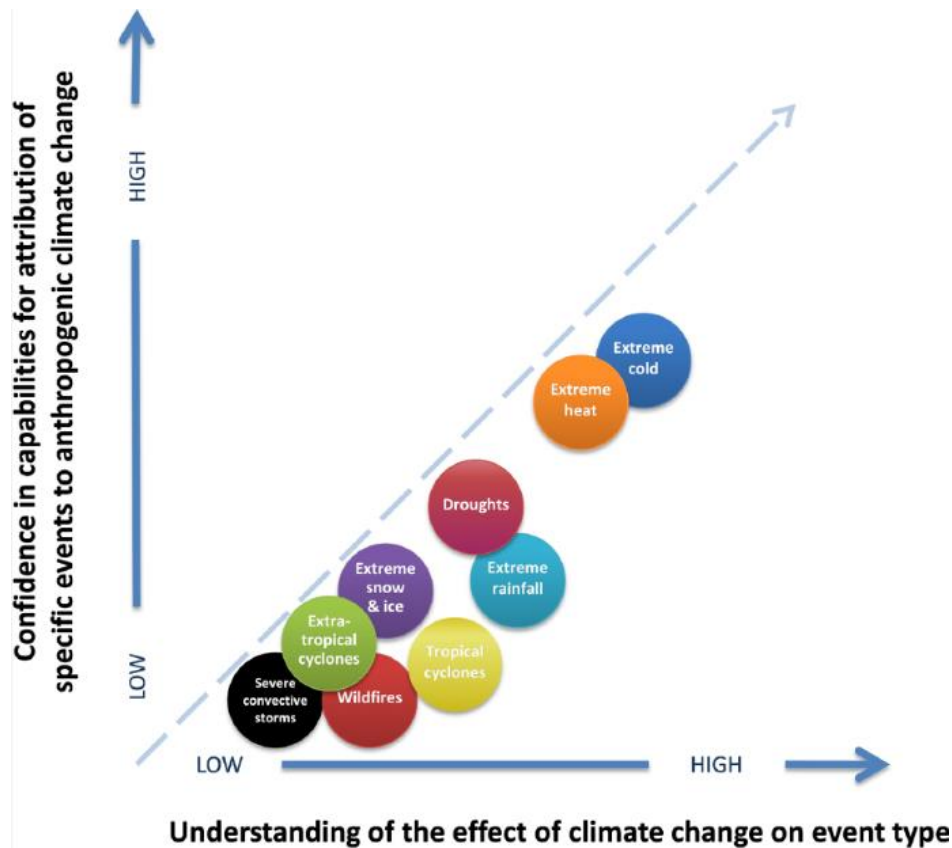
Figure 1.6: Schematic depiction for the state of attribution science for different event types. The horizontal position of each event type reflects an expert assessment of the level of understanding of climate change on the event type. The vertical position of each event type indicates an assessment of scientific confidence in current capabilities for attribution of specific events to anthropogenic climate change for that type of event. Figure reproduced from the US National Academy of Sciences (2016).

More recently, there have also been questions about what frameworks of analysis for extreme event attribution are most appropriate. For example, some scientists argue that identifying human-induced changes to the dynamical drivers of extreme events should not be attempted, as model uncertainty related to climate dynamics is too high, such that any (otherwise robust) thermodynamic signal may be lost as a consequence (Trenberth et al. 2015; Shepherd 2016). Other researchers have pointed out that quantifying thermodynamic changes to the likelihood of witnessing an extreme event in isolation will be much less helpful for planners and decision makers in the future, and that dynamical influences cannot be dismissed (Otto et al. 2016; Stott 2016; Stott et al. 2016).

Finally, there are clearly overlapping themes in the published literature between the time of emergence of mean climate signals, and the attributable change in likelihood of witnessing extreme events in the present day. Nevertheless, these two fields have primarily evolved independently, and to date there has been little exploration of a framework to formalise how these two fields of research relate to one another.

These research gaps all warrant further investigation, and hence motivate the analyses presented in this thesis. Specifically, I will focus on three research themes:

1) Investigating new approaches to event attribution for meteorological drought, through the use of novel statistical techniques.

2) Explore issues relating the sensitivity of attribution statements to choices about analytical frameworks

3) Examine robust differences in the emergence of signal-to-noise ratios in temperature for different populations of the world, and use this to reconcile event attribution techniques as part of a broader climate emergence framework.

1.6 Thesis outline

It is becoming increasingly clear that *absolute* measures of regional climate change have specific limitations. Instead, understanding how the magnitude of anthropogenic changes to a climatic variable compares with internal variability may prove more useful, especially when interpreting the potential impacts of climate change. This thesis aims to extend techniques relating to both the time of emergence of climate signals, as well as techniques related to probabilistic attribution of extreme events, before demonstrating how the two sub-fields can be understood as related elements of interpreting signal-to-noise ratios in a changing climate.

Chapter 2 reconciles known model uncertainties in the simulation of meteorological drought over New Zealand with a robust understanding of anticipated changes in large-scale synoptic weather patterns under climate change. A novel statistical approach is presented to evaluate the change in likelihood of specific drought events occurring in response to anthropogenic climate warming, using the 2013 New Zealand drought as a case study.

Chapter 3 considers the limitations of metrics commonly used to quantify the role of human influence in specific extreme events. A method is then proposed to demonstrate how the 'time of emergence' literature can be interpreted as a foundation for probabilistic event attribution, with particular relevance for real-time attribution approaches in the context of future climate services.

Chapter 4 focuses on a more well-understood type of extreme event: that of very hot days (NAS 2016). By focussing on a type of climatic extreme which is easier to characterise in climate models, I demonstrate a method of comparing the increasing frequency of very hot

days with concurrent model simulations of cumulative carbon dioxide emissions. By focussing on spatially aggregated patterns of change, robust patterns of differential population exposure to future climate impacts are found, despite model uncertainty at finer spatial scales.

Chapter 5 presents a discussion of some of the policy ramifications of considering climate change in a signal emergence framework. To illustrate such a point, a novel method is proposed for distributing a global carbon budget at a national level, under the assumption that those nations which will experience the most rapid and significant climate change emergence are entitled to emit proportionately more CO_2 emissions.

Chapter 6 explores some limitations associated with the time-evolution of extreme event attribution metrics, before explicitly demonstrating how signal-to-noise ratios and changes to the likelihood of extreme event occurrences can be reconciled with one another.

Finally, Chapter 7 discusses some important caveats when interpreting the results from the preceding chapters, before Chapter 8 outlines some key conclusions.

1.6.1 Models used in thesis

For each analysis presented in this theses, only models from the CMIP5 archive are used. The use of many different ESMs from modelling centres around the world will provide a good estimate of model uncertainty, particularly relative to using multiple simulations from only one type of model (Bellprat and Doblas-Reyes 2016). However, it is important to recognise that a CMIP5-based multi-model ensemble is an 'ensemble of opportunity' and the full space of possible model uncertainty can never be sampled in its entirety (Sanderson et al. 2015). This is a long-recognised issue for any climate modelling framework, and such a limitation should be considered carefully when interpreting the results presented in this thesis. Furthermore, not all models contributing to the CMIP5 ensemble are independent from one another, owing to the fact that constituent components of different modelling frameworks often have a shared genealogy (Knutti et al. 2013). Adjusting any conclusions using a multi-model ensemble to account for this inherent degeneracy is extremely difficult, and remains a source of active research (Masson and Knutti 2011; Sanderson et al. 2015). A CMIP5-based approach nevertheless remains the best available method of estimating model uncertainty associated with any climate change-related analysis, and thus will be the approach hereafter

employed. Further discussion about the strengths and limitations of the models used in this thesis can be found in Sections 2.6 and 7.2.

For each chapter, all models are selected from the CMIP5 archive which have output available (1) for the required variables (either total precipitation, sea level pressure, or surface air temperature), (2) at the required temporal resolution (either monthly or daily timescales), and (3) for the experiments required for calculations (either 'Historical', 'RCP2.6', 'RCP4.5' or 'RCP8.5' runs). Only a single ensemble member is used from each model available, so as to ensure the multi-model estimates of statistical uncertainty are not artificially reduced (Knutti et al. 2009).

1.6.2 Relation to published work

This thesis primarily consists of individual, though related, pieces of analysis, with most chapters corresponding to an original research article which has either been published in the peer review literature or is in preparation.

Chapter 2 has been published in *Journal of Geophysical Research: Atmospheres* (doi: 10.1002/2016JD025602), and was written by myself, with Peter Gibson (PG), Sam Dean (SD), Suzanne Rosier (SR), Daniel Mitchell (DM) and Dave Frame (DF) also contributing authors. I conceived the ideas for the paper, performed the analysis and wrote the manuscript; PG assisted with the preliminary processing of the data used; SD, SR, DM and DF provided helpful comments on earlier draft versions of the paper.

Chapter 3 has been published in *Climatic Change* (http://dx.doi.org/10.1007/s10584-017-1906-3); I am the sole author on this piece of work.

Chapter 4 has been published in *Environmental Research Letters* (http://dx.doi.org/10.1088/1748-9326/11/5/055007) as well as being the subject of a commentary piece within the same journal (http://dx.doi.org/10.1088/1748-9326/11/9/091002). While I was lead author on this work, co-authors include Dave Frame, Manoj Joshi (MJ), Ed Hawkins (EH), Erich Fischer (EF) and Chris Jones (CJ). Specifically, I conceived the ideas for the paper, performed the analysis and wrote the manuscript; CJ provided model data on cumulative CO_2 emissions used in the paper; EH and EF offered feedback on the framework of uncertainty analysis within the manuscript, while MJ and DF offered helpful comments to strengthen the final version of the manuscript.

The work presented in Chapters 5 and 6 is in the final stages of preparation, with the intention of also being submitted to *Environmental Research Letters*.

Alongside the work presented explicitly within this thesis, I have published closely related work in the *Bulletin of the American Meterological Society* (http://dx.doi.org/10.1175/1520-0477-95.9.S1.1), and contributed to related analysis which was led by Andrew King and published in *Geophysical Research Letters* (doi: 10.1002/2015GL067448).

Chapter 2

Investigating event-specific drought attribution using self-organizing maps

The work described in this chapter follows Harrington et al. (2016b),published in *Journal of Geophysical Research: Atmospheres*.

Previous studies evaluating anthropogenic influences on the meteorological drivers of drought have found mixed results owing to (1) the complex physical mechanisms which lead to the onset of drought, (2) differences in the characteristics and timescales of drought for different regions of the world, and (3) different approaches to the question of attribution. For a mid-latitude, temperate climate like New Zealand, strongly modulated by oceanic influences, droughts last on the order of three months, and are less strongly linked to persistent temperature anomalies than continental climates. Here, we demonstrate the utility of a novel approach for characterizing the meteorological conditions conducive to extreme drought over the North Island of New Zealand, using the January-March 2013 event as a case study. Specifically, we consider the use of self-organizing map (SOM) techniques in a multi-member coupled climate model ensemble to capture changes in daily circulation, between two 41-year periods (1861-1901 and 1993-2033). Comparisons are made with seasonal pressure and precipitation indices. Our results demonstrate robust (>99% confidence) increases in the likelihood of observing circulation patterns like those of the 2013 drought in the recent-climate simulations when compared with the early-climate simulations. Best-guess estimates of the fraction of attributable risk range from 0.2 to 0.4, depending on the metric used and threshold considered. Contributions to uncertainty in these attribution statements are also discussed.

2.1 Introduction

The mechanisms which contribute towards the onset of a drought event are complex, and require the understanding of coupled land-atmosphere feedback processes in order to be accurately characterised (Dai 2011a; Seneviratne 2012). The wide range of metrics employed in the peer-reviewed literature to quantify the severity of drought (Heim 2002; Zargar et al. 2011) is indicative of the difficulty faced by the research community in

characterising drought events. Given this difficulty in even quantifying the magnitude of an observed drought event, attempting to approach the question of attribution is exceptionally challenging.

2.2.1 Previous research

Previous approaches towards drought attribution have found differing results for the same "event", depending on which meteorological aspects of the observed drought event were considered for analysis. The majority of studies which have considered changes to the likelihood of extreme precipitation deficits (often linked to a subsequent drought event), for example observed events in Australia (King et al. 2014); Brazil (Otto et al. 2015a); the Horn of Africa (Marthews et al. 2015) or the Central US (Rupp et al. 2015), did not detect any statistically significant differences between model simulations which included observed anthropogenic influences on the climate system and those which did not. Two independent studies did however find human-induced increases in the likelihood of exceptionally low seasonal precipitation accumulations over the Middle East region (Bergaoui et al. 2015; Kelley et al. 2015), a result which can be explained as a unique regional response to the weakening of the Mediterranean storm tracks with climate warming, resulting in enhanced Eastern Mediterranean drying through circulation changes (Hoerling et al. 2012).

While precipitation deficits are a dominant factor in droughts (Williams et al. 2015), there are nevertheless other critical non-linearities and couplings in the system that are more difficult to characterise, including coupled land-atmosphere feedbacks on small scales (Seneviratne et al. 2010). Some studies have looked at the co-occurrence of extreme temperatures and precipitation deficits: both King et al. (2014) and Rupp et al. (2015) consider this link in the context of whether increased dryness exacerbates the likelihood of concurrent hot years occurring, but find no discernible change in the likelihood of annual precipitation deficits. Diffenbaugh et al. (2015) meanwhile notes the record dryness over California for the 2012-2014 drought event coincides with record high temperature anomalies and subsequently demonstrates the increased chance of these extremes co-occurring have led to an increased probability of extreme drought over the region. Williams et al. (2015) goes one step further and decomposes the relative contributions of changing temperatures and precipitation patterns towards observing the exceptionally low recorded values of the Palmer Drought Severity Index (PDSI), again with respect to the 2012-2014 California drought.

Other approaches have used multi-thousand member initial-condition ensembles to constrain an estimate of changes to simulated precipitation minus simulated evaporation (P-E), and found no attributable change in the likelihood of record low values for the 2014/15 Brazil drought (Otto et al. 2015a). Interestingly, the use of a coupled land surface model in an otherwise equivalent model ensemble approach by Rupp et al. (2015) actually showed a decrease in the likelihood of exceptional soil moisture deficits over the summer months for Texas, despite the large (at least ten-fold) anthropogenic increase in the probability of record hot temperatures also occurring.

Simulating the causal chain of processes which leads to the occurrence of severe hydrological or agricultural drought is a non-trivial exercise (Seneviratne et al. 2010; Orth et al. 2015), and the ability of coupled climate models to realistically simulate the coupled mechanisms which contribute to a severe drought is a fundamentally limiting constraint (Burke and Brown 2008; Dai 2011a; Lorenz et al. 2016). Whilst only looking at precipitation deficits as a key meteorological determinant of drought is unsatisfactory, simulating a more complex measure of drought severity will likely lead to mischaracterising uncertainty when ascribing a change in event probability to anthropogenic influences. An alternative and novel proposal suggested here is to primarily focus on the large-scale circulation patterns which occur over the course of a given drought event. Without making prior assumptions about which specific circulation types are most conducive to drought, we assess how close modelled summers come to replicating the exact frequency and persistence of the circulation regimes which were observed during a specific drought event.

2.2.2 Event-specific attribution using self-organizing maps: framing the question

When considering probabilistic event attribution, there is a question of utility in how specifically tailored towards an observed event the analysis should be, in order for the result and any subsequent attribution statement to be suitably unbiased (NAS 2016; Stott 2016; Stott et al. 2016). The most common approach is for a specific event to actually be generalised as a class of event, whereby some metric is used to characterise the type of event in question (extreme daily temperatures for example), and probabilistic estimates of the fraction of attributable risk (hereafter FAR, (Allen 2003)) are made by then considering occurrences or exceedances of a threshold which was observed for the actual event (Shepherd

2016). This way, attribution statements can have relevance for the occurrence of similar types of events in the future.

While explored less frequently, one can also consider the in-depth characteristics specifically related to the observed event only (Hannart et al. 2016) – examples include using analogue approaches in an attempt to understand differences between dynamical and thermodynamic effects on a given event probability (Yiou et al. 2007). In this study, we propose to use self-organizing map (SOM) techniques to characterise the synoptic circulation of each day over the 3-month period of January to March 2013 (herein JFM13), during which the most severe drought on record occurred over the New Zealand region (Harrington et al. 2014). Each day is assigned one of twelve possible circulation regime types (hereafter SOM nodes), based on which pattern most closely resembles the modelled circulation for that day: these twelve SOM nodes are identified using reanalysis data over the period 1979-2014, and are intended to characterise the range of typical summer circulation patterns. We will then look at how closely each individual summer in an ensemble of fully coupled climate model simulations from the Coupled Model Intercomparison Project Phase 5 (CMIP5, (Taylor et al. 2012)) resembles the specific frequency and persistence of each SOM node which was observed for the JFM13 observed event. Hence, rather than evaluating the probabilistic change to some metric which is intended to characterise drought over the North Island, this approach evaluates whether the specific sequence of synoptic circulation patterns observed over the summer of JFM13 were made more or less likely to occur, as a result of human influences on the climate system.

2.2 The 2013 North Island drought case study: event description

Beginning in early January 2013, the region of the southwest Pacific over New Zealand experienced an unusual sequence of persistent, or ''blocking'', high-pressure anticyclones (Harrington et al. 2014). Over the subsequent three-month period to the end of March, the northern half of the country experienced the most extreme drought over 41 years of observations, both in terms of the severity of soil moisture deficit (when characterised using a potential evapotranspiration deficit metric, or PED) and the areal extent over which record dryness was experienced (Harrington et al. 2014; Figure 2.1). At the peak of the event, it was estimated that only 20 days of water supply remained for the capital city of Wellington (http://bit.ly/1sQwD6u). Subsequent economic analysis suggests the drought reduced the

country's annual gross domestic product by approximately 0.6 percentage points, equating to at least US$1.3 billion in lost revenue (New Zealand Treasury 2013).
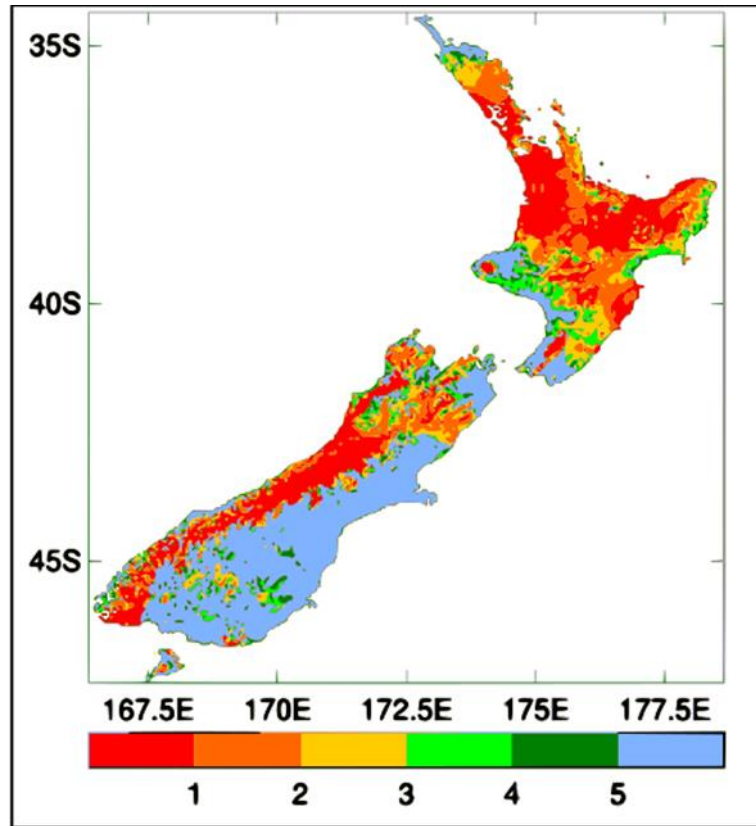
Figure 2.1: Ranking of July 2012−May 2013 potential evapotranspiration deficit (PED) relative to the previous 40 years in the National Institute of Water and Atmospheric Research Virtual Climate Station Network (VCSN) data set (Tait et al. 2006, doi: 10.1002/joc.1350). Rank 1 (red) means 2012-2013 was the highest PED over that area in 41 years; orange refers to second-highest PED and so on. Light blue colours signify the PED was not in the top five.

Monthly mean sea level pressure (MSLP) data, using ensemble mean data from the 20th Century Reanalysis Project version 2c (hereafter 20CR, (Compo et al. 2011)), shows anomalously high pressure situated over the New Zealand region for the summer of 2013 (Figure 2.1a), though not record breaking (Harrington et al. 2014). Despite this, the self-calibrated Palmer Drought Severity Index (PDSI, Dai 2011b) shows the summer of 2013 exhibited record low levels of soil moisture over the North Island region (Figure 2.1b), consistent with the results using nationwide PED estimates. For subsequent analysis of MSLP and precipitation deficits in Section 3.2, we use the regions marked as the green (37°-45°S, 163°-178°E) and blue (34°-40°S, 172-178°E) boxes in Figure 2.1a respectively – these correspond to the regions where anomalies of each corresponding variable were most extreme over the duration of the event (Figure 2.2).
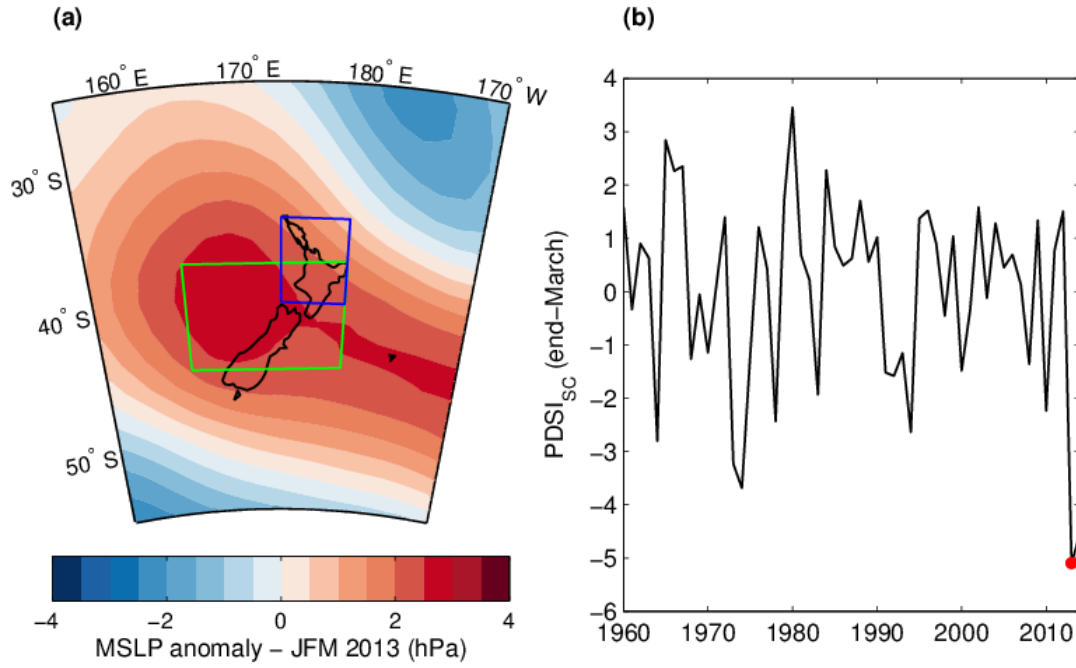
Figure 2.2: (a) MSLP anomalies over the New Zealand domain for January - March 2013. Anomalies are taken from the 20th Century Reanalysis Project v2c with respect to a 1981-2010 climatology. The domain size corresponds to that used for subsequent self-organizing map analyses. Blue and green rectangles show, respectively, the sub-regions used for the area-averaged MSLP and precipitation analysis in Section 3.2.2. **(b)** Monthly self-calibrated PDSI values (see Dai (2011a) for details), averaged over the blue rectangle in panel (a), taken at the end of March for each calendar year. Negative values indicate soil moisture deficits. Red circle indicates value for 2013.
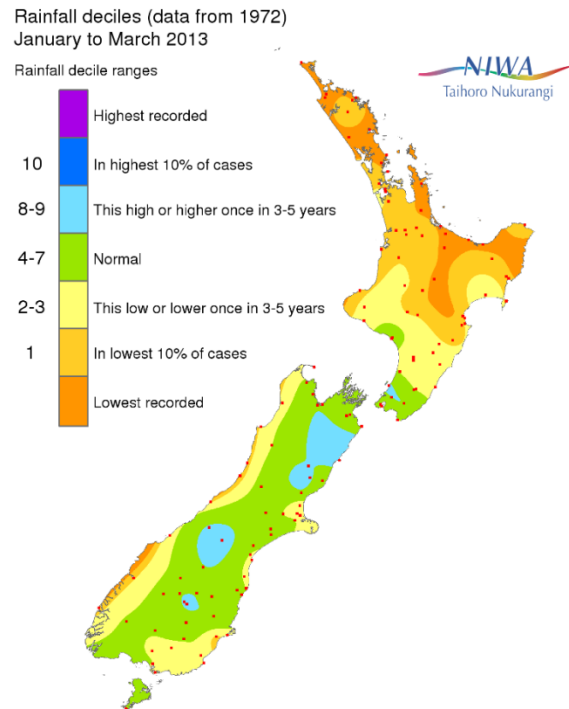
Figure 2.3: Precipitation deciles for the summer of January to March 2013 using rain-gauge based observations. Figure obtained with permission from the National Institute of Water and Atmospheric Research.

2.3 Methods

2.3.1 Self-organizing maps

Self-organizing maps (SOMs) are a powerful technique to characterise the synoptic climatology of a given region, and provides a method of quantifying how well climate models represent day-to-day variability in typical synoptic situations. In this study, a sequential SOM is used with a random initialization scheme, following Hewitson and Crane (2002) : a random initialization scheme was chosen over other methods (such as in Reusch et al. (2005)) to minimise computational memory requirements whilst having a negligible effect on the resultant SOM configurations (Gibson et al. 2016b).

The SOM training procedure involves the presentation of input data cases (in this case the daily MSLP data for the region), randomly chosen and presented individually, to iteratively update each node. The Euclidean distance measure between data cases and each node is then used to determine the "winning node", consistent with previously published approaches (e.g. Cassano et al. (2006); Alexander et al. (2010)). Both the winning node and surrounding nodes are updated in this process using a Gaussian neighbourhood update function. The learning rate and radius parameter in the SOM were set to decrease linearly to zero and one, respectively. Sensitivity testing determined that the starting values of these parameters had negligible influence on the final maps produced. The reader is referred to Hewitson and Crane (2002) for further discussion and justification of methodological choices of the SOM procedure.

The end result of this process is a map of 'self-organized' nodes where the order of these nodes on the two-dimensional array map relates to a degree of similarity between nodes (Figure 2.4), though we emphasise the equidistant configuration of these nodes in grid form should not be interpreted as a quantitative measure of similarity between neighbouring nodes. The MSLP patterns shown across the nodes are designed to preserve and span the typical range of daily spatial patterns found in the data. An attractive feature of this methodology is that node patterns trained from reanalysis data can then be used in the evaluation of climate models. For example, at each daily time step the MSLP field in a particular model is allocated to one of the SOM nodes trained through reanalysis.

We calculate SOMs over a spatial domain around New Zealand (25°-55°S, 155°-190°E) using daily 20CR MSLP data over the months of January to March, for the period 1979-2014:

only training the SOM nodes with summertime MSLP data enables a more in-depth consideration of features specific to the 2013 drought event. We have chosen to use 20CR reanalysis data for this analysis, primarily because Gibson et al. (2016b) have demonstrated that 20CR is a reliable reanalysis product for studying large scale circulation features over the New Zealand domain for the period of analysis considered. Comparisons with alternative reanalysis products are also briefly considered in Section 2.3.3.

A configuration of 12 possible SOM nodes is selected for subsequent analysis, though alternative configurations of 20 and 30 nodes demonstrate very similar results (see Appendix). We then consider three metrics to represent the collective properties of synoptic circulation over each summer season: node frequency, average lifetime and maximum lifetime. Node frequency is defined as the total sum of days when a given node occurs, expressed as a percentage; average lifetime is defined as the average duration (days) each node persists for; maximum lifetime is defined as the longest continuous number of days over which the same node persists.
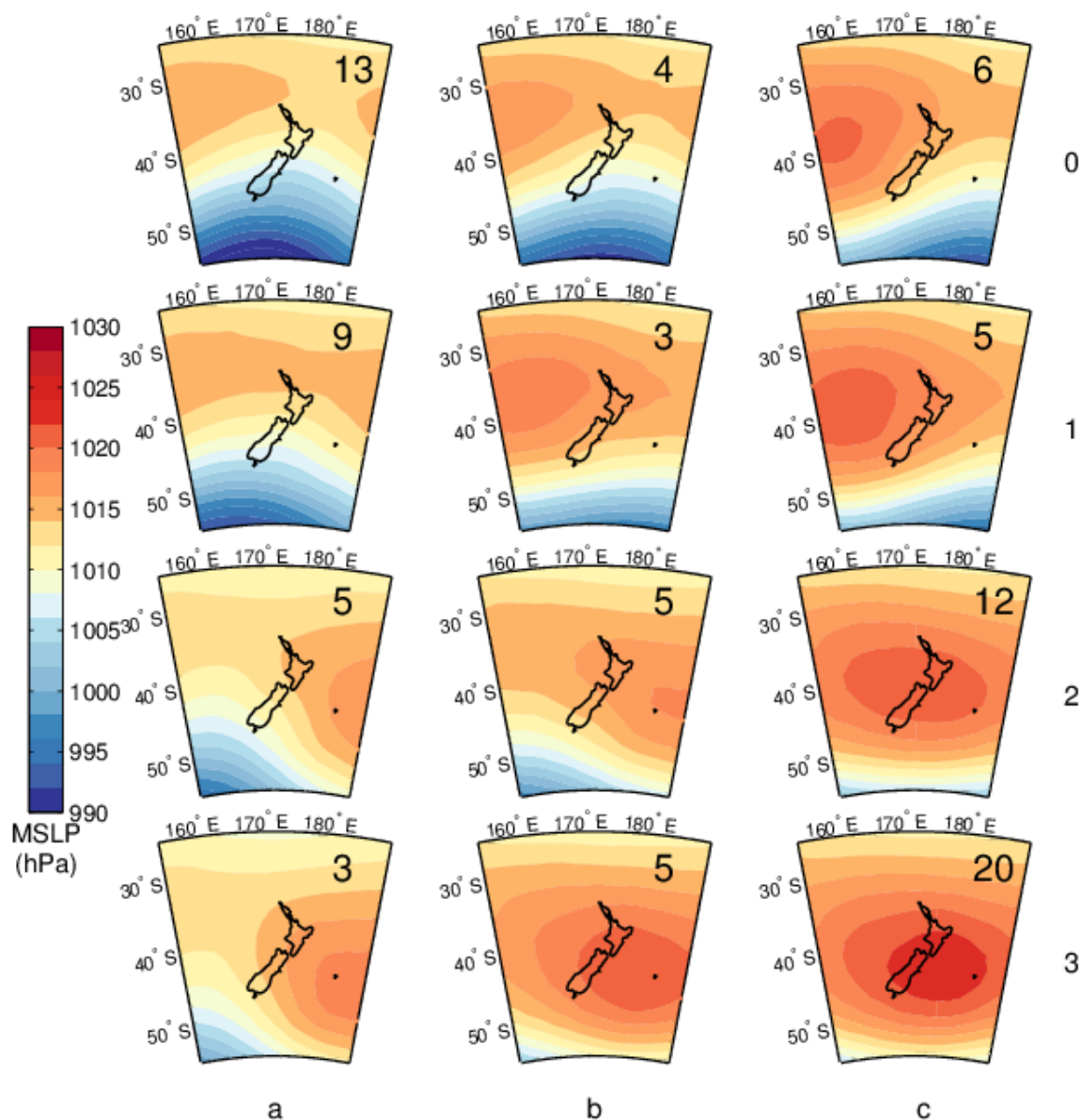
Figure 2.4: SOM nodes trained from 20CR daily MSLP for the months January to March over the period 1979–2014. In terms of referencing, node 'a3' refers to the node in the lower left corner of the SOM plane. Inset numbers refer to the number of days each node was observed over the duration of the JFM13 drought.

2.3.2 The 2013 drought in the context of self-organizing maps

Figure 2.5 reveals how the SOM properties for the summer of 2013 compare from a climatological perspective. The box-and-whisker plots present statistics based on the aforementioned SOM metrics for each summer from 1979-2014: in each box, the centre is the median, the boxes are the upper and lower quartiles, while the maximum whisker length is 1.5 times the inter-quartile range (following Tukey (1977)). The blue line in panels (a) and (b) represent the SOM metrics when all summer days from all years are aggregated together. When compared with the climatological average, we see that the JFM13 (red circles)

25

exhibited anomalously high frequencies of nodes a0, a1, c2 and c3, though none of these individual frequencies were the highest on record. In terms of persistence metrics, the average lifetime scores for 2013 were broadly consistent with the climatological average for all nodes – while node c3 represented an exception, there were still several other years which recorded exceptionally longer average lifetimes. The patterns which exhibited anomalously high (low) node frequencies also showed similar statistics for maximum node lifetime.
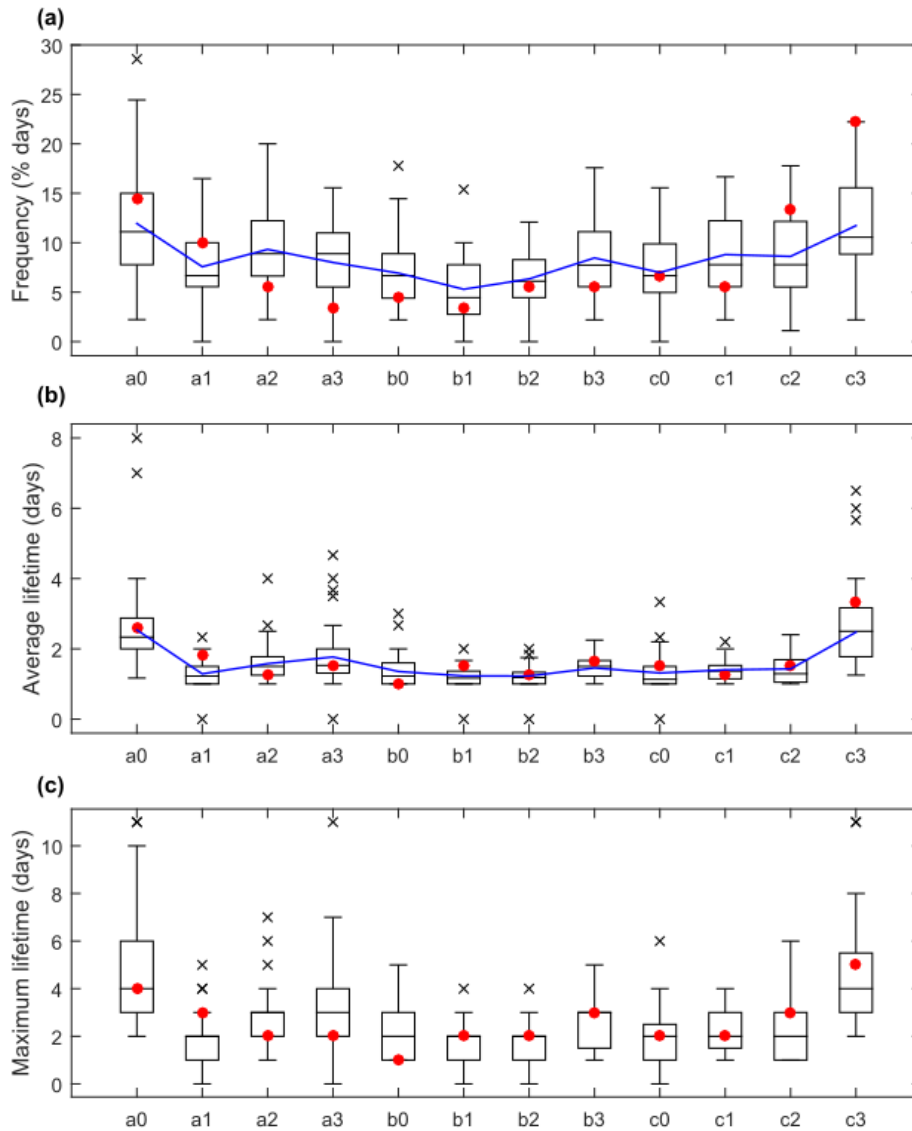
Figure 2.5: Box plots showing, for each SOM node, the observed (a) frequency, (b) average lifetime and (c) maximum lifetime for all summers between 1979 and 2014. In each box, the centre is the median, the boxes are the upper and lower quartiles, while the maximum whisker length is 1.5 times the inter-quartile (IQ) range. All outliers beyond this 1.5*IQ range are shown as black crosses. Red filled circles show the observed statistics for JFM2013. Blue lines correspond to the climatological mean.

2.3.3 Evaluating CMIP5 model fidelity for attribution using SOMs

A more recent application of self-organizing maps has been to assess which models in the CMIP5 archive best simulate the observed synoptic climatology over a given region (eg. Gibson et al. (2016b)). Here, we consider how this could be used to assess the climatology of summertime circulation over the New Zealand region for each model.

Model Number	CMIP5 Model Name	Approximate horizontal resolution (latitude° x longitude°)
1	ACCESS1-0	1.25 x 1.875
2	ACCESS1-3	1.25 x 1.875
3	bcc-csm1-1-m	1.125 x 1.125
4	bcc-csm1-1	2.8125 x 2.8125
5	CanESM2	2.8125 x 2.8125
6	CCSM4	0.9 x 1.25
7	CESM1-CAM5	0.9 x 1.25
8	CNRM-CM5	1.41 x 1.41
9	CSIRO-Mk3-6-0	1.875 x 1.875
10	GFDL-CM3	2.0 x 2.5
11	GFDL-ESM2G	2.0 x 2.5
12	GFDL-ESM2M	2.0 x 2.5
13	HadGEM2-ES	1.25 x 1.875
14	IPSL-CM5A-LR	1.875 x 3.75
15	IPSL-CM5A-MR	1.25 x 2.5
16	MIROC5	1.41 x 1.41
17	MIROC-ESM	2.8125 x 2.8125
18	MIROC-ESM-CHEM	2.8125 x 2.8125
19	MPI-ESM-LR	1.8 x 1.8
20	MPI-ESM-MR	1.8 x 1.8
21	MRI-CGCM3	1.125 x 1.125
22	NorESM1-M	1.875 x 2.5

Table 2.1: CMIP5 model runs used for analysis. Only the 'r1i1p1' runs from each model were used.

For 22 models from CMIP5 (Table 2.1), we compile daily MSLP data over the same 36-year period as the available reanalysis data, using 'Historical' data from 1979-2005 concatenated with 'RCP8.5' simulations over the period 2006-2014. For stability of the SOM procedure, a second order conservative remapping scheme was employed to regrid all models and reanalyses to a common $1.5° \times 1.5°$ grid. By considering only the three months of JFM over which the SOM nodes were trained, corresponding SOM time series are calculated for each model by identifying the closest resemblance to each SOM node for a given day. Based on this data, we consider 1) the total number of days each node is observed between each

CMIP5 model and the reanalysis over the climatology period, as well as the 2) average and 3) maximum lifetime of each node, and calculate the mean absolute error (MAE) for each of these three metrics. MAE is defined as follows,

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |f_i - y_i|$$
Eq. 2.1

where n denotes the number of SOM nodes under consideration (default = 12), y_i refers to the frequency (or lifetime) of node i in the reanalysis and f_i denotes the corresponding frequency (or lifetime) for the CMIP5 model of interest. An MAE of zero would indicate a perfect comparison in the frequency/lifetime of SOM nodes between a CMIP5 model and reanalysis. We have chosen to focus on mean absolute errors as the metric of choice for subsequent analysis, as this formula provides an absolute measure of model error and gives equal weight to errors from all SOM node types. While Pearson correlations and Spearman rank correlations have also been considered, results for subsequent applications in Section 2.4.1 were found to be less consistent across different SOM node configurations (see Figure S2.1-S2.3).

It is important to clarify that the absolute value of the MAE should be interpreted only in the context of each individual metric. For example, absolute MAE values for node frequency should not be compared with MAE values for average lifetime, as node frequency is measured as a percentage of days while lifetime metrics are measured in number of days. For example, a $MAE_{frequency}$ of 5 and a $MAE_{avg.\ life}$ of 0.2 does not mean a given CMIP5 model simulates the climatology of average node lifetimes 25 times better than node frequencies. Instead, this indicates the mean difference in average node lifetime between the reanalysis and the CMIP5 model is 0.2 days, while the mean difference in node frequency between the model and reanalysis is 5 percent.
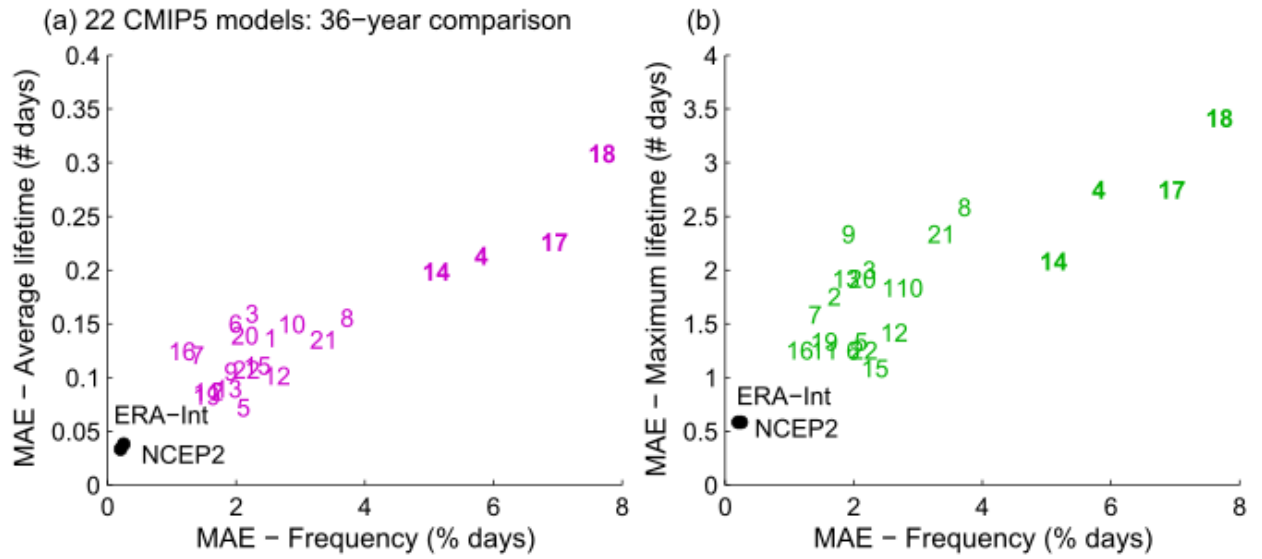
Figure 2.6: Spread in CMIP5 model performance based on the MAE scores, comparing the climatological node frequency with **(a)** average and **(b)** maximum persistence comparisons with 20CR data. A MAE of zero indicates a perfect match in the chosen metric between the model and reanalysis. Numbers refer to those CMIP5 models listed in Table S1. Black circles correspond to comparison between 20CR and two reanalysis datasets: ERA-Interim and NCEP2.

Figure 2.6 demonstrates this approach used to compare the 20CR reanalysis with each CMIP5 model, as well as two alternative reanalysis products (ERA-Interim, (Dee et al. 2011); NCEP2, (Kanamitsu et al. 2002)), across the full 36-year climatology. Based on this figure, it is apparent that some models do a better job than others at simulating climatological node characteristics of 20CR, though no models perform as well as the alternative reanalysis products. This latter point is expected, given that the CMIP5 simulations only have constrained boundary conditions (like external forcing rates) and are otherwise freely evolving, while the reanalysis datasets are, like 20CR, constrained using actual observations over the entire period. It is also noted that the combination of (1) very low and (2) highly similar MAE scores found for the two alternative reanalysis products shows the choice of reanalysis product has a negligible influence on the results of this SOM-based analysis.

An important step of any attribution study is to determine which models are a reasonable surrogate for reality and thus may be considered suitable for subsequent analysis. Based on the results of Figure 2.6, it might be argued that some models perform poorly relative to the model ensemble as a whole, and therefore may warrant exclusion from subsequent analysis of anthropogenic changes in circulation. However, the MAE scores of Figure 2.6 aggregate each model across the full climatology period and using all nodes, and thus may not necessarily translate to consistent model biases in the frequency distribution for each individual node. Indeed, Figure 2.7 compares, for each CMIP5 model, the distribution of

node frequencies of each summer across the 1979-2014 climatology period, for *each individual* SOM node. It is evident that no single model accurately emulates the 20CR frequency distribution for all individual SOM nodes; but more importantly, no one model performs equally poorly across all 12 SOM nodes. Because of this lack of sub-node consistency in model bias, we determine that no individual model can be justifiably excluded based on any of the individual validation criteria considered. For these reasons, and because of how we intend to utilise SOMs in an attribution context (Section 2.3.4), all models have been included for subsequent analysis.
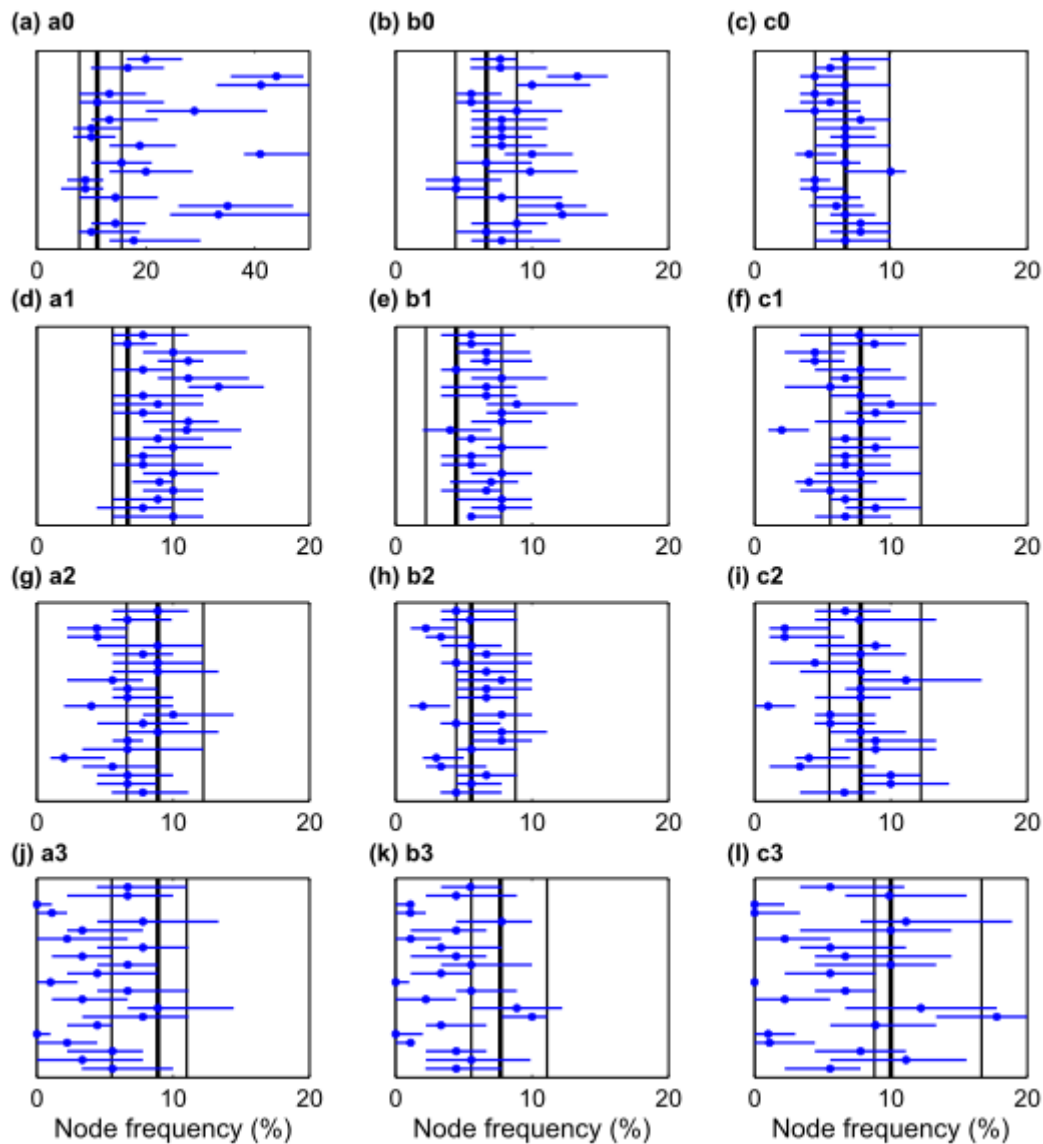
Figure 2.7: Comparison of the node frequency distributions between each CMIP5 member (blue lines) with 20CR for all summers over the 1979-2014 climatology period, and for each individual SOM node. Blue dots correspond to the median node frequency over the 36-year period; blue lines show the inter-quartile range; solid black line and thin black lines in each panel show the median and inter-quartile range of node frequencies for 20CR. Models 1 to 22 are presented in ascending order along the y-axis.

2.3.4 Application of SOMs in the context of event-specific attribution

Evaluating the fidelity of a model's synoptic climatology is only one application of self-organizing maps. Instead of aggregating all available years to compare climatological frequencies of each node type with the observed climatology of the same region, MAE scores can also be computed to compare each individual summer of a model with an observed summer in the real world. In addition to the MAE score obtained by comparing the summertime climatology of 20CR over 1979-2014 with that of model '12' (GFDL-ESM2M), one can also calculate the MAE between the node frequency/lifetime of *each* summer in the model, and the summer of 2013 in the reanalysis (filled circles shown in Figure 2.8). These individual MAE scores can be used to quantify the likelihood of a model simulating a summer with synoptic circulation features analogous to those of JFM13, and thus also be used to investigate *changes* in the probability of simulating summers like that of JFM13 in response to anthropogenic influences.
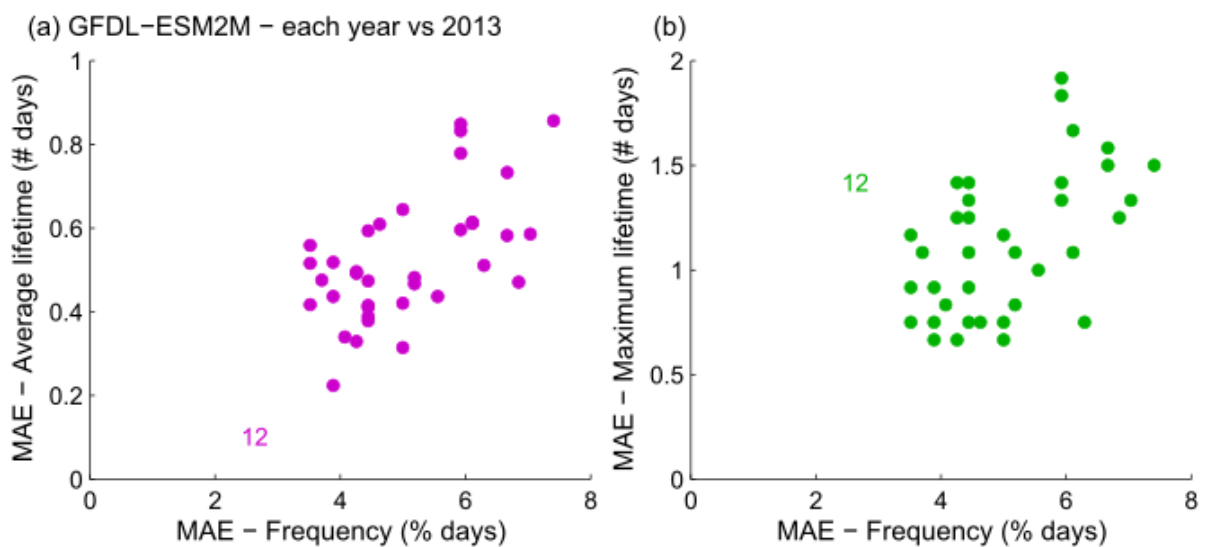
Figure 2.8: As in Figure 2.6, but showing the climatological correlation scores for a single model (number '12', denoting GFDL-ESM2M) decomposed into the correlation scores (filled circles) between each summer of the model over 1979-2014 and the summer of 2013 in the reanalysis.

To consider changes to the probability of the summer of 2013 occurring in response to anthropogenic influences on the climate system, we now consider the output from the same CMIP5 simulations (under the same 'Historical' and 'RCP8.5' experiments) for two time periods: 1861-1901 (hereafter 'early-climate') and 1993-2033 (hereafter 'recent-climate'). The more recent time period, centred on the summer of 2013, is treated as a representative analogue of the present-day climate, whilst the earlier period is intended to represent a

climate with a much smaller influence from anthropogenic factors. We assume that possible differences in natural forcings between the two periods will be negligible when smoothed over a 41-year timescale, and therefore consider differences between the two distributions as attributable to human influences on the climate system. The combination of using a 41-year aggregation period, and considering many different CMIP5 models, each of which are freely evolving, enables a further assumption that possible modes of decadal variability (Henley et al. 2015) will have negligible influence on these results. Finally, it is important to note that the relative influences of ozone depletion versus greenhouse gas increases cannot be easily distinguished using the model simulations available, and we therefore choose to consider 'anthropogenic' changes to the climate as the aggregate signal of both effects.

For each model summer (JFM), we calculate the frequency, average lifetime and maximum lifetime of each SOM node, and compare with the corresponding characteristics for the observed summer in 2013. For each summer in each model, MAE scores are then computed for each of the three metrics, by comparing the value for each of the 12 nodes with the corresponding observed value in the summer of 2013. By calculating MAE scores to compare each modelled summer with JFM13 for the 22-model subset, we obtain a 41x22 element probability distribution (PDF) of MAE scores for each of the two periods considered.

To test the statistical significance of changes to each probability distribution between the two periods, we apply a two-sample Kolmogorov-Smirnov (K-S) test, and where appropriate, also calculate the estimated FAR for a given threshold as FAR = $(p_1-p_0)/p_0$, where p_0 is the probability of exceeding a prescribed threshold in the early-period ensemble and p_1 denotes the probability of exceedance in the recent-period ensemble (Allen 2003; Stott et al. 2004). We also repeat all calculations of FAR by resampling an equal number of data points from the distribution, with replacement, 10000 times, and present the best-guess (50th percentile) estimate, as well as the 10th and 90th percentile of these bootstrapped distributions.

2.4 Results

2.4.1 Changes to the likelihood of the JFM13 drought from a daily circulation perspective

Figure 2.9 shows the aggregated PDFs of model MAE scores for the 'early-climate' and 'recent-climate' ensembles, for each of the three SOM metrics, presented using a kernel

smoothing function. The distribution of frequency MAE scores appears to be near-Gaussian, while the distributions of average lifetime and maximum lifetime MAE exhibit heavier tails.
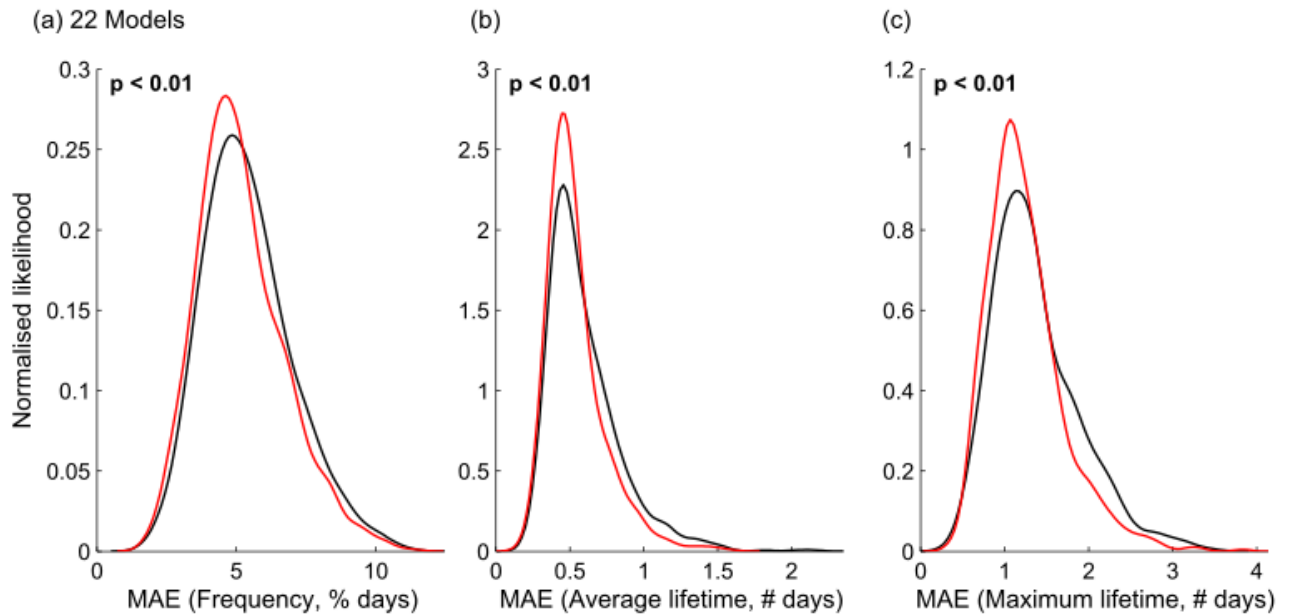
Figure 2.9: Probability distributions of the **(a)** frequency; **(b)** average lifetime and **(c)** maximum lifetime correlation scores for the 22-model ensemble. PDFs are estimated using a kernel density smoother. Black lines correspond to 1861-1901 ensemble; red lines correspond to 1993-2033 ensemble. Inset p-values are found using a two-sample K-S test.

The key question of this SOM-based approach to event-specific attribution is how these probability distributions have changed when comparing the 'early-climate' model ensemble with the 'recent-climate' ensemble centred on 2013. We find statistically significant ($p<0.01$) shifts towards lower MAE scores for all three of the SOM metrics when comparing the recent-climate PDF (red) with the early-climate PDF (black). Similar results are also found when considering different SOM node configurations (see Appendix). This is a significant result, and demonstrates that an anthropogenic signal can be found in the daily-scale properties of circulation related to a drought event. Specifically, these results suggest the sequence and properties of circulation patterns coincident with the 2013 New Zealand drought were more likely to occur in response to anthropogenic influences on the climate system.

While the frequency MAE scores show a systematic shift in the entire distribution, the lifetime metrics instead show a smaller distribution tail and a more pronounced statistical mode. Understanding how these distribution changes can be interpreted in the context of a more traditional FAR-based framework is considered in more detail in Section 2.5.

2.4.2 Comparison with alternative attribution techniques

For both completeness, and to understand how these SOM-based results compare with more orthodox approaches to event attribution, here we present alternative assessments of human influence on the 2013 New Zealand drought, using methods that have been previously used in the peer-reviewed literature. Specifically, we consider the frequency of only specific SOM node types, under the assumption that the notable prevalence of some node types during the drought event may suggest they are more conducive to drying than others. We also compare the daily-scale circulation changes with changes in the distribution of an area-averaged, seasonal circulation metric and precipitation deficit metric.

2.4.2.1 Changes to the frequency of anomalous SOM nodes

Some previous studies have investigated anthropogenic influences on changing circulation patterns by counting the number of occurrences of a particular circulation regime type, especially if this circulation pattern has been demonstrated as being anomalously prevalent during an extreme event of interest. For example, Schaller et al. (2016) considers the number of days in the UK wet winter of 2013/14 when a south-westerly flow regime occurred, as record-breaking occurrences of this regime type coincided with record-breaking flooding over the region. Here we present an analogous approach for the 2013 drought, by counting the number of days which exhibited either SOM node 'c2' or 'c3'. As mentioned in Section 2.3.2, the combined frequency of these two nodes was the highest over the 36-year reanalysis period in the summer of 2013.

Figure 2.10 shows a systematic shift in the recent-climate distribution to summers with more frequent occurrences of these node patterns, which were anomalously frequent in the summer of 2013 and represent anticyclonic MSLP patterns situated over the North Island (Figure 2.4). Corresponding FAR estimates for a series of arbitrary thresholds (chosen between the 80th and 98th percentile of the 'early-climate' distribution) show a bootstrapped 10th percentile of FAR of 0.3 for moderately extreme thresholds (>25-30% of days per summer); however no discernible signal is apparent at the very tail of the distribution.
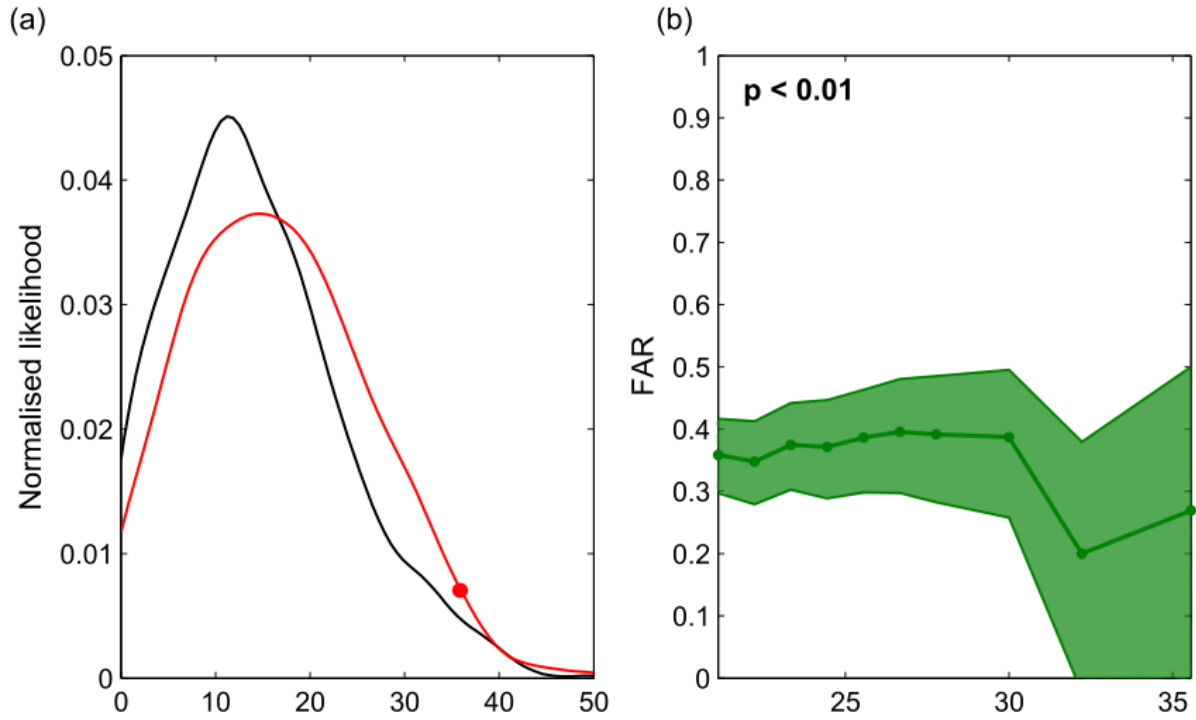
Figure 2.10: (a) PDFs of the percentage of days in each summer which correspond to either SOM node 'c2 or 'c3'. Black lines correspond to the 1861-1901 model ensemble; red lines correspond to the 1993-2033 ensemble. The red circle corresponds to the 2013 value of 35.6%. **(b)** FAR estimates for thresholds corresponding to between the 80th and 98th percentile of the 1861-1901 distribution, sampled at 2-percentile intervals. The green shading indicates an 80% confidence interval based on a 10000-sample bootstrap. Inset p-value corresponds to a two-sample KS test between the early and late ensembles: bold font indicates $p<0.05$.

2.4.2.2 Comparison with 3-month area-average approaches

For the two 41-year periods using the same 22-model ensemble, we extract model MSLP taken as the January-March mean value averaged over the region which exhibited a maximum anomaly for the 2013 drought event (green box in Figure 2.2). In addition, we calculate the total precipitation accumulation (mm) over the same 3-month period, averaged over the region which exhibited the most severe low-precipitation anomalies during the summer of 2013 (blue box in Figure 2.2). While it is noted that dry day frequencies have also been considered as a relevant metric to evaluate the 2013 New Zealand drought (Harrington et al. 2014), we choose to focus on 3-monthly cumulative rainfall totals, as this is the most common approach in the peer-reviewed literature for assessing precipitation changes in a drought attribution context (Lott et al. 2013; Barlow and Hoell 2015; Bergaoui et al. 2015; Diffenbaugh et al. 2015; Funk et al. 2015; Marthews et al. 2015; Otto et al. 2015a; Rupp et al. 2015; Swain 2015)

Figure 2.11a reveals a weak negative correlation between the 3-monthly total precipitation of a model summer and the corresponding MSLP averaged over the same period, suggesting

that higher summertime MSLP is more likely than not to result in lower-than-normal precipitation over the North Island. There is also clear evidence for a systematic shift towards higher summertime MSLP in the more recent time period, consistent with our present understanding of changes to MSLP over the New Zealand region in response to human-induced climate change (Gibson et al. 2016a). An increase in variance of 3-month precipitation totals also translates to a small subsequent increase in the likelihood of exceptionally low precipitation totals. K-S tests show statistically significant differences ($p < 0.05$) exist between the early- and recent-climate PDFs for the area-average MSLP but not for precipitation.

Interrogating these results further, FAR estimates are considered for a series of arbitrary thresholds at the tail of the early-period distribution for each variable. Figure 2.11b confirms a statistically robust increase in the likelihood of exceptionally high 3-monthly MSLP is found in the 'recent-climate' model ensemble, with the 10th percentile of FAR estimates exceeding 0.3 for all MSLP thresholds considered, except at the very tail of the distribution. Meanwhile, Figure 2.11c shows a small but detectable anthropogenic increase in low precipitation totals across all moderately extreme percentiles, with the bootstrapped ensemble of FARs showing a more distinct anthropogenic increase in the likelihood of 3-month totals below 450 mm. These results are consistent with similar analyses which found an increased likelihood of summers with exceptionally high dry day frequencies over the North Island due to anthropogenic influences (Harrington et al. 2014).
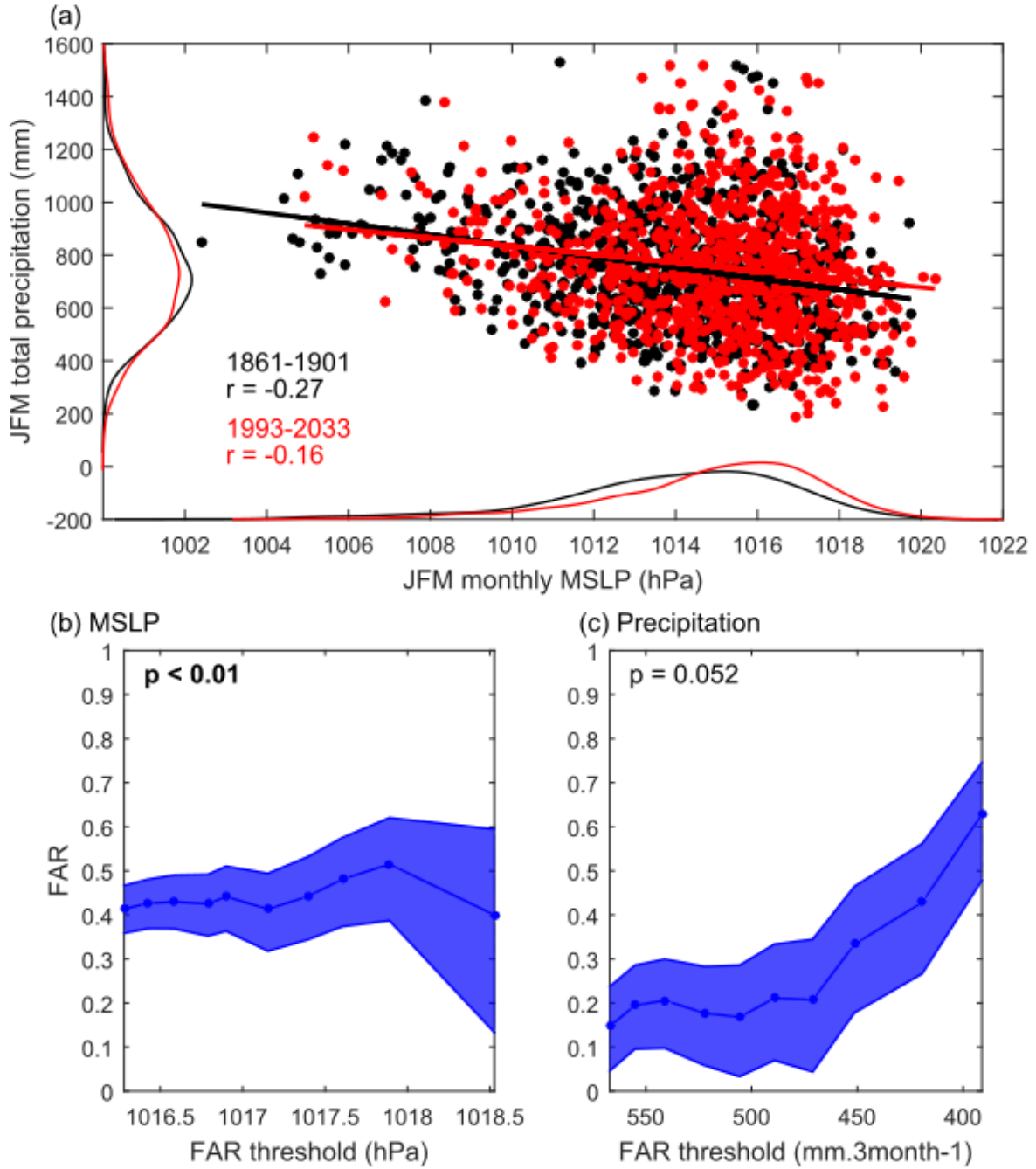
Figure 2.11: (a) Relationship between 3-monthly MSLP and 3-monthly total precipitation averaged over the regions shown in Figure 1. Black and red circles show each model summer in the early (1861-1901) and recent (1993-2033) time periods respectively. Least-squares linear regressions have been applied to each ensemble; r-values correspond to the Pearson's correlation score. Bottom panels show the best-guess (median) fraction of attributable risk (FAR) calculated for **(b)** MSLP and **(c)** precipitation. FAR thresholds were considered between the 80th and 98th percentile of the 1861-1901 distribution, sampled at 2-percentile intervals for MSLP (circles), while thresholds pertaining to the 20th to the 2nd percentile were sampled for precipitation accumulations. The blue shading indicates an 80% confidence interval based on a 10000-sample bootstrap. Inset are p-values of a two-sample KS test between the early and late ensembles of each variable: bold colour indicates p<0.05.

2.6 Discussion

By considering probability distributions of MAE scores, the traditional determination of FAR estimates (by counting the exceedances of a metric threshold based on the observed event) is not possible, since by definition, the observed event corresponds to the lower bound of the probability distribution (an MAE score of zero). Further, it is not immediately clear how best to compare human-induced changes to the distribution tails of the three MAE-based metrics with the area-averaged distributions of seasonal precipitation accumulations and MSLP, as well as the 'c2+c3' node counts. To circumvent these issues, we select values corresponding to every 2nd percentile between the 80th and 98th percentiles of the 'early-period' distributions for each metric, thereby enabling a comparable sequence of thresholds to consider FAR estimates. It is noted that we actually consider quantile thresholds ranging from the 20th down to the 2nd percentile for the case of 3-monthly precipitation totals and the MAE scores, but label them to be the same as the other metrics. The FAR statements shown in Figure 2.12 thus demonstrate attributable probability increases in the tails of each distribution, rather than providing a single FAR estimate for the 2013 event.
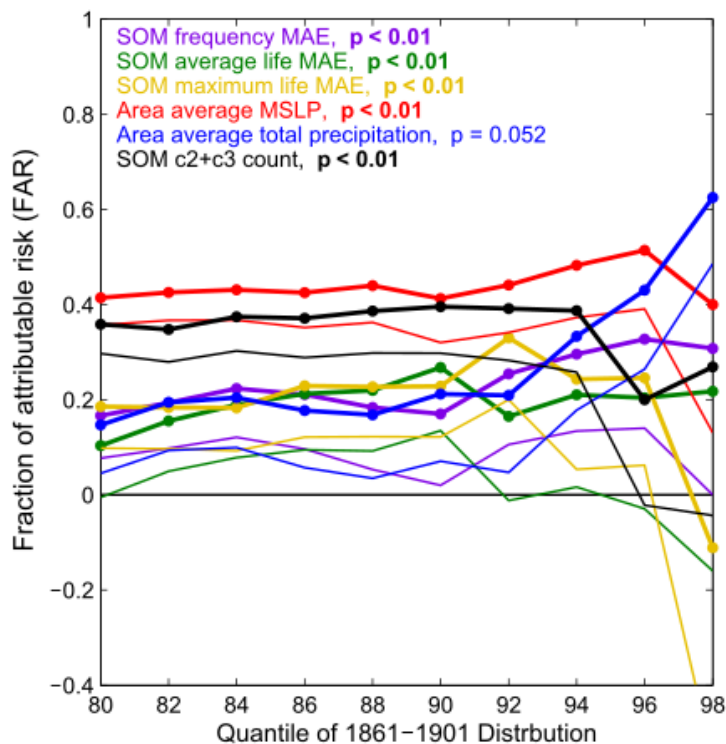
Figure 2.12: FAR estimates calculated at percentile thresholds of the early (1861-1901) time period comparable between the six metrics used to consider the 2013 New Zealand drought. The quantile thresholds actually correspond to the 20th to the 2nd percentile for precipitation totals and the MAE scores. Bold lines with filled

It is evident that the best-guess estimates of FAR for comparable thresholds along the tail of the distribution for each of the six metrics differ, but all are above zero – this is consistent with the fact that five out of six metrics show statistically significant differences between the early-climate and recent-climate PDFs when based on K-S tests alone (inset p-values). The bootstrapped 10th percentile estimate of FAR for the area-averaged MSLP index (thin red line) remains consistently above 0.33, suggesting at least a 50% increase in the likelihood of exceptionally high seasonal mean sea level pressures over the New Zealand domain between the early- and recent-climate model runs. Similar FAR values are also found for the node count-based metric (solid black lines), which is expected given the two metrics represent related phenomena.

It is important to emphasise that these estimates of FAR for each metric consider fundamentally different characteristics of the 2013 drought, and the variety of estimates found in Figure 2.12 should therefore be expected. Moreover, metrics which consider variables over seasonal timescales, like that of the area-averaged pressure and precipitation, have inherently lower variance when compared with daily-scale distributions and are thus more likely to experience the earlier emergence of a FAR signal (Christiansen 2015; King et al. 2016).

The estimates of FAR for thresholds between the 80th and 90th percentile are consistent for all six metrics – of most interest is that all three MAE-based metrics demonstrate a similar best-guess FAR of approximately 0.2. However, as the FAR thresholds considered extend beyond the 90th percentile, this homogeneity breaks down, and the metrics with formerly robust signals show a lower-bounded FAR which overlaps with zero (suggesting the possibility of no change in likelihood) while the precipitation PDF begins to exhibit a robust increase in the likelihood of exceptionally dry summers. This divergence is indicative of the relatively small sample sizes associated with attribution using a CMIP5 model ensemble - it is therefore difficult to make robust statements about the fraction of attributable risk for very extreme thresholds in the absence of a much larger model ensemble.

There are some further caveats that require consideration when interpreting the results presented in this study:

1) When several metrics on different timescales are being concurrently assessed, there are open questions as to the most appropriate methods of model validation for attribution analysis (Mitchell et al. 2015; Mitchell 2016). In this study, we consider the climatological SOM statistics to determine which models replicate the real-world frequency and persistence of summertime circulation patterns, but choose to include all models for subsequent attribution, as determining which models 'performed poorly' depended on the validation method used. It is also important to note that the number of models which reliably simulate the observed distributions of seasonal precipitation and MSLP over the regions used for the area-average metrics may differ from those which meet a validation criteria based on daily-scale circulation characteristics (Masson and Knutti 2011), and this must be considered when comparing subsequent FAR estimates.

2) As demonstrated in Section 2.4.2.1, there is a robust anthropogenic increase in the likelihood of observing those SOM nodes which occurred frequently during the 2013 drought (like nodes 'c2' and 'c3'). However, it is also important to recognise that, due to the complex topography of New Zealand, each of the circulation regimes characterised by the twelve SOM nodes in this analysis can enhance drying in some parts of the country, while potentially increasing rainfall in other regions: this is illustrated by the composite precipitation anomalies for each SOM node shown in Figure 2.13a.
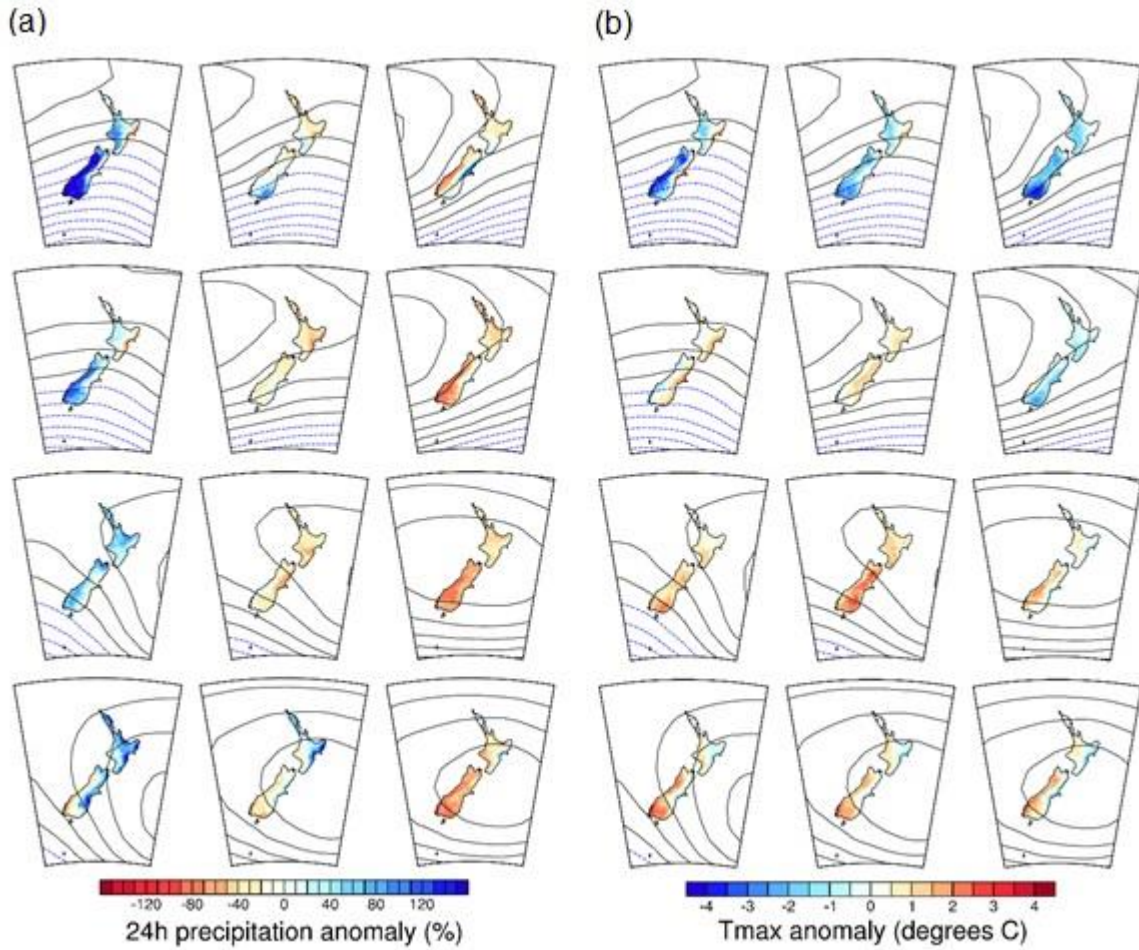
Figure 2.13: Composite maps of daily **(a)** total precipitation and **(b)** daily maximum temperature for all JFM observations over 1979-2014. Each individual map represents a different SOM node, following the same configuration as for Figure 2 of the main manuscript – for example, node a0 in top-left, and c3 in bottom-right. To obtain the composite anomalies for each node, the relevant VCSN observations (Tait et al. 2006, doi: 10.1002/joc.1350) of each day during January-March, when that node was observed, were subtracted from the mean climatology of that month over 1979-2014. The anomalies for all days with the same node were then averaged together.

In addition, the spatial patterns of drought for the 2013 event were not completely homogeneous, with record-breaking dryness observed over the western South Island and the majority of the North Island, while the more traditionally drought-prone eastern regions of the country experienced only moderate soil moisture deficits (Figure 2.1). Because of these complexities in the spatial evolution of the drought, assuming that one or two particular SOM nodes were the cause of the events' severity, as is implied in Section 2.4.2.1, could lead to an overly simplistic characterisation of the unique properties of the 2013 event. It is for this reason that we emphasise the MAE-based results of this analysis, as this technique assigns equal weight to the difference in node frequency (or persistence) when comparing model summers against JFM13, and thus enables a truly event-specific method of attribution.

3) The Kolmogorov-Smirnov test is a common approach used to establish whether two probability distributions can be interpreted as being statistically different from one another, or more accurately, whether they originate from the same continuous distribution (Steinskog et al. 2007). However, it is evident that changes between the 'early-climate' and 'recent-climate' distributions to the number of years with very high MAE scores will be less informative than evaluating changes to the very low end of the MAE distribution, especially when asking the question of whether a summer like that of 2013 was more likely to occur in response to anthropogenic climate change. Christidis and Stott (2015), for example, only subsampled model winters which obtained a Pearson correlation score greater than 0.6, when comparing the upper-level circulation characteristics with the observed wet UK winter of 2013/14. The SOM-based FAR estimates in Figure 2.12 are intended to be analogous to this fixed threshold approach, and thus provide supplementary insight to the full PDFs considered in Figure 2.9.

2.7 Summary

Any statement regarding the fraction of attributable risk for a specific extreme event needs to be carefully interpreted in the context of the question being asked (van Oldenborgh et al. 2015). In this study, we introduce a new approach to consider whether and to what extent the specific daily circulation characteristics observed over the New Zealand region during the record-breaking drought of early 2013 were more or less likely to occur, and break that down further to separately consider the frequency, average persistence and maximum persistence of different circulation types. The introduction of SOMs in an attribution context represents a novel contribution towards an emerging focus on event-specific attribution techniques in the research community (Hannart et al. 2016). It also represents a method of circumventing difficulties in evaluating anthropogenic changes in drought likelihood, particularly for those locations where precipitation-temperature coupling mechanisms are less significant (like New Zealand).

This SOM-based approach fills a gap in the toolset available for event-specific attribution. While different techniques remain more suited for extremes which occur on daily or multi-day timescales, the approach suggested in this study will be useful for seasonal-to-annual scale extremes which cannot be sufficiently characterised by temporally averaging a single variable (such as MSLP) over the length of the event. Whilst the analogue approach (Stott et al. 2016) represents a complementary method, applying such a technique to studies of

drought still requires the original assumption that a single, specific measure of drought (such as soil moisture) could be adequately simulated by climate models over time, and also be representative of the underlying mechanisms which contribute to event severity. Our approach instead focuses only on changes to the properties of daily-scale circulation over the duration of the extreme event.

The final aim of this study was to illustrate how different approaches towards event attribution can yield a variety of results even when considering the same event. For the New Zealand drought case study, our analysis indicates there was a small but discernible anthropogenic increase in the probability of observing daily circulation characteristics like those of the summer of 2013, with best-guess FAR estimates of 0.2 being found for all three SOM-based metrics when using moderate percentile thresholds. There were also robust increases in the likelihood of exceptionally high seasonal MSLP patterns over the regions which were most extreme during the 2013 drought, with a best-guess estimate of a 50% increase in likelihood. We emphasise that these results are sensitive to the threshold at which FAR was calculated, and discernible changes to likelihood in the extreme tail of the distribution could not be reliably identified using the model ensemble available.

Future work is needed to explore the potential utility of self-organizing maps in the context of event attribution, as well as to determine how an event-specific framework could best complement other common attribution techniques. In Chapter 3, I further explore the sensitivity of attribution statements to the analytical framework employed, with a particular emphasis on differences between event-specific and event-as-class frameworks.

Chapter 3

Investigating differences between event-as-class and probability density-based attribution statements with emerging climate change

The work described in this chapter follows Harrington (2017), published in *Climatic Change*.

As discussed extensively in Chapter 2, there is significant public and scientific interest in understanding whether and to what extent the severity and frequency of extreme events have increased in response to human influences on the climate system. As the science underpinning the field of event attribution continues to rapidly develop, there are growing expectations of faster and more accurate attribution statements to be delivered, even in the days to weeks after an extreme event occurs. As the research community looks to respond, a variety of approaches have been suggested, each with varying levels of conditioning to the observed state of the climate when the event of interest has occurred – each of these frameworks will thus produce slightly different answers. One such approach to utilise unconditioned multi-model ensembles requires pre-computing estimates of the change in probability of occurrence for a wide range of possible 'events'. Given the longer-term ambition of operationalising event attribution, the original assumptions regarding how to quantify the human contribution to changing event likelihoods need to be re-evaluated accordingly.

In this study, I employ simplified examples to demonstrate issues relating to the attribution of human influence on a wider class of event thresholds, not just record-breaking extremes. Specifically, I consider differences between event-as-class attribution statements with changes in the probability density of the distribution at the event threshold of interest. For the majority of extreme event attribution studies, it is likely that the two metrics are comparable once uncertainty estimates are considered. However, results show these two metrics can produce divergent answers from each other for moderate climatological anomalies if the present-day climate distribution experiences a substantial change in the underlying signal-to-noise ratio. As the emergent signals of climate change becomes increasingly clear, this study highlights the need for clear and explicit framing in the context of applying pre-computed

attribution statements, particularly if attribution perspectives are to be included within the framework of future climate services.

3.1 Introduction

Over the last decade, efforts to understand whether and to what extent extreme weather events have become more likely in response to human influence have been explored under the new research field of 'extreme event attribution'. The primary focus of this scientific sub-discipline concerns specific extreme weather events, and commonly those which have been unprecedented in severity over the location of interest. The original proposal of the fraction of attributable risk (hereafter FAR) by M.R. Allen (Allen 2003) was suggested as a means of quantifying the potential liability for climate change damages following epidemiological practice: this involved calculating the relative increase in the probability of observing a given extreme event by comparing the likelihood of occurrence in simulations of the present-day climate with some modelled estimate of a counterfactual climate which has experienced no anthropogenic influences (Stone and Allen 2005).

In the decade since the first successful application of event attribution techniques by Stott and colleagues (2004), and following subsequent developments in both process understanding and model capability (NAS 2016; Stott et al. 2016), the number of studies concerning the attribution of human influence for specific extreme weather and climate events have increased dramatically. This is most evident in the annual special issue of the *Bulletin of the American Meteorological Society* (Peterson et al. 2013; Herring et al. 2014, 2015) which focuses on attribution studies of events having occurred in the previous calendar year. Despite only four iterations being published, there has been a clear expansion in the range of event types considered, methodological innovations and in the geographic spread of the events being assessed (Otto 2016; Stott 2016).

As the need for more frequent and timely attribution assessments continues to rise, a variety of approaches have been suggested to address this demand. Here, we briefly review several different approaches, before highlighting some fundamental considerations when interpreting statements concerning the role of anthropogenic influences on the likelihood of a given event. For the purposes of this study, we limit our focus to only model-based attribution approaches - several complementary statistical techniques are also available for locations with high-quality observational records (e.g. Vautard et al. 2015; van Oldenborgh et al. 2015).

3.1.1 Differing approaches towards attribution

Nearly all model-based approaches to probabilistic event attribution consider the likelihood of some climatic 'event' occurring in a simulation of the present-day climate, when compared with a simulated climate where either some or all anthropogenic influences have been removed. Understanding what specific differences exist between these factual and counterfactual scenarios – and particularly the level of conditioning on the observed state on the climate system - is paramount for interpreting any subsequent statement about the change in likelihood of observing the event of interest (NAS 2016).

If the boundary and/or initial conditions of model simulations for both distributions are highly constrained on observed features of the event, this enables an in-depth look at the role of human influence on specific physical mechanisms which may have exacerbated the severity of an extreme event. For example, Meredith and colleagues (2015) considered only the human-induced increases in sea surface temperatures and otherwise utilised the same observational constraints as input to a high-resolution Weather and Forecasting Research (WRF) model: this enabled a discernible anthropogenic signal to be detected for an extreme convective precipitation event, something which would not have been possible with a coarse-resolution global climate model. By contrast, Stott and colleagues (2004) focussed on the overall anthropogenic influence on the likelihood of the 2003 European heatwave occurring by considering a large spatial domain and using only freely-evolving model simulations. Among other reasons, this approach was most pragmatic because (1) these types of model simulations were shown to be capable of simulating the relevant physical mechanisms and (2) using a high-resolution model over such a large spatial domain would have been computationally cost-prohibitive.

3.1.2 Unconditioned attribution using coupled climate models

Every technique considered in the field of event attribution has different advantages and disadvantages, particularly for the implementation in a near real-time context. For example, when analyses utilise many simulations from only a single climate model, there needs to be high confidence that the model being used is capable of simulating the relevant physical mechanisms which contribute to the onset of a particular extreme event (Bellprat and Doblas-Reyes 2016). To circumvent this reliance on the quality of a single model, one alternative approach is to utilise pre-existing simulations from many different coupled climate models,

such as those which contributed to the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor et al. 2012).

In any model-based approach, care is needed to ensure the relevant statistics of the simulated climate can be considered a reasonable surrogate for the real world, particularly for a region over which an event of interest occurs. In the methodology proposed by Christidis and colleagues (2015, hereafter C15), well-known optimal fingerprinting techniques (Hegerl and Zwiers 2011) are used to quantify observationally-constrained estimates of the signal of anthropogenic climate change for individual models. Measures of variability in historical observations over a given region are then compared against pre-industrial control model simulations: for those models which adequately simulate the variance-covariance structures of the observational record, a re-sampled estimate of internal variability is then combined with the scaling factors to produce factual and counterfactual probability distributions for each model.

A key benefit of utilising a fully unconditioned approach to rapid event attribution is that no observations are required, beyond identifying the event anomaly of interest. This means, for a given region and given type of event (seasonal heat or heatwaves or extreme rainfall), prior analysis can be performed to identify which models adequately simulate the observed climatology of the real world, by comparing against historical observations (King et al. 2015b). For those models which have been identified as adequate surrogates for the real world, well-understood optimal fingerprinting techniques can then be used, as proposed in the method by C15, to compute factual and counterfactual distributions which include and exclude the role of anthropogenic forcings respectively. Once these simulations have been performed, one could immediately calculate the relative change in likelihood of a real-world anomaly having occurred soon after the fact. Similar techniques have also been demonstrated by King and colleagues (2016), but rather than considering individual CMIP5 models separately, all models which adequately simulate the relevant climatology for locations with high-quality observational records are aggregated together into a single multi-model distribution.

3.2 Quantifying attribution statements for climate services: questions about a dramatically warmed distribution

The introductory remarks of Section 3.1 were framed in the context of 'orthodox' event attribution studies, which commonly focus on those climatic events which result in severe societal impacts. However, several recent studies have also emphasised an increasing demand for routine statistics from weather (and climate) service providers to be presented in the context of a changing climate (Hewitt et al. 2012; Gregow et al. 2015; Brasseur and Gallardo 2016; Goddard 2016). For example, most climate or weather service providers for a given country (or city) present what the average temperature was for the just-completed month (or season, or year), as well as how that temperature compares to the climatological average. Alongside the communication of these statistics in absolute terms and with respect to the historical climatology of the region, there could be an additional question of *how much more/less likely was that particular monthly/seasonal/annual temperature anomaly to be witnessed today, as a result of anthropogenic climate change?* This question could be successfully answered with an unconditioned modelling framework using the methods outlined in Section 3.1.2, and would help to compliment other proposed approaches, including those fully conditioned with observational constraints (Hannart et al. 2016).

As a conceptual example, we suppose a local stakeholder is interested in summer mean temperatures over a pre-specified spatial domain with a high-quality observational network, such as Central England for example (King et al. 2015b). The relevant model validation and optimal fingerprinting techniques have been performed, such that the pre-computed factual and counterfactual distributions are already available in an unconditioned multi-model framework. For illustrative purposes, assume that both distributions are Gaussian probability density functions (PDFs, Figure 3.1). For simplicity, it is assumed that no changes in variance have occurred in these factual distributions, and distributional shifts can be considered in units of standard deviations. Now relative to the counterfactual distribution (blue), suppose the factual distribution has warmed by half a standard deviation in the present day (magenta). We also consider an equivalent distribution which represents the projected climatology of 2080 under a high emissions scenario (red); this second distribution has warmed by three standard deviations relative to the counterfactual scenario.

Now suppose the preceding boreal summer has just ended, and the seasonal-mean temperature over this region of interest exhibits a $+1\sigma$ anomaly relative to the counterfactual

distribution (black line in Figure 1). Any traditional approaches to assess whether the observed anomaly of this just-completed summer was more or less likely to occur involve comparing the cumulative density of these counterfactual and factual distributions, with the observed anomaly as the lower bound, and infinity as the upper bound. When using this approach for Figure 3.1, it is clear that the likelihood of observing an anomaly equal to or greater than the event anomaly has increased for both factual distributions, and substantially so for the '2080' scenario.

However, an equally valid question to ask is whether or not an anomaly of that *specific* magnitude was more or less likely to occur: a pragmatic approach may therefore be to consider the ratio of the probability densities of each distribution at the specified event threshold. The coloured circles in Figure 3.1 reveal that while there was an increased likelihood of observing that specific anomaly in the 'present-day' PDF when compared with the 'pre-industrial' simulations, the probability density at this specific threshold is in fact much lower for the '2080-scenario' PDF, suggesting that this specific $+1\sigma$ anomaly is actually less likely to occur in the future. This is because the second factual distribution has warmed so dramatically, the mode of the Gaussian distribution has shifted well beyond the anomaly threshold, hence actually resulting in a comparatively lower chance of occurrence.

The probability ratio, or equivalent fraction of attributable risk, are the most useful and widely applicable metrics to quantify the influence of an event class occurring, especially for the distribution tails. However, it is also important to understand how the probability densities at the event threshold might compare between the factual and counterfactual distributions. The motivation of this study is thus to understand how these two types of metric compare, when systematically considering a range of different event thresholds and signal-to-noise ratios.
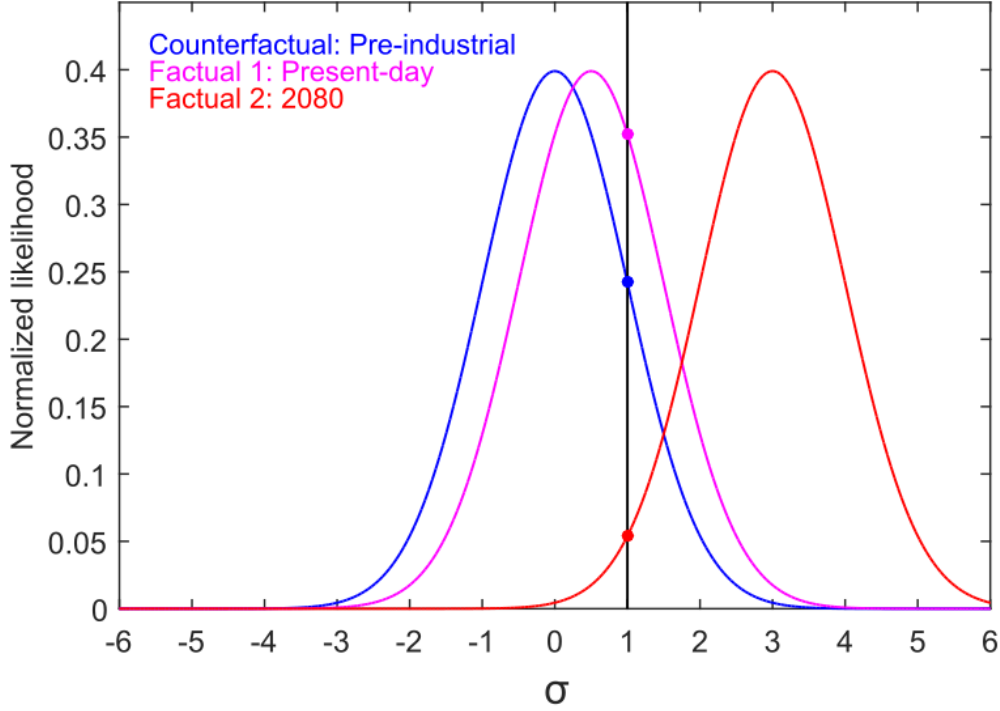
Figure 3.1: Schematic illustrating the differences between the changing likelihood of exceeding an event threshold (black line) versus observing that specific event threshold with warming. The probability density functions represent, respectively, area- and seasonal-averaged temperatures under a counterfactual scenario (blue); a factual scenario with a small (+0.5σ) distributional shift (magenta); and a factual scenario with a more dramatic (+3σ) shift in the distribution mean (red).

3.3 Systematic exploration of differences in attribution statements

To develop a more formal framework for analysis, Figure 3.2 presents two Extreme Value Type I (hereafter Gumbel) distributions, f_0 and f_1, according to the formula,

$$f_n(x) = \frac{1}{\sigma_n} \exp\left(\frac{(x - \mu_n)}{\sigma_n}\right) \exp\left(-\exp\left(\frac{(x - \mu_n)}{\sigma_n}\right)\right) \qquad \text{Eq. 3.1}$$

where μ and σ respectively denote the location and scale parameter of the Gumbel distribution. These PDFs could, for example, represent the maximum daily temperatures of each summer season, averaged over Central England. For simplicity in this demonstration, f_0 has a fixed location parameter of $\mu_0=0$, while the scale parameter of both distributions is set to 1. Next, we define the signal-to-noise ratio of f_n as,

$$S_n = \frac{\mu_n}{\sigma_n} \qquad \text{Eq. 3.2}$$

A variety of possible values are considered for the value of μ_1. Since we have made $\sigma_0 = \sigma_1 = 1$, we consider S hereafter in general units of σ. We can therefore now consider f_1 as a model

ensemble of simulations representing the present-day climate with an anthropogenic signal, S, while f_0 represents an ensemble of counterfactual model simulations where anthropogenic influences have been removed (since we have defined $S_0 = 0$).
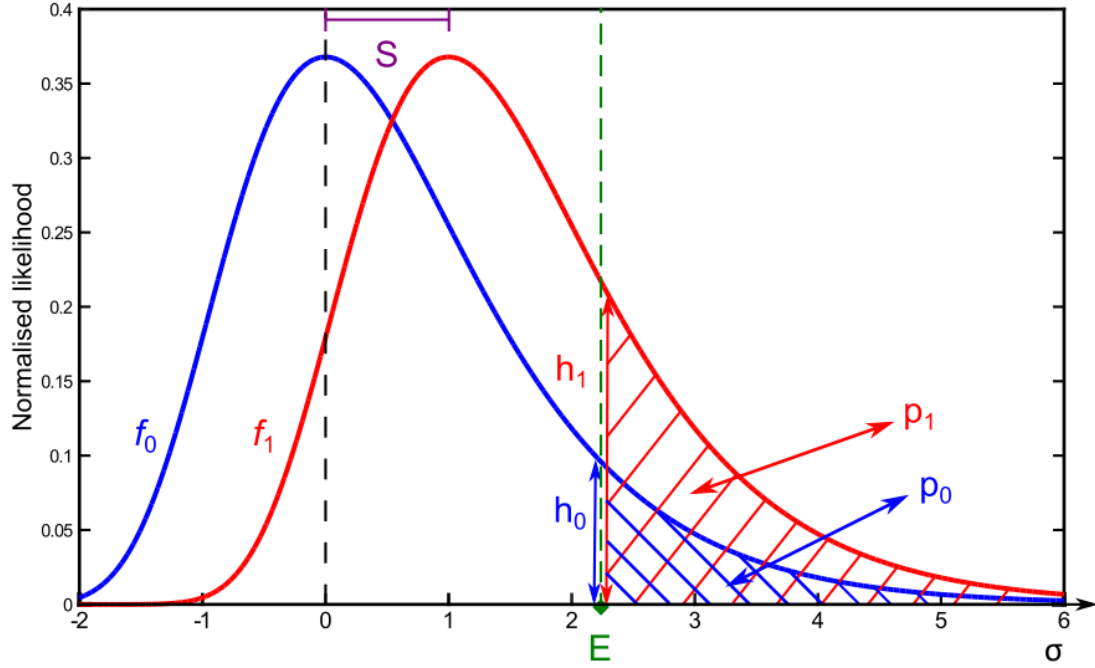
Figure 3.2: Schematic presenting the idealised counterfactual (blue) and factual (red) Gumbel distributions considered for the main analysis, labelled as f_0 and f_1 respectively. Also presented are the key variables introduced in the main text.

Suppose a given event, E, occurs in the real world – in the example of seasonal maximum temperatures, this could be the spatially-averaged, observed maximum temperatures for the season just completed. To understand the change in probability of seeing an event *like E* as a result of anthropogenic changes to the climate system, one common approach is to assess the probability ratio, which we define as the event-as-class PR (hereafter PR_{EAC}), according to the formula,

$$PR_{EAC}(E) = \frac{p_1}{p_0}$$

Eq. 3.3

$$p_n = \int_{x=E}^{\infty} f_n(x)$$

Eq. 3.4

where p_0 and p_1 represent the likelihood of experiencing a seasonal temperature equal to *or greater than* the threshold E in the counterfactual and factual model ensembles, respectively. However, to understand how the odds of observing a summer *exactly like that of E* have

changed due to human influences, we also define the probability density-based probability ratio (hereafter PR$_{PD}$) as

$$PR_{PD}(E) = \frac{h_1}{h_0}$$ Eq. 3.5

$$h_n = f_n(E)$$ Eq. 3.6

This time h_0 and h_1 denote the respective probability densities at the specific threshold E in these idealised counterfactual and factual probability distributions.

Next, we consider pre-defined probability distributions to demonstrate how the values of PR$_{EAC}$ and PR$_{PD}$ compare for a range of event thresholds, and specifically consider how this relates to the corresponding magnitude of the anthropogenic signal-to-noise ratio found in the 'factual' distribution. We then utilise the uncertainty associated with having only a finite-sized model ensemble available, to quantify when the value of these two metrics diverge significantly from one another.

3.3.1 Methods

To consider how differences between the two PR metrics vary as a function of the underlying anthropogenic signal, the event threshold and the size of the ensemble used, the following methodology is proposed:

(1) For a prescribed anthropogenic signal, S, we define a new factual distribution, f_1. The counterfactual distribution, f_0, remains unchanged throughout.

(2) We then consider an event, E, which occurs between the 50th and 99.99th percentile of the counterfactual distribution, f_0. Because we have explicitly defined the two PDFs, the value of PR$_{PD}$ for the specified threshold is then calculated.

(3) By artificially imposing the restrictions of having only a model ensemble of finite size through the use of a bootstrapping procedure, a range of possible PR$_{EAC}$ estimates is calculated for the same event threshold. To do this, we sample N times, with replacement, from each of the two factual and counterfactual PDFs. For each of these distribution subsets, an estimate of PR$_{EAC}$ can be calculated. We then repeat this process 1000 times and extract the median estimate of PR$_{EAC}$ for the specified threshold E, as well as a 5-95% uncertainty range. This is repeated for a variety of different sample sizes, N.

(4) Steps 2 and 3 are then repeated for all possible event thresholds between the 50th and 99.99th percentile of f_0. The entire sequence is then repeated for a series of different factual distributions, f_1, each with a different anthropogenic signal considered, spanning the range S = $[0, 4\sigma]$.

Hence, the three independent variables considered are S, the anthropogenic signal of the factual distribution; E, the magnitude of the event considered; and N, the number of samples used to obtain an estimate of PR_{EAC}.

3.4. Results

Figure 3.3 shows, for three variations of f_1 with differing anthropogenic signals (S=0.5σ, 1.0σ, and 2.0σ), how the value of the two PR metrics compare as a function of the event magnitude, E. We present results using both a 100- and 10000-member ensemble: this represents the range of model ensemble sizes that might be found for both a CMIP5-based attribution study (King et al. 2015a) or with a very large, initial-condition ensemble (Massey et al. 2015; Black et al. 2016).

It is evident that for a given value of E, both the median PR_{EAC} and PR_{PD} values increase for larger anthropogenic signals, as expected. However, it is clear that the threshold-specific PR_{PD} metric is much more sensitive to the magnitude of the underlying signal, when compared to PR_{EAC}. The PR_{EAC} metric exhibits a particularly convenient property, in that a similar probability ratio is found across a surprisingly wide range of possible event thresholds.

As the event threshold considered in Figure 3.3 becomes more extreme, the two PR metrics converge to the same value. This is an expected result, since by definition,

$$\lim_{E \to \infty} PR_{PD} \equiv PR_{EAC}$$

A more interesting concept is to instead consider at what lower event threshold the two PR metrics begin to diverge. Since sampling uncertainty exists whenever a finite ensemble size is used to quantify PR_{EAC}, as is the case for the majority of event attribution studies, this uncertainty range can be used as a method of quantifying when these two PR metrics are no longer similar. It is worth emphasising here that the use of a bootstrap should not be interpreted as a robust measure of statistical uncertainty, but represents a simple method of

inferring how PR_{EAC} and PR_{PD} differ from one another, and how this divergence varies with the size of the model ensemble being used.
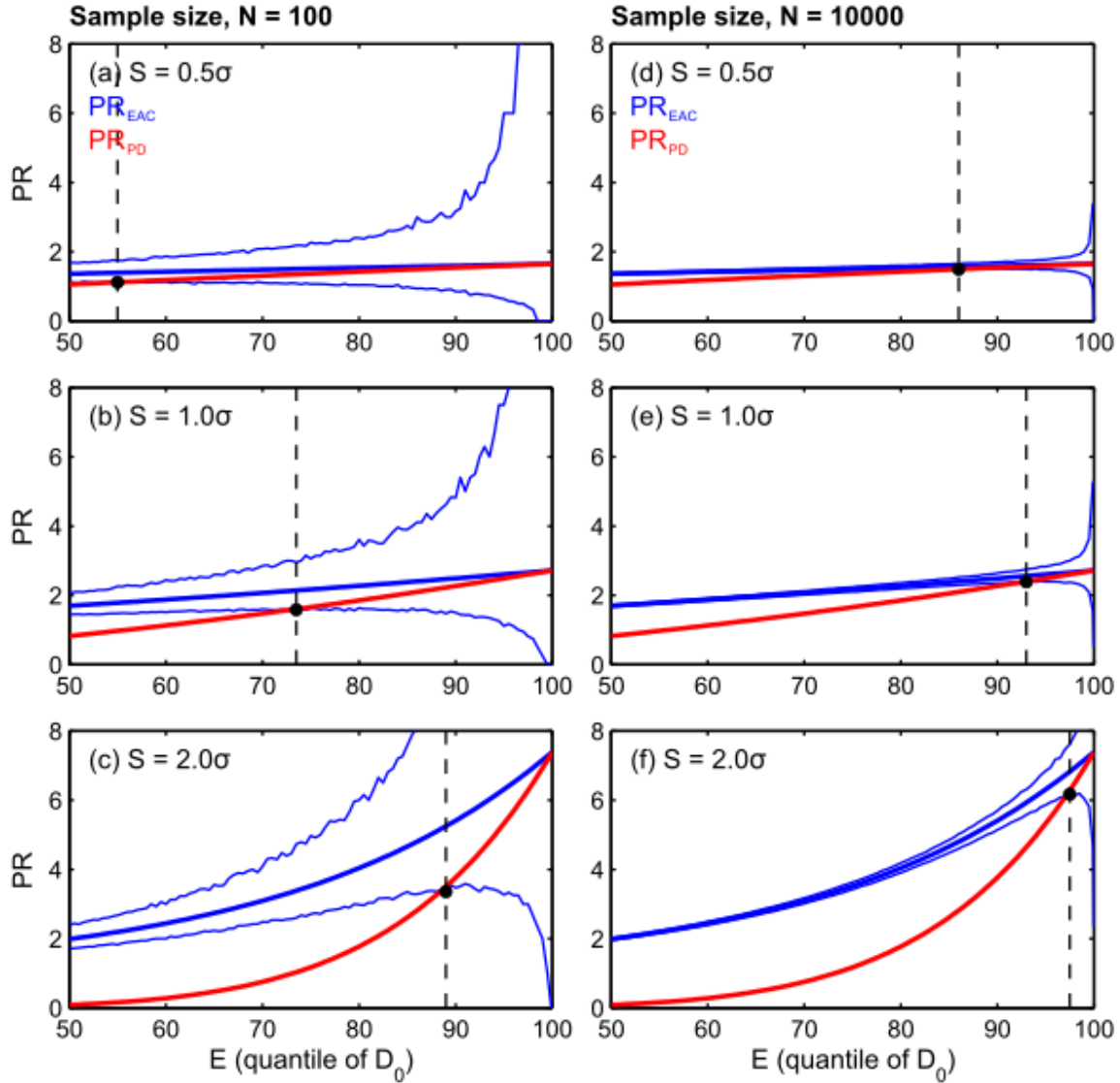
Figure 3.3: Comparison of PR metrics for different 'event' thresholds and 'ensemble' sizes for the idealized Gumbel distributions. Left-hand panels compare the 90% confidence range of PR_{EAC} (blue) with PR_{PD} (red) when considering a factual ensemble f_1 with an anthropogenic signal S of **(a)** 0.5σ, **(b)** 1.0σ and **(c)** 2.0σ, over a range of possible event thresholds, E, between the 50th and 99.99th percentile of the counterfactual distribution. Right-hand panels consider the same anthropogenic signals, but with a bootstrap sample size of 10000. Black dashed lines in each panel show the threshold of divergence for that particular factual distribution, such that the PR_{PD} for values of E greater than this quantile threshold lie inside the range of possible PR_{EAC} estimates.

We thus define the threshold of divergence, or TOD, as the minimum event threshold at which the value for PR_{PD} occurs within the 90% confidence interval of the corresponding PR_{EAC} estimate. This means that for a given event threshold, if $E \geq TOD$, there is no significant difference between the corresponding values of PR_{EAC} and PR_{PD}, whilst the opposite applies if $E < TOD$. In Figure 3.3, the black dashed lines correspond to the TOD for each prescribed value of N and S. As S increases, the range of event thresholds where there is

no discernible difference between the threshold-specific and event-as-class attribution statements reduces to eventually only the most extreme events; or using the proposed nomenclature, the TOD increases as a function of S.

To explore this concept further, Figure 3.4a illustrates how the threshold of divergence changes as a function of the anthropogenic signal in f_1 for a range of possible ensemble sizes. Specifically, the signal-event threshold space has been separated into three categories: (1) $PR_{EAC} \approx PR_{PD}$ according to our proposed criteria; (2) the distinction is dependent on the size of the model ensemble; and (3) where PR_{PD} does not overlap with PR_{EAC}. It is important to consider the relative range of E occupied by these three zones as a function of S. For values of S near zero, which may be analogous to short-duration extreme rainfall events or events spanning small spatial scales such that a detectable anthropogenic signal is only just apparent, the two PR metrics are deemed equivalent for nearly the entire range of event thresholds considered. There is then a transitional zone where determining whether or not PR_{EAC} and PR_{PD} are comparable is highly sensitive to the size of the model ensemble being used.

As the signal of anthropogenic climate change continues to become more pronounced, the range of event thresholds over which there are statistically significant differences between PR_{EAC} and PR_{PD} (red shading in Figure 3.4a) becomes increasingly larger. This is because the sampling uncertainty associated with an event-as-class PR estimate decreases as the underlying magnitude of the fraction of attributable risk increases. As this spread in PR_{EAC} narrows, the value of PR_{PD} continues to overlap with this uncertainty range for events only at the very tail of the distribution (see also Figure 3.3).

As increases in computational efficiency and greenhouse gas emissions are both expected to continue into the future, corresponding increases may also be expected in the sample size of the model ensemble used to assess a given event, and the underlying anthropogenic signal of the factual distribution representing the 'present-day climate'. Based on the results shown in Figure 3.4a, this potential convergence to higher values of S and N suggests there will be a wider range of event thresholds over which PR_{EAC} and PR_{PD} will not be comparable, when based on the criteria proposed in this analysis. To provide some context for the expected changes in S, recent research has shown signal-to-noise ratios of annual mean temperatures, relative to 1986-2005, are expected to emerge beyond $+2\sigma$ for over half the global surface area by 2050 under a moderate warming scenario (Frame et al. 2017).

While the results presented in this study are shown for Gumbel distributions, we reiterate the ease with which this framework could be considered with other distribution types. Based on the additional example of a Gaussian distribution, presented in the supplementary information, it appears that the threshold of divergence increases more slowly for distributions with heavier tails.
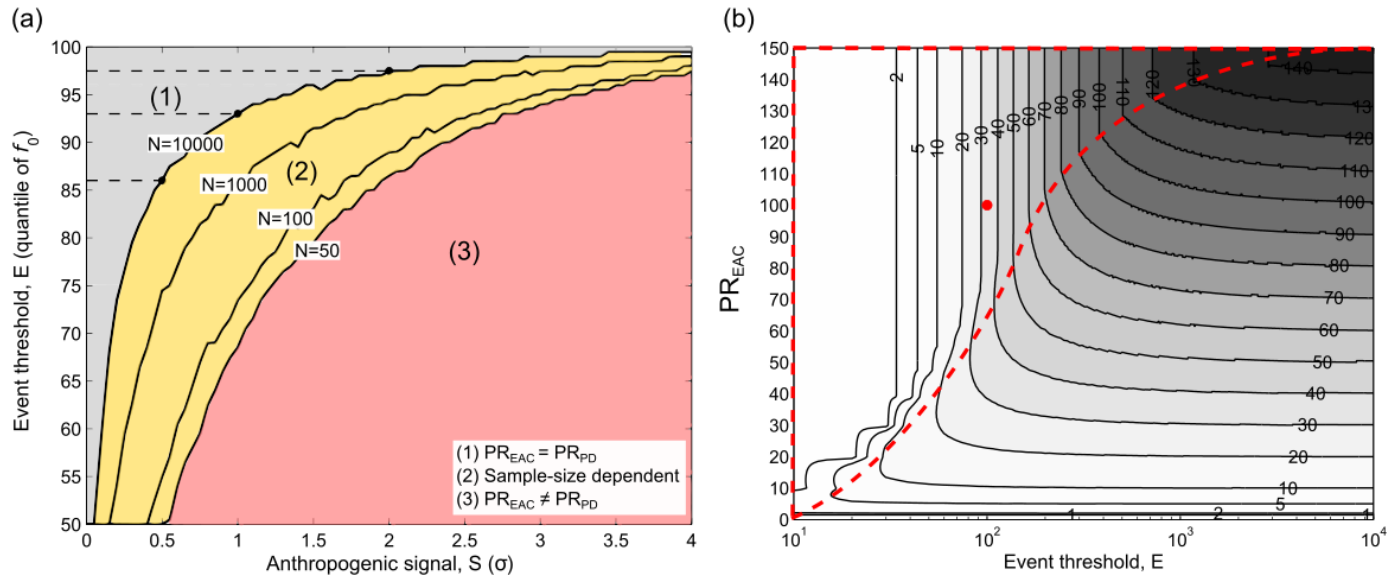
Figure 3.4: (a) The difference between event-specific and event-as-class PR metrics as a function of event threshold, E, the underlying anthropogenic signal of the factual distribution, S, and the sample size of the model ensembles used, N. Black solid lines correspond to the TOD as a function of S for each labelled sample size. The thin dashed lines show the TOD for specific values of S (0.5, 1.0, 2.0) using an ensemble size of 10000, as presented in the right-hand panels of Figure 3. **(b)** Quantifying PR_{PD} (shaded contours) as a function of event threshold, E (here, presented as return periods) and the median PR_{EAC}, which represents a proxy for the anthropogenic signal, S in the left-hand panel. Red dashed lines surround region where PR_{PD} diverges significantly from PR_{EAC}. The red circle corresponds to a 100-year return period and $PR_{EAC} = 100$.

3.5. Discussion

3.5.1 Implications for 'orthodox' extreme event attribution studies

While presenting the threshold of divergence as a function of anthropogenic signal-to-noise ratios enables the reader to understand systematic differences between PR_{EAC} and PR_{PD} with emergent climate change, it may be difficult to interpret the relevance of these results in the context of regular attribution analyses.

There are fairly robust statistical relationships between PR_{EAC} and PR_{PD} for differing anthropogenic signal-to-noise ratios and event thresholds. As such, Figure 3.4b instead presents the estimated value of PR_{PD} as a function of (1) the return level of the event considered (again, with respect to the counterfactual climatology), and (2) PR_{EAC} (as a proxy

for the anthropogenic signal in the factual distribution), for the idealized PDFs defined in Section 3.

When framed in this way, it is clear to see that in order for there to be substantive differences between the two attribution metrics (inside red dashed lines), you would need to be, for example, evaluating a 1-in-100 year event *and* also identify a 100-fold increase in likelihood (red circle). For the vast majority of extreme event analyses, the focus is on an event which has resulted in severe impacts. Therefore, it should be unusual for an event analysis concerning a relatively modest 1-in-100 year event (based on the counterfactual distribution) to result in an estimated $PR_{EAC} > 100$: by definition, this high a probability ratio means the factual distribution has shifted so dramatically, there should be more severe event thresholds for researchers to be focused on.

So while it is worthwhile to confirm that no significant divergence exists between the two attribution metrics, these sorts of high probability ratios for low event thresholds are not commonly the focus for extreme event attribution studies, and an *a priori* assumption that PR_{EAC} and PR_{PD} are comparable is probably valid for most cases.

3.5.2 Implications for near real-time attribution in the context of climate services

As mentioned in Section 3.2, this presumption of equivalence between PR_{PD} and PR_{EAC} will not hold in the context of incorporating attribution perspectives into core climate services of national weather/climate providers. If all possible climatological anomalies over a region of interest will be assessed as they are observed month-to-month or year-to-year, then PR_{PD} would likely be a more useful metric to employ, particularly when using the methods outlined in C15 and highlighted in this study. If both attribution metrics are chosen to be communicated, reasons for their differences should also be discussed and justification should be made about which metric is more relevant for planners or decision makers and why.

3.6 Summary and Outlook

In model-based probabilistic event attribution, an analysis becomes more event-specific when the boundary conditions used to simulate the factual and counterfactual distributions are conditioned to the observed state of the ocean or atmosphere at the time of the event. The subsequent statement about 'role of human influence' in an extreme event will also change as

a consequence (as the specific physical mechanisms which are being evaluated are no longer the same). There are important reasons for having this spectrum of options, in terms of conditioning on the event of interest: no conditioning requires the entire climate system to be simulated as a whole, but (1) compounding uncertainties (in terms of process understanding, and model uncertainties) may overwhelm any anthropogenic signals as a result, and (2) the computational cost for analysing a given spatial resolution will also be very high (since the explicit simulation of the entire climate system is required). More conditioning on the observations of the event will result in an attribution statement with higher confidence (as some possible sources of uncertainty will have been eliminated (Shepherd 2016)), but it will have less relevance to other extreme events which may occur in the future (Otto et al. 2016), and may only quantify the human influence on one part of a causal chain of physical phenomena contributing to the severity of a given event. From the perspective of an in-depth attribution study, multiple perspectives using varying levels of conditioning may therefore be complimentary.

As Figure 3.5 shows, these aforementioned choices about the extent of observational conditioning could be interpreted as an 'upstream' consideration on the framing of the analysis: from a very simplistic perspective, these decisions act to modify the characteristics of the 'counterfactual' and 'factual' distributions which are being compared. Once these decisions have been made, and the two PDFs to compare have been found, there is then a need to quantify the change in likelihood of witnessing the 'event' of interest. The most common approach has been to calculate at the cumulative density function of the two distributions, with the event threshold being the lower bound, and infinity as the upper bound (the traditional PR and FAR metrics). However, this study highlights that there are *also* choices about the attribution metric that can be used. Just like making choices about the extent of conditioning: (1) understanding both metrics will be more valuable than focusing on one in isolation; (2) each type of metric will answer a slightly different question, and thus will be of different value to different stakeholders; and (3) there are difficulties and uncertainties in quantification which are slightly different for each of the two metrics.
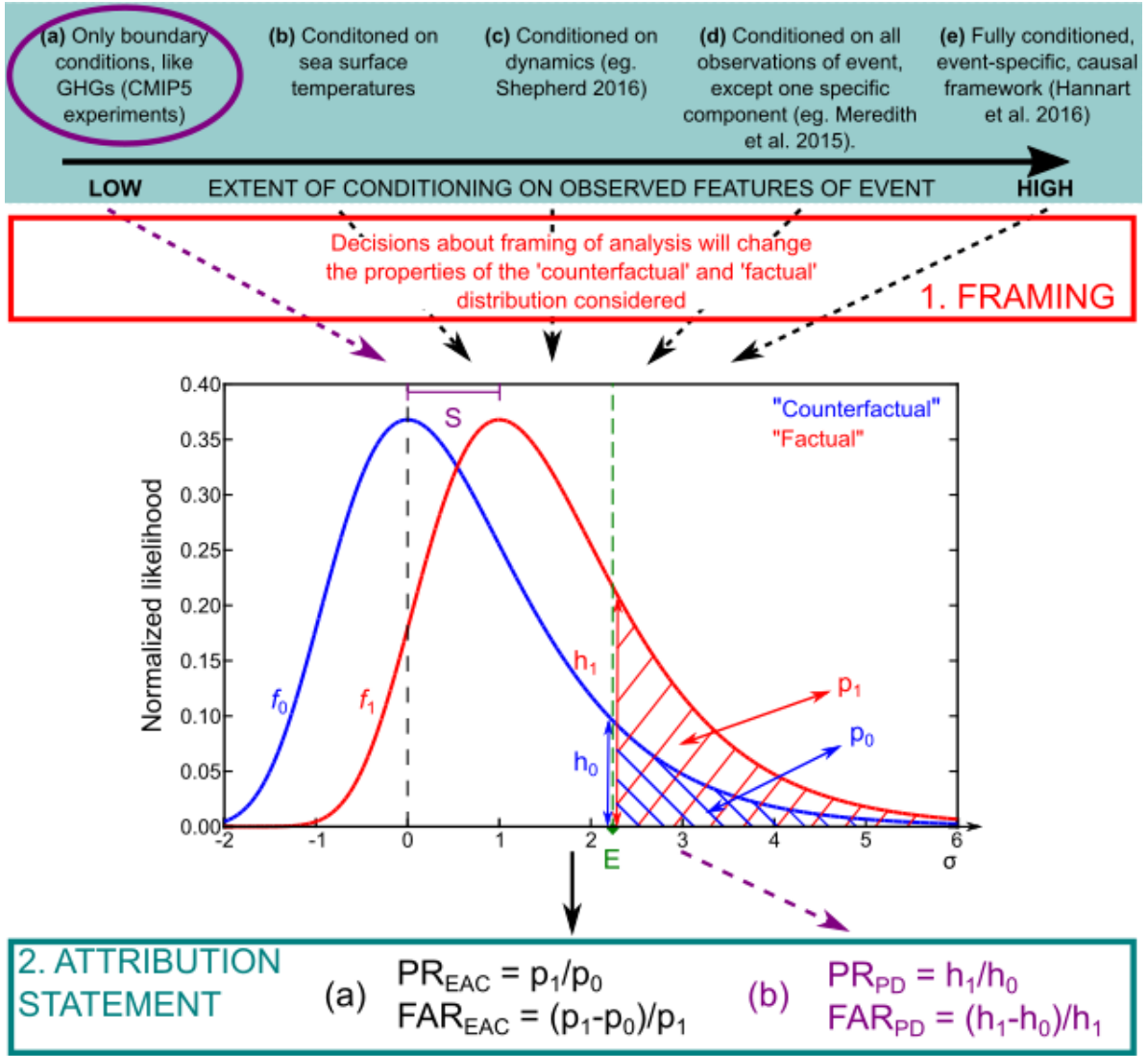
Figure 3.5: Schematic illustrating the role of choices about observational conditioning (top) influence the properties of the 'counterfactual' and 'factual' distributions investigated in an attribution analysis. There are also choices as to how to interpret the change in likelihood of an event threshold occurring. The purple text and arrows indicates the framework suggested in this study to be well-suited to providing attribution perspectives in the context of routine climate services.

This analysis has demonstrated that for the most extreme event thresholds and low signal-to-noise ratios, it can be a valid assumption that both PR_{EAC} and PR_{PD} are comparable to one another, when taking into account uncertainty estimates for a finite model ensemble. However, this prior assumption of equivalence will not always hold: specifically, for more modest climatological event thresholds, and under high signal-to-noise ratios. This is most relevant in the context of rapid attribution assessments proposed using an unconditioned model framework in Section 3.1.2. If seasonal- or annual-mean temperatures were to be evaluated (for example, in the context of climate services), not all positive anomalies will continue to be 'more likely to occur' in a continuously warming world: once the mode of the

factual distribution shifts beyond the anomaly of interest, the factual probability density at that specific threshold will become smaller than the counterfactual probability density soon thereafter. It is within this framework that understanding the differences between PR_{EAC} and PR_{PD} becomes important – both values should be calculated, and the reasons for their differences clearly articulated.

As the emergent signals of anthropogenic climate change become increasingly clear, different types of attribution statement will be of value to different stakeholders, depending on their priorities for decision making. This study further highlights the need for clarity when framing the interpretation of attribution metrics, particularly if attribution perspectives are to be incorporated into future climate services.

Chapter 4

Emergent population exposure to daily temperature extremes with increasing cumulative CO_2 emissions

The work described in this chapter follows Harrington et al. (2016a), published in *Environmental Research Letters*.

Understanding how the emergence of the anthropogenic warming signal from the noise of internal variability translates to changes in extreme event occurrence is of crucial societal importance. As with the drought case study presented in Chapter 2, most recent research which has attributed changes in extreme events to human influence have often focussed on the change in likelihood of witnessing a specific extreme event which occurred in the present day and produced damaging impacts. These event-specific studies interrogate extreme events with specific spatio-temporal characteristics, and commonly focus only on net anthropogenic changes to the climate system in the present day.

To instead obtain a more complete understanding of how changes to extreme weather manifest themselves through time, some studies have assessed the time-evolution of changes to extreme climate indices, often comparing against corresponding increases in global mean temperature. In this study, we extend this alternative approach to consider the exposure of different populations of the world to changes in extreme daily temperatures. This chapter presents a further novel extension on previous studies, by attempting to link the coincident increases in global mean temperature with the primary cause of these changes: cumulative carbon dioxide emissions.

By utilising simulations of cumulative carbon dioxide (CO_2) emissions and temperature changes from eleven earth system models, we demonstrate that the inherently lower internal variability found at tropical latitudes results in large increases in the frequency of extreme daily temperatures (exceedances of the 99.9th percentile derived from pre-industrial climate simulations) occurring much earlier than for mid-to-high latitude regions. Most of the world's poorest people live at low latitudes, when considering 2010 GDP-PPP per capita; conversely the wealthiest population quintile disproportionately inhabit more variable mid-latitude climates. Consequently, the fraction of the global population in the lowest socio-economic

quintile is exposed to substantially more frequent daily temperature extremes after much lower increases in both mean global warming and cumulative CO_2 emissions.

4.1 Introduction

To understand how detectable anthropogenic influences on the climate system will proliferate with time, a large body of literature has considered the question of when the signal of climate change emerges from the background noise of internal variability. This 'time of emergence' concept has been considered in both observational records and climate model simulations, for a range of climate indices such as temperature (Hawkins and Sutton 2012; Diffenbaugh and Scherer 2011; Joshi et al. 2011; Mahlstein et al. 2011; Hawkins et al. 2014), precipitation (Giorgi and Bi 2009; Maraun 2013), the hydrological cycle (Sedláček and Knutti 2014), sea level rise (Lyu et al. 2014) and even transitions between different ecosystem regimes (Mahlstein et al. 2013).

There has also been intense public and scientific interest in recent years as whether and to what extent the severity and frequency of extreme weather events have increased in response to anthropogenic climate warming (Peterson et al. 2012, 2013, Herring et al. 2014, 2015). However, determining how the 'time of emergence' concept can be applied in the context of extreme events continues to be a developing area of research. Recent work (Fischer and Knutti 2015) has quantified the fraction of all moderate heat extremes and precipitation extremes globally which could be attributed to anthropogenic climate change in the present-day as well as for future climate change scenarios, while the emergence of statistically significant changes to specific climate extreme indices have also been demonstrated (King et al. 2015a, 2016).

A key interpretation of studies on the time of emergence of climate change indicators suggests that, on seasonal timescales (Diffenbaugh and Scherer 2011; Mahlstein et al. 2011; Kirtman et al. 2014; Mora et al. 2013), climate signals may emerge from the 'noise' of internal climate variability more quickly for low latitude regions than elsewhere around the world. This has been convincingly demonstrated for mean climate indicators (Diffenbaugh and Scherer 2011; Mahlstein et al. 2011), but these implications have not been explored when considering the exposure of different parts of the global population to the emergence of extremes, particularly temperature extremes, which have been found to occur much earlier than precipitation-related extremes (King et al. 2015a; Fischer et al. 2014).

When considering the regional-scale impacts which may occur in response to a changing climate, the majority of results are communicated with respect to corresponding changes in global mean temperature – this comparison is useful in the context of international climate targets, which are also framed in relation to global mean temperature anomalies (Knutti et al. 2016). However, progress in recent years quantifying the near-linear correlation between cumulative carbon dioxide (CO_2) emissions with corresponding global temperature anomalies (Allen et al. 2009; Matthews et al. 2009; Meinshausen et al. 2009) have permitted a more in-depth consideration of how the regional impacts of climate change may respond directly to the emission of long-lived greenhouse gases, with potentially important policy implications (Frame et al. 2014).

In this study, we examine evidence for spatial heterogeneity in the time-evolution of extreme temperature exceedances between different regions of the world aggregated according to local population and income characteristics, using for the first time, a direct comparison with the accumulation of simulated CO_2 emissions.

4.2 Data and methods

4.2.1 Probability Ratios

Following the methodology of previous work by Fischer and Knutti (2015), we employ the use of the 'probability ratio' (PR) metric, defined as PR=P_1/P_0 where P_0 is the probability of exceeding a certain quantile during the pre-industrial control period and P_1 is the likelihood of exceedance in a more recent period, for example the last 30 years. PR can be interpreted in a climate modelling framework as the increased likelihood of an extreme temperature threshold being exceeded, when comparing a more recent time period with a representation of a climate in the absence of human interference (Allen 2003; Stott et al. 2004).

Using models from the Coupled Model Intercomparison Project Phase 5 (CMIP5), we concatenate 'Historical' simulations for the period 1901-2005 with corresponding Representative Concentration Pathways (RCPs) for the period 2006-2100, and then consider time-varying PR values using moving 30-year windows. In this analysis, we choose to define P_0 as the 99.9th percentile of daily temperatures (corresponding to a 1-in-1000 day temperature extreme) based on 200 years of pre-industrial simulations. By focussing on changes to the number of exceedances above a fixed, well-defined percentile threshold over running 30-year intervals, our analysis does not require any assumptions about the shape and

type of the underlying statistical distribution, and thus avoids recent concerns relating to the use of parametric analysis when considering a non-stationary time-series for the analysis of climate extremes (Sardeshmukh et al. 2015; Sippel et al. 2015). It is also noted that a 1-in-1000 return period is sufficiently high for the signal-to-noise ratios considered in this analysis, such that any potential deviations between 'event-as-class' and 'event-specific' probability ratios (as discussed in Chapter 3) are negligible for the results hereafter presented in this chapter.

4.2.2 Cumulative CO_2 emissions

One of the most important results to emerge from the latest assessment report of the Intergovernmental Panel on Climate Change (IPCC) was the near-linear correlation identified between cumulative CO_2 emissions and corresponding increases in global mean temperature. This relationship has helped to demonstrate the influence of different mitigation measures, and the comparison of potential risks between possible future emissions pathways.

In order to improve confidence in temperature projections by reducing model uncertainty associated with carbon cycle feedbacks, the RCP scenarios have been created with prescribed concentrations rather than emissions (Moss et al. 2010). Consequently, the only previous studies considering the physical climate impacts of cumulative emissions from RCP scenarios have inferred cumulative CO_2 based on a best-guess linear scaling of global mean temperature anomalies (Seneviratne et al. 2016). Here, we instead utilise estimated RCP emissions calculated by Jones and colleagues (2013), whereby the time-evolution of atmospheric carbon and corresponding simulated exchange of carbon with land and ocean sinks for a smaller subset of earth system models (Table 4.1) have been used to infer CO_2 emissions that are compatible with each prescribed concentration pathway. Thus, all subsequent calculations of cumulative emissions corresponding to extreme climate impacts include the uncertainty in translating global warming anomalies to cumulative emissions as well as the uncertainty in linking global temperatures to probability ratios (Figure 4.1).

CMIP5 Models used	Hist-RCP2.6	Hist-RCP4.5	Hist-RCP8.5
CanESM2	1	1	1
GFDL-ESM2G	1	1	1
GFDL-ESM2M	1	1	1
HadGEM2-CC	-	1	1
HadGEM2-ES	1	1	1
IPSL-CM5A-LR	1	1	1
IPSL-CM5A-MR	1	1	1
IPSL-CM5B-LR	-	1	1
MIROC-ESM	1	1	1
MIROC-ESM-CHEM	1	1	1
MPI-ESM-LR	1	1	1

Table 4.1: Subset of CMIP5 models and experiments with available estimates of simulated cumulative CO_2 emissions

Figure 4.1 demonstrates the correlation between cumulative carbon emissions and global warming anomalies, using the compatible emission pathways calculated for the eleven CMIP5 models with available model output (see Jones et al. (2013) for a full description of the methodology). Here, we are considering the concatenated 'Historical-RCP8.5' simulations from 1901-2100. Global mean warming anomalies are taken with respect to 1861-1880 while cumulative carbon emissions are taken relative to 1870.

It is evident that a linear relationship between CO_2 accumulations and temperature indeed exists with relatively good agreement between all models. When comparing with the best-guess estimate of observed global warming and cumulative emissions in 2014 (green dashed lines; Peters et al. (2015)), it may appear that the rate of warming per PgC of cumulative CO_2 is faster than expected in the models. While previous research has found stronger feedbacks in the land carbon cycle resulted in higher warming projections for emissions-driven scenarios, when comparing with concentration-driven simulations (Friedlingstein et al. 2014), it is important to emphasise that multiple additional influences contribute to the observed warming anomaly other than just cumulative CO_2 emissions (including non-CO_2 greenhouse gases, aerosols, land use changes, internal variability), and an in-depth analysis by Gillett et al. (2013) found no systematic model bias in the warming response to cumulative carbon emissions only.
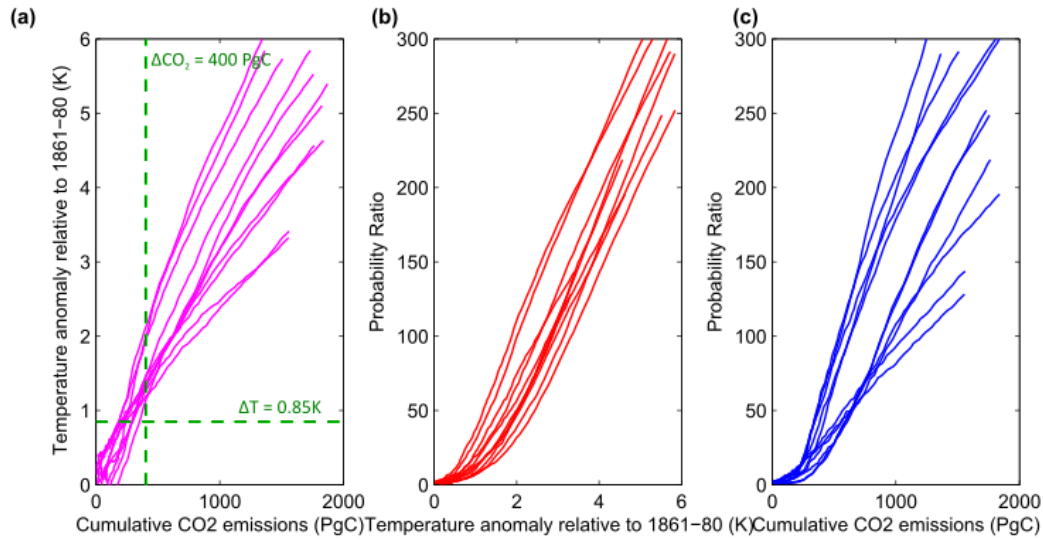
Figure 4.1: Comparison of the relationship between modelled cumulative CO_2 emissions (since 1870), global temperature increases (relative to pre-industrial simulations) and the globally averaged increase in probability of 1-in-1000 day temperature extremes (PR). Data is taken over 1901-2100 using Historical and RCP8.5 simulations for eleven CMIP5 models. Panels compare **(a)** Cumulative CO_2 versus global mean temperature increases; **(b)** Global mean temperature versus PR; **(c)** Cumulative CO_2 versus PR changes. Present-day estimates of cumulative CO_2 emissions and global mean warming are shown as green dashed lines in panel (a).

4.2.3 Defining the Cumulative Emissions of Probability Ratio Emergence (CEPRE) metric

Previous work has considered the temporal evolution of PR (or similarly, risk ratios) averaged either globally or over large regions (Figure 4.2a). Here, we choose to focus instead on when specific PR thresholds are exceeded at a single grid scale. Framing the emergence of temperature extremes in this manner also enables a consistent method of comparison between individual locations, as well as between different RCP scenarios. We define the cumulative emissions of probability ratio emergence (CEPRE$_X$) as the simulated accumulation of CO_2 emissions (since 1870) corresponding to the central year of the 30-year period over which a specific PR threshold (X) is exceeded at a specific grid cell – we note that a PR threshold is only considered to be exceeded when all subsequent values remain above the same threshold for the remainder of the time series available (2100). For example, figures 4.4a, 4.4c and 4.4e show, respectively, the CEPRE corresponding to when the PR exceeds 2, 10 and 50 for each grid point globally. Hawkins et al (2014) highlighted possible endpoint effects with this type of analysis, but sensitivity tests have shown this is negligible in this study, due to the low variability in PR: this is discussed further in Section 4.4.3.
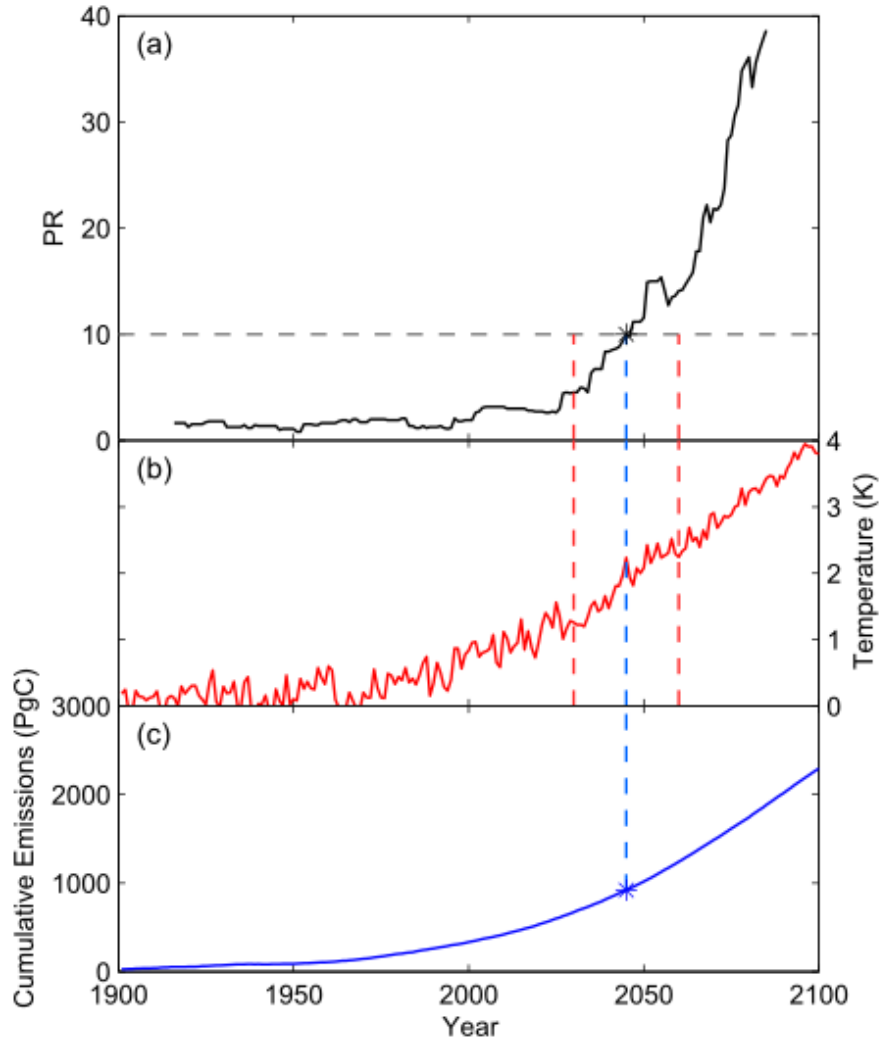
Figure 4.2: Schematic illustrating how the time-evolution of high-temperature probability ratios are converted to the CEPRE metric. **(a)** Time-evolution of PR for an individual grid point, presented at the central year of the 30-year period (dashed red line given as example) over which each PR value was calculated. **(b)** Corresponding global mean temperature anomalies (relative to 1861-1880) over the period 1901-2100. **(c)** Compatible cumulative CO_2 emissions over the same period, measured as a total since 1870. The black asterisk corresponds to the year when a PR threshold of 10 (grey dashed line) is first exceeded, and continues to remain so, with the corresponding blue asterisk indicating the $CEPRE_{10}$ value for that particular grid box.

4.2.4 Population and socio-economic scenarios

To consider how these emergent changes in PR are experienced by the populated regions of the world, we overlay global population and gross domestic product data, accounting for purchasing power parity (GDP-PPP), prepared for the year 2010 by the Global Carbon Project (http://www.cger.nies.go.jp/gcp/population-and-gdp.html). Population and socio-economic data have been provided at a 0.5 x 0.5 degree spatial resolution (Figure 4.3): to calculate corresponding GDP per capita data, total GDP-PPP is simply divided by the corresponding population total in each grid box. To obtain estimates of CEPRE aggregated

according to the highest and lowest socio-economic quintiles, the spatial distribution of PR for each ESM are calculated on their native spatial resolution, and each model is then downscaled to the same 0.5 x 0.5 degree resolution, using a bilinear interpolation approach.

Future changes to population and economic growth will have clear influences on these results - we have chosen to focus on fixed data of the present-day to ascertain emergent risks for present populations, and consider future projections of population and economic data in Section 4.4.2. While internal variability in temperature is greater over smaller spatial scales, the aggregation of grid cells over areas comparable in size to populated regions of the world have previously demonstrated discernible shifts in the probability distributions of temperature and precipitation extremes (Fischer et al. 2013, 2014), and thus will be suitable for analysing the emergence of high-temperature probability ratios.
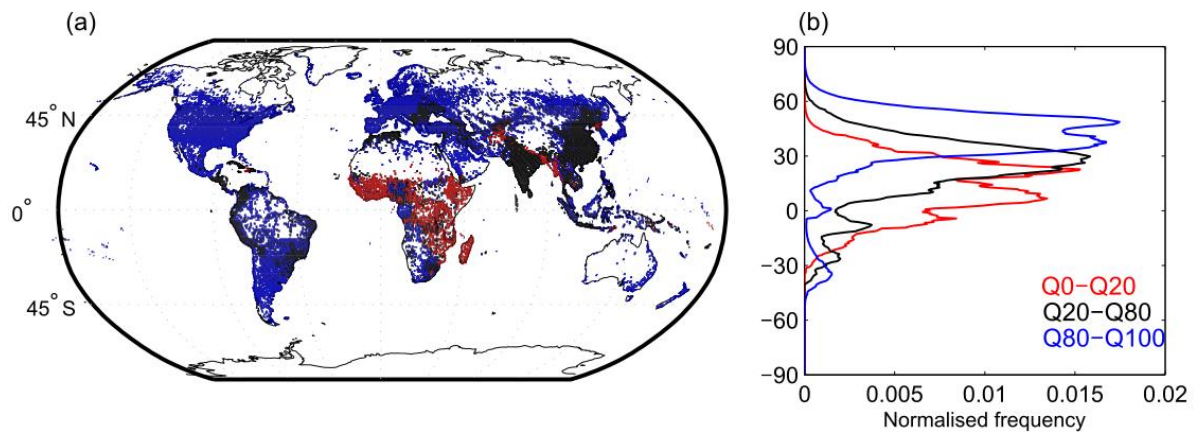
Figure 4.3: (a) Map showing population distribution data for 2010. Blue cells correspond to the wealthiest quintile of the global population; red cells correspond to the poorest quintile of the global population, while the remainder of black cells show the rest of the global population distribution. **(b)** Normalised zonal distribution of the global population, aggregated according to the poorest quintile (red), wealthiest quintile (blue) and the middle 60% of the population (black). Frequencies have been smoothed using a 5 degree filter.

4.3 Results

4.3.1 Spatial patterns

Figure 4.4 illustrates the cumulative emissions required for PR thresholds of 2, 10 and 50 to be exceeded in different regions of the world. The maps show the median model CEPRE for each PR threshold using RCP8.5 simulations, while the corresponding timing of exposure of the poorest and wealthiest quantiles of the global population are also presented. Considering the spatial distribution of CEPRE for each PR threshold, it is clear that fewer cumulative emissions are required for the continual exceedance of these PR thresholds to occur for lower

latitudes, compared with higher latitudes, while oceanic regions also generally appear to experience a more rapid time of emergence than corresponding land surfaces nearby. These results are consistent with previous research (Diffenbaugh and Scherer 2011; Fischer and Knutti 2015) and suggests that although land and high latitude regions experience larger warming signals than the global mean, the variability in daily temperatures over ocean and low latitude regions are significantly lower, thereby resulting in the earlier emergence of more frequent high-temperature extremes.
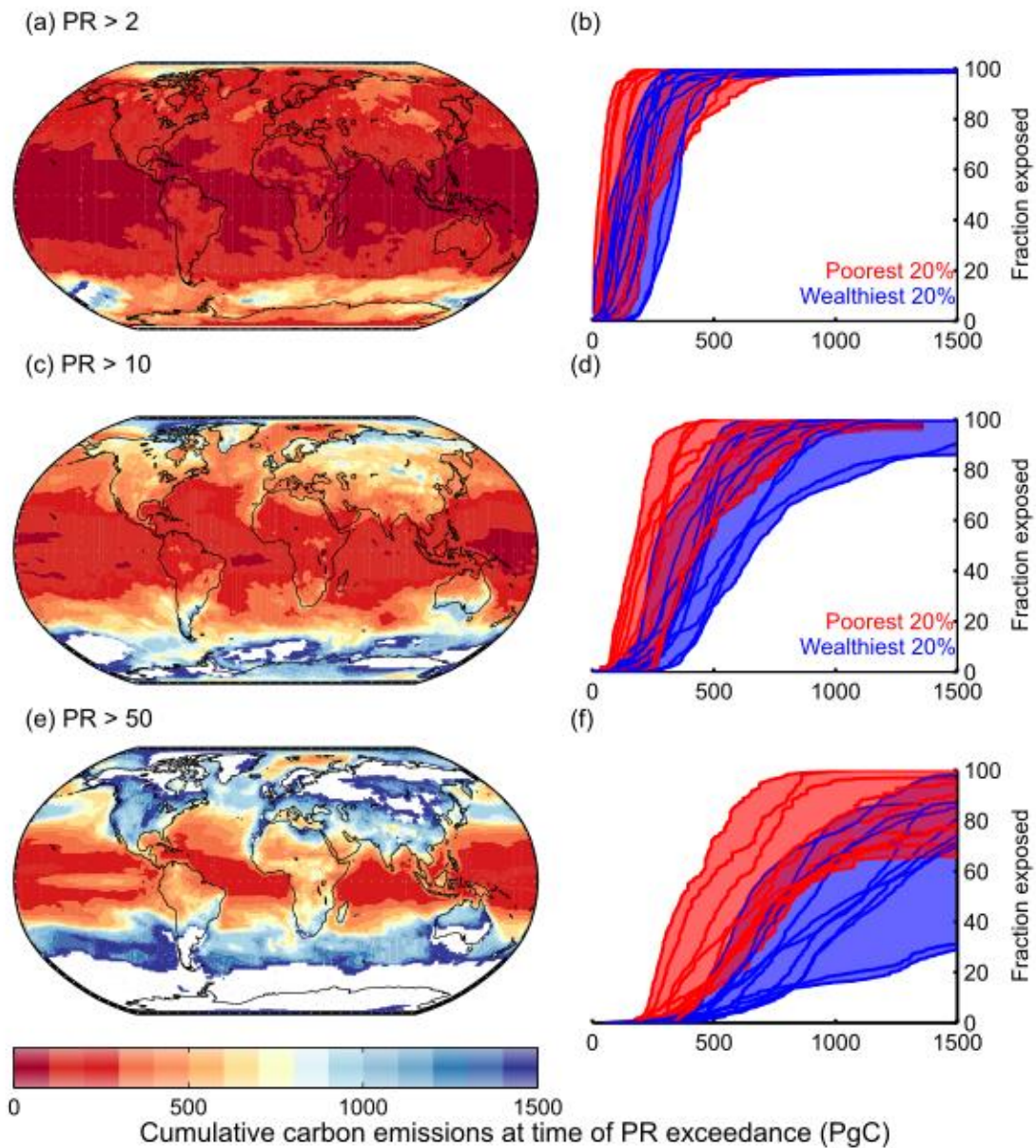
Figure 4.4: Left-hand panels show the global distribution of the median model estimate of the cumulative emissions of probability ratio emergence (CEPRE). This is defined at the simulated cumulative emissions of carbon dioxide (in Petagrams of carbon, with respect to 1870) at the central year of the 30-year period when the probability ratio for one-day temperature extremes permanently exceeds a threshold of **(a)** 2; **(c)** 10 and **(e)** 50 at each individual grid point. White regions indicate where the median model response shows no exceedance of the

71

respective PR threshold by the end of 2100. Corresponding right-hand panels show the cumulative fraction of the poorest (red) and wealthiest (blue) quintiles of the global population (based on 2010 GDP-PPP per capita) which are exposed to the specified PR threshold exceedances, also as a function of cumulative carbon emissions, for eleven CMIP5 models over the period 1901-2100, using 'Historical' and 'RCP8.5' simulations.

Considering the differences over quintiles in GDP-PPP per capita, we find increasingly large differences in CEPRE between the wealthiest and poorest populations, as the PR threshold increases. For example, comparing the emissions required to exceed a PR of 2 (Figure 4.4b) shows minor differences in the evolution of population exposure with warming between the two economic groupings. By contrast, considering the cumulative emissions required for population exposure to a much higher PR threshold of 50 reveals a much wider gap in the time of emergence between the two sub-populations, albeit with large variability between models. It is also noted that $CEPRE_2$ is less than 400PgC for nearly 100% of both sub-populations, which corresponds to present-day estimates of global cumulative emissions (Peters et al. 2015). Interestingly, when considering the CEPRE for RCP4.5 and RCP2.6 simulations instead (Figures 4.5 and 4.6), the differences in the timing of PR exposure for the lower latitude locations are negligible. Instead, it is exceedances of the higher PR thresholds which are not reached for mid-to-high latitudes by the end of the century.

(a) PR > 2

(b)

(c) PR > 10

(d)

(e) PR > 50

(f)

Poorest 20%
Richest 20%

Fraction exposed

Cumulative carbon emissions at time of PR exceedance (PgC)
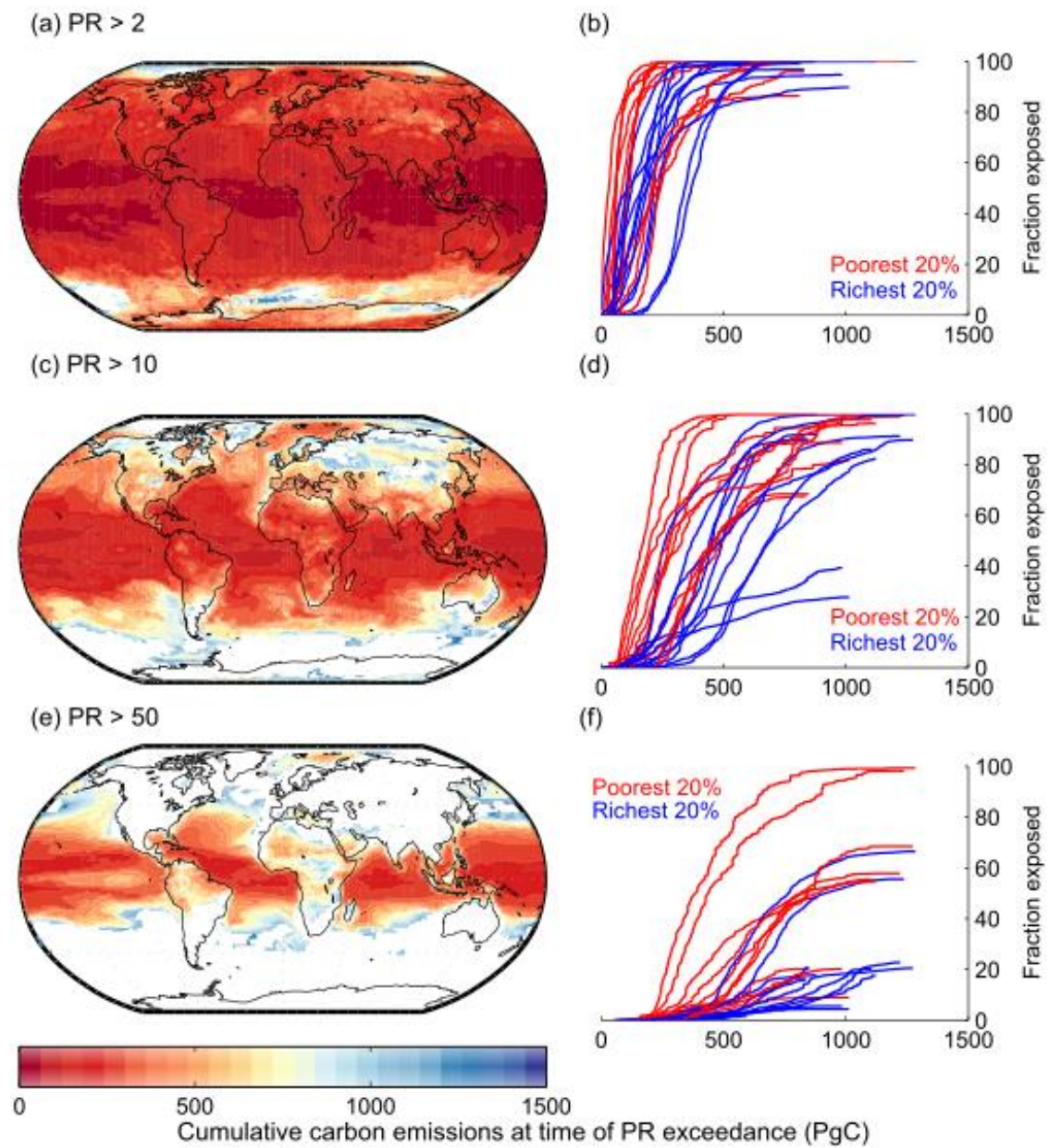
Figure 4.5: Same as Figure 4.4, but using 'Historical' and 'RCP4.5' simulations have over the period 1901-2100.

73
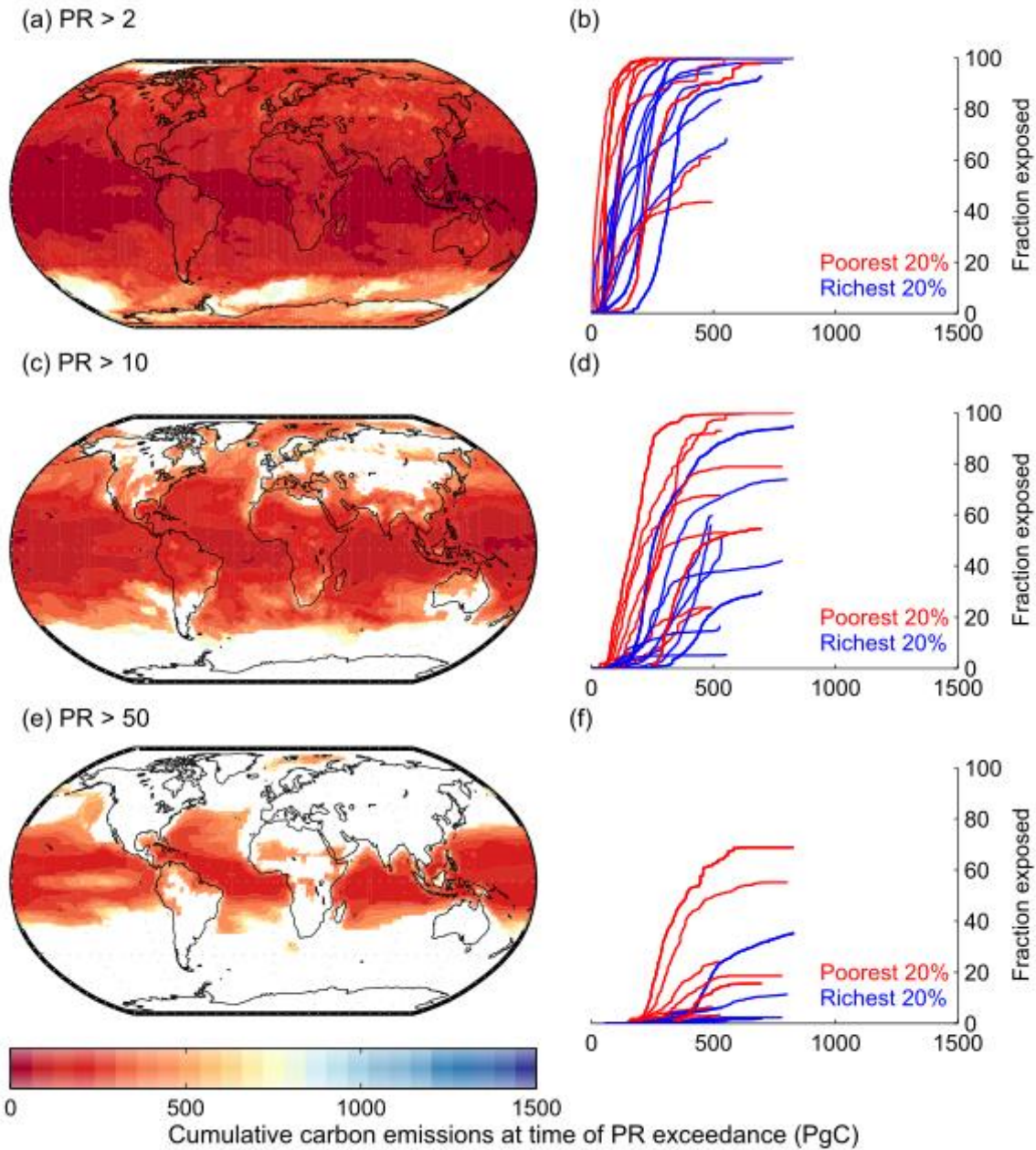
Figure 4.6: Same as Figure 4.4, but using 'Historical' and 'RCP2.6' simulations have over the period 1901-2100.

4.3.2 Differences in cumulative emissions required for emergent population exposure

Comparing, for each individual model, the difference in cumulative emissions required for 50% of the poorest population quintile to exceed each PR threshold versus the equivalent to occur for 50% of the wealthiest quintile, Figure 4.7 reveals that nearly all models across all RCP scenarios show high-temperature probability ratios occurring earlier for the poorest populations than for their wealthier socio-economic counterparts. Under an RCP8.5 scenario

for example, models suggest that between -10 and 560 PgC of additional carbon would be emitted between the time when 50% of the poorest members of society continually experience a 50-fold increase in 1-in-1000 day hot extremes and the time when exposure occurs for an equal number of citizens within the wealthiest population quintile, thereby emphasising the contrasting time horizons available for adaptation measures.
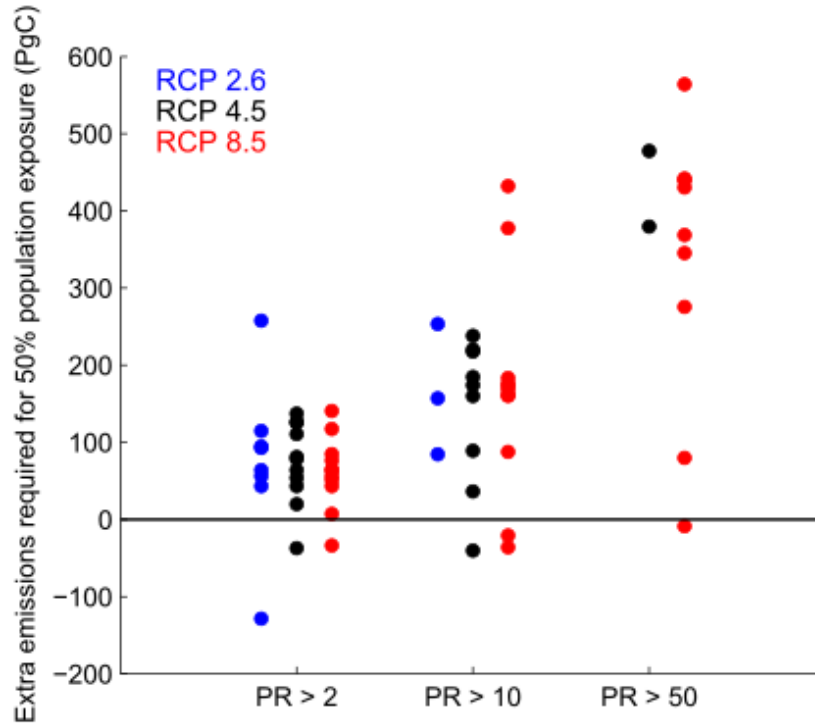
Figure 4.7: The difference in cumulative emissions between the central year (of the 30-year period) when 50% of the poorest population quintile experiences the continuous exceedance of a given PR threshold, and the corresponding central year when 50% of the wealthiest population quintile exceeds that same threshold. Only model values are shown where greater than 50% cumulative exposure occurs for both sub-populations by 2100. Blue, black and red circles correspond to each model using RCP2.6, RCP4.5 and RCP8.5 simulations respectively.

Further interrogating the differences in the cumulative distribution of CEPRE between the wealthiest and poorest socio-economic quintiles, Figure 4.8 shows the differences in the fractions of each population quintile which have experienced emergence of each PR threshold, as a function of global cumulative CO_2 emissions. This figure demonstrates that (1) after a given level of cumulative emissions, up to an additional 60% of the poorest members of society cross each of these PR thresholds than corresponding wealthy populations; and (2) these patterns of unequal population exposure in response to accumulating CO_2 emissions occur consistently across all three RCP scenarios. This shows that the differences in the timing of emergence of temperature extremes between low latitude

and high latitude regions are insensitive to the rate of temperature change over the twenty-first century.
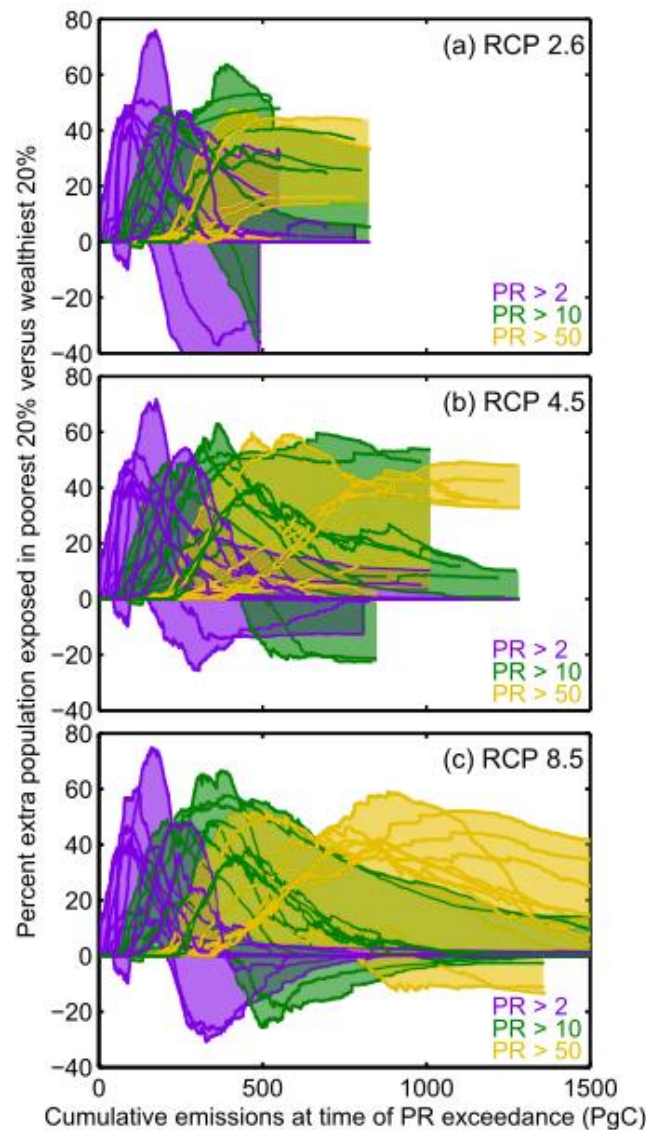
Figure 4.8: Difference in the fraction of population exposure to specified PR thresholds (colours) between the poorest 20% and the wealthiest 20% of the global population, as a function of CEPRE. Positive values indicate that a greater fraction of the poorest population quintile is experiencing permanent exceedances of each specified PR threshold, compared with the wealthiest quintile, after a given quantity of cumulative carbon emissions.

In 2013, the United Nations Framework Convention on Climate Change established the Warsaw International Mechanism to address the potential loss and damage from climate change impacts for developing countries (James et al. 2014). The policy-relevance of our result lies in the disparity between richer and poorer people in terms of their exposure to the timing of emergence of temperature extremes. Whilst exposure to higher probability ratios does not result in higher vulnerability, the adaptive capacity of a region can generally be

considered to scale with 1) climate variability and 2) income - certainly across the range implied by considering the richest and poorest quintiles (Grambsch and Menne 2003; Hayden et al. 2011). These results do therefore suggest, *ceteris paribus*, earlier and more significant relative vulnerability to temperature extremes among the world's poor and are thus of potential importance to policymakers.

4.4 Sensitivity analysis

While the compatible emissions profiles used in this study have been shown to accurately replicate the original Integrated Assessment Models used in developing the Representative Concentration Pathways (Jones et al. 2013), it has been demonstrated that emission-driven simulations overestimate warming projections when compared with concentration-driven simulations (Friedlingstein et al. 2014). We therefore choose to avoid specifying absolute cumulative emission targets for preventing the emergence of specific PR thresholds. However, even with the added uncertainty of considering the end-to-end link between cumulative emissions and extreme temperatures, the key differences in fractional exposure of emergent high-temperature probability ratios between the wealthiest and poorest global population quintiles remain clear, as evidenced by the equivalent results found when considering the link with rises in global mean temperatures directly (Figure 4.9).
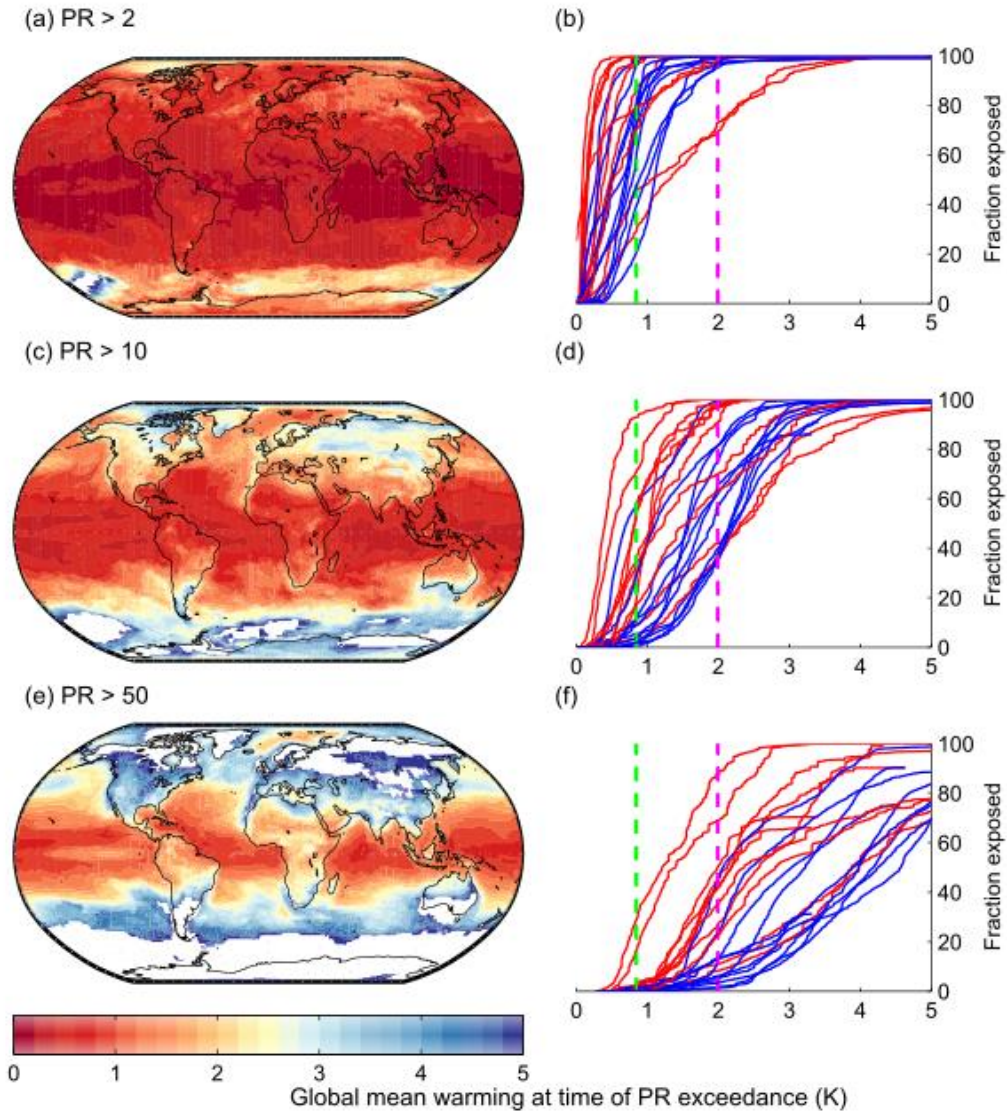
Figure 4.9: Same as Figure 4.4, but considering the global mean warming anomaly over the 30-year period when each respective PR threshold is continually exceeded at each grid point, rather than cumulative carbon emissions. 'Historical' and 'RCP8.5' simulations have been used over the period 1901-2100. Temperature anomalies are taken with respect to 1861-1880. Green and magenta lines correspond to global mean warming of 0.85K and 2K respectively.

4.4.1 Interpreting model differences in population exposure to PR thresholds

To investigate how the model spread in Figure 4.7 relates to the transient climate response to cumulative carbon emissions (TCRE, Gillett et al. (2013)) of the model ensemble, we define TCRE as the global mean warming anomaly of each model after 1000 PgC (1EgC) of cumulative carbon dioxide emissions (Figure 4.1a). Figure 4.10 compares each models' corresponding TCRE with the difference in cumulative emission requirements for 50% exposure of the two population quintiles to each PR threshold (as per Figure 4.7),

concatenating values across all RCP scenarios. Interestingly, it appears that those models with the largest TCRE show the smallest difference in the emissions required for exposure to PR thresholds between the wealthiest and poorest population quintiles, while the models with the lowest TCRE show the largest differences between the sub-populations.
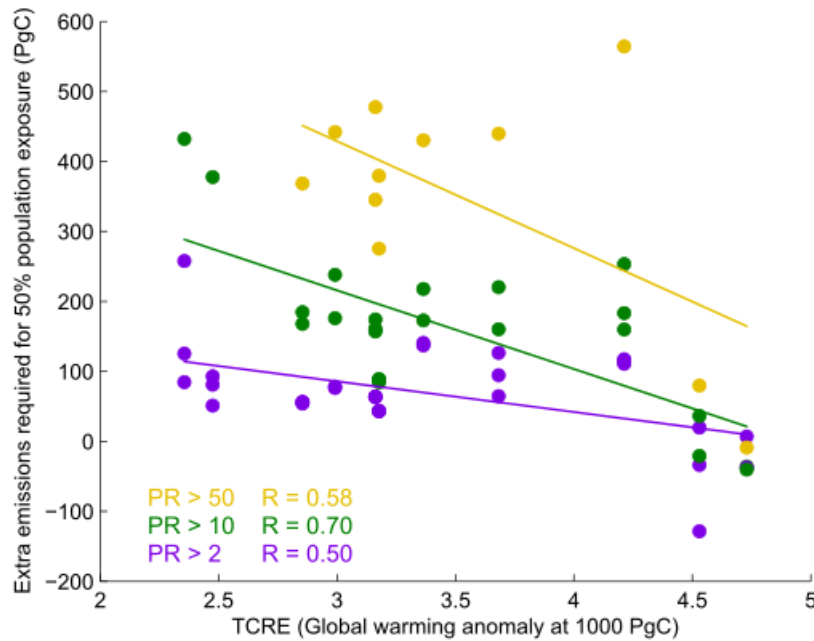
Figure 4.10: Comparing model TCRE with the difference in cumulative emissions between the central year (of the 30-year period) when 50% of the poorest population quintile experiences the continuous exceedance of a given PR threshold, and the corresponding central year when 50% of the wealthiest population quintile exceeds that same threshold. Only model values are shown where greater than 50% cumulative exposure of a given PR threshold occurs for both sub-populations by 2100.

Considering the individual model response of the two models with the lowest (Figure 4.11) and highest (Figure 4.12) warming responses to 1EgC of cumulative CO_2, it becomes evident that rapid increases in probability ratios occur in tropical latitudes for all models, while corresponding PR threshold exceedances in higher latitudes only occur quickly for the high-sensitivity models, subsequently resulting in smaller differences in CEPRE between the two sub-populations. This result provides further evidence of the potentially counter-intuitive results which occur when considering more policy-relevant metrics of change, such as the difference in timing between different population quintiles, rather than simply the globally aggregated measures of change for individual models like TCRE.
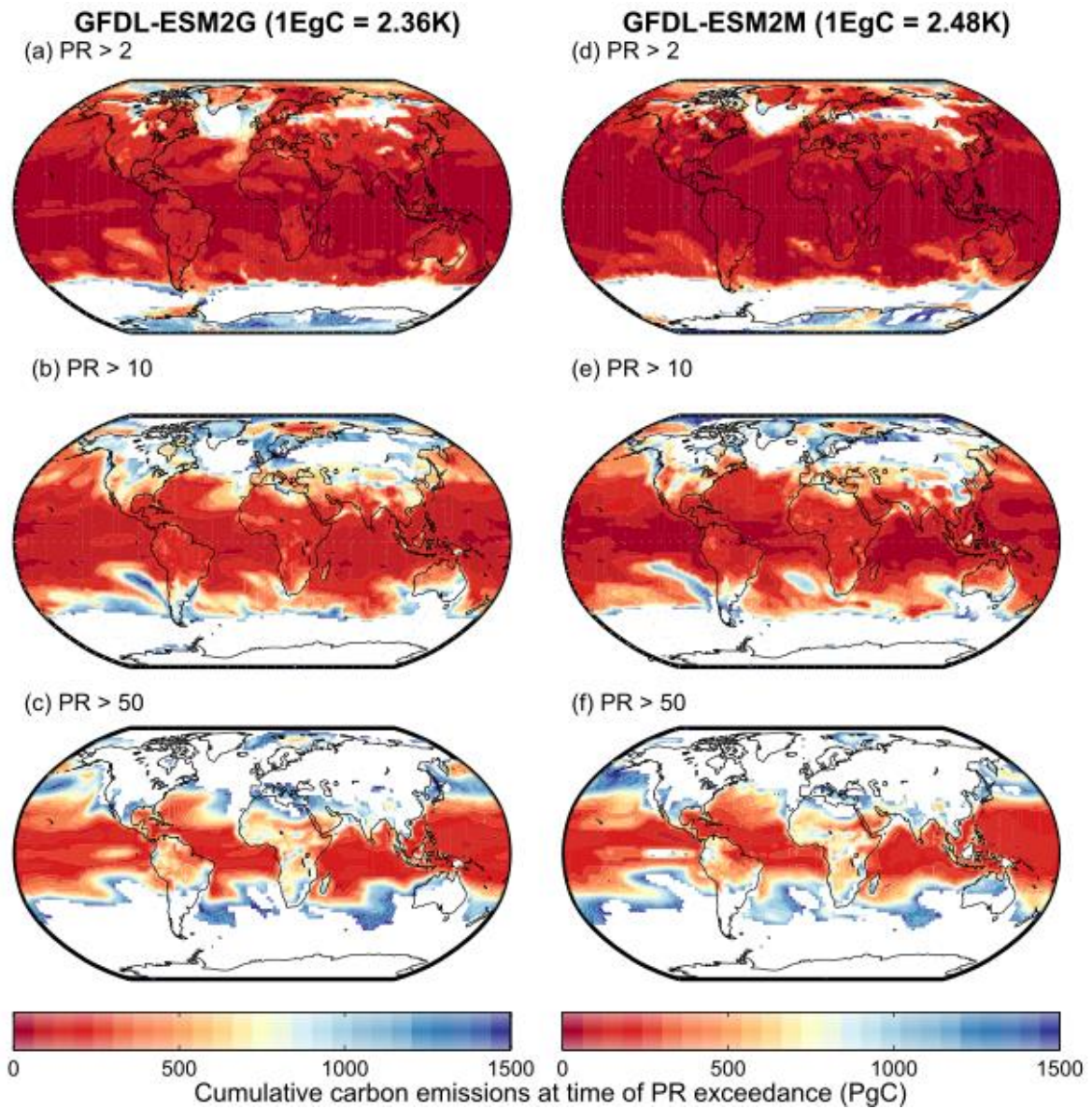
Figure 4.11: Same as left-hand panels of Figure 3.4, but showing the individual model responses of (left) GFDL-ESM2G and (right) GFDL-ESM2M. Corresponding values of empirical TCRE are listed next to each model.
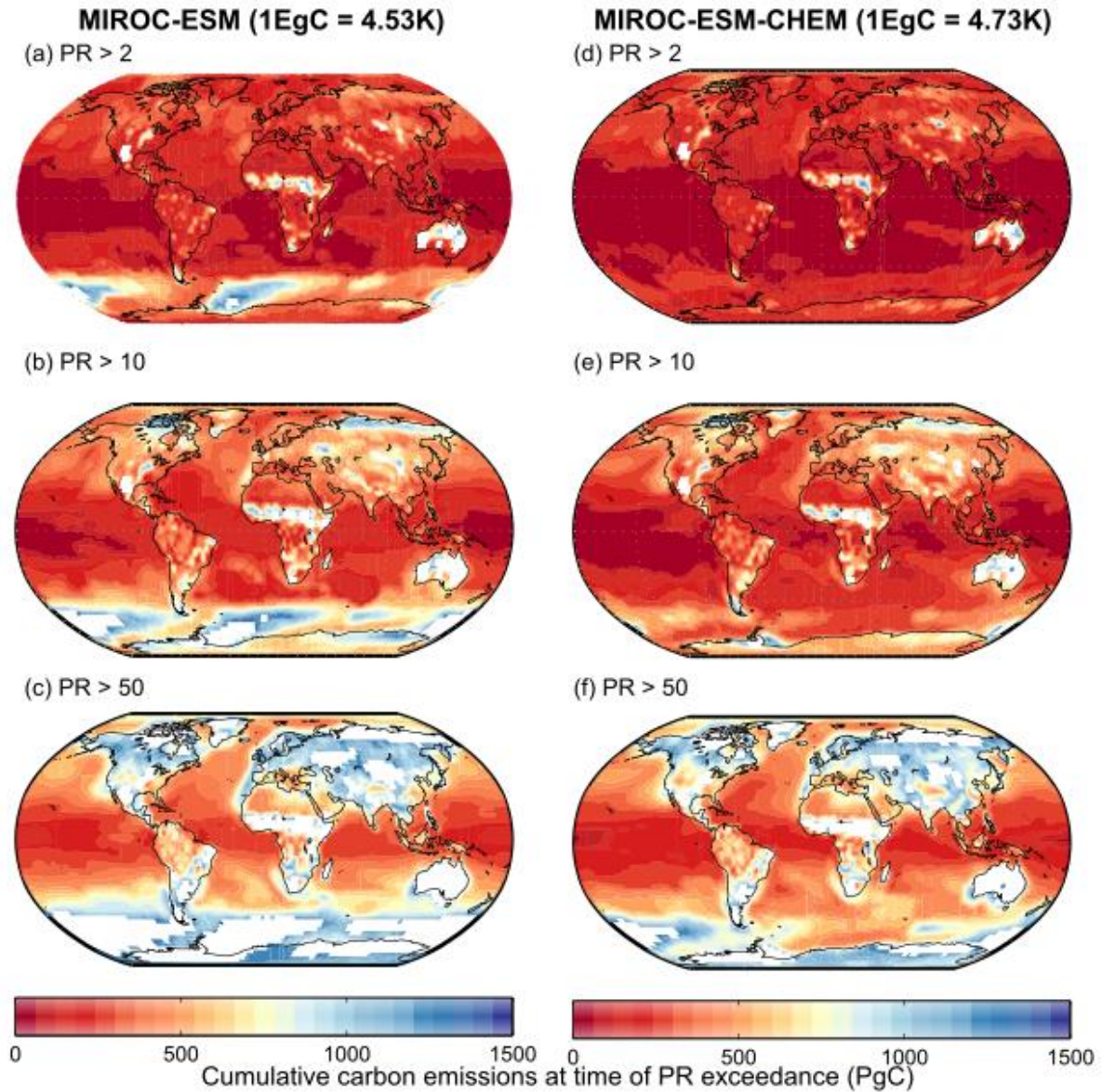
Figure 4.12: Same as Figure 4.11, but for (left) MIROC-ESM and (right) MIROC-ESM-CHEM.

4.4.2 Sensitivity to the present-day and future areal distribution of population quintiles

It is clear from Figure 4.3a that the wealthiest 20% of the global population is dispersed over a land area larger than the area covered by the poorest socio-economic quintile (by a factor of approximately seven). To understand the extent to which this difference in areal distribution influences the results presented in Section 4.3, we reproduce the results for Figure 4.7, but this time resampling 1000 times, at random, an aggregated area equal to the poorest quintile from all grid points which contribute towards the wealthiest socio-economic quintile. Figure 4.13 shows that the extra cumulative emissions required to reach 50% exposure for the

wealthiest population is mostly similar when considering the full quintile (coloured circles) and the median result of the equal-area resampling approach of grid cells encompassing the wealthiest quintile (black diamonds) for each model. The corresponding error lines represent the 90% confidence range of this resampling procedure for each model. This demonstrates that, even when considering regions of equal land area, an equivalent fraction of the poorest populations continues to experience earlier emergence of all three PR thresholds than their wealthier counterparts.

The patterns of when the global population will experience the emergence of hot days with unprecedented frequency under future warming scenarios will also be sensitive to how the spatial distribution of the wealthy and poorer sectors of the global community will change with time. While it is impossible to predict the future evolution of global population and GDP-PPP, we can make use of the population scenarios provided for Shared Socioeconomic Pathway 2 (SSP2), which contributed towards the Integrated Assessment Model used for to develop the RCP8.5 scenario (see http://bit.ly/1LfFLY9 for more information). Repeating once again the results of Figure 4.7, but this time using projected population and GDP data for the years 2050 and 2090, the squares and triangles (respectively) in Figure 4.13 reveals very little difference in the rapidity of emergence of the poorest quintile relative to the wealthiest quintile. In fact, the model spread and magnitude of extra emissions for 50% population exposure to each PR threshold is nearly identical between the future population scenarios and the default 2010 data, indicating that the emergence of high-temperature extremes continues to occur after much fewer CO_2 emissions for poorer nations when compared with the wealthiest nations, even when considering future socio-economic scenarios.
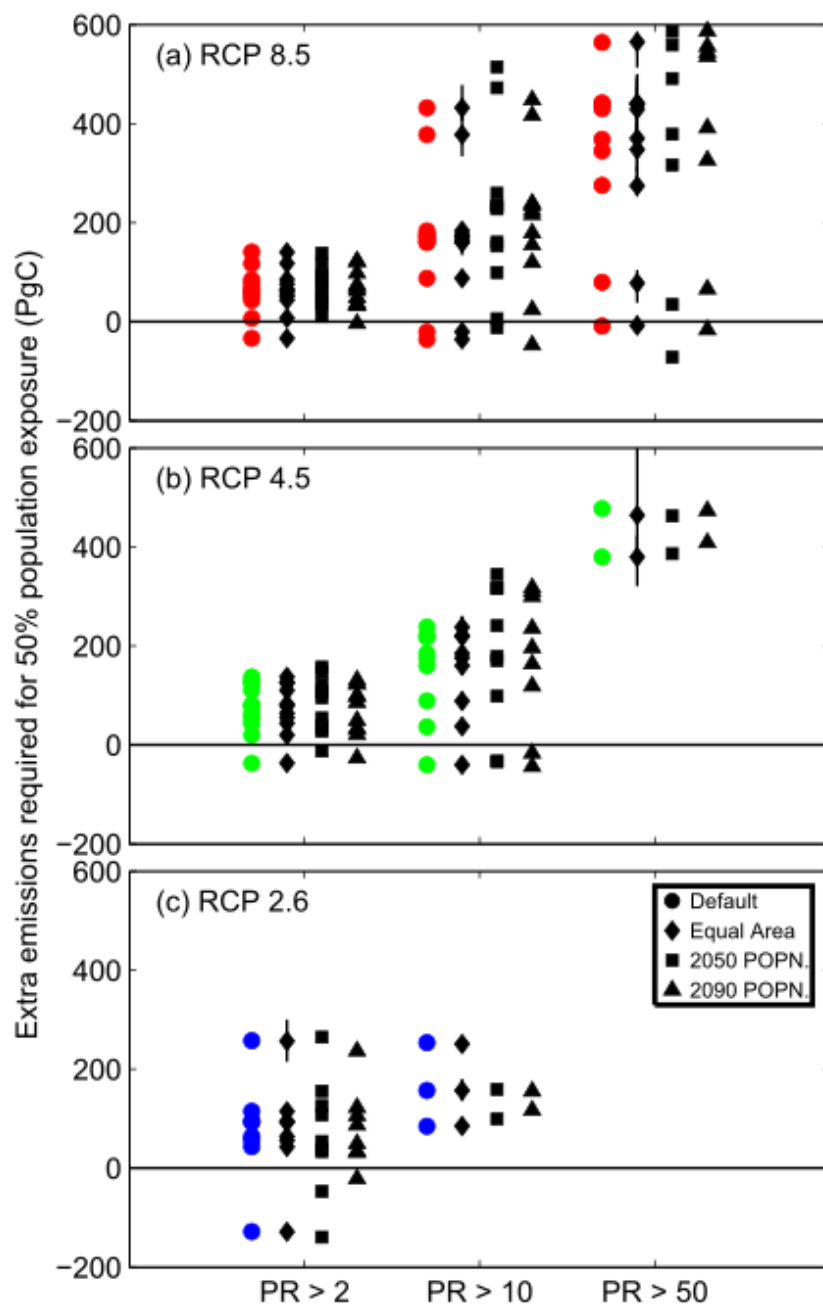
Figure 4.13: Same as Figure 4.7, but comparing the values of differential population exposure using default population and socio-economic scenarios (coloured circles) with alternative scenarios: black diamonds correspond to using equal areas for the poor and wealthy sub-populations; black squares (triangles) correspond to using projected population and GDP data for the year 2050 (2090).

4.4.3 Sensitivity to 'end-point' effects

When interpreting when the emergence of a given climate signal, the potential influence of 'end-of-simulation' effects must be carefully considered (Hawkins et al. 2014) in order to prevent overconfident statements about when a signal has emerged beyond a given threshold

of internal variability (Mora et al. 2013). The methodology presented in this analysis specifically defines the 'emergence' of a given PR threshold as only corresponding to the central year of the 30-year period when a PR threshold is exceeded, and *continues* to be exceeded for all subsequent 30-year periods in the available time series (the final 30-year period corresponding to 2071-2100) – we do not make any inferences about a 'permanent' exceedance of any PR threshold since we have no information available to consider beyond 2100.

To test whether the specific end-point of 2100 has any major influence on the population exposure CDF's shown in Figure 4.4, we repeat the analysis, but this time artificially restricting the time period considered for each simulation such that the final 30-year period is limited to 2051-2080. As shown in Figure 4.14, there are no significant differences in the median model CEPRE response when compared to Figure 4.4. However, it is important to note that some of the model population exposure CDFs do differ slightly from Figure 4.4 (aside from reaching lower cumulative emissions totals) for the PR>50 thresholds, where the emergence over some grid points appear to be sensitive to the length of period considered. Nevertheless, we conclude that the key results presented in the main analysis are largely insensitive to potential 'end-point' influences.
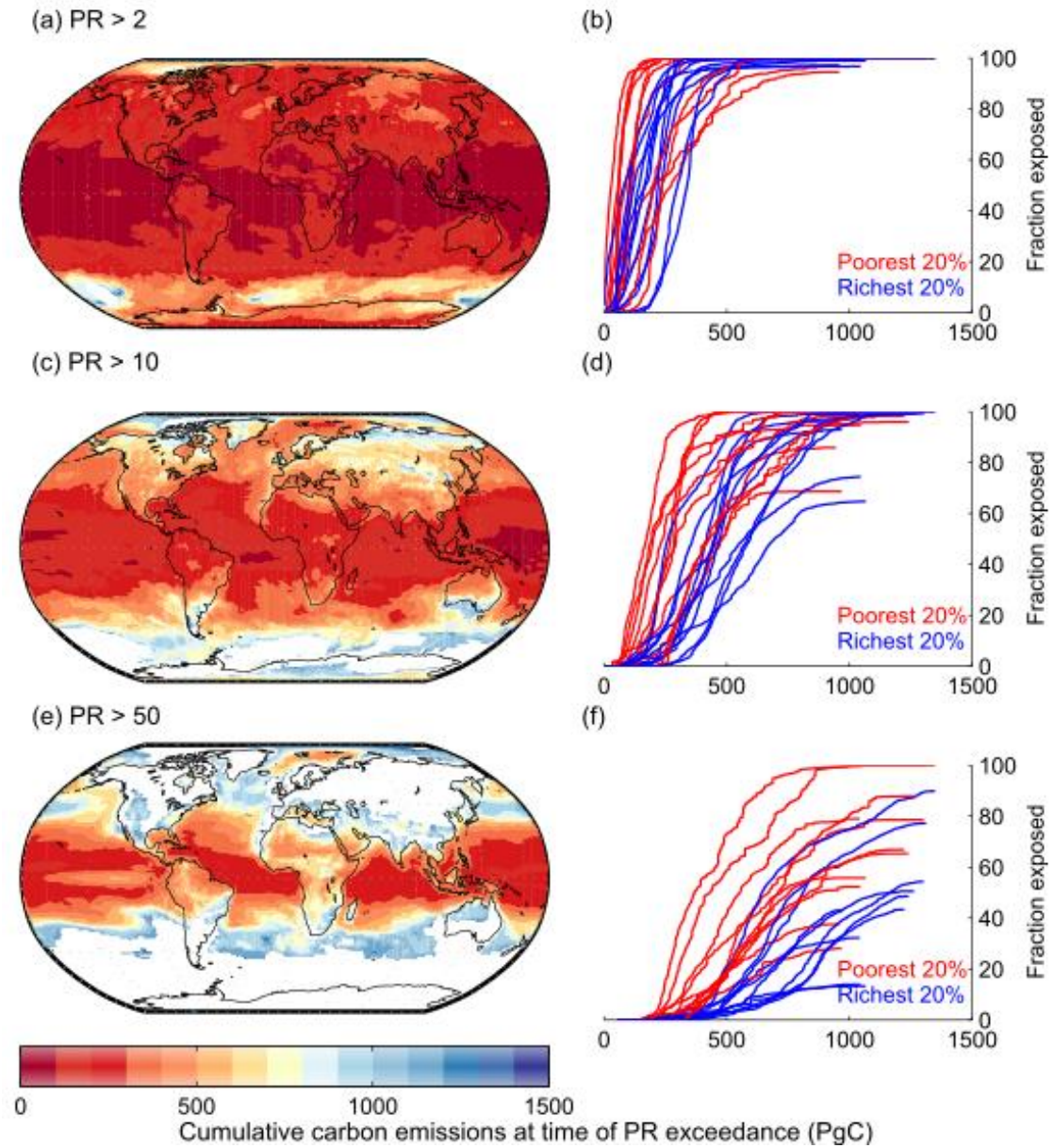
Figure 4.14: Same as Figure 4.4, but with PR exceedances only considered out to 2080.

4.5 Discussion

In this study, we have for the first time simulated the direct influence of cumulative carbon emissions on the time of emergence of daily extreme temperatures, by considering the CEPRE metric at specific PR thresholds. In considering only the persistent emergence of given PR thresholds in the context of fractional population exposure, our results are less sensitive to internal variability, despite the smaller spatial scales considered. Moreover, previous research has found the emergence of more frequent heat extremes averaged over a

longer time period (such as 5-day or 30-day anomalies) occurs even earlier when compared with 1-day extremes (Fischer and Knutti 2015).

By assessing emergent probability ratio increases as a function of cumulative CO_2 emissions, this approach serves well to evaluate the relative changes in heat extremes between different regions in the world, as well as across scenarios, and could therefore be of use to those working in vulnerability, impacts and adaptation, and to integrated assessment modellers, with the caveat that the continuing exceedance of specific PR thresholds could be interpreted as a proxy for heat-related damages (Dunne et al. 2013; Burke et al. 2015; Hansen and Sato 2016).

Even if emissions are towards the low end of the range considered by the RCP scenarios, this analysis shows that the pattern of changes in frequency of daily temperature extremes remains robust: daily temperature extremes emerge to more frequently affect the poorest 20% of the global population, when compared with the wealthiest 20% of the global population for all RCP scenarios, for a range of probability ratios. This result can be explained primarily due to the fact that the poorest people in the world densely populate lower latitude regions, where the low variability in temperature enhances the pace of emergence of a given signal-to-noise ratio when compared with higher latitude regions (Hawkins and Sutton 2009; Diffenbaugh and Scherer 2011; Mahlstein et al. 2011; Hawkins et al. 2014; Hansen and Sato 2016).

4.6 Summary

As global cumulative carbon dioxide emissions continue to increase, the fractional gap in population exposure between the poorer and wealthier members of society will only widen for exponentially higher PR thresholds (Figure 4.8). While all populated regions around the globe will enter a new regime of temperature extremes with no observed historical precedence if cumulative emissions continue to increase at current rates, the impacts, in terms of frequency of heat extremes, will become significantly worse for poorer nations when compared with their wealthier counterparts. We therefore argue that, even though our results show the emergence of more severe temperature extremes will always occur for poorer populations first, the potential prevention of crossing extreme PR thresholds means that the poorest members of the global community will always be the greatest beneficiaries of action towards a low-carbon pathway.

Chapter 5

On the implications of spatial consistency in the future emergence of regional climate signals

Throughout Chapter 4, we have developed extensive arguments to illustrate the differences in population exposure to emergent increases in extreme hot days with continuing climate change. An increasing body of evidence also suggests there are substantive disparities between those countries which are expected to experience the most rapid emergence of climate change, and those countries which are responsible for the majority of cumulative CO_2 emissions to date. In this study, we demonstrate how patterns of spatial heterogeneity in the emergence of climate change are not only robust between models, but insensitive to future scenario uncertainty. As a result, we specify which nations are already experiencing climate change emergence significantly faster (or slower) than the global population as a whole. We subsequently define a new joint proxy for national circumstance, called the emergence-emissions index, which captures patterns in relative contributions towards anthropogenic climate change, coupled with the relative severity of expected changes to regional climate. The potential implications of this emergent regularity in the climate change problem, as well as the limitations of our approach, are also discussed.

5.1 Introduction

Spatial heterogeneity in the physical response to anthropogenic climate change is becoming increasingly recognised as an emerging issue of climate equity and justice (Althor et al. 2016; Davis and Diffenbaugh 2016; Green 2016). However, identifying a readily quantifiable metric of climate change which sufficiently captures the spatial complexity of the emergent impacts of climate change, as well as enabling an assessment of changes under future scenarios, remains difficult. Althor et al (2016) presented, for individual nations, a comparison between greenhouse gas emissions and measures of present-day vulnerability to climate change, using DARA's Climate Vulnerability Monitor (CVM). While representing an aggregate measure of 22 metrics of environmental vulnerability, this CVM metric is inherently limited by its (1) complexity, (2) qualitative framework, and (3) inability to be calculated in the context of future climate scenarios. In this study, we present evidence to

support the use of emergent temperature signal-to-noise ratios as an alternative estimate of climate vulnerability, specifically in the context of considering spatial inequity in future climate change.

Recent research has increasingly focused on the presentation of anthropogenic climate change in the context of local variability. Previous studies have shown, with similar spatial patterns of results, when significant increases in the likelihood of exceptional or unprecedented heat occur, relative to some benchmark distribution (such as the 20th century): examples have been demonstrated across annual (Diffenbaugh and Scherer 2011; Mahlstein et al. 2011; Lehner and Stocker 2015; Hawkins and Sutton 2012), seasonal (Anderson 2011, 2012; Diffenbaugh and Scherer 2011; Mahlstein et al. 2011; Hansen and Sato 2016; Mueller et al. 2016), monthly (Sippel et al. 2015; Coumou and Robinson 2013) and daily (Fischer et al. 2013, 2014; Fischer and Knutti 2015; Harrington et al. 2016a; King et al. 2015a; Angélil et al. 2014, 2016) timescales, as well as for a variety of heatwave metrics (for example, Russo *et al* 2016). These studies share the characteristic of framing climatic emergence in the context of pre-existing local variability.

We choose to focus on the emergence of signal-to-noise ratios in temperature in this study. Some slow-emerging impacts, like changes to ecosystem zones (Mahlstein et al. 2013) and more-frequent precipitation extremes (King et al. 2015a), will not be well captured with a focus on temperature S/N ratios. Nevertheless, this choice is supported by previous results which highlight the links between increasing heat extremes and reduced crop yields (Lobell and Burke 2008; Battisti and Naylor 2009; Asseng et al. 2015a; Lobell et al. 2011; Liu et al. 2016), as well as impacts on ectotherms (Deutsch et al. 2008), and the similar spatial patterns evident when comparing the aforementioned emergence studies with more comprehensive climate vulnerability indices (Althor et al. 2016).

5.2 Data and methods

For this analysis, we reproduce the method of quantifying signal-to-noise ratios in interannual temperatures as defined in Hawkins & Sutton (2012). We consider only those 25 models from the Coupled Model Intercomparison Project Phase 5 (CMIP5, Taylor et al 2012) which have monthly surface air temperature data available for the 'Historical', 'RCP2.6', 'RCP4.5' and 'RCP8.5' experiments. Only a single ensemble member ('r1i1p1') from each model is considered. We concatenate 'Historical' simulations for the period 1901–2005 with

corresponding representative concentration pathways (RCPs) for the period 2006–2098 (some models only provided data out to 2099).

To compute the climate change signal for this model-based analysis, temperature increases for each model at each grid point are regressed against corresponding increases in global mean temperature for simulations across the twentieth and twenty-first century, and then normalized with respect to a 1986-2005 climatology (see Hawkins and Sutton (2012) for further details). The measure of noise is defined as the interannual standard deviation of annual means in temperature across the pre-industrial control simulations of each climate model. For simplicity we hereafter present, for each model, estimates for the signal-to-noise ratios in temperature (hereafter S/N) averaged over the period 2069-98, under each RCP scenario. All models have signal-to-noise ratios calculated on their native spatial grids and then interpolated to a common 2.5° x 2.5° spatial resolution.

5.3 Results

5.3.1 Robust, scenario-independent, spatial heterogeneity in the patterns of emergent climate change

Figure 5.1 presents the multi-model median end-of-century S/N ratios at each grid point, for the 3 RCP scenarios considered in this analysis. It is immediately evident that the patterns of signal emergence are spatially consistent across all three RCP scenarios. Figure 5.1d demonstrates the remarkably robust pattern correlations between each RCP scenario, with the multi-model median exhibiting a Pearson's coefficient of determination (R^2) value greater than 0.95 for all comparisons. While pattern scaling relationships have been thoroughly investigated for absolute temperature changes under future warming scenarios (for example, Tebaldi and Arblaster (2014)), this consistency in the spatial patterns of signal emergence have not been previously identified in such stark terms.
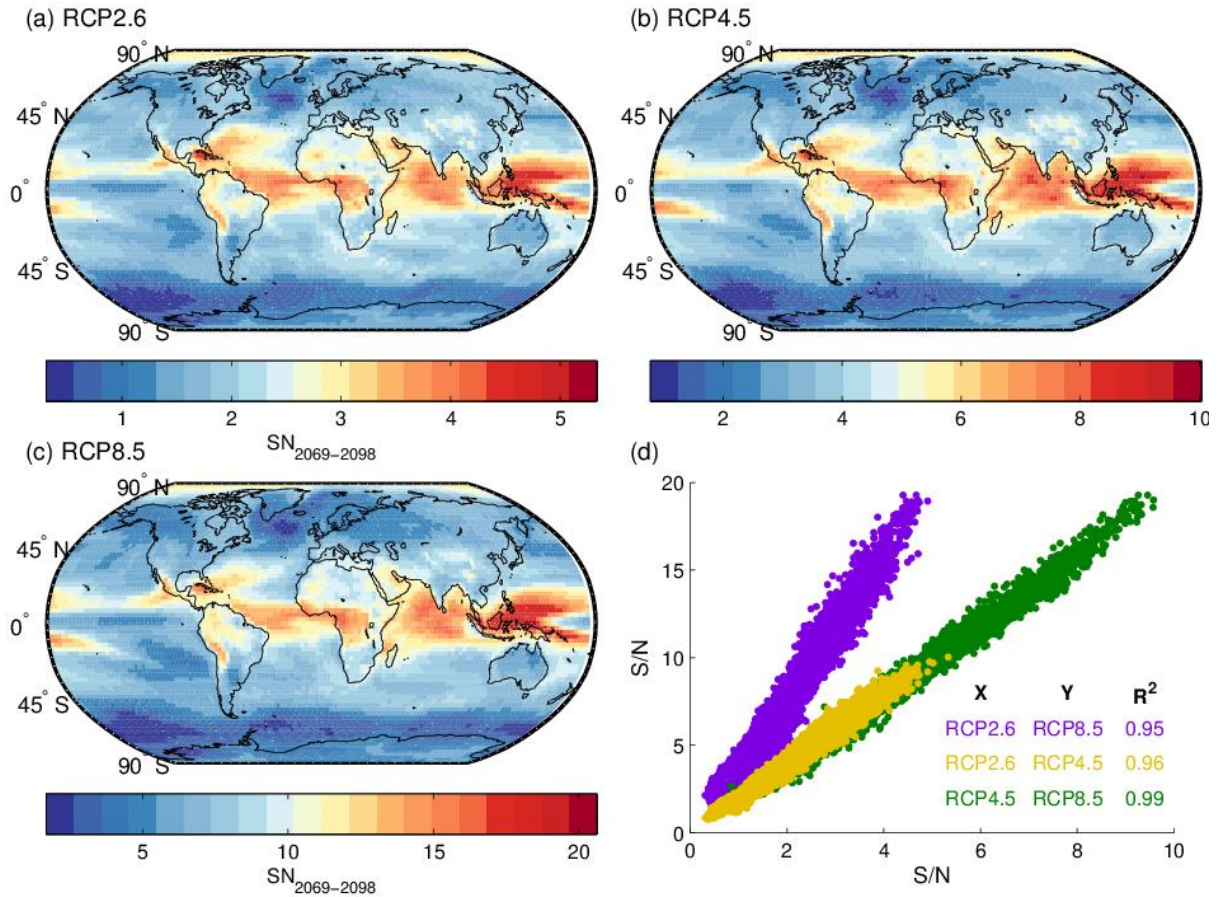
Figure 5.1: Global patterns of multi-model median signal-to-noise ratios, relative to 1986-2005, expected by 2069-2098, under an **(a)** RCP2.6; **(b)** RCP4.5 and **(c)** RCP8.5 scenario. Panel **(d)** shows the pattern correlation scores found when making grid-scale comparisons of the expected S/N ratios between each of the RCP scenarios.

The general patterns of lower latitudes and ocean regions exhibiting larger emergent signal-to-noise ratios than higher latitudes and land regions, respectively, is entirely consistent with previous results (Mahlstein et al. 2011; Diffenbaugh and Scherer 2011; Hawkins and Sutton 2012; King et al. 2015a; Fischer and Knutti 2015; Harrington et al. 2016a). Interestingly, when all models are considered in this aggregated median pattern response, there is remarkable regularity in the magnitude of difference between different scenarios: end-of-century signal-to-noise ratios are twice as severe under RCP4.5 compared with RCP2.6, and twice as high under RCP8.5 when compared with RCP4.5. The authors do however emphasise that this doubling relationship is unlikely to be robust when considering individual models, or if mid-century signal-to-noise ratios were instead presented.

In light of this robust and scenario-independent pattern scaling in temperature signal-to-noise ratios, we next consider how individual countries experience emergent temperature change when compared with the global population as a whole.

5.3.2 Country-level emergence relative to the global population median

5.3.2.1 Aggregations according to income level

We now consider how these robust spatial patterns of S/N ratios translate to the exposure of different members of the global population, disaggregated into individual countries (see Appendix for details of data sources and processing techniques for socio-economic data). For each model in each RCP scenario, the S/N ratio experienced by the median global citizen is quantified (using gridded population data for the year 2015 at 0.25° x 0.25° resolution). The population of all countries in each income grouping (as specified by The World Bank) are then aggregated together to produce a distribution of corresponding S/N ratios. This distribution of S/N values are then normalized with respect to the median value of the global population, under that scenario and for that specific model. This is repeated for all models in a given RCP scenario - the distributions of normalized S/N for each model are then aggregated together according to these same population groupings and the S/N values corresponding to 10%, 50% and 90% cumulative population exposure are identified. This is then repeated for each RCP scenario.

When aggregated according to income level, Figure 5.2 shows there are robust patterns in exposure to different signal-to-noise ratios for the different groups of countries: those nations grouped as low income are more likely to be exposed to S/N ratios higher than the global population as a whole. Similarly, high income countries are exposed to relatively smaller changes, in terms of signal-to-noise ratios, by the end of the century. In fact, the median normalized S/N ratio demonstrates the remarkable linearity when considering each income group, with an additional $\approx +0.12 \pm 0.01\ \sigma\,\sigma^{-1}$ in relative emergence found when moving from high income countries to successively lower income groupings. It is also important to highlight the significantly higher upper bounds in normalized S/N evident in the lower-middle and low income groups (1.7-1.8), when compared with the upper-middle and high-income groups (1.2-1.4).

Most importantly, these results demonstrate that these patterns of normalized signal emergence are extremely consistent between the three RCP scenarios, even when considering 25 different model realizations. This means patterns of spatial heterogeneity in temperature signal emergence are insensitive to scenario uncertainty, and will therefore also apply to patterns of signal-to-noise ratios in the present day.
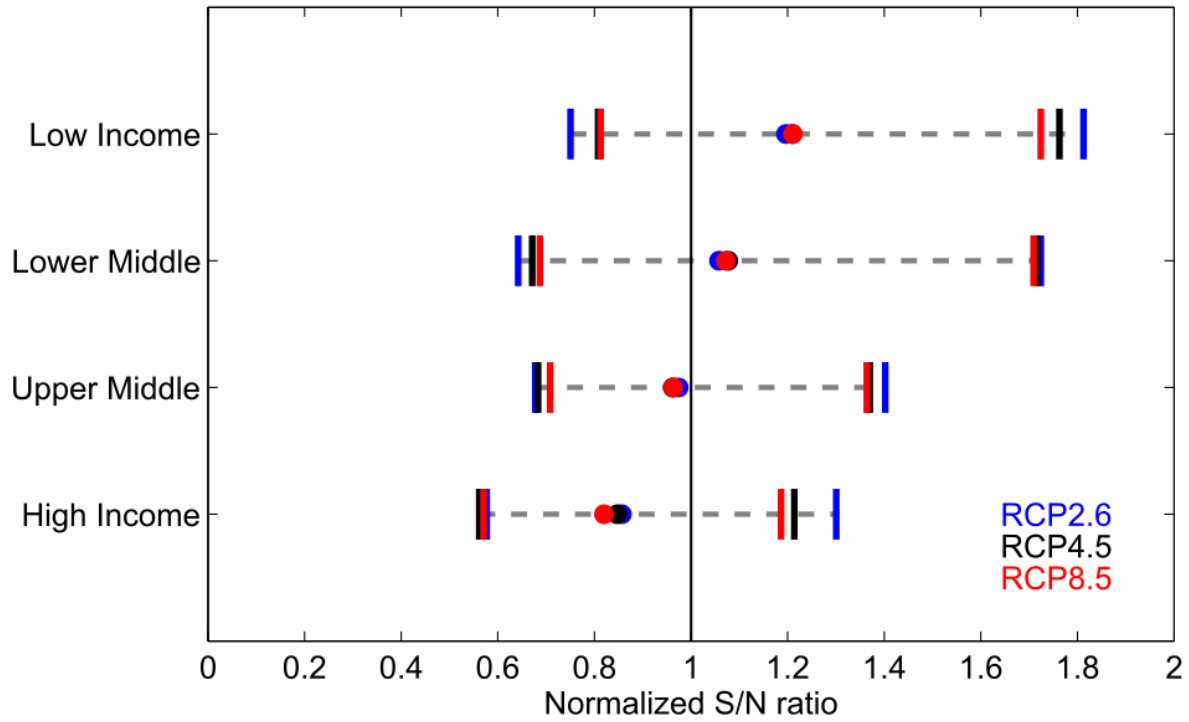
Figure 5.2: Patterns of normalized signal-to-noise ratios in temperature by the end of the century. S/N ratios are considered for each model, for populations in countries aggregated according to income level, and normalized with respect to the S/N ratio for the median global citizen in that same model. The blue, black and red circles represent the median normalized S/N ratio for each income group for all models under an RCP2.6, RCP4.5 and RCP8.5 scenario respectively. The lower and upper bars represent $10^{th} - 90^{th}$ percent confidence intervals.

5.3.2.2 Aggregations according to individual countries

Since these spatial patterns of uneven signal emergence are robust irrespective of scenario uncertainty, those individual countries which experience faster-than-average or slower-than-average emergent temperatures in the future can be assumed to also exhibit these properties in the present day. To identify these countries, we apply the criteria that 90% of the population of a given nation are expected to experience an emergent climate signal higher (or lower) than the median citizen of the global population.

Consistent with Figure 5.2, of the 19 countries which very likely lower-than-average S/N ratios, 15 are high income countries. Meanwhile, over half of the 38 countries which experience faster-than-average S/N ratios are classified as low or lower-middle income countries. Interestingly, the combined number of people which experience exceptionally fast or slow emergence, when compared with the collective global population, are similar, with approximately 500 million people included in each category. All remaining countries experience signal-to-noise ratios which are consistent with the global median, when model uncertainty in the physical response of the climate system is taken into consideration.

Table 5.1: Countries which will very likely (>90% likelihood) experience emergent signal-to-noise ratios in temperature either lower (left column) or higher (right column) than the global median citizen.

Countries with normalized S/N ratio very likely below 1 (n = 509 million)	Countries with normalized S/N ratio very likely above 1 (n = 506 million)	
Austria[H]	Angola[UM]	Burundi[L]
Belgium[H]	Benin[L]	Bahrain[H]
Belarus[UM]	Bahamas[H]	Barbados[H]
Canada[H]	Central African Republic[L]	Côte d'Ivoire[LM]
Czech Republic[H]	Congo[LM]	Comoros[L]
Denmark[H]	Cuba[UM]	Dominican Republic[UM]
Estonia[H]	Gabon[UM]	Gambia[L]
Finland[H]	Ghana[LM]	Guinea[L]
Germany[H]	Guinea-Bissau[L]	Equatorial Guinea[UM]
Iceland[H]	Haiti[L]	Indonesia[LM]
Kazakhstan[UM]	Jamaica[UM]	Liberia[L]
Lithuania[H]	Martinique[H]	Oman[H]
Latvia[H]	Papua New Guinea[LM]	Qatar[H]
Mongolia[LM]	Rwanda[L]	Senegal[L]
Netherlands[H]	Sierra Leone[L]	Singapore[H]
Norway[H]	Solomon Islands[LM]	Somalia[L]
Poland[H]	Sri Lanka[UM]	Suriname[UM]
Paraguay[UM]	Togo[L]	Trinidad and Tobago[H]
United Kingdom[H]	Western Samoa[LM]	United Arab Emirates[H]

These results, in terms of income inequality in exposure to rapidly emerging climate change, are consistent with previous studies (eg. Harrington et al. 2016). They are an important extension of work on pattern scalings in absolute temperatures (such as by Tebaldi et al 2014), and help to demonstrate that differences between countries in the expected emergence of severe climate change are not sensitive to different possible emission scenarios, and are therefore also valid for the present day.

5.4 Towards a joint metric of national circumstance: defining the emergence-emissions (EE) index

Given this newly-recognised consistency in the spatial patterns of emergent climate change, we now consider how the normalized S/N ratios for each country compare with that country's contribution towards anthropogenic climate change. While previous studies have highlighted the importance of many different radiative forcing mechanisms (Ward and Mahowald 2014; Fuglestvedt and Kallbekken 2016) when quantifying country-level contributions towards observed increases in global mean temperature (Matthews 2014; 2016), we hereafter consider

cumulative CO_2 emissions as a simple and versatile proxy of an individual nation's contribution towards collective anthropogenic warming.

Figure 5.3a highlights that it is mostly high income countries which represent a disproportionately high fraction of cumulative CO_2 emissions globally, while Figure 5.3b reveals that those same high income countries are experiencing, on average, proportionately less severe climate emergence. The converse of these results are also true for low income countries, as illustrated in Figure 5.2, and also in previous studies (Mahlstein et al. 2011; Hansen and Sato 2016).
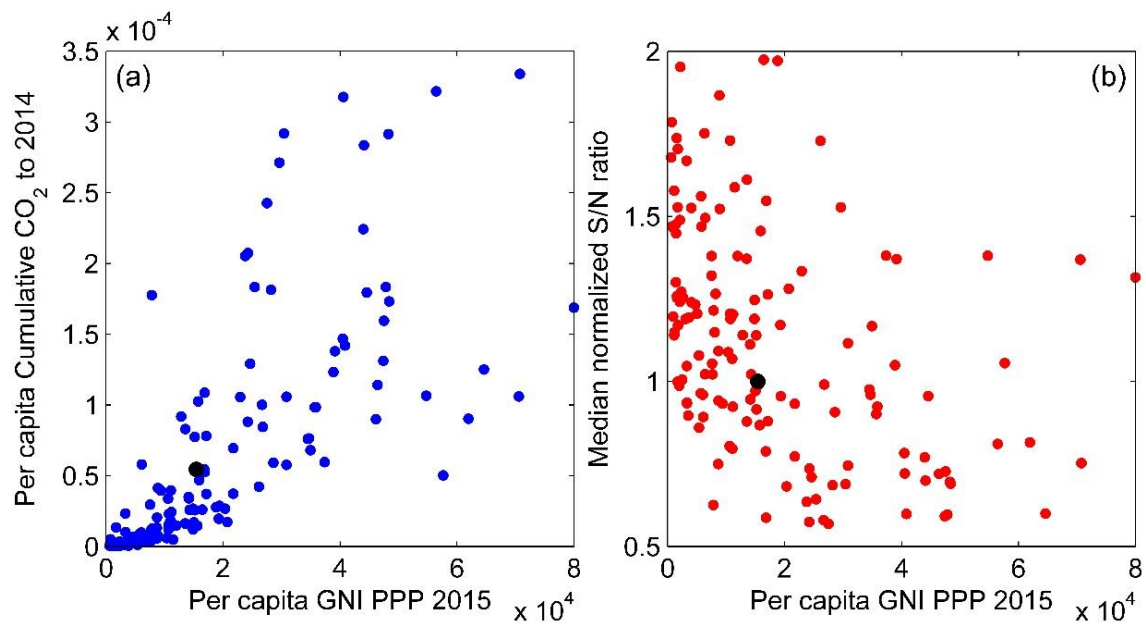
Figure 5.3: Scatter graphs showing individual country comparisons between per capita Gross National Income in 2015, adjusted for purchasing power parity (GNI-PPP), with **(a)** cumulative CO_2 emissions to 2014 per person; and **(b)** the multi-model median S/N ratio expected by 2069-98, normalized with respect to the global median citizen and aggregated across all RCP scenarios. Blue and red circles show 170 individual countries assessed in this analysis; black circle shows the estimated global mean.

5.4.1 Defining the EE index

To highlight these amplified patterns of inequality in the physical patterns of climate change, and the contributions to the problem, we propose a combined metric of national circumstance, called the emissions-emergence metric. Specifically, to quantify whether a country's fractional contribution to global emissions is *consistent* with their expected relative severity of climate change impacts, we define the emergence-emissions (EE) index, for a given country, or grouping of countries, *i*, as follows:

$$EE_i = \left[\frac{C_i}{C_G} \times \frac{P_G}{P_i}\right] \div \frac{(S/N)_i}{(S/N)_G} \qquad \text{Eq. 5.1}$$

where C_G, P_G and $(S/N)_G$ denote the carbon emissions, population and median signal-to-noise ratio associated with the entire global population. The term C can represent total CO_2 emissions over any period of time (a benefit of considering CO_2 only), as long as it is self-consistent for both C_i and C_G. An EE value above (below) unity means that the fractional contribution of a country or group of countries to global emissions is more (less) than what is consistent with the severity of climate change they will subsequently experience.

The part of the equation in square brackets is equivalent to previous proposals (Gignac and Matthews 2015; Matthews 2016) to quantify historical carbon debts and credits: in essence, how much more or less have a country contributed to global emissions compared with an equal-per-capita share. The additional consideration presented here is the spatial heterogeneity in the physical manifestation of climate change, expressed in terms of emerging signal-to-noise ratios in annual mean temperature.

5.4.2 Country-level EE estimates using historical emissions

To demonstrate the utility of the EE index, results are presented for 170 countries in Figure 5.4, using shares of historical CO_2 emission fractions as an illustrative example, including both annual emissions in 2014 and cumulative emissions from 1751-2014. Consistent with the patterns highlighted in Figure 5.3, there are robust linear relationships between EE and income per capita (when presented using a log-log scale) with R>0.87 for not only historical cumulative emissions, but also annual emissions in 2014. This linearity with income is expected, as the EE index effectively captures the coupled disparity of low (high) income nations experiencing faster (slower) than average emergence whilst making lower (higher) than average contributions to global emissions.

As presented in the right-hand panels of cumulative population exposure, historical EE values are weighted to more countries being below and severely below the line, than above it. This is a critical point, and represents a clear demonstration of the disparity in those countries which have contributed to the problem and those which are experiencing the impacts. In the context of considering income as a predictive measure of historical EE, a low income country is not just experiencing more rapid climate change emergence than it has contributed towards, but

this relationship scales non-linearly (for example, the global median income per capita corresponds to an EE index well below 1).

There are also notable differences between historical EE indices when comparing 2014 annual emissions with cumulative emissions since 1751: most importantly, the population distribution mapped to EE is skewed even further towards lower values when considering cumulative emissions. Nearly 20% of the global population has contributed less than 1% towards cumulative emissions, relative to what is consistent with their normalized S/N ratios, while 79% of the global population has an EE<1 when considering cumulative emissions. These statistics are slightly improved when evaluating annual emissions for 2014: there exists a less severe gradient in the EE-income relationship, indicating less severe disparities in emissions in the context of emergent signal-to-noise ratios.
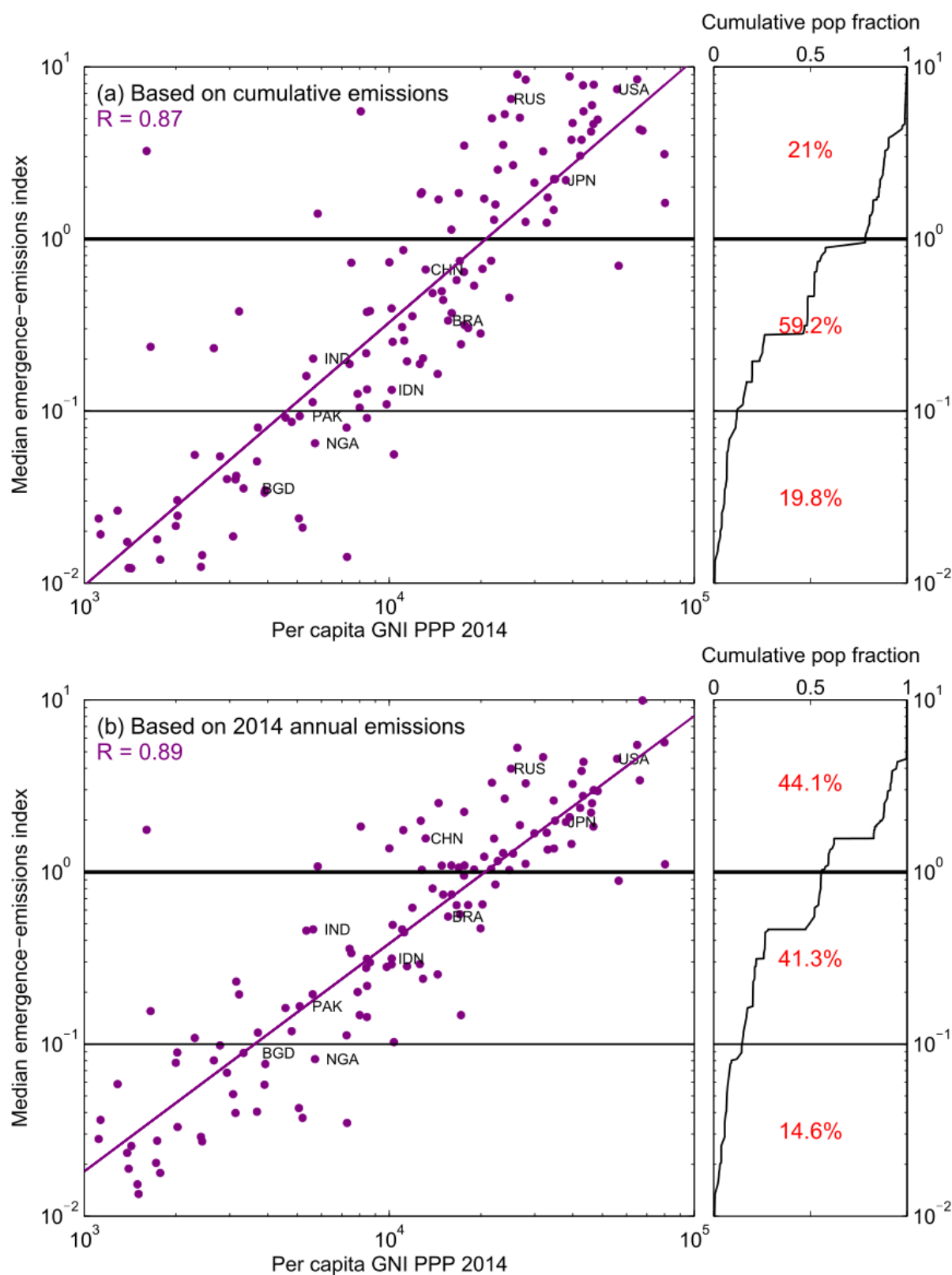
Figure 5.4: Scatter graphs comparing Gross National Income per capita for the year 2014, adjusted for purchasing power parity (GNI PPP), with the emissions-emergence index for the median citizen of 170 individual countries. EE index results are shown using **(a)** cumulative CO_2 emissions up to 2014; and **(b)** annual CO_2 emissions for the year 2014. Right-hand panels illustrate the cumulative fraction of the global population compared against each countries' EE index. R-value is based on a least-squares linear regression. Labels are shown for the 10 largest countries, by population, in 2014.

5.5 Discussion

In the context of RCP scenarios in the CMIP5 framework, the collective success or failure of future climate mitigation efforts are expressed as scenario uncertainty (Hawkins and Sutton 2009). Our analysis has demonstrated that the patterns of spatial heterogeneity in relative climate change emergence are, to first order, insensitive to such scenario uncertainty. This section briefly presents some potential implications of this emergent regularity in the patterns of future climate change.

5.5.1 Informing region- and country-specific progress towards 'equitable' carbon budgets

In light of the strong linear correlation between cumulative CO_2 emissions, and subsequent global warming, there has been invigorated focus on the concept of a carbon budget in recent years. Gignac and Matthews (2015) presents the consideration of splitting up future carbon budgets as a spectrum of possibilities, with the two end-points being an equal per-capita fraction and using present-day fractions of emissions (otherwise known as a grandfathering approach). Raupach *et al* (2014) quantifies this concept by introducing a sharing index, w, where $w=0$ implies splitting a budget according to grandfathering while $w=1$ represents an equal per-capita distribution of a future carbon budget. In this context, we suggest that the robust model uncertainty associated with the physical patterns of future climate change emergence could be used to provide a reasonable bridge between per-capita-based budgets and the grandfathering approach. Or more specifically, the model uncertainty in EE can be used to give a range of values of w for a given country that is consistent with that country's expected future exposure to emergent climate change.
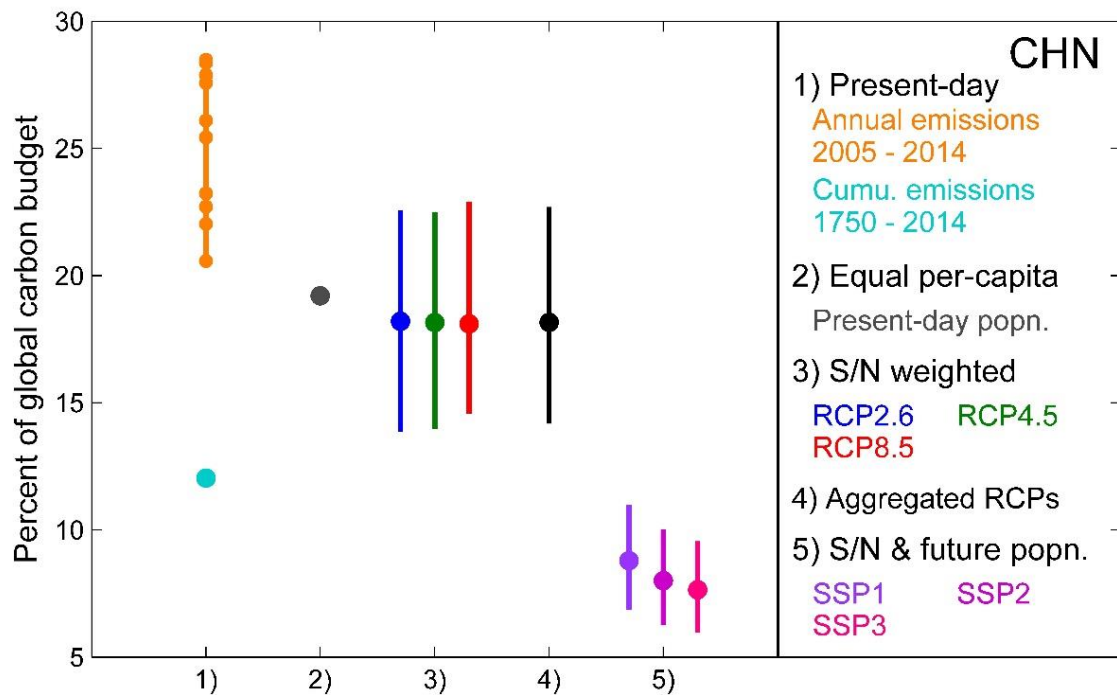
Figure 5.5: Fractional carbon budgets for China according to multiple criteria: 1) Historical emission fractions for each year between 2005 and 2014 (orange) as well as using cumulative emissions since 1751 (light blue); 2) fraction of global carbon budget consistent with an equal per-capita share; 3) per-capita fraction adjusted to normalized S/N ratios for each RCP scenario, with 10%-90% confidence range; 4) same as for 3 but aggregating normalized S/N ratios from all RCP scenarios together; 5) using the 10th-90th percentile range of per-capita budgets from 4, but adjusting to three future population (SSP) scenarios.

An interpretation could be considered as follows: within the context of uncertainty associated with (1) model projections of future climate and (2) different population scenarios is considered, the normalized S/N ratios of each country can be used to inform what equitable fractions of a global carbon budget would look like. Figure 5.5 presents such an example for China: orange circles show the range of 2005-2014 annual emissions, as a fraction of the globe (w=0); per-capita fractions using the present-day population (w=1) are shown in grey. The black markers present what fractional budgets are deemed consistent with future S/N ratios, while the purple and pink bars show that projected reductions in population (relative to the globe as a whole) means that China may be anticipated to represent an even smaller fraction of global emissions in the future. Table 5.2 presents similar comparisons for three regional groupings of countries: these result show that per-capita emission fractions, when weighted to account for patterns of normalized S/N, are actually higher for Least Developed Countries, when compared with both OECD countries (as of 1990) and ASEAN countries.

Table 5.1: Percentage of global carbon emissions for three regional groupings, based on (left) cumulative emissions up to 2014, and considering an EE of 1 (middle), with associated model uncertainty (right).

	Cumulative CO_2 1751-2014	**Per capita emissions consistent with climate emergence**	10th-90th percentile range
OECD90	54.03	**10.18**	7.03 – 14.34
ASEAN	2.11	**12.00**	6.86 – 19.42
LDC's	0.36	**14.03**	8.64 – 20.52

5.5.2 Using the EE index to inform national contributions to a Global Green fund

An additional benefit of the EE index being defined using CO_2 emissions only, is that it could be calculated for cumulative emissions since 1751, for a specific range of years (between 1990 and the present-day, for example), or only for a single year. As a corollary, the EE index could also be calculated for each country for each individual year in the future, based on that nation's fraction of global CO_2 emissions for that year, as well as their fraction of the global population at that time (thereby taking into account changes in population dynamics).

The EE index for each country for that year could be subsequently used to calibrate how much each individual country might contribute to some international climate finance mechanism, such as the UN-administered Green Climate Fund (GCF). For example, those countries which emit a lot in a given year while experiencing minimal impacts (high EE) would pay more than their per-capita share. The uncertainty associated with the model simulations of S/N ratios enables a specification of minimum and maximum limits that countries might be expected to pay, with each government making their own decision about specifically how much, thereby allowing them to maintain some element of year-to-year payment flexibility. One benefit of this approach is that it incentivises national mitigation actions, particularly for high-emitting nations: those countries which contribute less to global emissions for a given year would be expected to contribute proportionally less to the GCF. This suggestion is but one way of using the EE index, which considers both the inputs into the climate system in the form of emissions, and the outputs in the form of changes to local climate.

5.5.3 Limitations

There are several limitations to only focussing on historical CO_2 emissions in the definition of the EE index. Fuglestvedt and Kallbekken (2016) correctly point out that if all external forcing mechanisms are considered, then those countries which are contributing more or less towards some measure of collective anthropogenic warming (Otto et al. 2015b) may vary (due to aerosol emissions, for example). It is for this reason that any focus on 'historical responsibility' of individual nations (Matthews 2016) has been avoided, beyond demonstrating the utility of the emergence-emissions index in Figure 5.4. Instead, the robust spatial disparities in emerging climate change can be used to quantify appropriate contributions of individual nations towards *future* efforts to ameliorate the worst impacts of climate change. Furthermore, the suggestion of financial contributions to the Green Climate Fund is only practicable with the consideration of cumulative CO_2 fractions, which has the unique feature of being comparable across numerous timescales. Finally, we argue that any issues in presenting EE with respect to only cumulative CO_2 emissions, rather than some more complex measure of net anthropogenic forcings (Ward and Mahowald 2014), are circumvented by only presenting a $10^{th} - 90^{th}$ per cent confidence range, associated with model uncertainty in the normalized S/N estimates.

The suggestion presented in Section 5.5.2 is a novel approach to link the uneven physical manifestation of the climate response with CO_2 emissions, when considering how to fund projects related to climate change mitigation, adaptation, or even loss and damages. However, we explicitly avoid any consideration as to how that money should be allocated to individual countries: such an issue is difficult to address at present, and may or may not require the probabilistic attribution of human influence to specific damaging climate-related events (Thompson and Otto 2015). Such discussions of how to effectively distribute any collective finance mechanism are beyond the scope of this study.

Third, the ideas suggested in this study do not address how *absolute* changes to global and regional emissions could be prioritised, nor what are considered equitable mitigation strategies. There are many different approaches to the concept of equity in the context of climate mitigation that have not been discussed here (du Pont et al. 2017), each of which prioritise different burden-sharing principles. The EE index merely represents a simple extension of the 'equal per capita' approach (Meinshausen et al. 2015; du Pont et al. 2017), by further incorporating a measure of uneven climate impacts.

Finally, it is acknowledged that metrics of normalized temperature emergence will not completely capture the full range of future climate change impacts expected for some regional communities around the world (low-lying coastal populations, for example). Future research is therefore required to develop a more holistic representation of climate impacts which could be simplified within the EE framework. Further discussion in this context is provided in Section 7.1.1.

5.6 Summary and implications

This is the first time a joint metric has been presented which captures the emergent regularity of climate change, both in terms of (1) contributions to anthropogenic warming, and the (2) expected physical manifestation of changes in local climate in response to future warming. The first term can, for the most part, be interpreted as a function of cumulative CO_2 emissions; the second term can be most effectively captured using emergent signal-to-noise ratios, which strongly aligns with corresponding changes to extreme temperatures on daily (Harrington et al 2016; Fischer and Knutti 2015), seasonal (Diffenbaugh & Scherer 2011; Mahlstein 2011; Anderson 2011), and annual (Anderson 2011; Diffenbaugh & Scherer 2011) temperature extremes, as well as for heatwave events (Russo et al 2016), and future crop yield declines (Battisti et al 2009; Liu et al 2016; Lobell and Burke 2011).

Most importantly, the analysis presented in Figure 5.2 reveals that the expected rates of climate emergence for individual countries (relative to the global population as a whole) remains fixed, irrespective of possible future emission pathways. Therefore, the only way for a country to be able to modify their EE index is on the 'emissions' side; this enables the EE index to become an effective metric to capture the success of national mitigation pathways for high-emitting nations. For example, if the payment to the GCF was proportional to a country's EE index for a specific year, there would be quantifiable and substantive reward incentives for efforts towards local mitigation, in terms of paying less into this fund.

The EE index is defined as a simple metric with built-in flexibility. It can be aggregated across different countries with shared interests, and can also be calculated across any timescale, be that for one year, ten years or all years back to 1750: this is the primary benefit of considering cumulative CO_2 emissions only. Uncertainty in the EE index may also be seen as a way of helping to facilitate the transition from a 'grandfathered' (present-day) share of emissions to an equal per-capita fraction (Raupach et al. 2014).

Irrespective of the severity of the climate experienced in the future, which is a function of potential mitigation measures that will be taken by different nations, this analysis reveals a robust understanding of how different countries will experience relative changes in temperature S/N ratios, when compared to the global population as a whole.

Chapter 6

Reconciling the temporal emergence of signal-to-noise ratios with the probability ratio metric for event attribution studies

When considering a policy-relevant and simple metric for emergent climate change, Chapter 5 uses annual-scale signal-to-noise ratios in temperature, while Chapter 4 focuses on the probability ratio metric for a pre-specified return period. This chapter explores the statistical relationships relevant to reconciling (1) the attribution of specific extreme events, with (2) the emergence of warming-driven signals from internal variability within a given climate distribution. Evidence is presented to show how the threshold-dependence of maximum probability ratios can lead to potential ambiguity for decision makers, before revealing how signal-to-noise ratios can provide a reasonable estimate of changing heat extremes with continued warming. As found in Chapters 4 and 5, the (relative) spatial patterns of change for different regions of the world are expectedly very similar between these two approaches.

6.1 Threshold-dependant variations in probability ratios

In probabilistic event attribution, both a doubling and a ten-fold increase in the likelihood of event occurrence (attributable to human influences) have been suggested as potentially important thresholds to reach before litigation efforts could be hypothetically considered for climate change-related damages (Adam 2011; Thornton and Covington 2016; FEL Otto, pers. comms. 2016). However, answering the question '*when will an X-fold increase in probability occur for a 1-in-n year extreme event*?' is inherently sensitive to the return period (n) of interest.
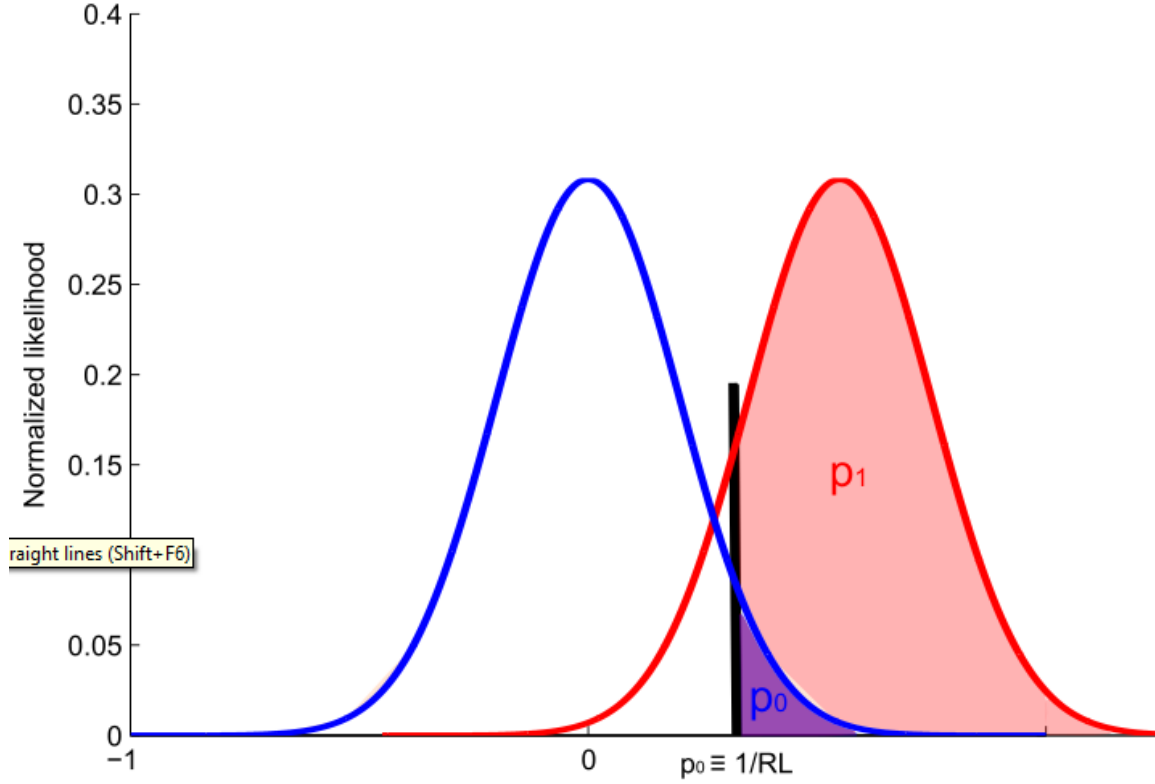
Figure 6.1: Schematic illustrating differences in the probability of exceeding an event in a counterfactual distribution (blue), compared with a distribution (red) which has warmed in the recent climate.

Consider, for example, an n-year return period with respect to a pre-industrial (or counterfactual) temperature distribution – an illustrative example of n=10 is presented in Figure 6.1. This means the probability of exceeding such an event, or p_0, is by definition:

$$p_0 = \frac{1}{n}$$

with the probability of exceedance in the factual distribution (red shading in Figure 6.1), or p_1, being a function of the distributional shift as a whole. The corresponding probability ratio, PR, is defined as:

$$PR = \frac{p_1}{p_0}$$

To consider what constitutes 'emergence into an unknown climate' in this context, we propose a definition as the point when the factual distribution has moved completely to the right of the event threshold, such that every observed temperature in the unprecedented climate is hotter than the event threshold of interest. In this circumstance, the maximum possible 'factual' probability is 1,

$$p_{1,MAX} \equiv 1$$

$$\therefore PR_{MAX} = \frac{p_{1,MAX}}{p_0}$$

$$= \frac{1}{1/n}$$

$$\therefore PR_{MAX} = n \hspace{4cm} \text{Eq. 6.1}$$

Therefore, the maximum possible probability ratio that can be obtained for a given event threshold is equal to the return period of that event. In the case of the examples demonstrated in Chapter 4, the maximum possible PR that could have been witnessed for the 1-in-1000 day extreme temperatures is 1000.

While this sounds like an altogether trivial demonstration, there are clear implications for the context of pursuing compensation for extreme event-related damages: for any assessment of changes in probability ratio through time, the subsequent answer will be highly sensitive to the event threshold chosen (by contrast, no such problem exists when evaluating the time of emergence of distributional signal-to-noise ratios, as with Chapter 5). However, uncertainty in quantifying the return period of any observed extreme event is large, and especially so for the most rare (and therefore most damaging) events. For example, estimates of the return period for the damaging UK floods in January 2014 were placed between 35 and 350 years (Schaller et al. 2016). Therefore in the context of assessing potential liability from damages associated with this event, the model-based estimate of probability ratios would need to be greater than two (or ten) across this *full range of possible return periods*: this is much less likely to occur (even with for the same underlying shift in the factual distribution) as a result of this threshold-dependence in PR_{MAX}.

To further explore this concept in a theoretical framework, Figure 6.2a presents how probability ratios proliferate as a function of the underlying signal-to-noise ratio for a theoretical Gaussian distribution (further details are provided in the Appendices related to Chapter 3). When evaluating these absolute probability ratios for a given shift in the underlying distribution, it is clear that there are dramatic differences in the PR that might emerge for a given shift in the distribution, depending on the event threshold considered. Consistent with the theoretical considerations presented above, as well as with previously published results (Fischer and Knutti 2015), it is clear that those very extreme return periods converge to higher probability ratios much more quickly when compared with more modest

event thresholds. However, if these probability ratios are in fact normalized according to PR_{MAX} for each individual return period, as shown in Figure 6.2b, then the changes in this normalized probability ratio (which in fact corresponds simply to p_1) are much more smooth and uniform as a function of event threshold. Most importantly, these values saturate beyond 0.99 for any S/N ratios greater than 6 (red dashed lines in Figure 6.2): this is a non-trivial result, and highlights that beyond some level of emergent climate change, the resultant probability ratios saturate to some fixed value as the new climate distribution eventually reaches a state of being entirely unrecognisable.
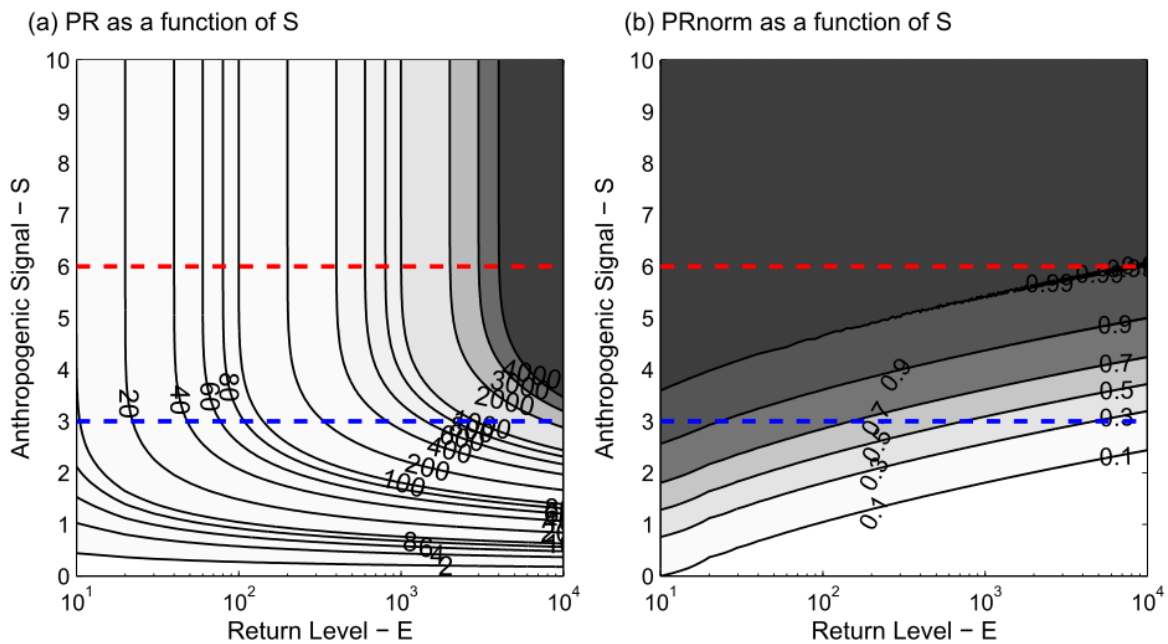
Figure 6.2: Contour plots, where shading denotes changes to **(a)** the probability ratio and **(b)** normalized probability ratio of witnessing an event equal to or exceeding an event threshold, E, when comparing a counterfactual distribution with a factual distribution which has experienced a shift by the anthropogenic signal-to-noise ratio, S. Here, both the counterfactual and factual distributions are assumed to be Gaussian. The blue dashed line denotes S=3, corresponding to when $PR_{NORM} = 0.5$ for a return period of 1000; the red dashed line denotes S=6, and represents the point when all return periods less than 10000 show $PR_{NORM} > 0.99$.

6.2 Using S/N ratios to inform threshold-specific probability ratios under future climates

To further consider the links between distributional signal-to-noise (S/N) ratios and probability ratios, particularly in the context of deducing when definitive emergence into new and unknown climates occur under future warming scenarios, end-of-century (2069-2098) S/N ratios are re-calculated following an identical methodology to that presented in Chapter 5. However, I now limit the models considered to those eleven ESMs which are used for the daily-scale analyses in Chapter 4. Alongside the signal-to-noise ratios focusing on annual

mean temperatures, corresponding S/N values are also presented for a daily timescale (by re-calculating the pre-industrial standard deviation evaluated at each grid cell using daily-resolution data).
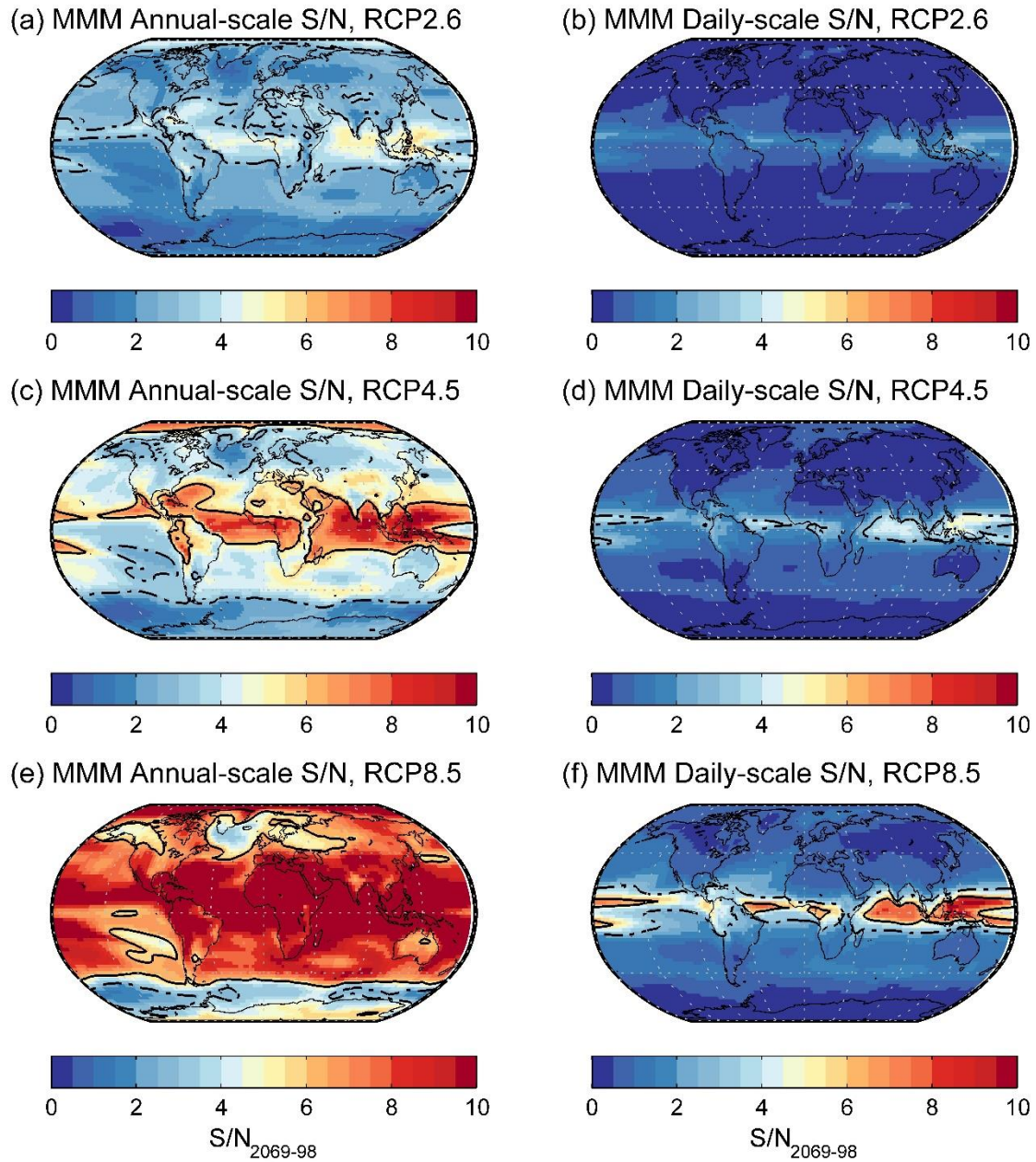
Figure 6.3: Global patterns of multi-model mean signal-to-noise ratios, relative to 1986-2005, expected by 2069-2098, under an RCP2.6 (top); RCP4.5 (middle) and RCP8.5 (bottom) scenario. Left-hand panels show this emergence for annual mean temperatures; right-hand panels consider daily temperatures. Dashed contour lines correspond to S/N = 3; solid contour lines correspond to S/N = 6.

Figure 6.3 shows multi-model mean maps of these end-of-century S/N ratios at annual (left) and daily (right) timescales for this 11-model ensemble, using common colour scales for all

three RCP scenarios. The dashed and solid black contour lines in each panel represent S/N ratios of 3 and 6, respectively.

When considering annual temperature emergence, it is found that S/N ratios greater than six emerge by the end of the century for most tropical latitudes under RCP4.5, and for nearly all locations under an RCP8.5 scenario (with the exception of Antarctica and the Southern Ocean). These results therefore suggest (subject to the assumption that annual mean temperatures approximate a near-Gaussian distribution) unprecedented emergence into entirely unknown climates will occur for these locations by the end of the century. However, no such exceedances of S/N = 6 occur anywhere in the world when considering a low-emissions RCP2.6 scenario; strong mitigation efforts can therefore prevent the most dramatic examples of future regional emergence.

These results also clearly demonstrate that the pace of signal emergence on annual timescales occurs several times more rapidly when compared with daily-scale emergence: in fact, annual-scale signal-to-noise ratios are much larger by the end of the century under RCP2.6 than the magnitude of daily-scale emergence under RCP4.5. Moreover, it appears that the rapidity of signal emergence becomes more concentrated over tropical locations for daily timescales, and particularly over oceanic regions – this is likely due to the fact that fluctuations in daily temperatures over these regions are inherently suppressed by (1) the heat capacity of the upper ocean preventing significant cooling of air temperatures, and (2) convective cooling mechanisms putting an upper limit on near-surface air temperatures (Clayson and Weitlich 2007; Laepple and Huybers 2014). Based on these results, it can be seen that exceedances of S/N =3 and S/N = 6 do also occur at daily timescale over tropical latitudes, particularly in the Western Pacific Warm Pool, under an RCP4.5 and RCP8.5 scenario respectively.

For illustrative purposes, and to understand what these numbers mean in more practical terms, I now transform these signal-to-noise ratios to probability ratios for a 1-in-1000 year (or day) event for each map, using the assumptions presented in Figure 6.2 (that these distributions are Gaussian and that changes in temperature variability through time are negligible). Figure 6.4 highlights the key issues discussed earlier in this section about the saturation of probability ratios at their maximum possible value: under an RCP8.5 scenario, it will be virtually certain that *all annual temperature anomalies* will be hotter than what would have been considered a 1-in-1000 year hot anomaly in 1986-2005; under RCP4.5, there will

also be at least 500-fold increases in witnessing a 1-in-1000 year event for most populated regions of the world between 60S and 60N. On daily timescales, there is saturation towards PR_{MAX} of 1000 for tropical latitudes under RCP8.5, while most other locations show a probability ratio of less than 50 when using this Gaussian transformation from model-based signal-to-noise ratios.

Figure 6.4: Global patterns of probability ratios for 1-in-1000 year (day) extreme temperature thresholds expected by 2069-2098, relative to 1986-2005, using a direct Gaussian transformation of the MMM signal-to-noise ratios in Figure 6.3. Left-hand panels show results for annual mean temperatures; right-hand panels

consider daily temperatures. Top panels show results under RCP2.6; middle panels under RCP4.5; bottom panels under RCP8.5.

6.3 Validity of a Gaussian transformation

There are clearly inherent limitations when applying a direct Gaussian transformation to model-based S/N data, in order to calculate changes in the likelihood of witnessing extreme events in the future (Sardeshmukh et al. 2015). Here, it is established whether or not this Gaussian transformation is in fact a fair representation of the daily temperature distributions simulated by the eleven climate models in question. Figure 6.5 compares the daily-scale probability ratios for a 1-in-1000 day temperature extreme using the Gaussian transformation (so the left-hand panels in Figure 6.5 are identical to the right-hand panels of Figure 6.4) with identical metrics, but this time *explicitly* calculated (for the same time periods) for each individual model using daily data and following the methodologies developed in Section 4.2. The right-hand panels of Figure 6.5 thus show the multi-model mean $PR_{99.9}$ at each grid point.
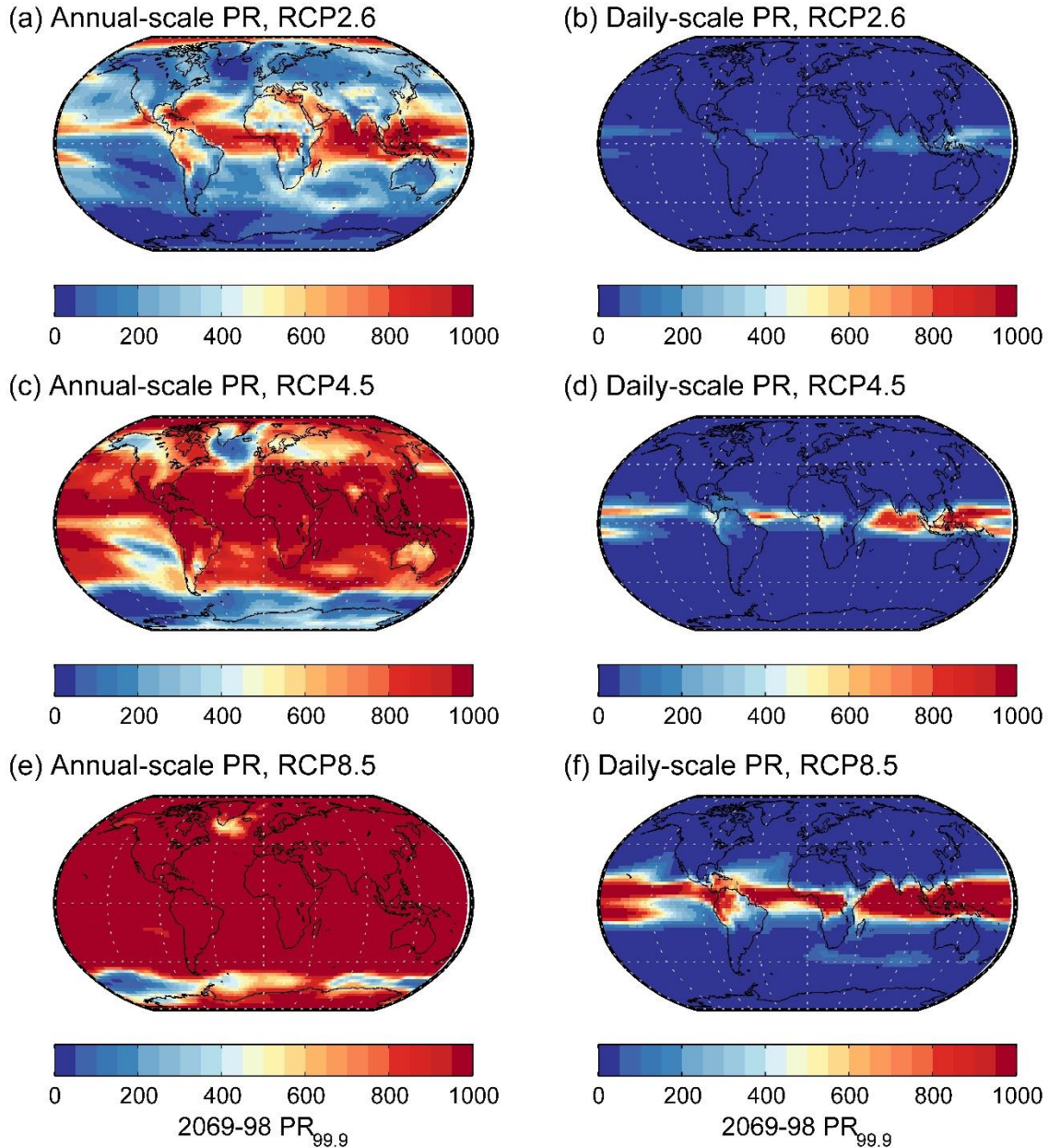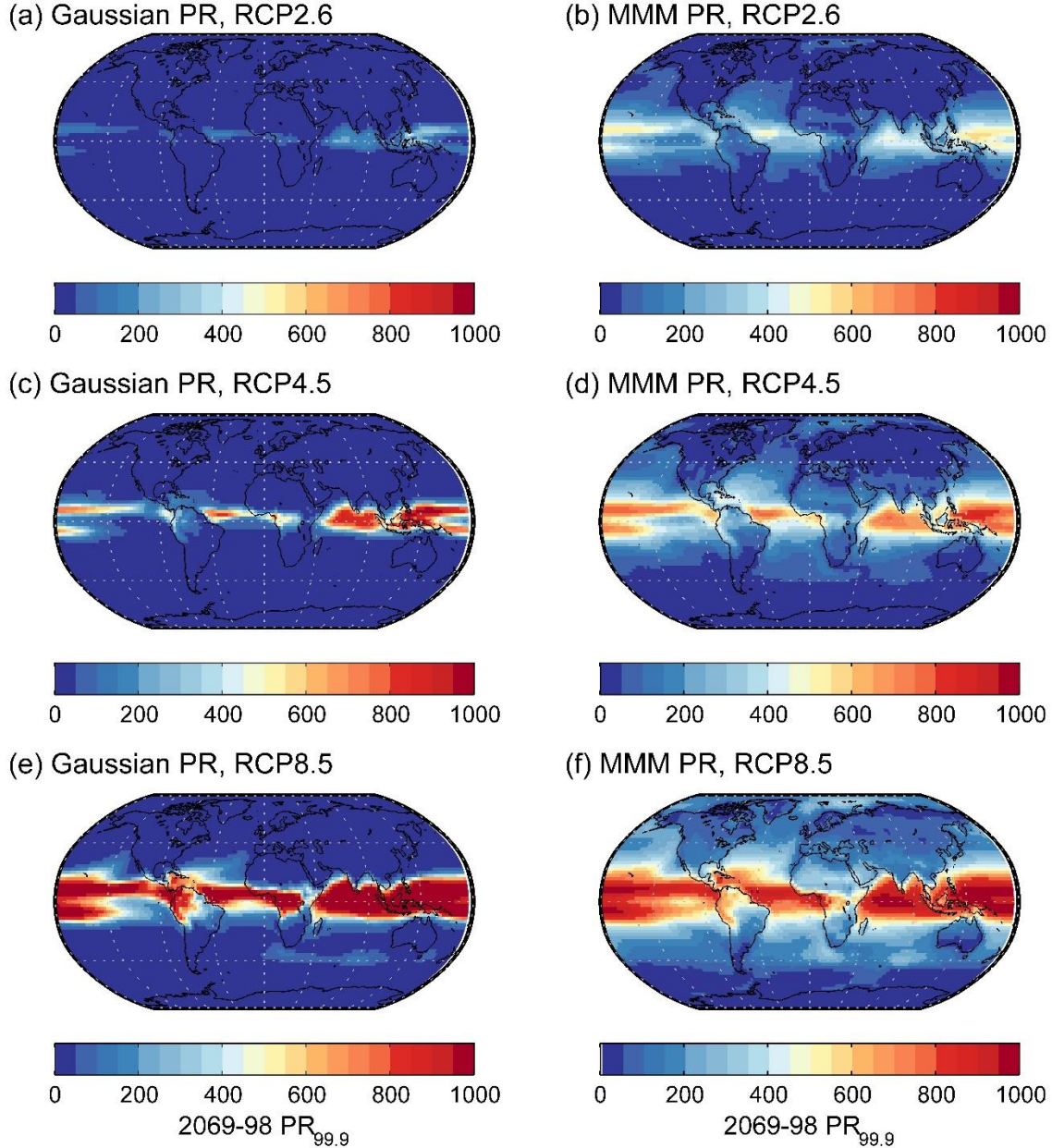
Figure 6.5: Left-hand panels show global patterns of probability ratios for 1-in-1000 day extreme temperatures expected by 2069-2098, relative to 1986-2005, using a direct Gaussian transformation of the MMM signal-to-noise ratios in Figure 6.3. Right-hand panels show results for the same periods, but using each individual model to explicitly calculated probability ratios from daily data, in accordance with the methodologies presented in Chapter 4, then calculating the multi-model mean at each grid point.

While the spatial patterns of PR$_{99.9}$ do mostly match between the theoretical and empirical results, a Gaussian approximation may slightly exaggerate the severity of the latitudinal gradient in PR that is witnessed in the actual model data. To evaluate these results in a more systematic manner, Figure 6.6 presents the statistical relationship between PR$_{99.9}$ and daily-scale S/N for the supposed Gaussian distribution (blue line) and for all grid points available across the results for 11 models under RCP8.5. There is some spread in the model-based

relationship between the two variables, but to first order, a Gaussian relationship appears to underestimate $PR_{99.9}$ for slower-emerging regions, yet represents a fair approximation for those fast-emerging regions which experience $PR_{99.9} > 500$ by the end of the century. This is consistent with other model-based studies which have considered whether Gaussian PDFs are an adequate characterisation of temperature distributions for the evaluation of changes to extremes (Loikith and Neelin 2015).
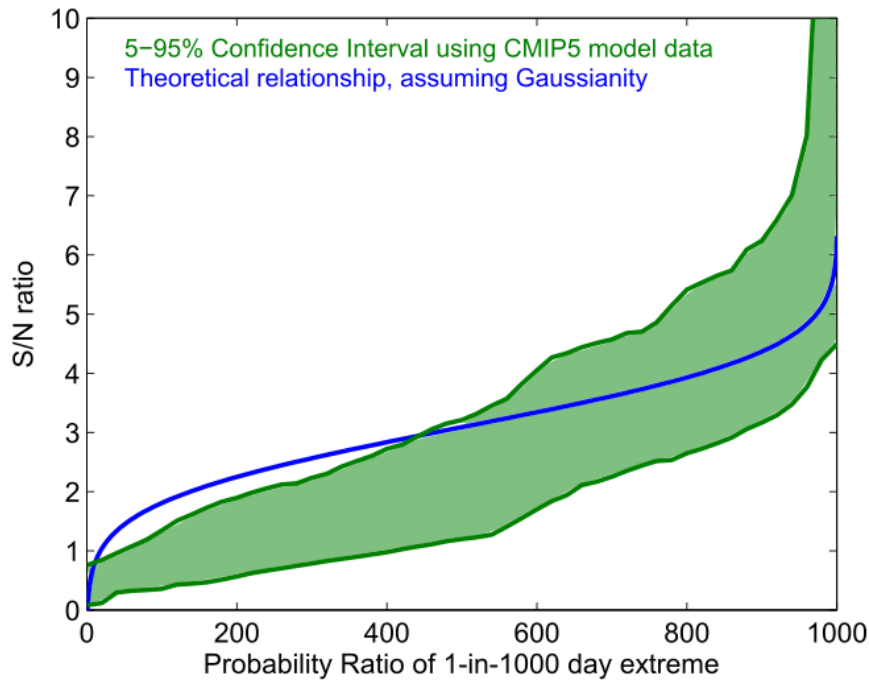
Figure 6.6: Scatter graph showing the relationship between a PR for a 1-in-1000 day temperature extreme and the underlying signal-to-noise ratio of the temperature distribution, comparing a theoretical Gaussian distribution (blue line) with model-based data from 11 CMIP5 models (green shading). All grid points from each model were used for end-of-century simulations under RCP8.5, and the 5% - 95% confidence interval of S/N was calculated for running values of PR.

6.4 Summary

This chapter illustrates the potential issues associated with maximum probability ratios varying as a function of event threshold. Simple Gaussian transformations of S/N ratios are then used to infer probability ratios for a variety of return periods. While the explicit calculation of changes to the likelihood of percentile exceedances (as with Chapter 4) remains preferable, these results show Gaussian transformations of signal-to-noise ratios are likely to be a reasonable surrogate for temperature extremes, especially for those fast-emerging locations which approach saturation at PR_{MAX}.

Chapter 7

Discussion

In this chapter, an evidence-based framework is proposed to understand how continuing cumulative CO_2 emissions will propagate to modify the likelihood of witnessing extreme events (of a specific severity) for different countries of the world. Discussions are then presented on (1) the strengths and weaknesses of considering only temperature as a proxy for other climate impacts which may emerge over the twenty-first century, and (2) the limitations of only utilising CMIP5 models to inform our estimates of future climate emergence.

7.1 A framework to interpret the role of cumulative carbon emissions in emergent changes to regional climate

Based on the sequence of analyses presented in Chapters 4, 5 and 6, an idealised framework can now be constructed, whereby the effect of continuing cumulative carbon dioxide emissions can be quantitatively traced towards the eventual change in extreme temperatures for specific countries or regions. Cumulative CO_2 emissions raise global average temperatures, which translates to an increase in the signal-to-noise ratio of annual temperatures for the median global citizen (Figure 7.1). Using the analytical framework presented in Chapter 5, this can then be translated to the regional patterns of temperature emergence for individual countries (or groups of countries). Corresponding changes to the likelihood of exceeding extreme temperature thresholds can then also be found, subject to the assumption of Gaussianity explained in Sections 6.2 and 6.3.
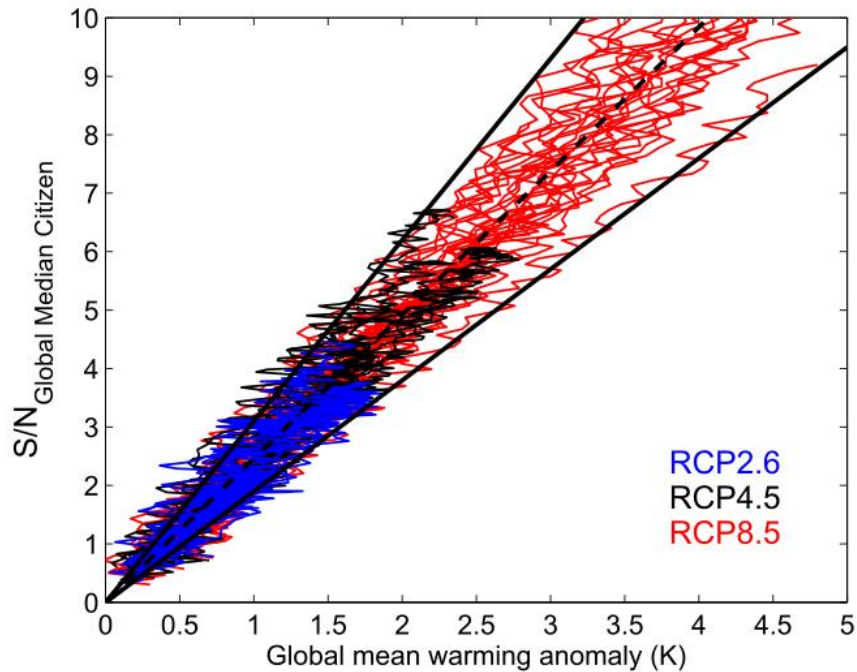
Figure 7.1: Relationship between increases in annual global mean temperature over the 21st century, and the corresponding signal-to-noise ratio of the median global citizen within each model. Anomalies are presented for each year in the period 2006-2098, with respect to the 1986-2005 mean, and shown for all models and all RCP scenarios used in Chapter 5. The dashed and solid lines denote best-guess and 5-95% confidence intervals of linear trends fitted to the data.

This framework is presented in the illustrative schematic shown in Figure 7.2. There is sufficient model-based uncertainty within each of these outlined steps that directly calculating the change in extreme temperatures for a given country per tonne of cumulative CO_2 would bring prohibitively wide uncertainty bounds (Schneider 2009). However, by isolating each process which links the steps 1 to 5 in the schematic, via the separate analyses within this thesis, robust uncertainty estimates can be provided for each individual section. Details related to the uncertainty bounds of each part of the schematic are as follows:

Cumulative CO$_2$ -> Global mean warming: While explicit calculations of the link between cumulative CO_2 emissions and global mean temperatures were presented in Figure 4.1, the uncertainty bounds of 1.3 – 3.0 K per TtC are in fact based on the results presented in the IPCC Summary for Policymakers (Figure 1.1). However, it is noted that the uncertainty range presented in Figure 4.1 fits within these upper and lower bounds. The best-guess estimate is based on the approach outlined in (Seneviratne et al. 2016), which yields a final answer of 2.35 K/TtC.

Global mean warming -> S/N ratio of median global citizen: Based on Figure 7.1, there also exists a robust, and scenario-independent, linear relationship between global mean

temperature increases and corresponding increases in the signal-to-noise ratio of annual temperatures for the median global citizen (following the methodology explained in Chapter 5). Using all 25 models that were used for the analysis in Chapter 5, a lower and upper uncertainty bound of 1.98 and 3.04 $\sigma.K^{-1}$, respectively, is found. The best-guess estimate is approximately $+2.45\sigma$ of warming for the median global citizen, per degree of global warming.

S/N ratio of median global citizen -> S/N ratios of specific country groupings: Presented in the schematic are the $10 - 90$ % uncertainty estimates of normalized S/N shown for all Low Income (Low Inc) and High Income (High Inc) countries, following Figure 5.2, with the best-guess answer shown as the median result for each grouping. Again, as emphasised in Chapter 5, this normalized S/N value can be calculated for the aggregated population of any individual country or group of countries with shared characteristics.

S/N ratios of specific country groupings -> Probability ratios of extreme hot years:

The calculations associated with this step have been explicitly outlined in Chapter 6, along with potential issues associated with the required Gaussian assumptions.
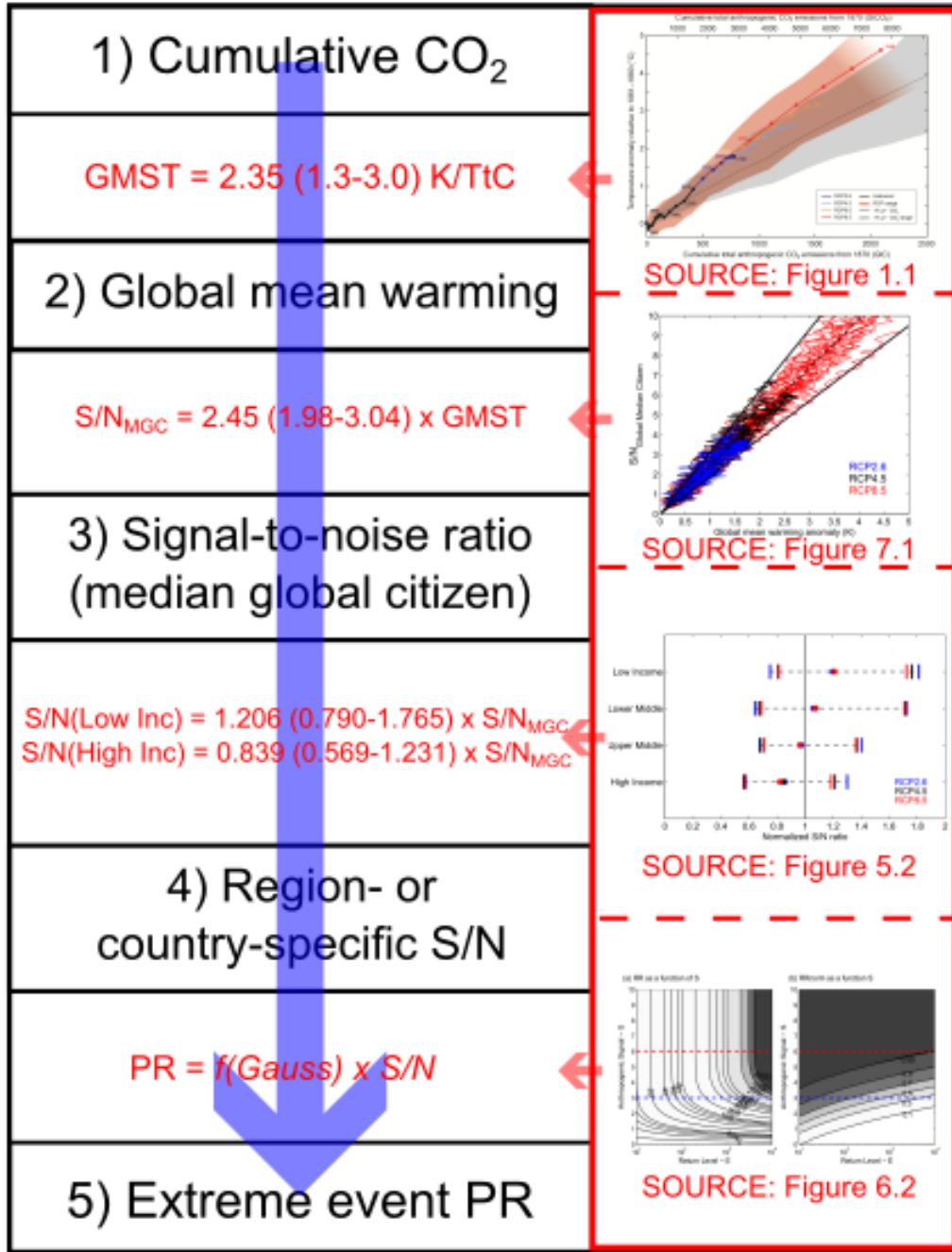
Figure 7.2: Schematic diagram demonstrating a complete framework to interpret how continuing carbon dioxide emissions translate to changes in regional temperature signal-to-noise ratios, and subsequent changes to the likelihood of witnessing extreme temperature exceedances. The left-hand panels denote the flowchart of constituent processes which link cumulative CO_2 emissions with changes to extreme temperatures; the red text between each numbered step provides quantitative estimates of the relationships between each variable. Where applicable, the best-guess estimate is listed with uncertainty bounds provided in brackets. The relevant sources for constraining these relationships are shown on the right-hand panels, and further details are provided in the main text.

While there are limitations associated with the simplicity of this framework (see Section 7.1.1), there nevertheless exists potentially important implications for interpreting the regional manifestation of future climate change. As an illustrative example, these results

reveal that an increase of +2°C in global mean temperatures will translate to a best-guess S/N ratio of +5.9σ (3.1-10.7) for low-income countries, and +4.1σ (2.3-7.4) for high-income countries. Since Section 6.1 suggests that exceeding a threshold of +6σ represents a shift into a completely unprecedented climate, a two-degree temperature target could indeed be an appropriate threshold for limiting 'dangerous' climate change for low-income countries. Equally, it can be seen that restricting global mean warming to 1.5°C instead of 2°C, as is suggested in the 2015 Paris Agreement (Rogelj and Knutti 2016), would prevent an additional +1σ - 1.5σ shift in the temperature distribution of the median global citizen, which could mark the difference between moderate and severe impacts for many regions around the world.

While exploring the full implications of this framework is beyond the scope of this thesis, future research in this area is planned.

7.1.1 Strengths and limitations of representing emergent climate change using a single, temperature-based metric

As highlighted in the introductory comments to Section 5.1, there are inherent limitations when restricting the focus of 'emergent climate change' to only consider the emergence of signal-to-noise ratios in annual-scale or daily-scale temperatures. Of course, the many and varied impacts of anthropogenic interference with the climate system cannot be simplified down to any singular metric. Some slow-emerging impacts, like changes to ecosystem zones (Mahlstein et al. 2013) or the tendency towards more-frequent precipitation extremes (King et al. 2015a; Min et al. 2011) will not be well represented, particularly since the spatial patterns which characterise such changes are fundamentally different to those which typify normalized temperature S/N ratios. Equally, those regions of the world which will experience enhanced drought severity in the future are not expected to follow a one-to-one mapping with corresponding changes to temperature or precipitation distributions. This is because those places with seasonally limited moisture availability, like the Mediterranean for example, will witness land-atmosphere feedbacks which will further exacerbate long-term drying trends (Greve et al. 2014; Allan 2014).

When attempting to distil the key patterns of emerging climate impacts to a (highly aggregated) level which will have some meaningful relevance to climate policy, an easy-to-understand, singular metric is clearly preferable (Ricke et al. 2016). So despite the obvious

limitations and caveats highlighted in the preceding paragraph, I have chosen to focus on the emergence of signal-to-noise ratios in temperature as representing such a metric in Chapter 5, for reasons explained hereafter.

As previously highlighted in Chapters 4 and 5, robust links exist between those locations which witness disproportionally faster increases in heat extremes, and those regions with substantive reductions in projected crop yield (Lobell and Burke 2008; Battisti and Naylor 2009; Asseng et al. 2015a; Lobell et al. 2011; Liu et al. 2016). Similar spatial patterns have also been found between normalized emergence metrics and the magnitude of warming-induced impacts on ectotherms (Deutsch et al. 2008; Beaumont et al. 2011), as well as when comparing temperature-based emergence metrics with more comprehensive climate vulnerability indices (Althor et al. 2016). Finally, I have explained in Section 6.1 why percentile-derived temperature extremes, including nearly every definition used to characterise heatwaves (Perkins and Alexander 2012; Russo et al. 2016), exhibit the same spatial patterns of normalized change as temperature S/N ratios under future warming scenarios (Perkins 2015). This could conceivably further extend the links between patterns of normalized S/N with heat-driven economic losses (Burke et al. 2015), labour capacity reductions (Dunne et al. 2013) and even heatwave-induced mortality increases attributable to climate change (Mitchell et al. 2016).

The fact that so many specific impacts follow patterns of emergent temperature change, including those which are influenced by multiple constituent physical processes (like crop yield) should not be considered surprising. Warming temperatures are the primary physical response to an enhanced greenhouse effect (Sutton et al. 2015; Millar et al. 2016; King et al. 2016), while secondary changes to the climate system (for example, heavy precipitation intensification) will inherently manifest themselves more slowly (for example, following the Clausius-Clapeyron relation (Allen and Ingram 2002; Fischer and Knutti 2016)). Understanding differences in the magnitude of primary versus secondary changes in the climate system can be most easily illustrated with Figure 7.3 (reproduced from Fischer and Knutti (2015), Figure 3): this reveals that large fractions of the global surface under 3°C of warming will witness more than 50-fold increases in the frequency of extreme hot days, while changes to corresponding precipitation extremes are more than an order of magnitude less dramatic. So while both rainfall patterns and temperature regimes are important contributors to determine where crops successfully thrive (Lobell and Burke 2008), the *relative* magnitude of signal-to-noise ratios in temperature dominate the overall geographic

pattern of crop yield losses projected under future climate scenarios (Asseng et al. 2015b; Liu et al. 2016).
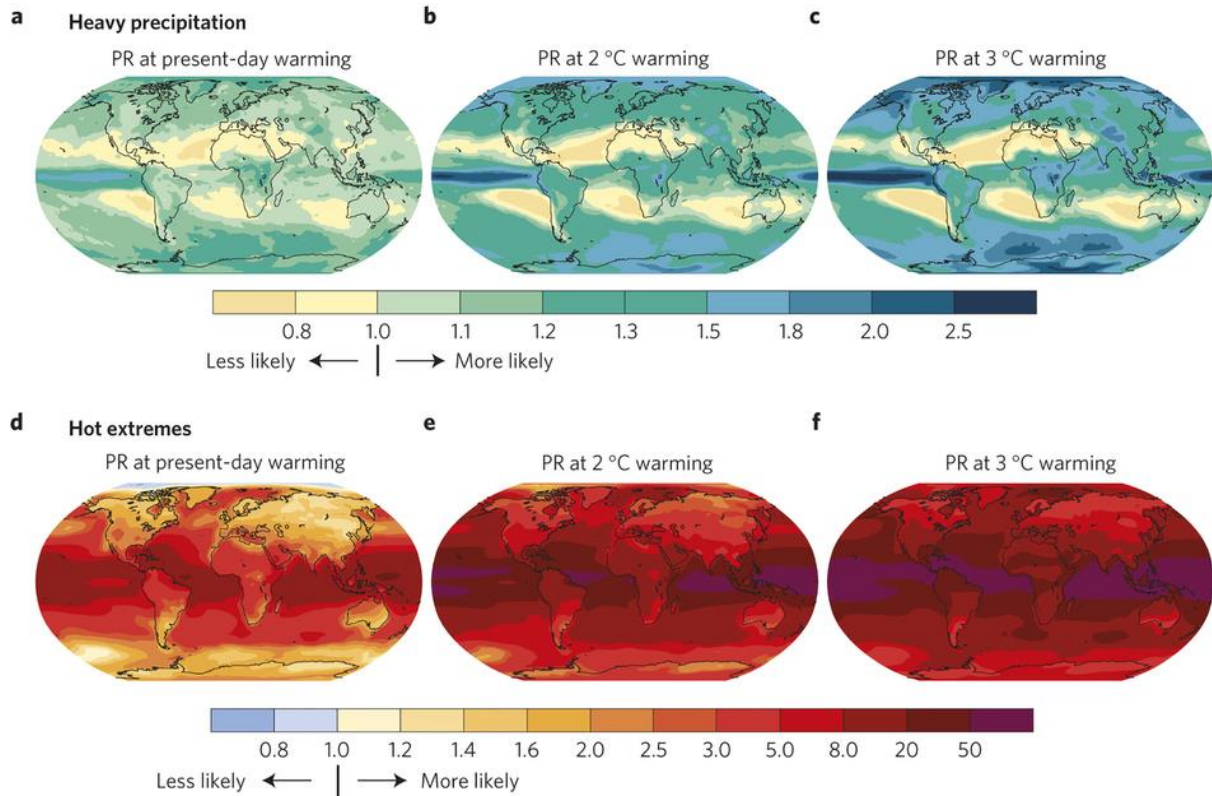
Figure 7.3: Multi-model mean probability of exceeding the pre-industrial 99th percentile of daily precipitation **(a–c)** and temperature **(d–f),** relative to pre-industrial. Ratios are shown for 30-year periods in which the global mean temperatures warmed 0.85°C (present-day) (a,d), 2°C (b,e) and 3°C (c,f) above pre-industrial conditions. Figure reproduced from Fischer and Knutti (2015), Figure 3.

It is nevertheless clear that further research efforts are needed to identify and quantify the emergence of signal-to-noise ratios for those variables which characterise different impacts. However, just as there are distinct *geographic* differences between absolute changes in warming and the regional patterns of signal-to-noise ratios in temperature, it is important to emphasise that any representation of warming-induced signals for other variables should always be presented in an *emergence* framework – that is, quantified in the context of existing variability. An example of such a next step, which benefits from both reliable process understanding and high-quality historical observations, would be the emergent influences of human-induced sea level rise on coastal flooding risk for different regions of the world (Lyu et al. 2014; Hallegatte et al. 2013).

Complementing the discussions about probability ratio saturation presented in Section 6.1, some studies have suggested that different climate impacts will likely present sigmoidal patterns of emergence, such that saturation towards some new 'state' will occur after a fixed

level of warming – one example is the complete destruction of coral reef systems (Ricke et al. 2016). Importantly, if multiple impacts exhibit these non-linear patterns of emergence, but saturate to a 'new normal' after *different* amounts of global warming, then there are potentially crucial implications about the distribution of beneficiaries for different levels of mitigation (Ricke et al. 2016). Such questions about the overlap of signal-to-noise ratios for different climate change impacts therefore represent an important focal point of future research for the climate modelling community (Rogelj and Knutti 2016).

7.2 Strengths and limitations of CMIP5 models

While the ability of any climate modelling framework to adequately simulate the real world is strongly correlated with the spatial resolution used for simulations (Ban et al. 2014; Gibson et al. 2016b; Giorgi et al. 2016), this comes at the cost of computational expense (further discussion is found in Section 3.1.1 and 3.5). In this thesis, I have chosen to only focus on analyses using global climate models from the CMIP5 archive, as opposed to any higher-resolution climate modelling frameworks, for several reasons:

1. CMIP5 data are readily available, and have a consistent set of experiments which have been produced by all of the key modelling centres around the world. Biases which may be the consequence of using only a single-model ensemble (Bellprat and Doblas-Reyes 2016) are therefore reduced.
2. The majority of higher-resolution modelling frameworks are limited only to specific regions of the world (such as the CORDEX regions). This was insufficient for components of this thesis, as I was interested in the spatial patterns of changes to extreme temperatures worldwide.
3. To address initial-condition uncertainty, other single-model frameworks have considered repeated simulations focussing on specific periods in time (such as the last decade). However, evaluating the time-evolution of changes to probability ratios through time (as with Chapter 4) required simulations over the length of both the twentieth and twenty-first century – a criterion suited to the use of CMIP5 models.

The models which are found in the CMIP5 archive operate on relatively coarse resolutions, which is necessary to enable the provision of explicit climate simulations globally. However, it is also important to recognise the limitations associated with such lower-resolution frameworks: for example, soil moisture deficits during the summer months over a small

country like New Zealand would not be adequately simulated (Samuel Dean, pers. comms 2014). For these reasons, I have chosen to focus only on the meteorological characteristics of drought over the New Zealand region, and specifically whether the large-scale circulation features which are conducive to summertime drought will become more likely in a warming world: such seasonal-scale synoptic weather patterns were already known to be sufficiently well-represented in CMIP5 models (Gibson et al. 2016a).

While the representation of some climate processes improve with model resolution (Hagos et al. 2015), the spatial scales at which CMIP5 models operate are mostly adequate when only considering temperature distributions (Allen and Stott 2003; Christidis et al. 2015), and further improvements with increasing model resolution are instead more relevant for those physical processes which are otherwise parameterised at low spatial scales. For example, convection-permitting models produce a much more accurate representation of summertime precipitation extremes over continental Europe (Ban et al. 2015; Prein et al. 2015). So while the assessment of changes to signal-to-noise ratios in precipitation regimes would be more difficult to establish using CMIP5 models only, we can be confident that such models are appropriate to determine the time of emergence of temperature signal-to-noise ratios with continuing anthropogenic climate change.

While the overall conclusions presented in the thesis remain robust – particularly with respect to disparities in the velocity of temperature emergence for low income countries when compared against high income countries – there are nevertheless uncertainties in the use of CMIP5 models which remain, such as the simulation of *regional* temperature variability on different timescales (Knutson et al. 2013; Laepple and Huybers 2014). Reducing uncertainty in this context is particularly difficult, owing to the difficulty in quantifying observational estimates of internal temperature variability (Knutson et al. 2013; Holmes et al. 2016), and further research is therefore required.

Chapter 8

Conclusions

It is the motivation of a doctoral thesis to present evidence justifying a novel perspective on a particular field of research. Here, two summary arguments are posited on the basis of the five pieces of original research presented in Chapters 2 to 6.

8.1 The need for two types of attribution statement for an observed extreme event

At present, an attribution statement related to a real-world extreme event seeks to serve two purposes: be specific enough that it accurately characterises the event which has just occurred, yet be general enough that it is applicable to similar 'types of event' that may occur in the future: as such, the application of any statement regarding the 'role of climate change' in influencing a specific extreme event requires inherently different frameworks for analysis. Therefore, I believe there should be two *types* of attribution statement presented for a given extreme event occurring in the future: one which is highly event-specific and only relevant to quantifying the role of climate change in modifying the likelihood of witnessing that specific event, and another statement which characterises a class of events with only one singular climatic variable being characterised, whereby the observed event would fit into this class.

The event-specific attribution statement can be used to understand what the implications are for the loss and damages caused in the real world by the observed event of interest; the other broader attribution statement will be applicable to similar types of extreme events in the future, can be used to evaluate the emergent changes in risk through time (or for different levels of warming) and will thus be of more value for both decision makers and planners. The recognition that these two types of attribution statement are demonstrating fundamentally different points is crucial, and it must be recognised that no singular attribution statement could, by definition, be capable of capturing the specific applications of these two independent metrics. Many approaches have been considered for both event-as-class frameworks (Stott 2016), while more ideas for event-specific analyses have also been posited in recent years (Yiou et al. 2007; Hannart et al. 2016; Vautard et al. 2016).

8.1.1 Supporting evidence

The analysis presented using the case study of the 2013 New Zealand drought in Chapter 2 provides a helpful demonstration of the argument outlined above: different answers to the question of '*how has human-induced climate change modified the likelihood of witnessing this event?*' will emerge depending on the framework of the analysis, and particularly in the definition of what the 'event' is. An example of an event-as-class framework in Chapter 2 was the use of the area-average, season-average MSLP metric. The FAR associated with this metric for the 2013 event will be similarly applicable to other summers in the future and may be therefore more useful for planners. However, one must recognise the inherent trade-offs with making this metric widely applicable: the more generic characteristics means that a record-breaking extreme value for this metric will not necessarily correlate with an equally severe drought event.

By contrast, the SOM-based metrics enable an assessment of the specific, daily-scale properties of the large-scale circulation patterns which passed over the New Zealand domain over the duration of the 2013 event. Consequently, the event-specific FAR values associated with the three SOM metrics could be useful to different stakeholders for different reasons: for example, Hawke's Bay winemakers could establish whether a similar type of growing season might be more common in the future. Indeed, this event-specific FAR metric may even be conceivably used for individual dairy farmers to claim compensation for some fraction of drought-related losses from the government, which could in turn be paid for using a global finance mechanism funded by high-emitting companies (or governments). It is extremely important to stress that this latter suggestion is purely speculative, and these sorts of theoretical suggestions are currently in a very formative state in the peer-reviewed literature (Parker et al. 2016; James et al. 2014; Thompson and Otto 2015).

These two case studies help to demonstrate the point that there is a fundamental trade-off between an event attribution statement being (1) specific to the characteristics of the extreme event observed in the real world, while also (2) general enough that the subsequent statement is of practical use for the planning and preparation for similar types of events in the future. For this reason, I believe that researchers should endeavour to quantify and present these two different types of metrics for a given attribution study. As with recent attribution studies when focussing exclusively on an event-as-class framework (Uhe et al. 2016; Eden et al.

2016), multiple approaches using independent methods should be considered for each of these two *types* of metric to improve confidence in the subsequent attribution statements.

8.2 Policy-relevant frameworks for understanding regional climate emergence should focus on signal-to-noise ratios

If we are to focus on the broader event-as-class framework for event attribution, which could be performed using only the modelling frameworks available from CMIP5 experiments, then Chapter 4 demonstrates an example of looking at absolute changes in probability ratio for a specific return period through time, and thus in response to progressive increases in anthropogenic forcings. However, as explained in detail in Chapter 6, interpretations of this approach could become be misleading, owing to the variability in maximum probability ratios which can be reached for a specified event threshold. Instead, a distributional signal-to-noise ratio should be the singular metric of focus for efforts to understand the time of emergence of changing risks for differing levels of global mean warming anomalies, or cumulative carbon dioxide emissions.

When focusing on signal-to-noise ratios for temperature, either looking on timescales of days or years, the *spatial patterns* of emergent climate change are clear: there is robust evidence in both Chapters 4 and 5 to demonstrate that the poorest countries of the world will be susceptible to substantially more rapid emergence for a given amount of cumulative emissions (or a given threshold of global mean warming) when compared with the wealthiest countries of the world. The recognition that there is minimal scenario uncertainty and robust agreement across models when quantifying these spatial disparities in the rapidity of emergent climate change is important: there is now sufficient scope to apply this information to inform expectations about future emission pathways for different countries of the world.

8.3 Summary

In this thesis, I have introduced a novel approach to circumvent the model uncertainty associated with drought quantification and attribution by utilising self-organizing maps in an event-specific attribution framework for the 2013 New Zealand drought. Chapter 3 highlighted the ongoing discussions in the peer-review literature about how the extent of conditioning on the observed features of an extreme event will result in different answers to the question of '*what is the role of anthropogenic climate change?*' on the likelihood of

witnessing that same event. Evidence is then presented to demonstrate that the same in-depth considerations should also be applied to different attribution *metrics*, especially if such methods are to be applied in the context of climate services in the future.

Chapter 4 demonstrates how the time-evolution of probability ratios in extreme daily temperatures can be discernibly linked to simulated cumulative CO_2 emissions, with robust differences found in the rate of population exposure to many-fold increases in the frequency of extreme hot days. Chapter 5 considers the implications of utilising relative metrics of regional climate change when identifying the expected 'worst-affected' regions under future warming scenarios. Following the demonstration of negligible scenario uncertainty in the future patterns of emergent signal-to-noise ratios for annual mean temperatures, I then propose a method of quantifying country-level carbon budgets which would be *consistent* with the corresponding impacts experienced by each nation under continued climate change. In Chapter 6, I discuss methods of reconciling detection and attribution techniques with underlying changes in signal-to-noise ratios with continued warming, before finally presenting in Chapter 7 a framework to quantify how future increases in cumulative carbon dioxide emissions can translate to changing heat extremes for different countries of the world.

References

Adam, D., 2011: Climate change in court. *Nat. Clim. Change*, **1**, 127–130, doi:10.1038/nclimate1131.

Alexander, L. V., P. Uotila, N. Nicholls, and A. Lynch, 2010: A New Daily Pressure Dataset for Australia and Its Application to the Assessment of Changes in Synoptic Patterns during the Last Century. *J. Clim.*, **23**, 1111–1126, doi:10.1175/2009JCLI2972.1.

Allan, R. P., 2014: Climate change: Dichotomy of drought and deluge. *Nat. Geosci.*, **7**, 700–701, doi:10.1038/ngeo2243.

Allan, R. P., B. J. Soden, V. O. John, W. Ingram, and P. Good, 2010: Current changes in tropical precipitation. *Environ. Res. Lett.*, **5**, 025205, doi:10.1088/1748-9326/5/2/025205.

Allen, M., 2003: Liability for climate change. *Nature*, **421**, 891–892, doi:10.1038/421891a.

Allen, M. R., and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419**, 224–232, doi:10.1038/nature01092.

Allen, M. R., and P. A. Stott, 2003: Estimating signal amplitudes in optimal fingerprinting, part I: theory. *Clim. Dyn.*, **21**, 477–491, doi:10.1007/s00382-003-0313-9.

Allen, M. R., and D. J. Frame, 2007: Call Off the Quest. *Science*, **318**, 582–583, doi:10.1126/science.1149988.

——, ——, C. Huntingford, C. D. Jones, J. A. Lowe, M. Meinshausen, and N. Meinshausen, 2009: Warming caused by cumulative carbon emissions towards the trillionth tonne. *Nature*, **458**, 1163–1166, doi:10.1038/nature08019.

——, J. S. Fuglestvedt, K. P. Shine, A. Reisinger, R. T. Pierrehumbert, and P. M. Forster, 2016: New use of global warming potentials to compare cumulative and short-lived climate pollutants. *Nat. Clim. Change*, **6**, 773–776, doi:10.1038/nclimate2998.

Althor, G., J. E. M. Watson, and R. A. Fuller, 2016: Global mismatch between greenhouse gas emissions and the burden of climate change. *Sci. Rep.*, **6**, 20281, doi:10.1038/srep20281.

Anderson, B. T., 2011: Near-term increase in frequency of seasonal temperature extremes prior to the 2°C global warming target. *Clim. Change*, **108**, 581, doi:10.1007/s10584-011-0196-4.

——, 2012: Intensification of seasonal extremes given a 2°C global warming target. *Clim. Change*, **112**, 325–337, doi:10.1007/s10584-011-0213-7.

Anderson, T. R., E. Hawkins, and P. D. Jones, 2016: CO2, the greenhouse effect and global warming: from the pioneering work of Arrhenius and Callendar to today's Earth System Models. *Endeavour*, **40**, 178–187, doi:10.1016/j.endeavour.2016.07.002.

Angélil, O., D. A. Stone, M. Tadross, F. Tummon, M. Wehner, and R. Knutti, 2014: Attribution of extreme weather to anthropogenic greenhouse gas emissions: Sensitivity to spatial and temporal scales. *Geophys. Res. Lett.*, **41**, 2150–2155, doi:10.1002/2014GL059234.

——, and Coauthors, 2016: Comparing regional precipitation and temperature extremes in climate model and reanalysis products. *Weather Clim. Extrem.*, **13**, 35–43, doi:10.1016/j.wace.2016.07.001.

Arrhenius, P. S., 1896: On the influence of carbonic acid in the air upon the temperature of the ground. *Philos. Mag. Ser. 5*, **41**, 237–276, doi:10.1080/14786449608620846.

Asseng, S., and Coauthors, 2015a: Rising temperatures reduce global wheat production. *Nat. Clim. Change*, **5**, 143–147, doi:10.1038/nclimate2470.

——, and Coauthors, 2015b: Rising temperatures reduce global wheat production. *Nat. Clim. Change*, **5**, 143–147, doi:10.1038/nclimate2470.

Ban, N., J. Schmidli, and C. Schär, 2014: Evaluation of the convection-resolving regional climate modeling approach in decade-long simulations. *J. Geophys. Res. Atmospheres*, **119**, 7889–7907, doi:10.1002/2014JD021478.

——, ——, and ——, 2015: Heavy precipitation in a changing climate: Does short-term summer precipitation increase faster? *Geophys. Res. Lett.*, **42**, 2014GL062588, doi:10.1002/2014GL062588.

Barlow, M., and A. Hoell, 2015: Drought in the Middle East and Central–Southwest Asia During Winter 2013/14. *Bull. Am. Meteorol. Soc.*, **96**, S71–S76, doi:10.1175/BAMS-D-15-00127.1.

Battisti, D. S., and R. L. Naylor, 2009: Historical Warnings of Future Food Insecurity with Unprecedented Seasonal Heat. *Science*, **323**, 240–244, doi:10.1126/science.1164363.

Beaumont, L. J., A. Pitman, S. Perkins, N. E. Zimmermann, N. G. Yoccoz, and W. Thuiller, 2011: Impacts of climate change on the world's most exceptional ecoregions. *Proc. Natl. Acad. Sci.*, **108**, 2306–2311, doi:10.1073/pnas.1007217108.

Bellprat, O., and F. Doblas-Reyes, 2016: Attribution of extreme weather and climate events overestimated by unreliable climate simulations. *Geophys. Res. Lett.*, **43**, 2015GL067189, doi:10.1002/2015GL067189.

Bergaoui, K., D. Mitchell, F. Otto, M. Allen, R. Zaaboul, and R. McDonnell, 2015: The Contribution of Human-Induced Climate Change to the Drought of 2014 in the Southern Levant Region. *Bull. Am. Meteorol. Soc.*, **96**, S66–S70, doi:10.1175/BAMS-D-15-00129.1.

Black, M. T., and Coauthors, 2016: The weather@home regional climate modelling project for Australia and New Zealand. *Geosci. Model Dev. Discuss.*, 1–28, doi:10.5194/gmd-2016-100.

Brasseur, G. P., and L. Gallardo, 2016: Climate services: Lessons learned and future prospects. *Earths Future*, **4**, 79–89, doi:10.1002/2015EF000338.

Burke, E., and S. Brown, 2008: Evaluating uncertainties in the projection of future drought. *J Hydrometeor*, **9**, 292–299.

Burke, M., S. M. Hsiang, and E. Miguel, 2015: Global non-linear effect of temperature on economic production. *Nature*, **527**, 235–239, doi:10.1038/nature15725.

Cassano, J. J., P. Uotila, and A. Lynch, 2006: Changes in synoptic weather patterns in the polar regions in the twentieth and twenty-first centuries, part 1: Arctic. *Int. J. Climatol.*, **26**, 1027–1049, doi:10.1002/joc.1306.

Christiansen, B., 2015: The Role of the Selection Problem and Non-Gaussianity in Attribution of Single Events to Climate Change. *J. Clim.*, **28**, 9873–9891, doi:10.1175/JCLI-D-15-0318.1.

Christidis, N., and P. A. Stott, 2015: Extreme Rainfall in the United Kingdom During Winter 2013/14: The Role of Atmospheric Circulation and Climate Change. *Bull. Am. Meteorol. Soc.*, **96**, S46–S50, doi:10.1175/BAMS-D-15-00094.1.

——, ——, and F. W. Zwiers, 2015: Fast-track attribution assessments based on pre-computed estimates of changes in the odds of warm extremes. *Clim. Dyn.*, **45**, 1547–1564, doi:10.1007/s00382-014-2408-x.

Ciais, P., and Coauthors, 2013: Carbon and Other Biogeochemical Cycles. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 465–570 www.climatechange2013.org.

Clayson, C. A., and D. Weitlich, 2007: Variability of Tropical Diurnal Sea Surface Temperature. *J. Clim.*, **20**, 334–352, doi:10.1175/JCLI3999.1.

Collins, M., and Coauthors, 2013: Long-term Climate Change: Projections, Commitments and Irreversibility. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1029–1136 www.climatechange2013.org.

Compo, G. P., and Coauthors, 2011: The Twentieth Century Reanalysis Project. *Q. J. R. Meteorol. Soc.*, **137**, 1–28, doi:10.1002/qj.776.

Coumou, D., and A. Robinson, 2013: Historic and future increase in the global land area affected by monthly heat extremes. *Environ. Res. Lett.*, **8**, 034018, doi:10.1088/1748-9326/8/3/034018.

Dai, A., 2011a: Drought under global warming: a review. *Wiley Interdiscip. Rev. Clim. Change*, **2**, 45–65, doi:10.1002/wcc.81.

——, 2011b: Characteristics and trends in various forms of the Palmer Drought Severity Index during 1900–2008. *J. Geophys. Res. Atmospheres*, **116**, D12115, doi:10.1029/2010JD015541.

Davis, S. J., and N. Diffenbaugh, 2016: Dislocated interests and climate change. *Environ. Res. Lett.*, **11**, 061001, doi:10.1088/1748-9326/11/6/061001.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.*, **137**, 553–597, doi:10.1002/qj.828.

Deutsch, C. A., J. J. Tewksbury, R. B. Huey, K. S. Sheldon, C. K. Ghalambor, D. C. Haak, and P. R. Martin, 2008: Impacts of climate warming on terrestrial ectotherms across latitude. *Proc. Natl. Acad. Sci.*, **105**, 6668–6672, doi:10.1073/pnas.0709472105.

Diffenbaugh, N. S., and M. Scherer, 2011: Observational and model evidence of global emergence of permanent, unprecedented heat in the 20th and 21st centuries. *Clim. Change*, **107**, 615–624, doi:10.1007/s10584-011-0112-y.

——, D. L. Swain, and D. Touma, 2015: Anthropogenic warming has increased drought risk in California. *Proc. Natl. Acad. Sci.*, **112**, 3931–3936, doi:10.1073/pnas.1422385112.

Donat, M. G., A. L. Lowry, L. V. Alexander, P. A. O'Gorman, and N. Maher, 2016: More extreme precipitation in the world's dry and wet regions. *Nat. Clim. Change*, **6**, 508–513, doi:10.1038/nclimate2941.

Dunne, J. P., R. J. Stouffer, and J. G. John, 2013: Reductions in labour capacity from heat stress under climate warming. *Nat. Clim. Change*, **3**, 563–566, doi:10.1038/nclimate1827.

Eden, J. M., K. Wolter, F. E. L. Otto, and G. J. van Oldenborgh, 2016: Multi-method attribution analysis of extreme precipitation in Boulder, Colorado. *Environ. Res. Lett.*, **11**, 124009, doi:10.1088/1748-9326/11/12/124009.

Edwards, P. N., 2011: History of climate modeling. *Wiley Interdiscip. Rev. Clim. Change*, **2**, 128–139, doi:10.1002/wcc.95.

Fischer, E. M., and R. Knutti, 2015: Anthropogenic contribution to global occurrence of heavy-precipitation and high-temperature extremes. *Nat. Clim. Change*, **5**, 560–564, doi:10.1038/nclimate2617.

——, and ——, 2016: Observed heavy precipitation increase confirms theory and early models. *Nat. Clim. Change*, **6**, 986–991, doi:10.1038/nclimate3110.

——, U. Beyerle, and R. Knutti, 2013: Robust spatially aggregated projections of climate extremes. *Nat. Clim. Change*, **3**, 1033–1038, doi:10.1038/nclimate2051.

——, J. Sedláček, E. Hawkins, and R. Knutti, 2014: Models agree on forced response pattern of precipitation and temperature extremes. *Geophys. Res. Lett.*, **41**, 2014GL062018, doi:10.1002/2014GL062018.

Flato, G., and Coauthors, 2013: Evaluation of Climate Models. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 741–866 www.climatechange2013.org.

Frame, D., M. Joshi, E. Hawkins, L. J. Harrington, and M. de Roiste, 2017: Population-based emergence of unfamiliar climates. *Nat. Clim. Change*, **7**, 407–411, doi:10.1038/nclimate3297.

Frame, D. J., and Coauthors, 2009: The climateprediction.net BBC climate change experiment: design of the coupled model ensemble. *Philos. Trans. R. Soc. Lond. Math. Phys. Eng. Sci.*, **367**, 855–870, doi:10.1098/rsta.2008.0240.

Frame, D. J., A. H. Macey, and M. R. Allen, 2014: Cumulative emissions and climate policy. *Nat. Geosci.*, **7**, 692–693, doi:10.1038/ngeo2254.

Friedlingstein, P., M. Meinshausen, V. K. Arora, C. D. Jones, A. Anav, S. K. Liddicoat, and R. Knutti, 2014: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks. *J. Clim.*, **27**, 511–526, doi:10.1175/JCLI-D-12-00579.1.

Fuglestvedt, J. S., and S. Kallbekken, 2016: Climate responsibility: Fair shares? *Nat. Clim. Change*, **6**, 19–20, doi:10.1038/nclimate2791.

Funk, C., S. Shukla, A. Hoell, and B. Livneh, 2015: Assessing the Contributions of East African and West Pacific Warming to the 2014 Boreal Spring East African Drought. *Bull. Am. Meteorol. Soc.*, **96**, S77–S82, doi:10.1175/BAMS-D-15-00106.1.

Fyfe, J. C., N. P. Gillett, and F. W. Zwiers, 2013: Overestimated global warming over the past 20 years. *Nat. Clim. Change*, **3**, 767–769, doi:10.1038/nclimate1972.

Gibson, P. B., S. E. Perkins-Kirkpatrick, and J. A. Renwick, 2016a: Projected changes in synoptic weather patterns over New Zealand examined through self-organizing maps. *Int. J. Climatol.*, n/a-n/a, doi:10.1002/joc.4604.

——, P. Uotila, S. E. Perkins-Kirkpatrick, L. V. Alexander, and A. J. Pitman, 2016b: Evaluating synoptic systems in the CMIP5 climate models over the Australian region. *Clim. Dyn.*, 1–17, doi:10.1007/s00382-015-2961-y.

Gignac, R., and H. D. Matthews, 2015: Allocating a 2 °C cumulative carbon budget to countries. *Environ. Res. Lett.*, **10**, 075004, doi:10.1088/1748-9326/10/7/075004.

Gillett, N., V. Arora, D. Matthews, and M. Allen, 2013: Constraining the ratio of global warming to cumulative CO2 emissions using CMIP5 simulations. *J Clim.*, doi:10.1175/JCLI-D-12-00476.1.

Giorgi, F., and X. Bi, 2009: Time of emergence (TOE) of GHG-forced precipitation change hot-spots. *Geophys. Res. Lett.*, **36**, L06709, doi:10.1029/2009GL037593.

——, C. Torma, E. Coppola, N. Ban, C. Schär, and S. Somot, 2016: Enhanced summer convective rainfall at Alpine high elevations in response to climate warming. *Nat. Geosci.*, **9**, 584–589, doi:10.1038/ngeo2761.

Goddard, L., 2016: From science to service. *Science*, **353**, 1366–1367, doi:10.1126/science.aag3087.

Grambsch, A., and B. Menne, 2003: Adaptation and adaptive capacity in the public health context. *Clim. Change Health Risks Responses McMichael AJ Campbell-Lendrum*

DH Corvalan CF Ebi KL Githeko Scheraga JD Al Eds Geneva World Health Organ., 220–236.

Green, D., 2016: The spatial distribution of extreme climate events, another climate inequity for the world's most vulnerable people. *Environ. Res. Lett.*, **11**, 091002, doi:10.1088/1748-9326/11/9/091002.

Gregow, H., and Coauthors, 2015: Worldwide Survey of Awareness and Needs Concerning Reanalyses and Respondents Views on Climate Services. *Bull. Am. Meteorol. Soc.*, **97**, 1461–1473, doi:10.1175/BAMS-D-14-00271.1.

Greve, P., B. Orlowsky, B. Mueller, J. Sheffield, M. Reichstein, and S. I. Seneviratne, 2014: Global assessment of trends in wetting and drying over land. *Nat. Geosci.*, **7**, 716–721, doi:10.1038/ngeo2247.

Hagos, S., L. R. Leung, Q. Yang, C. Zhao, and J. Lu, 2015: Resolution and Dynamical Core Dependence of Atmospheric River Frequency in Global Model Simulations. *J. Clim.*, **28**, 2764–2776, doi:10.1175/JCLI-D-14-00567.1.

Hallegatte, S., C. Green, R. J. Nicholls, and J. Corfee-Morlot, 2013: Future flood losses in major coastal cities. *Nat. Clim. Change*, **3**, 802–806, doi:10.1038/nclimate1979.

——, and Coauthors, 2016: Mapping the climate change challenge. *Nat. Clim. Change*, **6**, 663–668, doi:10.1038/nclimate3057.

Hannart, A., A. Carrassi, M. Bocquet, M. Ghil, P. Naveau, M. Pulido, J. Ruiz, and P. Tandeo, 2016: DADA: data assimilation for the detection and attribution of weather and climate-related events. *Clim. Change*, **136**, 155–174, doi:10.1007/s10584-016-1595-3.

Hansen, J., and M. Sato, 2016: Regional climate change and national responsibilities. *Environ. Res. Lett.*, **11**, 034009, doi:10.1088/1748-9326/11/3/034009.

Harrington, L. J., 2017: Investigating differences between event-as-class and probability density-based attribution statements with emerging climate change. *Clim. Change*, **141**, 641–654, doi:10.1007/s10584-017-1906-3.

——, S. Rosier, S. M. Dean, S. Stuart, and A. Scahill, 2014: The role of anthropogenic climate change in the 2013 drought over North Island, New Zealand. *Bull. Am. Meteorol. Soc.*, **95**, S45–S48, doi:10.1175/1520-0477-95.9.S1.1.

——, D. J. Frame, E. M. Fischer, E. Hawkins, M. Joshi, and C. D. Jones, 2016a: Poorest countries experience earlier anthropogenic emergence of daily temperature extremes. *Environ. Res. Lett.*, **11**, 055007, doi:10.1088/1748-9326/11/5/055007.

——, P. B. Gibson, S. M. Dean, D. Mitchell, S. M. Rosier, and D. J. Frame, 2016b: Investigating event-specific drought attribution using self-organizing maps. *J. Geophys. Res. Atmospheres*, **121**, 2016JD025602, doi:10.1002/2016JD025602.

Hartmann, D. L., and Coauthors, 2013: Observations: Atmosphere and Surface. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F.

Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 159–254 www.climatechange2013.org.

Haustein, K., and Coauthors, 2016: Real-time extreme weather event attribution with forecast seasonal SSTs. *Environ. Res. Lett.*, **11**, 064006, doi:10.1088/1748-9326/11/6/064006.

Hawkins, E., and R. Sutton, 2009: The Potential to Narrow Uncertainty in Regional Climate Predictions. *Bull. Am. Meteorol. Soc.*, **90**, 1095–1107, doi:10.1175/2009BAMS2607.1.

Hawkins, E., and R. Sutton, 2012: Time of emergence of climate signals. *Geophys. Res. Lett.*, **39**, L01702, doi:10.1029/2011GL050087.

Hawkins, E., and Coauthors, 2014: Uncertainties in the timing of unprecedented climates. *Nature*, **511**, E3–E5, doi:10.1038/nature13523.

Hayden, M. H., H. Brenkert-Smith, and O. V. Wilhelmi, 2011: Differential Adaptive Capacity to Extreme Heat: A Phoenix, Arizona, Case Study. *Weather Clim. Soc.*, **3**, 269–280, doi:10.1175/WCAS-D-11-00010.1.

Heffernan, O., 2016: The mystery of the expanding tropics. *Nat. News*, **530**, 20, doi:10.1038/530020a.

Hegerl, G., and F. Zwiers, 2011: Use of models in detection and attribution of climate change. *Wiley Interdiscip. Rev. Clim. Change*, **2**, 570–591, doi:10.1002/wcc.121.

Heim, R. R., 2002: A Review of Twentieth-Century Drought Indices Used in the United States. *Bull. Am. Meteorol. Soc.*, **83**, 1149–1165, doi:10.1175/1520-0477(2002)083<1149:AROTDI>2.3.CO;2.

Held, I. M., and B. J. Soden, 2006: Robust Responses of the Hydrological Cycle to Global Warming. *J. Clim.*, **19**, 5686–5699, doi:10.1175/JCLI3990.1.

Henley, B. J., J. Gergis, D. J. Karoly, S. Power, J. Kennedy, and C. K. Folland, 2015: A Tripole Index for the Interdecadal Pacific Oscillation. *Clim. Dyn.*, **45**, 3077–3090, doi:10.1007/s00382-015-2525-1.

Herring, S. C., M. P. Hoerling, T. C. Peterson, and P. A. Stott, 2014: Explaining Extreme Events of 2013 from a Climate Perspective. *Bull. Am. Meteorol. Soc.*, **95**, S1–S104, doi:10.1175/1520-0477-95.9.S1.1.

——, ——, J. P. Kossin, T. C. Peterson, and P. A. Stott, 2015: Explaining Extreme Events of 2014 from a Climate Perspective. *Bull. Am. Meteorol. Soc.*, **96**, S1–S172, doi:10.1175/BAMS-ExplainingExtremeEvents2014.1.

Hewitson, B. C., and R. G. Crane, 2002: Self-organizing maps: applications to synoptic climatology. *Clim. Res.*, **22**, 13–26, doi:10.3354/cr022013.

Hewitt, C., S. Mason, and D. Walland, 2012: The Global Framework for Climate Services. *Nat. Clim. Change*, **2**, 831–832, doi:10.1038/nclimate1745.

Hoerling, M., J. Eischeid, J. Perlwitz, X. Quan, T. Zhang, and P. Pegion, 2012: On the Increased Frequency of Mediterranean Drought. *J. Clim.*, **25**, 2146–2161, doi:10.1175/JCLI-D-11-00296.1.

Holmes, C. R., T. Woollings, E. Hawkins, and H. de Vries, 2016: Robust Future Changes in Temperature Variability under Greenhouse Gas Forcing and the Relationship with Thermal Advection. *J. Clim.*, **29**, 2221–2236, doi:10.1175/JCLI-D-14-00735.1.

Hope, P., G. Wang, E.-P. Lim, H. H. Hendon, and J. M. Arblaster, 2016: What Caused the Record-Breaking Heat Across Australia in October 2015? *Bull. Am. Meteorol. Soc.*, **97**, S122–S126, doi:10.1175/BAMS-D-16-0141.1.

Huang, J., H. Yu, X. Guan, G. Wang, and R. Guo, 2016: Accelerated dryland expansion under climate change. *Nat. Clim. Change*, **6**, 166–171, doi:10.1038/nclimate2837.

Huber, M., and R. Knutti, 2014: Natural variability, radiative forcing and climate response in the recent hiatus reconciled. *Nat. Geosci.*, **7**, 651–656, doi:10.1038/ngeo2228.

IPCC, 2013: Summary for Policymakers. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1–30 www.climatechange2013.org.

James, R., F. Otto, H. Parker, E. Boyd, R. Cornforth, D. Mitchell, and M. Allen, 2014: Characterizing loss and damage from climate change. *Nat. Clim. Change*, **4**, 938–939, doi:10.1038/nclimate2411.

Jones, C., and Coauthors, 2013: Twenty-First-Century Compatible CO2 Emissions and Airborne Fraction Simulated by CMIP5 Earth System Models under Four Representative Concentration Pathways. *J. Clim.*, **26**, 4398–4413, doi:10.1175/JCLI-D-12-00554.1.

Joshi, M., E. Hawkins, R. Sutton, J. Lowe, and D. Frame, 2011: Projections of when temperature change will exceed 2 °C above pre-industrial levels. *Nat. Clim. Change*, **1**, 407–412, doi:10.1038/nclimate1261.

Joshi, M. M., J. M. Gregory, M. J. Webb, D. M. H. Sexton, and T. C. Johns, 2008: Mechanisms for the land/sea warming contrast exhibited by simulations of climate change. *Clim. Dyn.*, **30**, 455–465, doi:10.1007/s00382-007-0306-1.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S.-K. Yang, J. J. Hnilo, M. Fiorino, and G. L. Potter, 2002: NCEP–DOE AMIP-II Reanalysis (R-2). *Bull. Am. Meteorol. Soc.*, **83**, 1631–1643, doi:10.1175/BAMS-83-11-1631.

Kay, J. E., and Coauthors, 2015: The Community Earth System Model (CESM) Large Ensemble Project: A Community Resource for Studying Climate Change in the Presence of Internal Climate Variability. *Bull. Am. Meteorol. Soc.*, **96**, 1333–1349, doi:10.1175/BAMS-D-13-00255.1.

KC, S., and W. Lutz, 2014: The human core of the shared socioeconomic pathways: Population scenarios by age, sex and level of education for all countries to 2100.

Glob. Environ. Change, doi:10.1016/j.gloenvcha.2014.06.004.
http://www.sciencedirect.com/science/article/pii/S0959378014001095 (Accessed
September 18, 2016).

Kelley, C. P., S. Mohtadi, M. A. Cane, R. Seager, and Y. Kushnir, 2015: Climate change in
the Fertile Crescent and implications of the recent Syrian drought. *Proc. Natl. Acad.
Sci.*, **112**, 3241–3246, doi:10.1073/pnas.1421533112.

King, A. D., and Coauthors, 2015a: The timing of anthropogenic emergence in simulated
climate extremes. *Environ. Res. Lett.*, **10**, 094015, doi:10.1088/1748-
9326/10/9/094015.

——, G. J. van Oldenborgh, D. J. Karoly, S. C. Lewis, and H. Cullen, 2015b: Attribution of
the record high Central England temperature of 2014 to anthropogenic influences.
Environ. Res. Lett., **10**, 054002, doi:10.1088/1748-9326/10/5/054002.

——, M. T. Black, S.-K. Min, E. M. Fischer, D. M. Mitchell, L. J. Harrington, and S. E.
Perkins-Kirkpatrick, 2016: Emergence of heat extremes attributable to anthropogenic
influences. *Geophys. Res. Lett.*, 2015GL067448, doi:10.1002/2015GL067448.

Kirtman, B., S. B. Power, J.A. Adedoyin, G.J. Boer, R. Bojariu, I. Camilloni, F.J. Doblas-
Reyes, A.M. Fiore, M. Kimoto, G.A. Meehl, M. Prather, A. Sarr, C., and Schär, R.
Sutton, G.J. van Oldenborgh, G. Vecchi and H.J. Wang, 2014: Near-term Climate
Change: Projections and Predictability. *Climate Change 2013 - The Physical Science
Basis*, Intergovernmental Panel on Climate Change, D. Qin, and G.-K. Plattner, M.
Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley, Eds.,
Cambridge University Press, Cambridge, 953–1028
http://ebooks.cambridge.org/ref/id/CBO9781107415324A031 (Accessed December
16, 2014).

Knutson, T. R., F. Zeng, and A. T. Wittenberg, 2013: Multimodel Assessment of Regional
Surface Temperature Trends: CMIP3 and CMIP5 Twentieth-Century Simulations. *J.
Clim.*, **26**, 8709–8743, doi:10.1175/JCLI-D-12-00567.1.

Knutti, R., and J. Sedláček, 2013: Robustness and uncertainties in the new CMIP5 climate
model projections. *Nat. Clim. Change*, **3**, 369–373, doi:10.1038/nclimate1716.

Knutti, R., and Coauthors, 2008: A Review of Uncertainties in Global Temperature
Projections over the Twenty-First Century. *J. Clim.*, **21**, 2651–2663,
doi:10.1175/2007JCLI2119.1.

Knutti, R., R. Furrer, C. Tebaldi, J. Cermak, and G. A. Meehl, 2009: Challenges in
Combining Projections from Multiple Climate Models. *J. Clim.*, **23**, 2739–2758,
doi:10.1175/2009JCLI3361.1.

——, D. Masson, and A. Gettelman, 2013: Climate model genealogy: Generation CMIP5 and
how we got there. *Geophys. Res. Lett.*, **40**, 1194–1199, doi:10.1002/grl.50256.

——, J. Rogelj, J. Sedláček, and E. M. Fischer, 2016: A scientific critique of the two-degree
climate change target. *Nat. Geosci.*, **9**, 13–18, doi:10.1038/ngeo2595.

Laepple, T., and P. Huybers, 2014: Global and regional variability in marine surface temperatures. *Geophys. Res. Lett.*, **41**, 2528–2534, doi:10.1002/2014GL059345.

Lehner, F., and T. F. Stocker, 2015: From local perception to global perspective. *Nat. Clim. Change*, **5**, 731–734, doi:10.1038/nclimate2660.

Liu, B., and Coauthors, 2016: Similar estimates of temperature impacts on global wheat yield by three independent methods. *Nat. Clim. Change*, **advance online publication**, doi:10.1038/nclimate3115. http://www.nature.com/nclimate/journal/vaop/ncurrent/full/nclimate3115.html (Accessed October 16, 2016).

Lobell, D. B., and M. B. Burke, 2008: Why are agricultural impacts of climate change so uncertain? The importance of temperature relative to precipitation. *Environ. Res. Lett.*, **3**, 034007, doi:10.1088/1748-9326/3/3/034007.

——, M. Bänziger, C. Magorokosho, and B. Vivek, 2011: Nonlinear heat effects on African maize as evidenced by historical yield trials. *Nat. Clim. Change*, **1**, 42–45, doi:10.1038/nclimate1043.

Loikith, P. C., and J. D. Neelin, 2015: Short-tailed temperature distributions over North America and implications for future changes in extremes. *Geophys. Res. Lett.*, **42**, 2015GL065602, doi:10.1002/2015GL065602.

Lorenz, R., and Coauthors, 2016: Influence of land-atmosphere feedbacks on temperature and precipitation extremes in the GLACE-CMIP5 ensemble. *J. Geophys. Res. Atmospheres*, **121**, 2015JD024053, doi:10.1002/2015JD024053.

Lott, F. C., N. Christidis, and P. A. Stott, 2013: Can the 2011 East African drought be attributed to human-induced climate change? *Geophys. Res. Lett.*, **40**, 1177–1181, doi:10.1002/grl.50235.

Lyu, K., X. Zhang, J. A. Church, A. B. A. Slangen, and J. Hu, 2014: Time of emergence for regional sea-level change. *Nat. Clim. Change*, **4**, 1006–1010, doi:10.1038/nclimate2397.

Maher, N., A. S. Gupta, and M. H. England, 2014: Drivers of decadal hiatus periods in the 20th and 21st centuries. *Geophys. Res. Lett.*, **41**, 5978–5986, doi:10.1002/2014GL060527.

Mahlstein, I., R. Knutti, S. Solomon, and R. W. Portmann, 2011: Early onset of significant local warming in low latitude countries. *Environ. Res. Lett.*, **6**, 034009, doi:10.1088/1748-9326/6/3/034009.

Mahlstein, I., J. S. Daniel, and S. Solomon, 2013: Pace of shifts in climate regions increases with global temperature. *Nat. Clim. Change*, **3**, 739–743, doi:10.1038/nclimate1876.

Maraun, D., 2013: When will trends in European mean and heavy daily precipitation emerge? *Environ. Res. Lett.*, **8**, 014004, doi:10.1088/1748-9326/8/1/014004.

Marthews, T. R., F. E. L. Otto, D. Mitchell, S. J. Dadson, and R. G. Jones, 2015: The 2014 Drought in the Horn of Africa: Attribution of Meteorological Drivers. *Bull. Am. Meteorol. Soc.*, **96**, S83–S88, doi:10.1175/BAMS-D-15-00115.1.

Massey, N., and Coauthors, 2015: weather@home—development and validation of a very large ensemble modelling system for probabilistic event attribution. *Q. J. R. Meteorol. Soc.*, **141**, 1528–1545, doi:10.1002/qj.2455.

Masson, D., and R. Knutti, 2011: Spatial-Scale Dependence of Climate Model Performance in the CMIP3 Ensemble. *J. Clim.*, **24**, 2680–2692, doi:10.1175/2011JCLI3513.1.

Matthews, H. D., 2016: Quantifying historical carbon and climate debts among nations. *Nat. Clim. Change*, **6**, 60–64, doi:10.1038/nclimate2774.

——, N. P. Gillett, P. A. Stott, and K. Zickfeld, 2009: The proportionality of global warming to cumulative carbon emissions. *Nature*, **459**, 829–832, doi:10.1038/nature08047.

McGuffie, K., and A. Henderson-Sellers, 2001: Forty years of numerical climate modelling. *Int. J. Climatol.*, **21**, 1067–1109, doi:10.1002/joc.632.

Meinshausen, M., N. Meinshausen, W. Hare, S. C. B. Raper, K. Frieler, R. Knutti, D. J. Frame, and M. R. Allen, 2009: Greenhouse-gas emission targets for limiting global warming to 2 °C. *Nature*, **458**, 1158–1162, doi:10.1038/nature08017.

Meinshausen, M., T. M. L. Wigley, and S. C. B. Raper, 2011: Emulating atmosphere-ocean and carbon cycle models with a simpler model, MAGICC6 – Part 2: Applications. *Atmos Chem Phys*, **11**, 1457–1471, doi:10.5194/acp-11-1457-2011.

Meinshausen, M., and Coauthors, 2015: National post-2020 greenhouse gas targets and diversity-aware leadership. *Nat. Clim. Change*, **5**, 1098–1106, doi:10.1038/nclimate2826.

Meredith, E. P., V. A. Semenov, D. Maraun, W. Park, and A. V. Chernokulsky, 2015: Crucial role of Black Sea warming in amplifying the 2012 Krymsk precipitation extreme. *Nat. Geosci.*, **8**, 615–619, doi:10.1038/ngeo2483.

Millar, R. J., Z. R. Nicholls, P. Friedlingstein, and M. R. Allen, 2016: A modified impulse-response representation of the global response to carbon dioxide emissions. *Atmos Chem Phys Discuss*, **2016**, 1–20, doi:10.5194/acp-2016-405.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl, 2011: Human contribution to more-intense precipitation extremes. *Nature*, **470**, 378–381, doi:10.1038/nature09763.

Mitchell, D., and Coauthors, 2016: Attributing human mortality during extreme heat waves to anthropogenic climate change. *Environ. Res. Lett.*, **11**, 074006, doi:10.1088/1748-9326/11/7/074006.

Mitchell, D. M., 2016: Attributing the forced components of observed stratospheric temperature variability to external drivers. *Q. J. R. Meteorol. Soc.*, **142**, 1041–1047, doi:10.1002/qj.2707.

——, and Coauthors, 2015: Solar signals in CMIP-5 simulations: the stratospheric pathway. *Q. J. R. Meteorol. Soc.*, **141**, 2390–2403, doi:10.1002/qj.2530.

Mora, C., and Coauthors, 2013: The projected timing of climate departure from recent variability. *Nature*, **502**, 183–187, doi:10.1038/nature12540.

Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, doi:10.1038/nature08823.

Mueller, B., X. Zhang, and F. W. Zwiers, 2016: Historically hottest summers projected to be the norm for more than half of the world's population within 20 years. *Environ. Res. Lett.*, **11**, 044011, doi:10.1088/1748-9326/11/4/044011.

Myhre, G., and Coauthors, 2013: Anthropogenic and Natural Radiative Forcing. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 659–740 www.climatechange2013.org.

National Academies of Sciences, Engineering, and Medicine, Committee on Extreme Weather Events and Climate Change Attribution, Board on Atmospheric Sciences and Climate, and Division on Earth and Life Studies, 2016: *Attribution of Extreme Weather Events in the Context of Climate Change*. National Academies Press, Washington, D.C., http://www.nap.edu/catalog/21852 (Accessed June 16, 2016).

New Zealand Treasury, *Economic Impacts of the Drought - Budget Economic and Fiscal Update 2013 — The Treasury - New Zealand*. http://www.treasury.govt.nz/budget/forecasts/befu2013/021.htm/009.htm (Accessed July 1, 2016).

van Oldenborgh, G. J., F. E. L. Otto, K. Haustein, and H. Cullen, 2015: Climate change increases the probability of heavy rains like those of storm Desmond in the UK – an event attribution study in near-real time. *Hydrol Earth Syst Sci Discuss*, **2015**, 13197–13216, doi:10.5194/hessd-12-13197-2015.

Orth, R., M. Staudinger, S. I. Seneviratne, J. Seibert, and M. Zappa, 2015: Does model performance improve with complexity? A case study with three hydrological models. *J. Hydrol.*, **523**, 147–159, doi:10.1016/j.jhydrol.2015.01.044.

Otto, F. E. L., 2016: Extreme events: The art of attribution. *Nat. Clim. Change*, **6**, 342–343, doi:10.1038/nclimate2971.

——, and Coauthors, 2015a: Factors Other Than Climate Change, Main Drivers of 2014/15 Water Shortage in Southeast Brazil. *Bull. Am. Meteorol. Soc.*, **96**, S35–S40, doi:10.1175/BAMS-D-15-00120.1.

——, D. J. Frame, A. Otto, and M. R. Allen, 2015b: Embracing uncertainty in climate change policy. *Nat. Clim. Change*, **5**, 917–920, doi:10.1038/nclimate2716.

——, G. J. van Oldenborgh, J. Eden, P. A. Stott, D. J. Karoly, and M. R. Allen, 2016: The attribution question. *Nat. Clim. Change*, **6**, 813–816, doi:10.1038/nclimate3089.

Parker, H. R., E. Boyd, R. J. Cornforth, R. James, F. E. L. Otto, and M. R. Allen, 2016: Stakeholder perceptions of event attribution in the loss and damage debate. *Clim. Policy*, **0**, 1–18, doi:10.1080/14693062.2015.1124750.

Perkins, S. E., 2015: A review on the scientific understanding of heatwaves—Their measurement, driving mechanisms, and changes at the global scale. *Atmospheric Res.*, **164–165**, 242–267, doi:10.1016/j.atmosres.2015.05.014.

Perkins, S. E., and L. V. Alexander, 2012: On the Measurement of Heat Waves. *J. Clim.*, **26**, 4500–4517, doi:10.1175/JCLI-D-12-00383.1.

Peters, G. P., R. M. Andrew, S. Solomon, and P. Friedlingstein, 2015: Measuring a fair and ambitious climate agreement using cumulative emissions. *Environ. Res. Lett.*, **10**, 105004, doi:10.1088/1748-9326/10/10/105004.

Peterson, T. C., P. A. Stott, and S. Herring, 2012: Explaining Extreme Events of 2011 from a Climate Perspective. *Bull. Am. Meteorol. Soc.*, **93**, 1041–1067, doi:10.1175/BAMS-D-12-00021.1.

——, ——, and ——, 2013: Explaining Extreme Events of 2012 from a Climate Perspective. *Bull. Am. Meteorol. Soc.*, **94**, S1–S74, doi:10.1175/BAMS-D-13-00085.1.

du Pont, Y. R., M. L. Jeffery, J. Gütschow, J. Rogelj, P. Christoff, and M. Meinshausen, 2017: Equitable mitigation to achieve the Paris Agreement goals. *Nat. Clim. Change*, **7**, 38–43, doi:10.1038/nclimate3186.

Prein, A. F., and Coauthors, 2015: A review on regional convection-permitting climate modeling: Demonstrations, prospects, and challenges. *Rev. Geophys.*, **53**, 2014RG000475, doi:10.1002/2014RG000475.

Randall, D. A., and B. A. Wielicki, 1997: Measurements, Models, and Hypotheses in the Atmospheric Sciences. *Bull. Am. Meteorol. Soc.*, **78**, 399–406, doi:10.1175/1520-0477(1997)078<0399:MMOHIT>2.0.CO;2.

Raupach, M. R., and Coauthors, 2014: Sharing a quota on cumulative carbon emissions. *Nat. Clim. Change*, **4**, 873–879, doi:10.1038/nclimate2384.

Reusch, D. B., R. B. Alley, and B. C. Hewitson, 2005: Relative Performance of Self-Organizing Maps and Principal Component Analysis in Pattern Extraction from Synthetic Climatological Data. *Polar Geogr.*, **29**, 188–212, doi:10.1080/789610199.

Riahi, K., and Coauthors, 2011: RCP 8.5—A scenario of comparatively high greenhouse gas emissions. *Clim. Change*, **109**, 33, doi:10.1007/s10584-011-0149-y.

Ricke, K. L., J. B. Moreno-Cruz, J. Schewe, A. Levermann, and K. Caldeira, 2016: Policy thresholds in mitigation. *Nat. Geosci.*, **9**, 5–6, doi:10.1038/ngeo2607.

Rogelj, J., and R. Knutti, 2016: Geosciences after Paris. *Nat. Geosci.*, **9**, 187–189, doi:10.1038/ngeo2668.

Rowlands, D. J., and Coauthors, 2012: Broad range of 2050 warming from an observationally constrained large climate model ensemble. *Nat. Geosci.*, **5**, 256–260, doi:10.1038/ngeo1430.

Rupp, D. E., S. Li, N. Massey, S. N. Sparrow, P. W. Mote, and M. Allen, 2015: Anthropogenic influence on the changing likelihood of an exceptionally warm summer in Texas, 2011. *Geophys. Res. Lett.*, **42**, 2014GL062683, doi:10.1002/2014GL062683.

Russo, S., A. F. Marchese, J. Sillmann, and G. Immé, 2016: When will unusual heat waves become normal in a warming Africa? *Environ. Res. Lett.*, **11**, 054016, doi:10.1088/1748-9326/11/5/054016.

Sanderson, B. M., R. Knutti, and P. Caldwell, 2015: Addressing Interdependency in a Multimodel Ensemble by Interpolation of Model Properties. *J. Clim.*, **28**, 5150–5170, doi:10.1175/JCLI-D-14-00361.1.

Sardeshmukh, P. D., G. P. Compo, and C. Penland, 2015: Need for Caution in Interpreting Extreme Weather Statistics. *J. Clim.*, **28**, 9166–9187, doi:10.1175/JCLI-D-15-0020.1.

Schaller, N., and Coauthors, 2016: Human influence on climate in the 2014 southern England winter floods and their impacts. *Nat. Clim. Change*, **advance online publication**, doi:10.1038/nclimate2927. http://www.nature.com/nclimate/journal/vaop/ncurrent/full/nclimate2927.html (Accessed April 12, 2016).

Schleussner, C.-F., and Coauthors, 2016: Science and policy characteristics of the Paris Agreement temperature goal. *Nat. Clim. Change*, **6**, 827–835, doi:10.1038/nclimate3096.

Schneider, S. H., 2009: *Climate Change Science and Policy*. Island Press, 545 pp.

Sedláček, J., and R. Knutti, 2014: Half of the world's population experience robust changes in the water cycle for a 2 °C warmer world. *Environ. Res. Lett.*, **9**, 044008, doi:10.1088/1748-9326/9/4/044008.

Seidel, D. J., Q. Fu, W. J. Randel, and T. J. Reichler, 2008: Widening of the tropical belt in a changing climate. *Nat. Geosci.*, **1**, 21–24, doi:10.1038/ngeo.2007.38.

Seneviratne, S. I., 2012: Climate science: Historical drought trends revisited. *Nature*, **491**, 338–339, doi:10.1038/491338a.

——, T. Corti, E. L. Davin, M. Hirschi, E. B. Jaeger, I. Lehner, B. Orlowsky, and A. J. Teuling, 2010: Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Sci. Rev.*, **99**, 125–161, doi:10.1016/j.earscirev.2010.02.004.

——, M. G. Donat, A. J. Pitman, R. Knutti, and R. L. Wilby, 2016: Allowable CO2 emissions based on regional and impact-related climate targets. *Nature*, **529**, 477–483, doi:10.1038/nature16542.

Shepherd, T. G., 2016: A Common Framework for Approaches to Extreme Event Attribution. *Curr. Clim. Change Rep.*, **2**, 28–38, doi:10.1007/s40641-016-0033-y.

Sippel, S., J. Zscheischler, M. Heimann, F. E. L. Otto, J. Peters, and M. D. Mahecha, 2015: Quantifying changes in climate variability and extremes: Pitfalls and their overcoming. *Geophys. Res. Lett.*, **42**, 2015GL066307, doi:10.1002/2015GL066307.

Steinskog, D. J., D. B. Tjøstheim, and N. G. Kvamstø, 2007: A Cautionary Note on the Use of the Kolmogorov–Smirnov Test for Normality. *Mon. Weather Rev.*, **135**, 1151–1157, doi:10.1175/MWR3326.1.

Stocker, T. F., and Coauthors, 2013: Technical Summary. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, T.F. Stocker et al., Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 33–115 www.climatechange2013.org.

Stone, D. A., and M. R. Allen, 2005: The End-to-End Attribution Problem: From Emissions to Impacts. *Clim. Change*, **71**, 303–318, doi:10.1007/s10584-005-6778-2.

Stott, P., 2016: How climate change affects extreme weather events. *Science*, **352**, 1517–1518, doi:10.1126/science.aaf7271.

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

——, and Coauthors, 2016: Attribution of extreme weather and climate-related events. *Wiley Interdiscip. Rev. Clim. Change*, **7**, 23–41, doi:10.1002/wcc.380.

Sutton, R., E. Suckling, and E. Hawkins, 2015: What does global mean temperature tell us about local climate? *Phil Trans R Soc A*, **373**, 20140426, doi:10.1098/rsta.2014.0426.

Swain, D. L., 2015: A tale of two California droughts: Lessons amidst record warmth and dryness in a region of complex physical and human geography. *Geophys. Res. Lett.*, **42**, 2015GL066628, doi:10.1002/2015GL066628.

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

Tebaldi, C., and J. M. Arblaster, 2014: Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Clim. Change*, **122**, 459–471, doi:10.1007/s10584-013-1032-9.

Thompson, A., and F. E. L. Otto, 2015: Ethical and normative implications of weather event attribution for policy discussions concerning loss and damage. *Clim. Change*, **133**, 439–451, doi:10.1007/s10584-015-1433-z.

Thomson, A. M., and Coauthors, 2011: RCP4.5: a pathway for stabilization of radiative forcing by 2100. *Clim. Change*, **109**, 77, doi:10.1007/s10584-011-0151-4.

Thornton, J., and H. Covington, 2016: Climate change before the court. *Nat. Geosci.*, **9**, 3–5, doi:10.1038/ngeo2612.

Trenberth, K. E., J. T. Fasullo, and T. G. Shepherd, 2015: Attribution of climate extreme events. *Nat. Clim. Change*, **5**, 725–730, doi:10.1038/nclimate2657.

Tukey, J. W., 1977: *Exploratory Data Analysis*. Addison-Wesley Publishing Company, 714 pp.

Uhe, P., F. E. L. Otto, K. Haustein, G. J. van Oldenborgh, A. D. King, D. C. H. Wallom, M. R. Allen, and H. Cullen, 2016: Comparison of methods: Attributing the 2014 record European temperatures to human influences. *Geophys. Res. Lett.*, **43**, 2016GL069568, doi:10.1002/2016GL069568.

Vautard, R., and Coauthors, 2015: Extreme Fall 2014 Precipitation in the Cévennes Mountains. *Bull. Am. Meteorol. Soc.*, **96**, S56–S60, doi:10.1175/BAMS-D-15-00088.1.

——, P. Yiou, F. Otto, P. Stott, N. Christidis, G. J. van Oldenborgh, and N. Schaller, 2016: Attribution of human-induced dynamical and thermodynamical contributions in extreme weather events. *Environ. Res. Lett.*, **11**, 114009, doi:10.1088/1748-9326/11/11/114009.

Vuuren, D. P. van, and Coauthors, 2011a: The representative concentration pathways: an overview. *Clim. Change*, **109**, 5, doi:10.1007/s10584-011-0148-z.

——, and Coauthors, 2011b: RCP2.6: exploring the possibility to keep global mean temperature increase below 2°C. *Clim. Change*, **109**, 95, doi:10.1007/s10584-011-0152-3.

Ward, D. S., and N. M. Mahowald, 2014: Contributions of developed and developing countries to global climate forcing and surface temperature change. *Environ. Res. Lett.*, **9**, 074008, doi:10.1088/1748-9326/9/7/074008.

Weber, E. U., 2010: What shapes perceptions of climate change? *Wiley Interdiscip. Rev. Clim. Change*, **1**, 332–342, doi:10.1002/wcc.41.

Williams, A. P., R. Seager, J. T. Abatzoglou, B. I. Cook, J. E. Smerdon, and E. R. Cook, 2015: Contribution of anthropogenic warming to California drought during 2012–2014. *Geophys. Res. Lett.*, **42**, 2015GL064924, doi:10.1002/2015GL064924.

Yiou, P., R. Vautard, P. Naveau, and C. Cassou, 2007: Inconsistency between atmospheric dynamics and temperatures during the exceptional 2006/2007 fall/winter and recent warming in Europe. *Geophys. Res. Lett.*, **34**, L21808, doi:10.1029/2007GL031981.

Zargar, A., R. Sadiq, B. Naser, and F. I. Khan, 2011: A review of drought indices. *Environ. Rev.*, **19**, 333–349, doi:10.1139/a11-013.

Appendix

Chapter 2: Supplementary information

Consideration of different error metrics and SOM node configurations

As discussed in Section 2.3, the mean absolute error (MAE) metric was chosen for this attribution analysis, as this formula provides an absolute measure of model error and gives equal weight to errors from all SOM node types. However, Pearson correlation scores and Spearman rank correlations were also considered. We also considered the results when interpreting the data using different-sized SOM configurations: 5x4 and 6x5 node configurations were tested as proposed alternatives.

As with Figure 2.9 of the main thesis, Figures S2.1-S2.3 show the probability distributions of each of the three SOM metrics (left to right: frequency, average lifetime, maximum lifetime) for each SOM configuration, for each of the three possible error metrics: MAE, Pearson R correlation and Spearman's rank correlation. It is evident from these figures that MAE provided the most consistent comparisons between the different SOM node configurations: those SOM metrics which demonstrated a statistically significant difference between the early-climate and recent-climate PDFs (bold p-values inset for each panel) under the 4x3 configuration using MAE also showed similar levels of statistical significance for the 5x4 and 6x5 configurations (with the exception of Figure S2.1d). By contrast, when using Pearson or Spearman correlations, whether or not statistically robust changes were found between the early- and recent-climate PDFs were dependant on the size of the SOM node configuration used: the p-values denoted displayed much greater variations.

It is for this reason that we have chosen to use MAE as the error metric of choice in the main article. Furthermore, we choose to focus on a 4x3 SOM node configuration, though Figure S2.4 shows the key results found in Figure 2.12 of the main article are comparable even when different SOM configurations are considered.
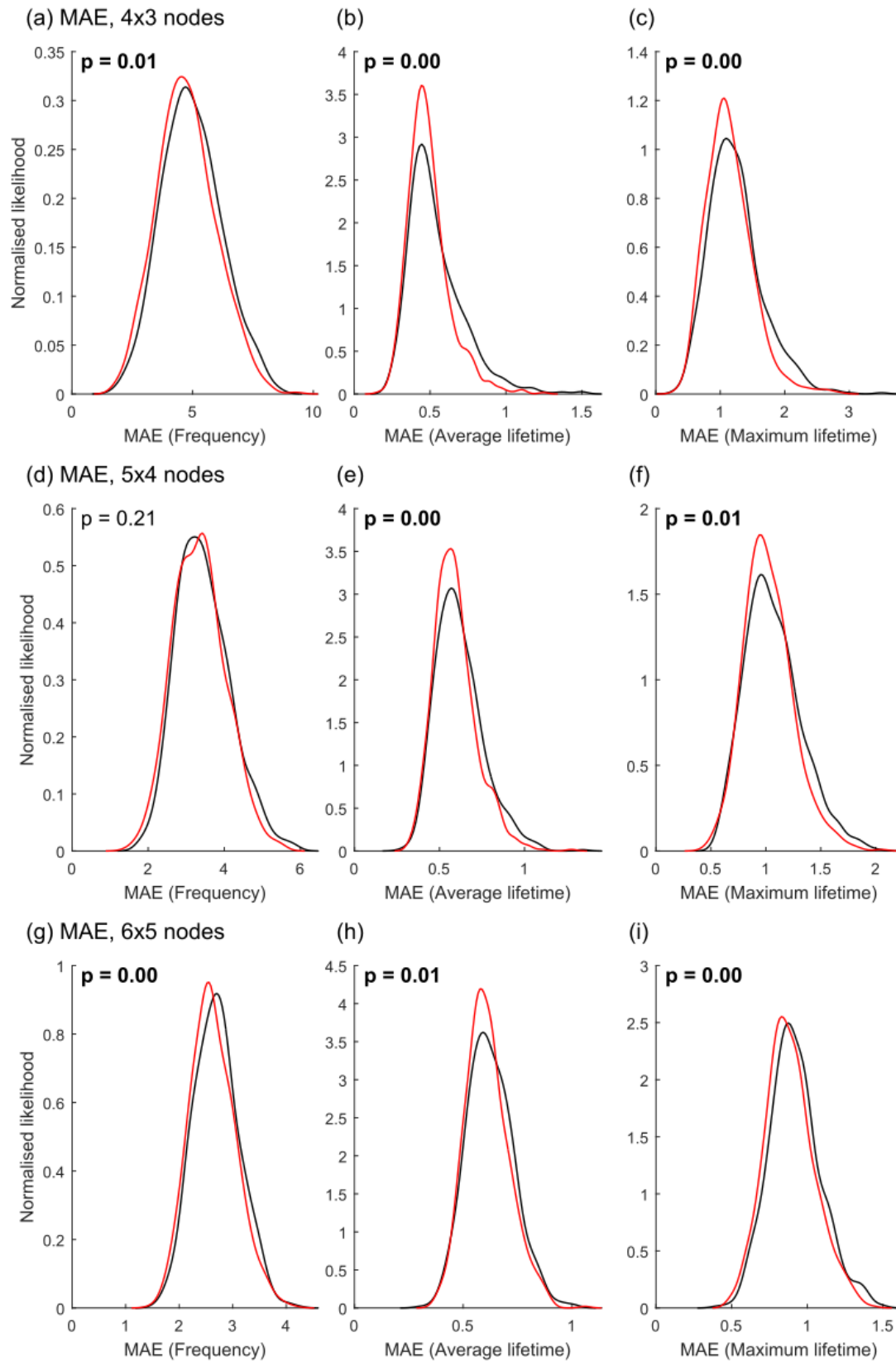
Figure S2.1: Same as Figure 2.9, but repeated for not only the default 4x3 SOM node configuration (top) but also a 5x4 (middle panels) and 6x5 (bottom) SOM node configuration. Inset: p-values corresponding to a two-sampled K-S test of the relevant metric, bold colours indicate p<0.05 denoting a statistically significant difference between the early- and recent-climate distributions.
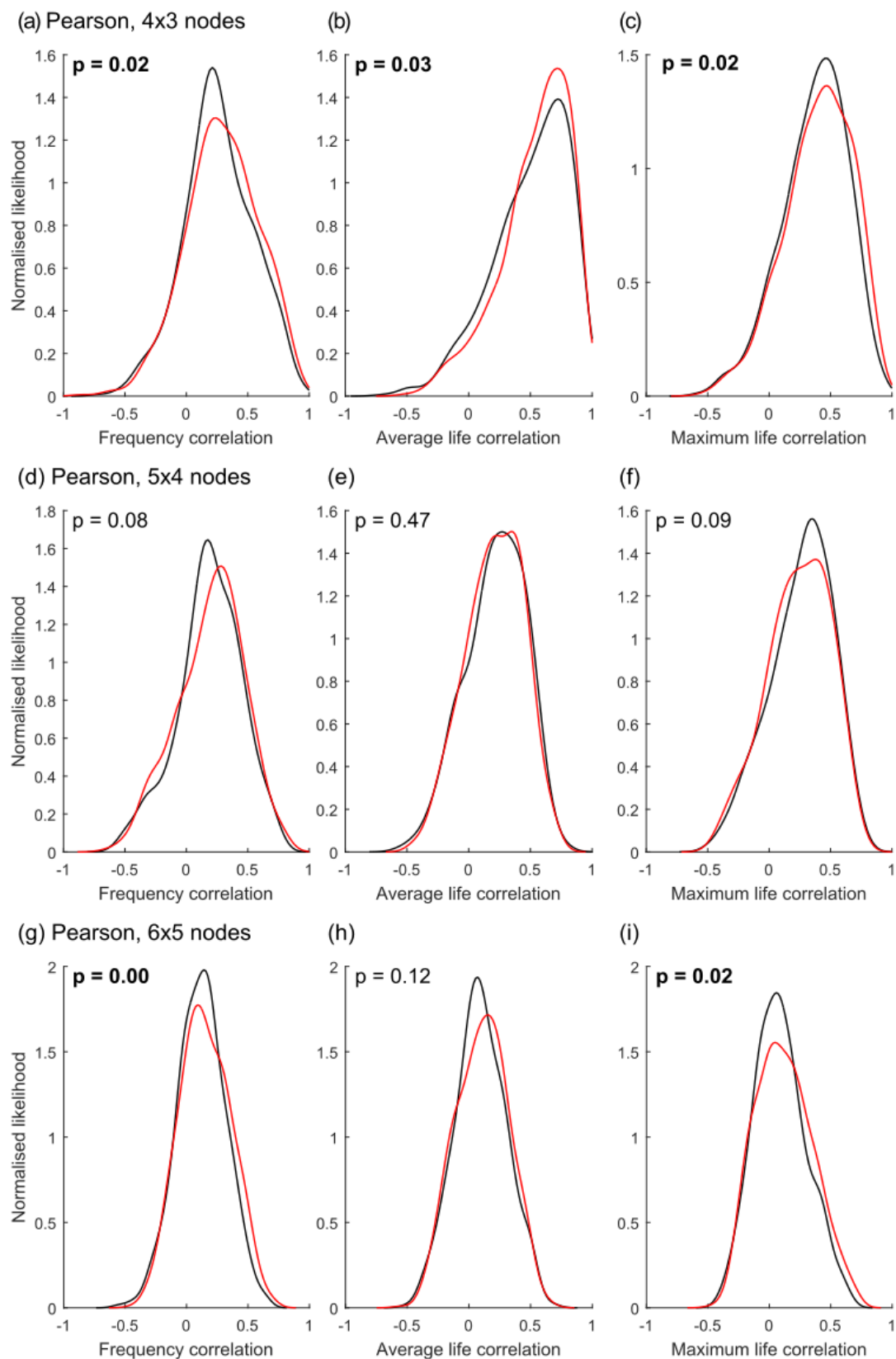
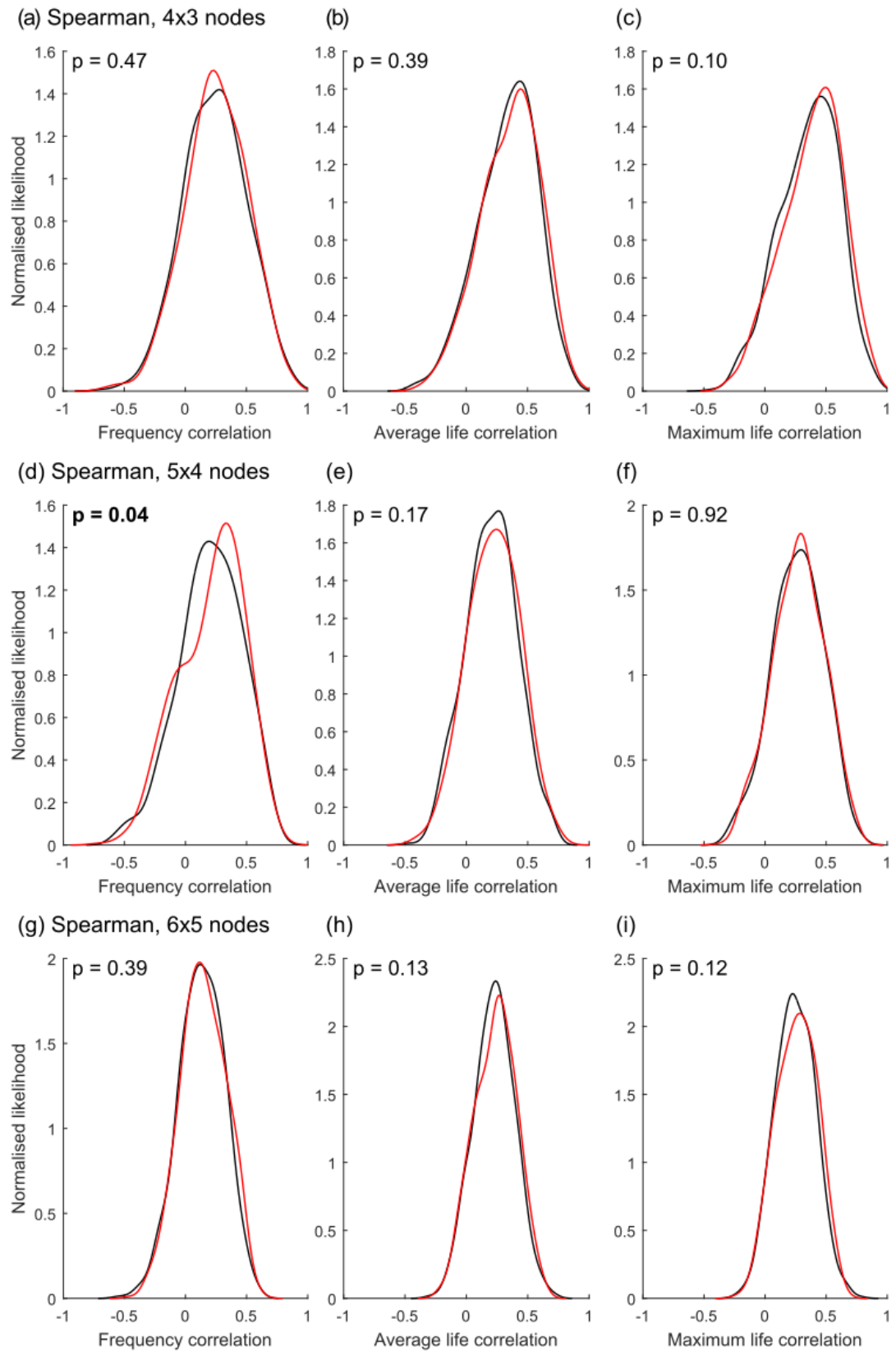**Figure S2.2**: Same as Figure S2.1, but using Pearson R correlation scores as the error metric of choice.

Figure S2.3: Same as Figure S2.1, but using Spearman rank correlations as the error metric of choice.
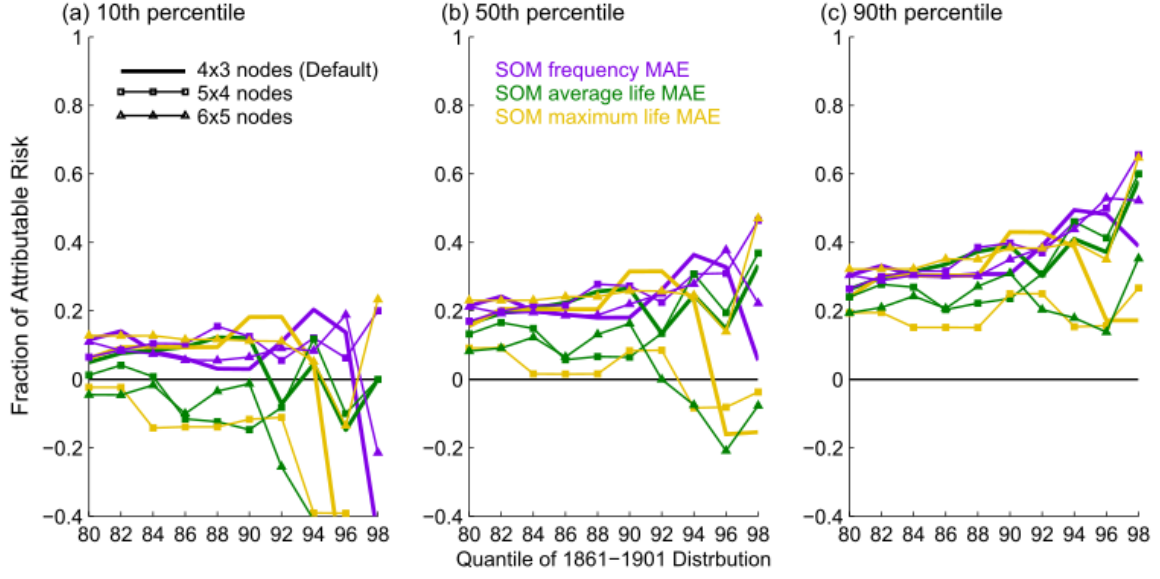
Figure S2.4: Same as Figure 2.12, but this time showing the bootstrapped 10th (left), 50th (middle) and 90th (right) percentiles of the bootstrapped FAR estimates, for the MAE-based metrics, with the analysis having been repeated using a 5x4 and 6x5 SOM node configuration.

Chapter 3: Supplementary information

The framework for analysis presented in Section 3.3 is herein repeated for a Gaussian distribution. Equations S3.1 and S3.1 are analogous to equations 3.1 and 3.2 in the main text. Similarly, Figure S3.1 presents two Gaussian distributions, f_0 and f_1, according to the formula,

$$f_n(x) = \frac{1}{\sigma_n\sqrt{2\pi}} \exp\left(\frac{-(x - \mu_n)^2}{2\sigma_n^2}\right) \qquad \text{Eq. S3.1}$$

where μ and σ respectively denote the mean and standard deviation of the Gaussian distribution. f_0 has a fixed mean of $\mu_0=0$, while the standard deviation of both distributions are set to 1 for simplicity. Thus, the signal-to-noise ratio of f_n for these Gaussian distributions are defined as,

$$S_n = \frac{\mu_n}{\sigma_n} \qquad (S2)$$

A variety of possible values are considered for the value of μ_1. Since we have made $\sigma_0 = \sigma_1$, we consider S hereafter in general units of σ. We can therefore now consider f_1 as a model ensemble of simulations representing the present-day climate with an anthropogenic signal, S, while f_0 represents an ensemble of counterfactual model simulations where anthropogenic influences have been removed (since we have defined $S_0 = 0$).

All other definitions of PR_{EAC}, PR_{PD} and TOD remain unchanged from the definitions presented in Section 3.3 of the main thesis.
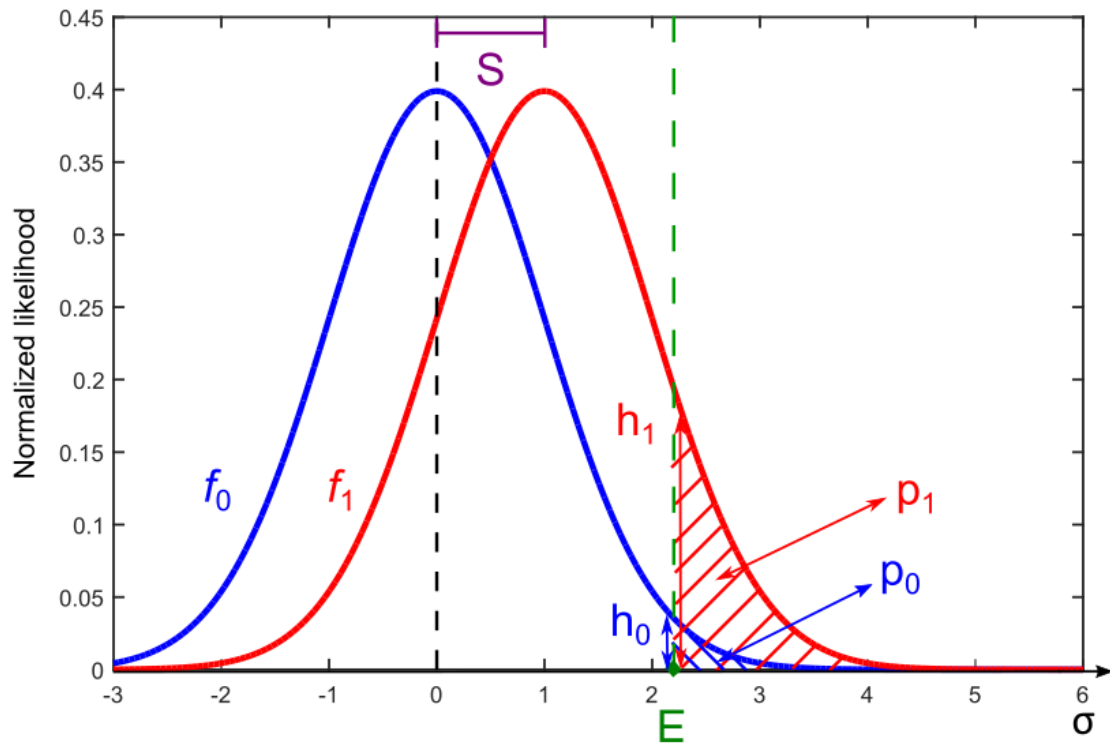

Figure S3.1: Same as Figure 3.2 in the main text, but presenting a counterfactual (f_0) and factual (f_1) PDFs of Gaussian data instead of Gumbel distributions.
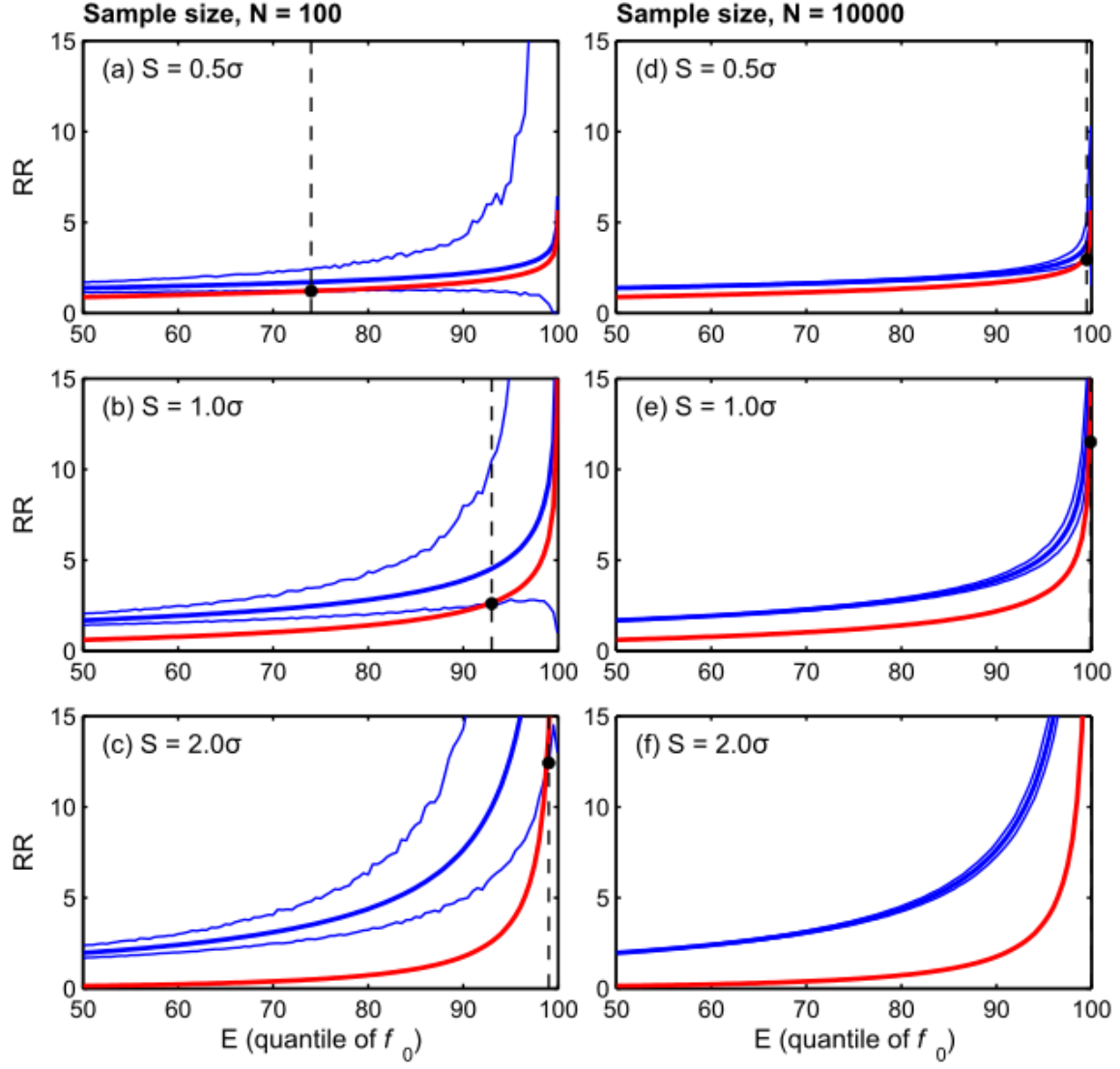
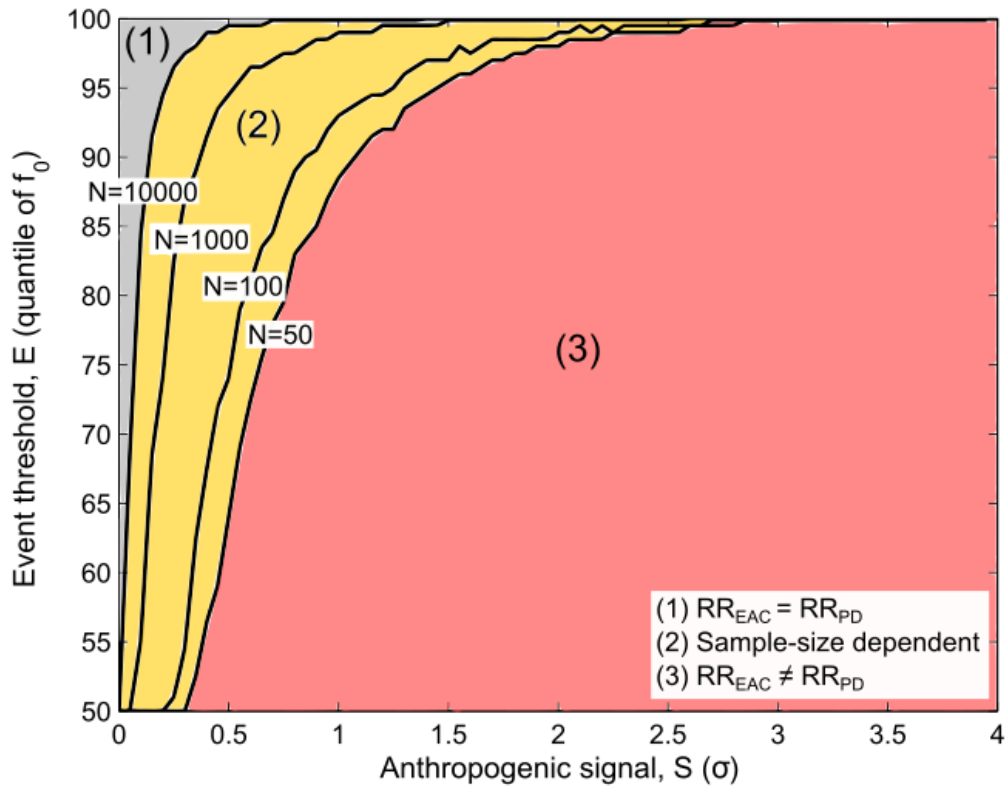**Figure S3.2**: Same as Figure 3.3 of the main text, but instead considering the Gaussian PDF examples.

Figure S3.3: Same as Figure 3.4a of the main text, but instead considering the Gaussian PDF examples.

Chapter 4: Supplementary information

Consideration of biases as a result of quintile selection

While the results presented in Chapter 4 show strong differences between the wealthiest and poorest quintiles of the global population, it is important to check that these results were not just some unusual artefact of the specific quantile ranges selected. Figure S4.1 compares the emergence of population exposure for all five socio-economic quintiles (Q0-Q20 denotes the poorest quintile; Q80-Q100 denotes the wealthiest quintile) with corresponding cumulative emissions, with respect to the timing of exposure of the entire global population as a whole. This figure not only reaffirms the results of the main article, demonstrating that the differences in risk emergence are most distinct between the wealthiest and poorest populations, but also clarifies that the timing of population exposure is slower as you consider each progressively wealthier quintile.
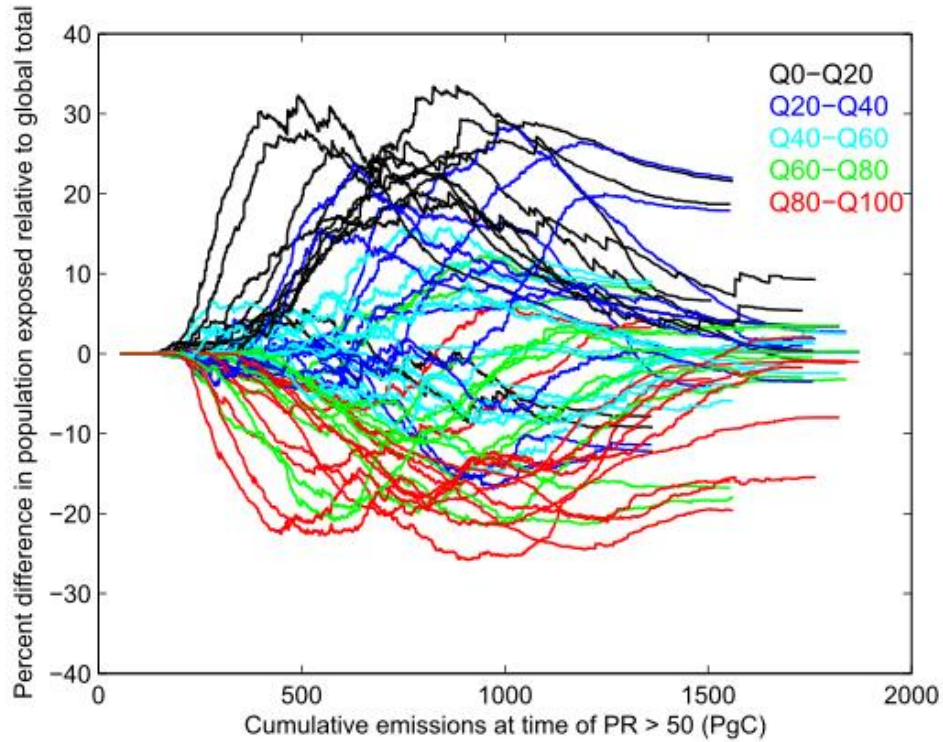
Figure S4.1: Same as Figure 4.8c in main text, but showing the differences in exposure CDF's of all five population quintiles with respect to the global total population

Differences between models, RCP scenarios, and temperature versus CO_2

Figures S9 helps to demonstrate variations in the timing of PR threshold exceedances between different models (Fig S9). Figure S9 shows the difference between the third-highest and third-lowest model CEPRE response at each grid point, so to demonstrate the magnitude of inter-model variability implicit in the results presented in Figure 4.4. It is evident that there is stronger agreement between models on CEPRE estimates for lower latitudes than higher latitudes across all three probability ratio thresholds, which is consistent with the findings of previous studies (Hawkins *et al.* 2014, Lyu *et al* 2014).

(a) PR > 2

(b) PR > 10
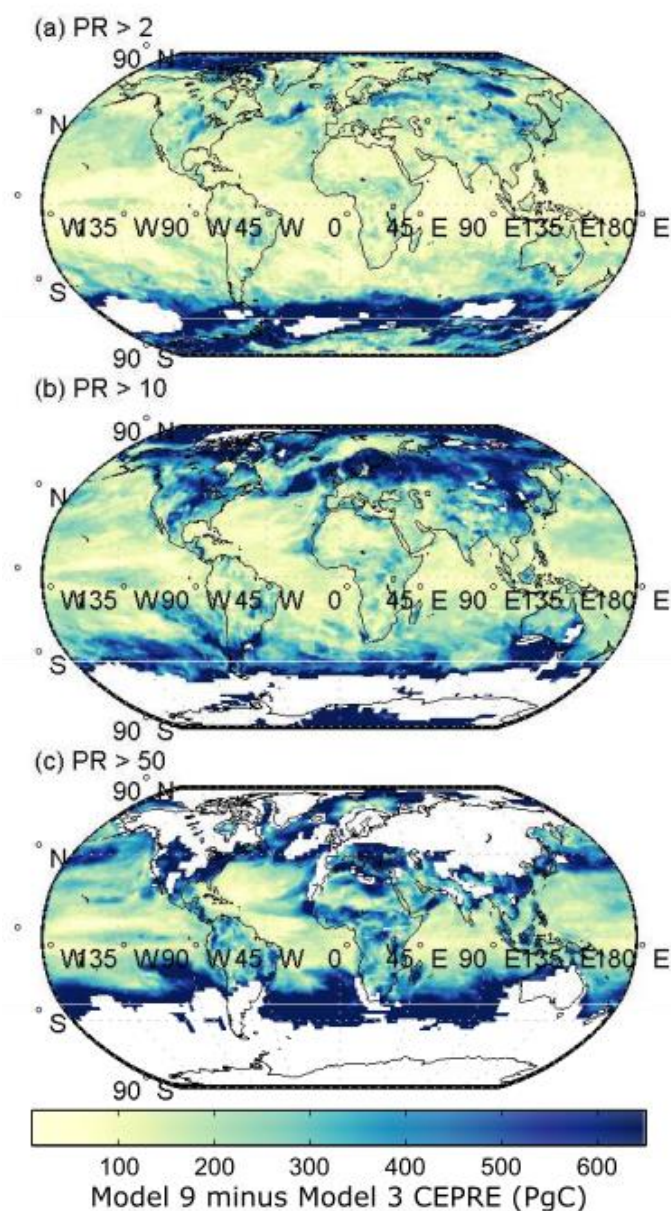
(c) PR > 50

Model 9 minus Model 3 CEPRE (PgC)

Figure S4.2 - Same as the left hand panels of Figure 4.4 in the main article, except instead of displaying the median model response (number 6 of 11) of CEPRE at each grid cell, this figure presents the difference between the third-highest and third-lowest model CEPRE value (corresponding to models 9 and 3 out of 11, respectively). This could be interpreted as an analogue for the interquartile range of the CEPRE model estimates. White regions indicate where at least 3 out of 11 models show no exceedance of the respective PR threshold by 2100.

Chapter 5: Supplementary information

Hereafter, we present further details about the data sources and methodologies used for the analyses presented in Chapter 5.

S/N emergence data

To quantify the emergence of climate signals under future warming, we consider monthly surface air temperature data from only those 25 climate models which provided output for the

154

Coupled Model Intercomparison Project 5 (CMIP5, Taylor et al. 2012) over the period 2006-2100 for three Representative Concentration Pathways: RCP2.6, RCP4.5 and RCP8.5. Only 'r1i1p1' ensemble members are considered. To compute the climate change signal, temperature increases for each model at each gridpoint are regressed against the corresponding global mean increases in temperature for model simulations across the twentieth and twenty-first century ('Historical' simulations were used over 1901-2005), and then normalized with respect to a 1986-2005 climatology. The measure of noise was taken as the interannual standard deviation of annual means in temperature across the pre-industrial control simulations of each climate model. We present end-of-21st-century estimates for the signal-to-noise ratios in temperature by averaging over the years 2069-98 for each model. All models have the signal-to-noise ratios calculated on their native spatial grids and then interpolated to a common 2.5° x 2.5° spatial resolution.

Present-day gridded population

Population and national boundary data for the year 2015 are taken at 0.25° x 0.25° resolution from the Gridded Population of the World version 4 (GPWv4) dataset, administered by the NASA Socioeconomic Data and Applications Center (SEDAC) and hosted by the Center for International Earth System Information Network (CIESIN, doi: http://dx.doi.org/10.7927/H4X63JVC). The national boundary mask from CIESEN is overlaid on the 2015 population dataset to obtain the number of residents in each 0.25° grid box (both the population data and the national boundary data have this resolution). Where there are multiple countries within a dataset we apportion the population evenly between them. This introduces small inaccuracies, but these are negligible at the inter-national and normalized scales on which the paper focuses. We then aggregate the appropriate national population data onto the 2.5° x 2.5° resolution grid of the climate model output. We do this for each of the CMIP5 models for which data are available under the three scenarios considered here. This gives us both a number of people (in each climate model grid box) and an S/N ratios (for each climate model grid box) for each model and for each scenario.

Socio-economic data

For each country and where available, per capita gross national income (GNI) data was provided for the year 2015 by The World Bank. GNI data has been converted into current international dollars using purchasing power parity rates, as of August 10, 2016. Data are available at http://data.worldbank.org/.

CO₂ emissions data

Fossil fuel CO_2 emissions data for 1751-2011 are provided by the Carbon Dioxide Information Analysis Center (CDIAC) at Oak Ridge National Laboratory (10.3334/CDIAC/00001_V2015). Where available, the national estimates of CO_2 emissions reported by countries to the UNFCCC during 1990-2012 are overwritten. The 2012, 2013 and 2014 estimates (or 2013 and 2014 where UNFCCC data are available) are preliminary and based on energy statistics published by BP (http://www.bp.com/en/global/corporate/about-bp/energy-economics/statistical-review-of-world-energy.html).

Future population scenarios

We have taken country-specific population data for three Shared Socioeconomic pathways, available at the SSP database, Version 1 (https://secure.iiasa.ac.at/web-apps/ene/SSPDB). We take data for 2010 and 2100 for individual countries under SSP1, SSP2 and SSP3 scenarios, using the Wittgenstein Center projections (labelled as "IIASA-WiC Population V9" in the SSP-Database). SSPs 1-3 span the range of possible futures related to future population projections based on expert assessments and with the use of Integrated Assessment Models. For further information on the population projections used for this study, the reader is referred to (KC and Lutz 2014) for extensive details. We don't use the population estimates from the Integrated Assessment Models directly, but instead calculate the percentage change in population between 2010 and 2100 for each country under each SSP scenario, and then apply these fractional changes to the high-resolution 2015 CIESEN population dataset, to quantify future possible population ranges.

Countries considered in this analysis

We note that we only present results for those 170 countries for which there is information available on future population scenarios, socio-economic information and CO_2 emissions. The combined population of these countries represents 97.38% of the global population as a whole.