Improving Understanding of Precipitation with Medical Image Registration

Adam Aaron Lawrence Levy

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

2015

Acknowledgements

First and foremost I would like to thank Mark, Myles and William for their fantastic supervision. They have each provided me with huge support, resources, and encouragement in my studies, and have complemented each other superbly. It is no exaggeration to say that this thesis would have suffered immensely if any one of them had not been on board. I would also like to thank Hugo Lambert and Chris Huntingford for the invaluable additional supervision they have provided, not only helping me with difficult academic questions, but also providing much needed context for understanding the academic domain.

I would also like to acknowledge the invaluable advice I have received from a number of people. In particular, I thank Rachel James for her fantastic input on the simulation of African precipitation and perturbed physics ensembles; Richard Millar for his experience with the CPDN ensemble, as well as for forwarding me noteworthy manuscripts; Robin Chadwick and Gill Martin for their early input on the applications of our techniques; Eric Gilleland, Aurel Moise and Fabrice Chauvin for their thoughts on climatological precipitation warping; and Mihai Duta and Albert Solernou at Oxford's Advanced Research Computing (ARC) for their advice on the correct use of a cluster.

I have also benefited hugely from input I have received from reviewers on manuscript submissions; from attendees at seminars I have presented; and from my examiners for my transfer and confirmation of status (Laure Zanna, Anu Dudhia and David Andrews). Special thanks to my viva examiners, Lesley Gray and Gabi Hegerl, not only for their insightful comments, but also for making the effort to put me at ease and so make the examination process an enjoyable one.

In addition, I would like to thank Sarah Harrington, without whom it is unlikely that I would have been able to apply for this doctorate, let alone hand it in.

I cannot imagine that I would ever have done a doctorate without the fantastic education I received prior to it, that both educated and inspired me in Physics. This began with Peter Sammut, who taught me the intrinsic beauty of science, and continued to grow thanks to my fantastic undergraduate tutors: John March-Russell and Laura Lauro-Taroni.

I am also hugely thankful for the support of a great number of friends, whose words have been both encouraging and inspiring. There are too many to name here, and so I will name just two. I am indebted to Sam Greenbury for inspiring me to search for the truth as honestly as possible, and showing me how exciting that search can be. I am also grateful for the support Siân Robins-Grace provided through the highs and the lows of the last three years. Thanks to her, I never once felt alone on this journey.

Finally, I would like to express my eternal gratitude to my parents and brother. Thank you not only for supporting me through this degree, but for all that led up to it, and will follow from it.

Abstract

Changes to the hydrological cycle under global warming are expected to involve both intensification of the hydrological cycle and changes in the location of key features. Thus, differences between General Circulation Models (GCMs) in simulated local precipitation changes may arise in part from location biases in the models' mean climates. Combined with the large internal variability and short spatial scales of precipitation, as well as the limited observational record, such biases would exacerbate the difficulty of robustly evaluating forced regional precipitation changes. While many techniques exist to correct biases in local intensity distributions, few have dealt with errors in location.

In this thesis, therefore, we make use of a technique from medical image registration to remove location biases from simulated precipitation fields, and demonstrate that this reduces disagreement on projected changes. We therefore proceed to introduce a tool tailored towards the removal of location biases from GCMs. We use this tool not only to investigate historical precipitation changes - showing that it marginally increases the detectability of the historical anthropogenic forcing on precipitation - but also to investigate the physical origin of precipitation biases. In particular, we compare the transformations found for different GCMs, and investigate how location biases in the West African Monsoon vary across a perturbed physics ensemble. This study enables us to identify a potential mechanism relating the location of the convergence zone to the parameterisation of the GCM.

By demonstrating the utility of such techniques, we hope that they will continue to be developed and applied in the Atmospheric Sciences, and that other applications may prove valuable.

Contents

Chapter 1

Introduction

1.1 Global Precipitation Changes

It follows from the Clausius-Clapeyron equation that the water vapour content of the atmosphere increases exponentially with temperature, provided relative humidity remains fixed (e.g. *Allen and Ingram*, 2002). Changes in precipitation, however, are harder to constrain. Although extreme precipitation events are dependent on water vapour availability, as all the moisture in a column may be rained out in such an event (*Min et al.*, 2011; *Pall et al.*, 2007), changes in globally and annually averaged precipitation will be determined primarily by energetic constraints.

As latent heat is absorbed by evaporation and released by precipitation, the hydrological cycle effectively transfers energy from the surface to the upper troposphere, and so the net change in global precipitation will depend on the nature of the warming. As *O'Gorman et al.* (2012) notes, conservation of energy yields:

$$L\Delta P = \Delta R_{TOA} - \Delta R_{SURF} - \Delta U_{SENS} \tag{1.1}$$

where L is the latent heat of evaporation, P is the amount of global precipitation, R_{TOA} and R_{SURF} are the radiative fluxes at the top of atmosphere and the surface,

respectively, U_{SENS} is the sensible heat flux at the surface, and the changes are in response to an external forcing.

From this relationship, it becomes clear that precipitation changes will depend on how any warming is brought about, with marked differences between changes in the incoming short-wave, or outgoing long-wave radiation. For warming induced by carbon-dioxide, where outgoing long-wave radiation is absorbed and re-emitted in the troposphere, top of atmosphere fluxes will be reduced. This has the effect of reducing the precipitation change relative to that one would expect from the surface warming alone, leading to differing ratios of precipitation change to temperature change (the 'hydrological sensitivity') than are brought about by a change in the solar constant - either through a change in solar luminosity, or the direct effect of stratospheric aerosols (*Andrews et al.*, 2009). Such differences in hydrological sensitivity would have profound consequences for any geoengineering carried out to reduce the solar constant (e.g. using sulphate aerosols) to counter the warming resulting from greenhouse gases (*Trenberth and Dai*, 2007).

Several studies have stressed that the precipitation response to a forcing can be split into two components: a 'fast' response to the forcing agent itself; and a 'slow' response to the temperature change of the lower troposphere, which takes place on longer timescales due to the need for the oceans to equilibrate (e.g. *Lambert and Faull*, 2007). The fast response can be disentangled from the slow response using simulations with fixed sea surface temperatures (SSTs), and varies greatly with the nature of the forcing (*Andrews et al.*, 2010). The slow response per degree Kelvin of temperature change, on the other hand, is found to be broadly independent of the forcing agent.

Due to the timescales required for SSTs to equilibrate, as well as the large inter-decadal variability of precipitation (*Dai*, 2012), it is not straightforward to calculate the climate's equilibrium hydrological sensitivity to carbon-dioxide from observations.

Indeed, global precipitation observations are only available through satellite retrievals since 1979 (*Huffman et al.*, 2009). Nevertheless, estimates using this data imply a hydrological sensitivity of roughly 2% per K of warming, though this result is found to be strongly sensitive to the time period used (*Adler et al.*, 2008). Nevertheless, it is in agreement with simulations, which have consistently indicated a range of around 2-3% per K over the last two generations of GCMs (*Hegerl et al.*, 2014; *Lambert and Webb*, 2008).

1.2 Regional Precipitation Changes

While such global constraints are valuable for understanding the energy balance of the planet, they are of little use in informing adaptation or mitigation decisions, which rely on information regarding regional, seasonal changes. The energetic considerations used above can be extended such that they describe regional, rather than global changes, through the inclusion of changes in horizontal transports of energy into the region. (*Muller and O'Gorman*, 2011) demonstrate that such considerations allow regional precipitation changes to be reasonably approximated based on only changes in horizontal energy transport and changes to global mean precipitation amount.

On regional scales, however, it has been more common to consider changes in local imbalances between precipitation and evaporation. It is argued that if circulations remain unchanged, and yet the water vapour content of the atmosphere increases, more water vapour will be transported away from regions that already have more evaporation than precipitation, and vice versa (*Chou et al.*, 2009; *Held and Soden*, 2006). As a result - all else being equal - this will strengthen the contrast between regions with evaporation and precipitation imbalances, and so this process is dubbed the 'rich get richer' or 'wet get wetter' mechanism.

Although this mechanism adequately describes changes in zonal means (e.g. *Sea-*

ger et al., 2010) other effects may dominate in certain regions. Most obviously, this relationship will break down where it implies increased evaporation relative to precipitation over land, yet there is limited moisture availability. Indeed, *Greve et al.* (2014) shows that almost as much land area experiences the opposite pattern (i.e. 'wet get drier' and vice versa) as experience an intensification of they existing hydrological cycle. By decomposing precipitation changes into components related to different mechanisms, *Chadwick et al.* (2012) demonstrates that the spatial pattern of the rich-get-richer mechanism is largely cancelled out by weakening of the circulation in the tropics. The spatial pattern of precipitation change in this region is therefore dominated by shifts in convergence zones.

Over long enough time periods, with a sufficiently large forcing, such shifts can have dramatic effects. While many descriptions of precipitation changes assume regions respond broadly linearly to temperature changes, climatological shifts over time may be large enough that substantial non-linear responses are introduced. Indeed, under such circumstances, some regions that initially become wetter may subsequently dry (*Hawkins et al.*, 2014).

1.2.1 Projected Precipitation Changes

It is clearly desirable to investigate the extent to which such considerations of regional precipitation changes can be translated into meaningful, robust projections for future precipitation. While simulated projections imply that a warming of 2°C would affect the local water cycle for more than half of the world's population (*Sedláček and Knutti*, 2014), nearer term projections of precipitation are less robust.

Confidence in projections is often assessed by evaluating agreement between GCMs from differing institutions simulating the same forcing scenario (although there will be differences in the extent to which different GCMs can simulate all the forcing agents specified), as direct validation relative to observations cannot be achieved as

would be possible for weather forecasts, for example. The uncertainty in projected precipitation changes was highlighted by Figure 10.9 of *Meehl et al.* (2007), which simultaneously depicted the multi-model mean projections and agreement on the sign of the precipitation change. Much of the Figure lacked colour, indicating that there is still great uncertainty on whether many regions will become wetter or drier as the world continues to warm.

This uncertainty is partly due to physically-based disagreements between GCMs, either in their simulated changes or their mean climatology (see Section 1.3.1), as evidenced by different GCMs projecting large increases or decreases in precipitation for the same region (e.g. *Roehrig et al.*, 2013, - see Chapter 6). However, some disagreements on sign may be attributed to relatively small changes in some regions. In areas where projected changes are small relative to natural variability, one would not necessarily expect agreement on sign between simulations (*Levy et al.*, 2013; *Tebaldi et al.*, 2011). *Kirtman et al.* (2013) therefore considers uncertainty in projections relative to each region's natural variability, but again demonstrates a great deal of uncertainty in precipitation changes over the next few decades, especially in comparison to temperature changes. Further, *Knutti and Sedlacek* (2013) emphasises that agreement between GCMs has changed little in the six years between the Coupled Model Intercomparison Project Phase 3 (CMIP3) and CMIP5, in spite of a 60-fold increase in computational power over that time.

Scheff and Frierson (2012) notes that the nature of projected precipitation changes can be hard to discern due to differences between GCMs' climatologies, with different GCMs simulating the same features at different locations. They therefore assess projected changes across an ensemble by investigating each model's precipitation changes relative to that model's historical mean climate, allowing direct comparison between processes in the GCMs to be made. Using this technique they are able to decompose poleward shifts of precipitation bands from the wet get wetter mechanism,

and so are able to demonstrate that the former is primarily responsible for the robust projection of drying in the subtropics over the twenty-first century.

Fischer et al. (2014) show that much of the disagreement between GCMs arises due to natural variability. Once natural variability is taken into account, GCMs agree well on changes in projected precipitation extremes, with most Northern Hemisphere land seeing increases in the intensity of extreme precipitation events. Projections in mean precipitation are less robust than in extremes, however, due to the larger signal expected for extreme events, which roughly follow increases in water vapour concentrations (as discussed in Section 1.1).

1.2.2 Historical Precipitation Changes

While understanding projected precipitation changes under climate change is essential for informing mitigation and adaptation decisions, it is also important to investigate the nature of historical precipitation changes. Not only does study of historical changes allow us to investigate how an external forcing affects precipitation, but it also provides the opportunity to validate GCM simulations.

Investigation of such historical changes in precipitation has proved difficult. There are a number of reasons for this. Precipitation distributions vary over much shorter length scales than temperature. Unlike temperature, which has increased almost everywhere over the past century (*Hartmann et al.*, 2013), precipitation is expected to increase the contrast between wet and dry regions (*Chou et al.*, 2009) while shifting climatological features (*Scheff and Frierson*, 2012), and so any forced changes in precipitation will also vary over small length scales.

These problems, which are also faced when investigating projections, are exacerbated when investigating historical precipitation changes by the limited observational record for precipitation. In general, studies investigating historical precipitation must choose between relatively sparse land-based datasets (e.g. *Fuchs et al.*, 2007), or rel-

atively short satellite/gauge based observations (*Huffman et al.*, 2009).

Central to the analysis of the historical climate are efforts to detect an external influence on an observed field, and to attribute that influence to a particular forcing (*Stott et al.*, 2001). Such techniques were developed in *Hasselmann* (1979) and *Allen and Tett* (1999), and first used to demonstrate that observed patterns of warming were consistent with anthropogenic forcings (*Hegerl et al.*, 1997). Detection and attribution is generally carried out by projecting observed patterns onto expected patterns, or 'fingerprints', which provide an estimate of the changes expected due to a particular influence (see Section 5.3). *Zhang et al.* (2007) used a land-based dataset to investigate whether an anthropogenic influence could be detected and attributed for precipitation. By investigating zonally, annually averaged latitudinal bands, they demonstrated that human influence had indeed influenced precipitation beyond changes expected from natural forcing or internal variability of the climate system. *Polson et al.* (2013a) investigate further the extent to which this signal is detectable in various land-based datasets by investigating changes in each season independently, again in zonally averaged, latitude bands. This study shows that the signal of an external forcing can be detected for March - May and December - February, while providing evidence that some of this signal is driven by human influences.

Both of these studies make use of land-based datasets, as they provide long time-series. This is especially useful for the study of precipitation due to its large internal (i.e. unforced) variability relative to temperature, which is influenced by the Pacific Decadal Oscilation (PDO) (*Dai*, 2012) and - most notably - the El-Niño Southern Oscillation (ENSO) (*Dai and Wigley*, 2000). Land-based datasets therefore allow a higher signal to noise to be achieved. Nevertheless, some studies of past precipitation make use of satellite/gauge based datasets, which use remote sensing to estimate precipitation fluxes (*Huffman et al.*, 2009).

The Global Precipitation Climatology Project (GPCP - *Huffman et al.*, 2009) dataset has been extensively used for the study of historical precipitation. GPCP makes use of both infrared (since 1979) and passive microwave (since 1987) satellite retrievals, as well as rain gauge data. As such, therefore, it is not a pure satellite-based dataset, though for certain regions (most notably over ocean), gauge-based observations will be either limited or absent entirely. This issue is exacerbated at high latitudes, as satellite retrievals struggle to identify precipitation falling over snow and ice (*Hegerl et al.*, 2014). Although these issues will reduce realism (*Huffman et al.*, 1997), and such products are relatively short compared to land-based datasets, they provide global coverage, which is invaluable when trying to characterise the nature of historical precipitation changes.

Polson et al. (2013b), for example, use the GPCP dataset in order to investigate whether increased carbon-dioxide has indeed caused an increased contrast in precipitation minus evaporation between different regions. This study evaluates changes relative to particularly wet or dry regions, and so is able to overcome climatological differences between GCMs and observations. It finds that there is indeed a detectable external forcing on the intensity changes of the hydrological cycle, and that this is partly attributable to changes in greenhouse gases. Such a study would not easily be possible without taking ocean fields into account, as the rich get richer mechanism will break down over land where there is limited available moisture for evaporation, as discussed above.

Similarly, *Marvel and Bonfils* (2013) uses the GPCP dataset to create fingerprints of precipitation change that simultaneously characterise both intensifications to the hydrological cycle and poleward shifts. This is done using a similar approach to *Scheff and Frierson* (2012), by characterising changes relative to each GCMs' mean climate, rather than to strict geographical locations. By using fingerprints that are designed to capture the expected physical mechanisms of change, this study is able to demonstrate

that recent changes in precipitation climatology have been externally forced, and this forcing is likely to be anthropogenic. Approaches whereby precipitation change is evaluated relative to climatology have been established for some time, with *Allan and Soden* (2007) investigating patterns of precipitation change in ascending and descending regions.

As discussed, internal variability is large for precipitation, and so studies using relatively recent satellite/gauge based observations tend to have difficulties disentangling signal from noise. This is especially problematic as ENSO, the leading source of variability, will not be correlated between observations and the coupled General Circulation Models (GCMs) used to simulate climate. To increase the signal to noise ratio, ENSO can be removed - especially from observations (*Polson et al.*, 2013b), where ENSO indexes have been defined (*Wolter and Timlin*, 2011) - and ensemble means can be taken across GCM simulations to smooth out the variability (*Polson et al.*, 2013a,b; *Zhang et al.*, 2007). *Marvel and Bonfils* (2013) is able to improve signal-to-noise through the use of a fingerprint onto which ENSO projects poorly.

Intensification of the hydrological cycle can also be investigated through analysis of fields that are closely related to precipitation and evaporation, as in *Durack et al.* (2012) where sea surface salinities are used as a proxy for the time integrated difference between precipitation and evaporation. This allows changes in the hydrological cycle to be indirectly evaluated from as early as the 1950s, enabling the wet get wetter mechanism to be robustly verified. Interestingly, sea surface salinities imply that the water-cycle is intensifying at a rate of $8 \pm 5\%$ per degree Kelvin - double the response estimated by GCMs.

While studies of precipitation and the hydrological cycle often focus on the precipitation response to greenhouse gas forcings, or the overall anthropogenic forcing, the effect of anthropogenic aerosols has also played crucial roles in historical precipitation changes. Indeed, *Polson et al.* (2014) shows not only that aerosols have had

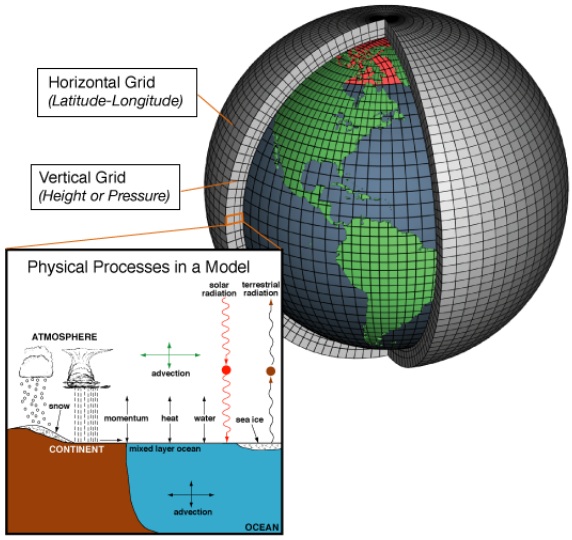

Figure 1.1: Idealised depiction of the structure of a General Circulation Model (GCM). The atmosphere and ocean are split into regular longitude-latitude grids with several vertical levels. The relevant physical processes are then simulated for each grid box, with processes taking place on scales smaller than the grid box requiring parameterisation. This requires small scale processes to be approximated using the large scale properties of the grid box. This illustration was provided by the National Oceanic and Atmospheric Administration (NOAA).

a detectable influence on monsoonal rainfall, but further demonstrates that this has been more influential than either natural or greenhouse gas forcings.

1.3 Simulating Precipitation

As discussed above, comparisons between different GCMs or between GCMs and observations are not only limited by uncorrelated variability. GCMs also have large and diverse biases in their simulated mean climate (*Dai*, 2006), which can create artificially large disagreements with each other and with observations on precipitation changes (*Allen and Ingram*, 2002; *Marvel and Bonfils*, 2013; *Scheff and Frierson*, 2012). Further, while GCM simulations of the mean climate improve with each phase of CMIP (*Knutti et al.*, 2013) - this has not given rise to better agreement on precipitation changes (*Knutti and Sedlacek*, 2013).

Biases arise in GCMs because - like all models - they must be run using limited computational resources, and so approximations must be made in numerically solving the appropriate physical equations. In the case of GCMs, the climate is typically simulated by dividing the globe into a uniform longitude-latitude grid (as depicted in Figure 1.1), with typical resolution of CMIP5 GCMs of around $2^\circ \times 2^\circ$. This implies grid boxes at the equator would have dimensions of approximately 200×200km. Processes that take place at scales smaller than that of grid boxes can not be simulated directly, and instead are parameterised based upon the large scale properties of the grid box (*Scaife et al.*, 2007). Though there are proposed modifications to this approach, such as simulating smaller scale processes to lower precision (*Palmer et al.*, 2014), such techniques remain a number of years from realisation.

There are many key climatological processes that take place on scales smaller than the typical GCM grid box, such as those governing cloud dynamics and convective processes. Not only do approximations in these processes lead to biases in precipitation (e.g. *Marsham et al.*, 2013), but they can also have large effects on the sensitivity of the climate to a forcing, as they will affect water vapour and cloud feedbacks (*Sanderson et al.*, 2008b). Here, however, we focus on errors in simulated precipitation.

Air must become saturated before clouds - and therefore precipitation - can form. This arises from moist air parcels being raised to the point where the water vapour begins to release latent heat into the surrounding air, and hence condense. The vertical motions required can be driven by a number of processes, including frontal motions, direct radiative heating of the Earth's surface, and air being driven onto orographic features (*Steiner and Smith*, 1998).

GCMs generally distinguish between 'large scale' precipitation and 'convective' precipitation, with coarse simulation of orography also contributing to precipitation biases (*Roe*, 2005). Large scale precipitation arrises from fluid dynamics and radiative

processes that can be resolved by the grid scale of the model, for example at a warm frontal system. Although simulating the processes leading to large scale precipitation will involve a cloud scheme, which will parameterise certain unresolved processes such as the formation of cloud ice (e.g. *Gregory and Morris*, 1996), the GCMs' ability to simulate the dynamics leading to the vertical motions generally means the majority of work has focussed on convective precipitation (*Randall et al.*, 2003).

Convection takes place on small time and length-scales, and is triggered by small scale instabilities, typically from the direct heating of the surface. For convective precipitation, cloud processes must again be parameterised, along with properties of the convecting plume, such as entrainment, where moist air in the convecting plume mixes with environmental air (e.g. *Gregory and Rowntree*, 1990).

The difficulties in parameterising the small scale processes of convection are highlighted by comparisons with experiments that simulate convective processes directly. Regional Climate Models (RCMs) are now capable of simulating with grid scales of the order of several kilometres, which are termed 'convection permitting', as they allow the direct simulation of some convective processes, such as large storms and mesoscale convective convective systems, though plumes and smaller scale rain showers must still be parameterised.

There is much evidence that resolving these processes increases realism. *Marsham et al.* (2013) show, for example, that simulation of West African Monsoonal rainfall is dramatically improved in terms of timing during the day, location and intensity. The improvements are seen even in a 12km simulation, which is too coarse to simulate many convective processes, though it is capable of resolving squall lines. *Kendon et al.* (2012) were able to simulate a region of the UK for 20 years using a 1.5km RCM, allowing properties of precipitation to be evaluated statistically, unlike other high resolution RCM simulations, which only capture several weeks or months, and so only permit the analysis of a particular event. The model simulates heavy rainfall

events with substantially more realistic spatial and temporal structures, simulates the timing of convection more accurately, and light rainfall events less frequently - all of which are common problems in GCM simulations of precipitation. Notably, the comparisons made for this analysis are relative to a 12km RCM, which still uses a substantially finer grid than those used by GCMs.

Indeed, the scales required to globally simulate convective processes or cloud dynamics directly would require substantial increases in computational resources that are likely to be several decades from realisation, even if all advances were dedicated to simulating on a finer grid. As a result, GCMs contain many large scale biases in their simulations of precipitation (*Dai*, 2006). Besides the issues listed above regarding intensity and timing, the spatial and seasonal distribution of climate scale features can also contain large errors, such as split Inter-Tropical Convergence Zone (ITCZ - *Li and Xie*, 2013), and the location and onset of monsoonal systems (e.g. *Roehrig et al.*, 2013; *Sperber et al.*, 2013).

The simulation of a split ITCZ has been extensively studied, as it is one of the most prominent biases in state-of-the-art climate models (*Li and Xie*, 2013). In the observed climate, the majority of tropical precipitation falls North of the equator due in part to the asymmetry of the coastal geometry. The effects of this asymmetry are amplified by the formation of stratus clouds over cold waters, which serve to further cool sea surface temperatures in the Southern Hemisphere (*Philander et al.*, 1996). GCMs frequently also simulate a precipitation maximum south of the equator (the 'double' ITCZ), due to excessive downward solar radiation in the Southern Hemisphere. This in turn is caused by insufficient cloud reflectivity during austral spring and summer (*Li and Xie*, 2013).

1.3.1 Bias Correction Techniques

Although it is essential that work continues to be undertaken to reduce biases in GCMs, it is also important that climate models' precipitation outputs are made as useful as possible. While, for some applications, using frameworks based around each GCMs' climatology can overcome differing biases in an ensemble (*Marvel and Bonfils*, 2013; *Polson et al.*, 2013b; *Scheff and Frierson*, 2012, see above,), many applications require some form of bias correction. These applications include regional climate (*Xu and Yang*, 2012), hydrological (*Ashfaq et al.*, 2010; *van Pelt et al.*, 2009) and disease (*Ermert et al.*, 2012) modelling. Traditionally, bias correction of GCM precipitation fields involves adjusting the grid-point intensity distribution according to observed grid-point intensities. Here, we refer to such techniques as local bias correction techniques.

Several local bias correction techniques - of varying complexity - exist, and generally involve mapping the intensity distribution of the climate model at each grid point onto that of the observed climate. This same mapping is then applied to the model's projected climate. *Piani et al.* (2010), for example, show the potential of such methods by deriving a GCM bias correction over Europe using data from 1961 to 1970, and applying it to model data from 1991 to 2000. Not only is the error in the mean of the intensity distribution reduced, as one would hope, but simulation of drought and heavy precipitation events also see improvements.

Kendon et al. (2008) note that in regions with a small signal to noise ratio, it is not possible to robustly derive corrections for the tails of a precipitation intensity distribution at the grid-point level. To overcome this, neighbouring grid points can be 'pooled' with the grid point of interest to artificially increase the sample size, and hence increase the signal-to-noise ratio. For such pooling, there is a trade off between signal-to-noise and regional specificity, with this study concluding that 3×3 blocks of grid boxes balance these concerns most effectively.

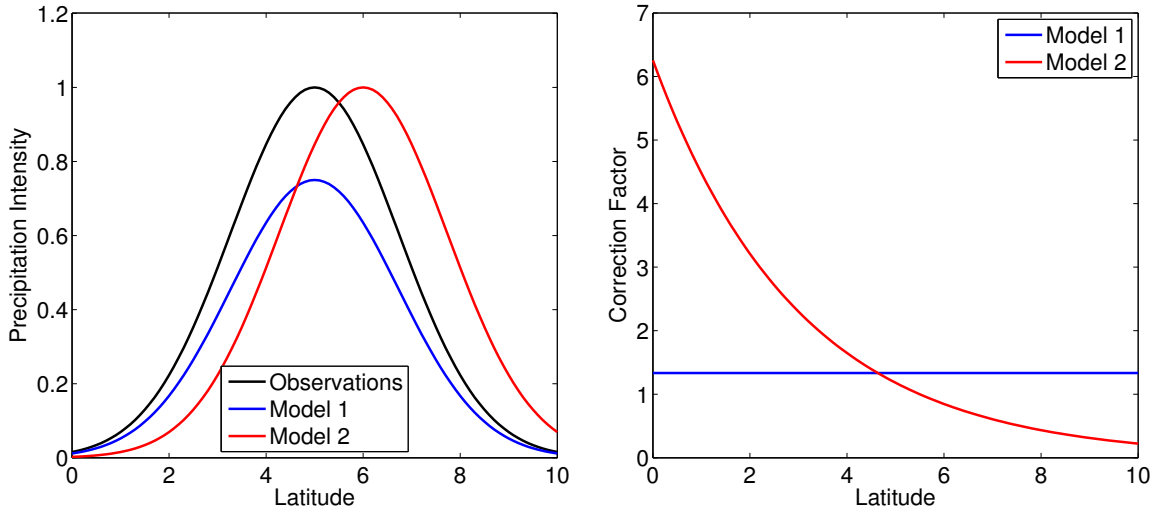

Figure 1.2: Left figure shows the latitudinal distribution of an idealised (gaussian) precipitation feature, as observed and as simulated by two GCMs. Model 1 simulates the precipitation feature in the correct location, although with insufficient intensity, whereas Model 2 simulates the correct - though displaced - latitudinal intensity distribution. The right figure shows the correction factors that would be required to correct these models using a multiplicative local bias correction technique.

As discussed above, however, there are biases in GCM simulations of precipitation that cannot be categorised as simple intensity biases. Local bias correction techniques will have little applicability, therefore, if there is not a physical relationship between the intensity distribution of the GCM and the observations (*Haerter et al.*, 2011; *White and Toumi*, 2013). If, for example, a climate feature such as the South Pacific Convergence Zone (SPCZ) is misaligned (either spatially or temporally), local bias correction techniques will make large adjustments to the intensity distribution of grid points, even if the intensity distribution of the feature were correct (but misaligned). Such erroneous corrections are depicted in Figure 1.2. While Model 1 would benefit from local bias correction, the location bias for Model 2 leads to very large correction factors.

Indeed, bias corrections of this form may lead to unphysical fields when applied to projections, as demonstrated by Figure 1.2. If the only change to a field is a uniform intensification (i.e. a feature becomes wetter by a fixed factor everywhere), then the

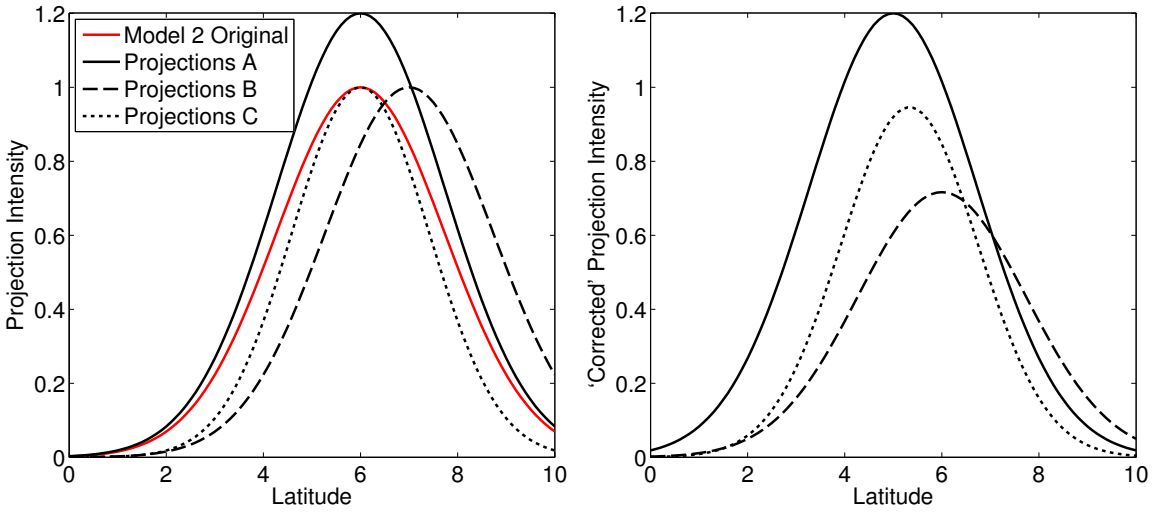

Figure 1.3: Projections based on Model 2 from Figure 1.2. The left hand figure shows three possible changes that could take place, involving an intensification (A), a shift (B) and a narrowing (C) of the precipitation distribution. Also shown, for reference, is the original precipitation distribution of Model 2. The right hand figure shows how these distributions are altered when they are 'corrected' using the local bias correction factors derived in Figure 1.2.

correction will show an intensified gaussian at the correct location, as multiplication is commutative (Projections A). If, however, the precipitation feature changes either in terms of its width or its location, the local bias correction is no longer applicable. Not only does it fail to shift the latitude of the feature by the desired amount in the case of the narrowing (Projections C), but it also alters the intensity distribution of the feature in both this case and for the shift (Projections B). It is clear, therefore, that even for this idealised precipitation feature undergoing idealised changes, local bias correction techniques have undesired effects where the original bias is one of location, not intensity.

It is desirable, therefore, to be able to correct location biases in GCMs, both in terms of spatial position and season. Such corrections would allow a GCM's simulated precipitation changes to be more directly compared either to those of other GCMs, or to observations. This would allow a more physical estimation of the uncertainty of precipitation changes to be estimated, which is not inflated by uncertainties arising

from differences in GCMs' mean climatologies (*Hargreaves and Annan*, 2014). Further, such a technique could be applied to GCM fields before they are fed into other models, potentially in conjunction with a local bias correction technique (*White and Toumi*, 2013).

Correcting location biases would allow for changes to be assessed such that differences in the mean climatology are accounted for. Other approaches have previously accounted for climatological differences, not by correcting biases, but by evaluating changes relative to the underlying climatology. For *Scheff and Frierson* (2012) and *Marvel and Bonfils* (2013) (described above) changes are evaluated relative to precipitation features themselves. Similarly, *Polson et al.* (2013b) and *Allan and Soden* (2007) investigate changes in the hydrological cycle relative to aspects of the climatology, either in terms of the ascending or descending regions, or local intensity of the hydrological cycle.

These approaches are able to overcome biases in precipitation climatology by evaluating change relative to the climatology itself, rather than relative to geographical location. Although this allows physical analysis of changes in the hydrological cycle, therefore, it can make it difficult to assess the precipitation changes experienced by particular regions. A technique capable of correcting location biases would therefore be useful for assessing precipitation changes directly, relative to the underlying geography, rather than climatology.

1.4 This Thesis

In this thesis, we therefore investigate the effect that location corrections - derived based on differences between the mean historical climates of a GCM and observations - have on simulated precipitation changes. While there will be disagreements between GCMs that arise from genuine physical differences in how changes are simulated (for

example, some models simulating a feature moistening, while others simulate it drying), it is possible that some disagreement is in fact the result of climatological differences (*Allen and Ingram*, 2002). Reductions in uncertainty would therefore imply that simulations of precipitation changes are more robust than has previously been estimated. Through such investigations, we hope not only to differentiate between sources of uncertainty in precipitation changes, but also to reduce the impact of one potential source of uncertainty - location biases.

While removing location biases in this way could provide insight into sources of uncertainty on changes, as well as allowing more robust estimations to be made, it could also enable the physical causes of location biases to be investigated. We therefore also explore how location biases vary between GCMs, and between different versions of the same GCM.

In order to remove and characterise location biases in a GCM's precipitation field, we wish to derive spatial and seasonal transformations to map the GCMs' mean historical climatology onto the observed mean climatology. This requires the derivation of a non-linear transformation, which can be achieved using image manipulation techniques. In this thesis we make use of techniques from medical image registration - first by using an existing tool, and then by developing a technique tailored to transforming precipitation - to remove location errors. These techniques are discussed in more depth in Chapters 2 and 3.

This thesis is arranged as follows. Chapter 2 investigates the potential of medical image registration techniques to reduce disagreement on projected precipitation changes, by applying a tool from medical image registration to remove location biases from 14 CMIP5 GCMs. Chapter 3 describes the mathematical framework, and broad behaviour of a new technique we have developed, tailored specifically to the removal of location biases from global precipitation fields. In Chapter 4, transformations are derived and selected for 21 CMIP5 GCMs, and properties of these transformations

are discussed in the context of physical differences between GCMs. Chapter 5 investigates whether these transformations can improve the realism of simulated historical precipitation changes and, subsequently, the extent to which an external forcing on precipitation can be detected. Chapter 6 applies our technique to a perturbed parameter ensemble, in order to characterise location biases in the West African Monsoon, in the hope of linking such biases to underlying physical properties of the perturbed GCM. Finally, Chapter 7 discusses the broad conclusions of this work, as well as potential areas in which it could be expanded.

Chapter 2

Proof of concept using *fnirt*[1]

This chapter investigates whether it is possible to improve GCM agreement on projected changes by reducing location biases. An idealised forcing experiment is used in this proof of concept study, as it can allow a greater signal to noise ratio than would be possible with historical precipitation simulations, while only involving a single forcing agent (carbon-dioxide).

Although several spatial mapping tools have been developed for investigating precipitation distributions, these have generally been used for validation (*Brown et al.*, 2012a) or data assimilation (*Hoffman and Grassotti*, 1996) of weather fields, and so are not well suited for climatological scales. Here, a new technique is proposed for the removal of location biases from simulated precipitation changes, based upon the errors in the GCMs' mean historical climate. This Chapter makes use of a publicly available[2] medical image registration tool, *fnirt* (FMRIB's Non-linear Image Registration Tool, part of the FMRIB Software Library: *Andersson et al.*, 2010; *Woolrich et al.*, 2009), which was developed for the transformation of brain images taken from Magnetic Resonance Imaging (MRI) scans onto a common grid, so that inter-patient comparisons can be made.

[1]This chapter is based upon (*Levy et al.*, 2013), for which I had primary responsibility. Figures 2.1 and 2.3 are new here.

[2]http://fsl.fmrib.ox.ac.uk/fsl/

By modifying the operation of *fnirt* so that it may be applied to climate fields, transformations are derived for 14 CMIP5 (Coupled Model Intercomparison Project Phase 5[3] - *Taylor et al.*, 2011) GCMs. These transformations shift each model's precipitation climatology to more closely match the observed distribution. They are then applied to each GCM's projected changes under a prescribed external forcing, to evaluate whether greater agreement is reached through this technique than with uncorrected model outputs.

This chapter is arranged as follows: Section 2.1 introduces the general techniques of medical image registration; Section 2.2 discusses how transformations are derived for GCMs based on their historical climate, and how they are applied to projected precipitation; and Section 2.3 discusses the results of this Chapter.

2.1 Medical Image Registration

Medical brain image registration operates by warping an input image to better resemble a reference image. In order to compare anatomical features in images of different patients' brains, this requires a non-linear, reversible (diffeomorphic) transform (*Beg et al.*, 2005). The transformation can therefore smoothly distort an image, but cannot introduce folds or tears.

Here, *fnirt* - a non-linear registration tool (*Andersson et al.*, 2010) - is used, having previously been extensively utilised in the medical field (e.g. *Filippini et al.*, 2009; *Westlye et al.*, 2010). *fnirt* was used both because it is a well established tool, and also because of the availability of guidance on its use, through the supervision of Mark Jenkinson.

fnirt operates in the small deformation framework of brain image registration (see Section 3.1.2 for more details on brain image registration), where the transformation is represented with basis functions (*Ashburner and Friston*, 1999; *Rueckert et al.*, 1999)

[3]http://cmip-pcmdi.llnl.gov/

and the coefficients (parameters) are found by numerically minimising an objective function of the form (*Bajcsy and Kovačič*, 1989):

$$\sum_i [\mathrm{C}(W_i(I), R_i) + \lambda \mathrm{O}(W_i)] \tag{2.1}$$

where the i indexes grid points, I and R are the input and reference images respectively, W is the warp, C is a cost function, which quantifies the difference between the reference image and the warped input image, λ is a tuneable parameter, and O to a regularisation function, penalising warps that are not spatially smooth (*Andersson et al.*, 2010). The tuning of λ therefore determines how smooth the warp will be, as it specifies how great a role O plays in the objective function.

In *fnirt* the cost function, C, is set to take the form of a sum of squared differences, and the regularisation term, O, is the mechanical bending energy of an elastic solid, subject to the warp W (*Ashburner and Friston*, 1999). The bending energy of a three dimensional solid is given by:

$$\mathrm{O}(W) = \sum_p \left[\sum_{a=1}^{3} \sum_{b=1}^{3} \sum_{c=1}^{3} \left(\frac{\partial^2 \vec{W}_a}{\partial x_b \partial x_c} \right) \right]_p \tag{2.2}$$

where the sum over p indicates a sum over all grid points, and the sums over a, b and c are over the dimensions of the geometry.

2.2 Applying *fnirt* to Climate Models

2.2.1 Deriving Warps

To investigate the potential of *fnirt* to reduce location biases in GCMs, it is applied to 14 CMIP5 models, listed in Table 2.1. These GCMs were chosen as, at

Table 2.1: List of CMIP5 GCMs used in study

Model Name(s)	**Modelling Group**
CNRM-CM5	Centre National de Recherches Meteorologiques
CanESM2	Canadian Centre for Climate Modelling and Analysis
GFDL-ESM2G, GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory
HadGEM2-ES	Met Office Hadley Centre
INM-CM4	Institute for Numerical Mathematics
IPSL-CM5A-LR, IPSL-CM5A-MR	Institut Pierre-Simon Laplace
MIROC-ESM, MIROC5	Japan Agency for Marine-Earth Science and Technology
MPI-ESM-LR, MPI-ESM-P	Max Planck Institute for Meteorology
MRI-CGCM3	Meteorological Research Institute
NorESM1-M	Norwegian Climate Centre

the time of investigation, they had precipitation fields from both historical and 1% CO_2 experiments (see below for more details on these experiments, and why they were chosen). Here, the aim is to derive transformations capable of removing location biases from GCMs. GCM precipitation is therefore warped onto observed precipitation fields, thereby characterising location errors in the GCM, with the monthly mean historical climatology of each GCM field taken as the input field (I in equation 2.1).

For the reference (R in equation 2.1) the monthly means of ERA-Interim's reanalysis (*Dee et al.*, 2011) are used, interpolated to each GCM's grid before warping. Reanalyses use historical observations to constrain a simulation, in order to obtain a set of coherent climate fields that are broadly representative of the historical climate. Reanalysis datasets do not precisely match observations (*Schneider et al.*, 2014), and it is important to note that ERA-Interim does not use precipitation observations in its analysis. However, these datasets are much closer to reality than GCMs, by construction. Here, a reanalysis dataset is used due to the convenience of its high resolution global coverage. Further, reanalyses provide an easier reference than observations, as they are somewhat 'model-like', and so are suitable for this proof-of-concept study.

For both GCM and ERA-Interim climatologies, only the years that the GCMs' simulations overlap with ERA-Interim (1979-2005) are used. By warping climatologies rather than individual years, the hope is to reduce only persistent location biases

in the GCM simulations. While there will be some sampling uncertainty, here it is assumed to be small relative to systematic errors for 27 year means, though this is investigated explicitly in Section 4.1. For each GCM, the warping process is repeated iteratively, until convergence is reached, for a range of different values of λ, varying from 25 to 200 $(\text{mm/day})^2$.

When registering an MRI image of a brain, the most important property to conserve is the image brightness (or intensity). *fnirt* therefore conserves intensities (equivalent here to precipitation fluxes, or rates) when warping, and so total rainfall is *not* conserved. This will lead to changes in the amount of global, annual precipitation, and so will introduce inconsistencies into both a model's moisture budget and tropospheric energy balance. In an attempt to circumnavigate this limitation, transformations are derived by applying *fnirt* iteratively. After each iteration, the warped GCM climatology is modulated using the transformation's Jacobian determinants, which characterise the change in area of any grid point, and so can be used to keep integrated precipitation amounts approximately constant. It was found, however, that this approach failed to converge upon a solution, making it unsuitable for the derivation of transformations. We therefore investigate only transformations that conserve intensities (as is the default in medical image registration), in this study.

An additional restriction of *fnirt* arises as it is designed to warp medical images of patients' brains, and so naturally operates on a three dimensional Cartesian geometry with no periodicity. This limits our application of the technique to climatological fields, which require periodic boundary conditions both longitudinally and seasonally, in order to allow features to move across the prime meridian, as well as between late and early seasons (e.g. between December and January).

For a preliminary investigation of the potential of these methods, the longitudinal, latitudinal, and temporal dimensions are therefore aligned along the rectilinear x, y, and z dimensions, respectively. In this grid, one month is defined as equivalent

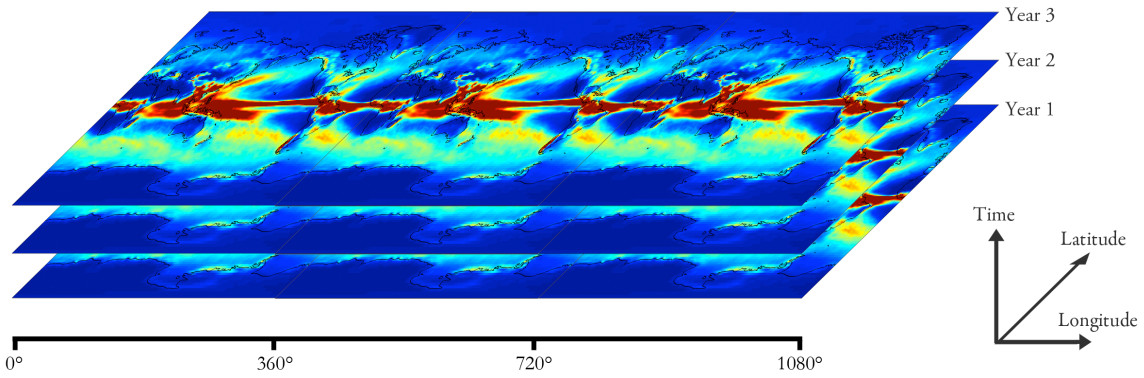

Figure 2.1: Illustration of the padding process, which is designed to allow features to move across the prime meridian, as well as between December and January. This operates by repeating the spatial component of the precipitation climatology at the longitudinal boundary, and the seasonal component at the beginning and end of the year. The data is therefore increased in size by a factor of 9. The latitude range depicted is from 90°S - 90°N, and longitude ranges from 0° - 1080°, due to the longitudinal padding. Note that each of the three global surfaces depicted is used to represent one year's entire monthly mean climate.

to ten degrees of latitude or longitude, chosen as giving the greatest improvement (not shown) to the climatology of HadCM3 (a CMIP3 GCM - *Gordon et al.*, 2000). Each climatology is padded with copies of itself at the longitudinal and temporal boundaries, as depicted in Figure 2.1. This allows the warp to move features across these boundaries, approximating the behaviour of true periodic boundary conditions. As there is little precipitation at high latitudes, the polar boundary condition is not addressed. While this padding allows movement across the prime meridian and the December-January boundary, it imposes no restrictions on how precipitation is moved across these boundaries. As a result, it would be possible for translation vectors to enter from both the Easterly and Westerly boundaries simultaneously, for example, which (with correct boundary conditions) would imply a discontinuity.

For simplicity, the same value of λ was selected for all GCMs. The warps derived using a λ of 150 $(\text{mm/day})^2$ were chosen, as (on average) they maximised the correlation between the models and ERA-Interim's climatology, whilst preventing features from entering on both the Eastern, *and* Western boundaries, and ensuring that the

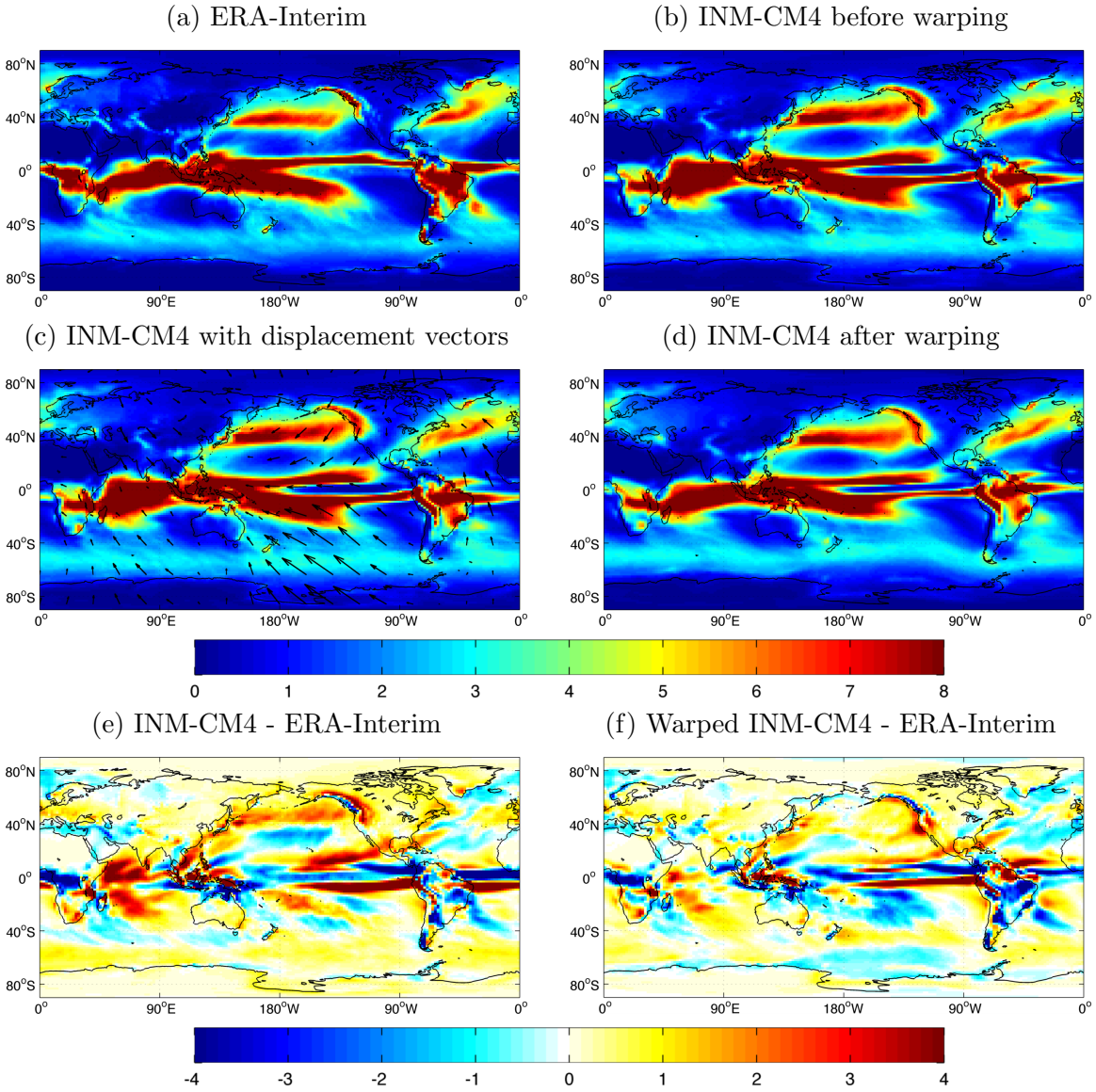

Figure 2.2: Figure exemplifying the warping process for an example GCM (INM-CM4) being warped to the ERA-Interim reanalysis with $\lambda = 150$ (mm/day)^2. The GCM data is reformatted so that *fnirt* medical imaging software may be applied to it, and location errors may be reduced. All images depict January, though all months of the climatology are used, and all units are in mm/day. Note that the warping process can also move features temporally, which is not illustrated in this figure.

warp's Jacobian determinant is between 0.1 and 10 everywhere[4].

Figure 2.2 illustrates how one month (January) of one model (INM-CM4) is modified under the registration process. One conspicuous error in this model (Figure 2.2b) is the 'split ITCZ' - a common error in GCMs (*Dai*, 2006). However, as the warp can

[4]Note that the determinant of a transformation's Jacobian quantifies the extent to which regions are stretched and compressed.

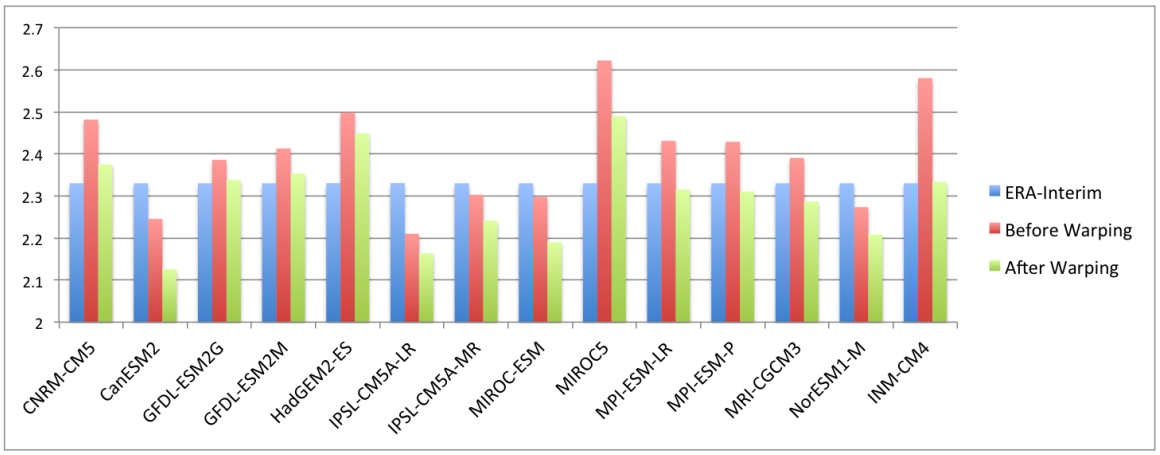

Figure 2.3: Globally averaged annual mean precipitation for ERA-Interim (blue bars) and each GCM, both before (red bars) and after (green bars) warping. Values are evaluated for the historical climatology (1979-2005), and are given in mm/day.

stretch and distort, but cannot cause regions to vanish, the warp is limited to shifting the lower branch West, as well as compressing it latitudinally (Figure 2.2c). The value of λ imposes a smoothness on the warp, which restricts its ability to shift the Northern ITCZ East. Another conspicuous error, the over-extensive region of heavy rain over the Indian Ocean shown in Figure 2.2b, is more amenable to warping, and is shrunk down to closely resemble the ERA-Interim pattern (as illustrated by the reduction in red for this region in 2.2f relative to 2.2e). The level of smoothness required can, however, lead to increased discrepancies with reanalysis in some areas, as seen in the North Atlantic storm tracks.

As discussed above, *fnirt* conserves image intensities (equivalent to precipitation fluxes) upon warping, and so it is expected that total global precipitation would be altered by the warping process. This could allow the potential for the technique to remove global biases from the GCMs, making models that are too dry wetter, and vice versa. To investigate this, the globally averaged, annual mean precipitation is evaluated for each GCM both before and after warping.

These average precipitation amounts are shown in Figure 2.3, alongside the value for ERA-Interim, for comparison. The values are calculated without area weightings,

so that changes can be interpreted in relation to the *fnirt* technique, which operates in a three-dimensional Cartesian geometry, as described above. It is found that all GCMs' mean precipitation is changed, with INM-CM4 changing by the largest fraction (a decrease of 9.6%).

In fact, all 14 GCMs' mean precipitation amount is found to decrease upon warping, regardless of whether they were initially too wet or dry (as evaluated in the Cartesian geometry). This provides evidence that the correction of individual features takes precedence over correcting the global, annual mean. Further, it suggests that features are generally simulated over too broad an area, and so the registration tool frequently reduces their size, thus losing precipitation. This can clearly be seen in the case of INM-CM4 (Figure 2.2), where many regions of high rainfall are condensed to cover a smaller area, after warping.

The extent to which the pseudo-periodic boundary conditions deviate from true periodic boundaries is also evaluated. To do this, the mean of the Jacobian determinant is calculated across the precipitation field (the central section of the padded input field - see Figure 2.1). Any deviation from unity implies the GCM's precipitation climatology is being compressed or expanded overall, implying the boundary conditions are breaking down. It is found that - with a λ of 150 $(\text{mm/day})^2$ - the GCMs are compressed by an average of 2.2%. We also find that the lower the value of λ, the greater this compression. While this figure implies only a small fraction of the field is affected by the lack of true periodic boundaries, it does demonstrate that there is a non-negligible lack of realism in the transformations' behaviour.

Next, the extent to which errors in the GCMs' climatologies are reduced upon warping is investigated. It is found that warping the models' mean monthly climatologies to ERA-Interim with a λ of 150 $(\text{mm/day})^2$ reduces the error (difference from 1) in their correlation with the reanalysis by 35%, on average. Figure 2.4 shows this improvement in the zonal means for January. It can be seen that after warping, the

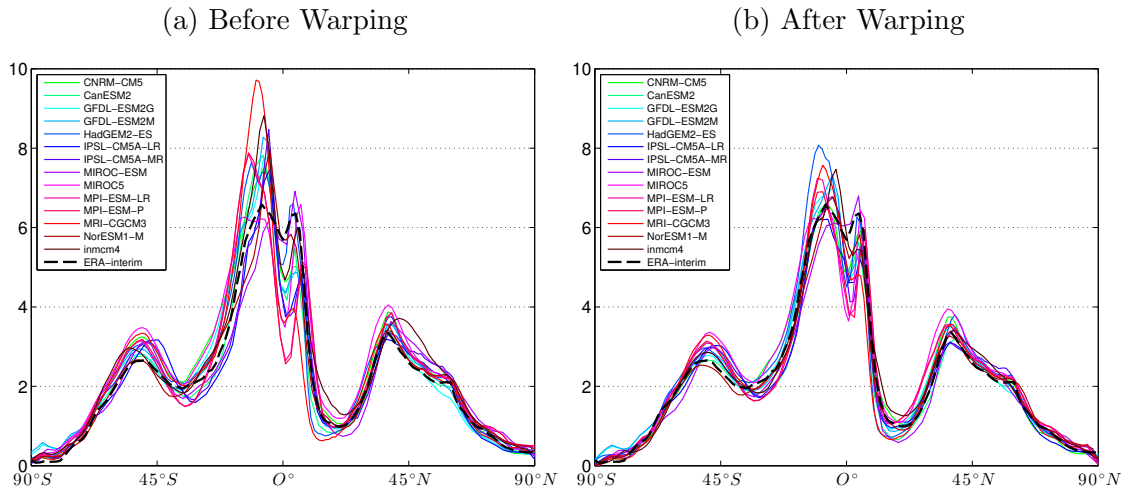

Figure 2.4: Historical zonal mean precipitation for January for all GCMs used (see legend), before (a) and after (b) warps are applied. Shown are the results for individual models, as well as ERA-Interim (dashed black lines). Means are taken across all years where there is overlap between the GCM historical simulations and the ERA-Interim reanalysis (1979-2005). Units are in mm/day.

GCMs agree better both with each other, and with ERA-Interim (dashed black line). These improvements are particularly pronounced in areas of high rainfall, such as the ITCZ and the storm tracks, as the warp will be driven to do most in areas with the largest discrepancies in grid point intensities.

This result verifies that applying the warping technique to GCMs *does* shift their mean climatologies closer to the observations, as it should by construction. However, to investigate how physically useful these corrections are, they are next applied the derived transformations to simulated precipitation changes.

2.2.2 Applying Warps to Projections

To investigate the potential of this technique to reduce uncertainty associated with climatological biases from projections, we now consider the change in precipitation predicted by the models under a prescribed external forcing. Errors of location will not remain unchanged through climate change, and indeed shifts of precipitation features

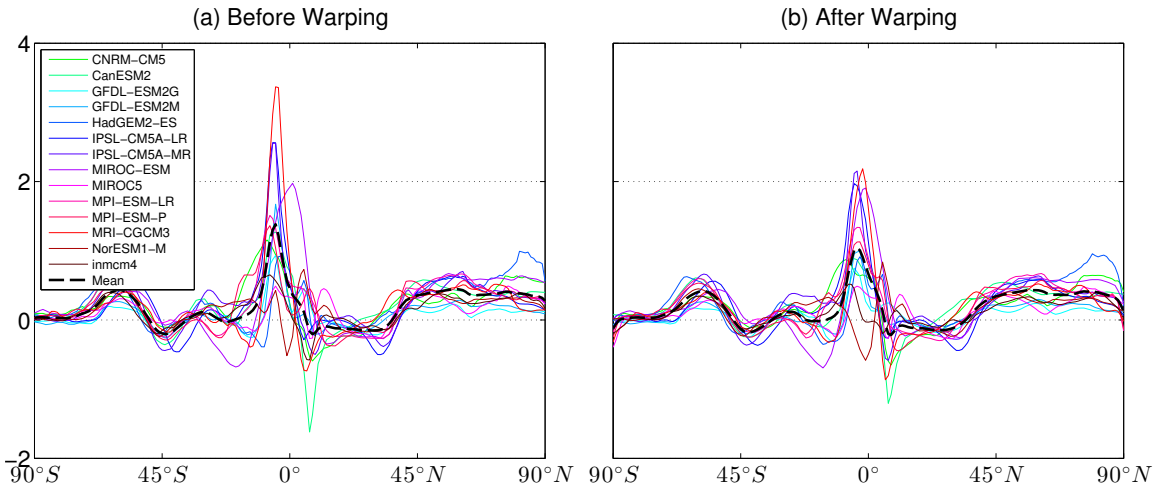

Figure 2.5: Zonal mean change in precipitation for January for all GCMs used (see legend) under the 1% CO_2 experiment, before (a) and after (b) warps are applied. Shown are the results for individual models, as well as the ensemble mean (dashed black lines). Changes are evaluated as the difference between the first and last 30 years of the experiment, with offsets separated by 110 years. All models are found to have a reduced root mean square difference from the multi-model mean. Units are in mm/day.

are expected (*Scheff and Frierson*, 2012), some of which may be large enough to lead to local non-linear responses (*Hawkins et al.*, 2014). In this study, the hope is that any substantial error in the present-day simulation will be physically persistent enough that such a correction will continue to remove more error than it introduces.

The CMIP5 1% CO_2 experiment - in which carbon dioxide levels are increased by 1% each year until they have reached quadruple their initial values (140 years) - was chosen for investigation. This experiment is used, as such a strong forcing will lead to a high signal to noise ratio. Further, by only varying carbon dioxide, the experiment avoids complexities arising from the simulation of other forcing agents, such as methane and aerosols. For each GCM, the monthly mean climatology is evaluated for the first and last thirty years of the experiment (a separation time of 110 years), and the difference is taken. Each GCM's warp is then applied, having been derived purely from the historical climate, and any changes to the agreement between models are then investigated.

Inter-model agreement on changes is quantified using the standard deviation and range at each point on the globe, after the models have been resampled to a 1° grid for utmost detail, although spatial smoothing can be used to increase the signal-to-noise ratio (*McSweeney and Jones*, 2013). The standard deviation and range are both reduced by 15%, after the historically derived warps are applied.

This increase in agreement is exemplified by the January zonal mean changes, shown in Figure 2.5. For all GCMs, the zonal means are more complex than those in Figure 2.4, indicating the complexity of precipitation changes over time. Changes on these timescales indicate that most of the globe will receive more precipitation, though there is evidence that global warming will increase the contrast between regions' evaporation minus precipitation (*Held and Soden*, 2006). The GCMs' zonal means therefore generally agree on sign both before and after warping, though there are clear improvements in the agreement on the magnitude of these changes. Most noticeable are improvements in the tropics, arising from the corrections to the GCMs' simulated ITCZs.

Also shown in Figure 2.5 are the ensemble means across the GCMs, before and after warping. Differences between GCMs in their simulated changes can lead to a weaker response in the ensemble mean, as changes can 'cancel out' (e.g. *Marvel and Bonfils*, 2013). Although a strengthening of the ensemble mean is not seen here, better consistency is found, with all GCMs closer to the ensemble mean after warping.

Improvements in agreement cover much of the globe, with 71% of the world seeing a reduction in standard deviation. There is also substantial improvement in some regions, with the transformations more than halving standard deviation for 5% of the Earth's surface.

These improvements in agreement contrast with those achieved using a crude local bias correction. For this local bias correction, the mean precipitation at each grid point of each month is scaled to match the local value of the reanalysis - the

only correction possible when working with monthly mean fields. We find that this approach increases inter-model standard deviation unless the scaling is limited to a very restricted range, where it is changing intensities by a factor of 1.5 or less. It is found that the global mean reduction never exceeds 2.6%. More sophisticated local techniques might perform better, but remain inappropriate if a feature, such as a convergence zone, is in the wrong place, rather than of the wrong magnitude (*Haerter et al.*, 2011), as this may introduce spuriously large intensity changes.

In Figure 2.6, both the multi-model mean projected changes for precipitation, as well as a measure of the agreement between GCMs, are shown. In Figure 10.9 of *Meehl et al.* (2007) agreement between GCMs is evaluated based on the *sign* of changes across GCMs. *Tebaldi et al.* (2011) argue that quantifying the inter-model agreement relative to an absolute value fails to allow for the range in natural variation across the globe, and that comparing it to the signal itself can be misleading when the signal is small, while *McSweeney and Jones* (2013) additionally stress the importance of such figures' legibility. In *Kirtman et al.* (2013), a combination of hatching and stippling is used to indicate regions where changes are small or large relative to internal variability in that region, respectively.

The agreement thresholds are defined at each geographical location with respect to the historical inter-annual and estimated inter-decadal variability. These thresholds are chosen as the size of variation in the local climate is intuitive to those living in that region (*Tebaldi et al.*, 2011). These thresholds are taken to be ERA-Interim's inter-annual standard deviation and $1/\sqrt{10}$ of this value, respectively. This is equivalent to criterion (1) of *Tebaldi et al.* (2011). This technique is used here to represent both boreal summer and winter, before and after the warp is applied (Figure 2.6).

Interestingly, Figure 2.6 illustrates that despite warping of the 14 GCMs shown in Table 2.1, the ensemble mean remains broadly similar to that calculated before warping. Indeed, the general pattern of wetting and drying remains, though in some

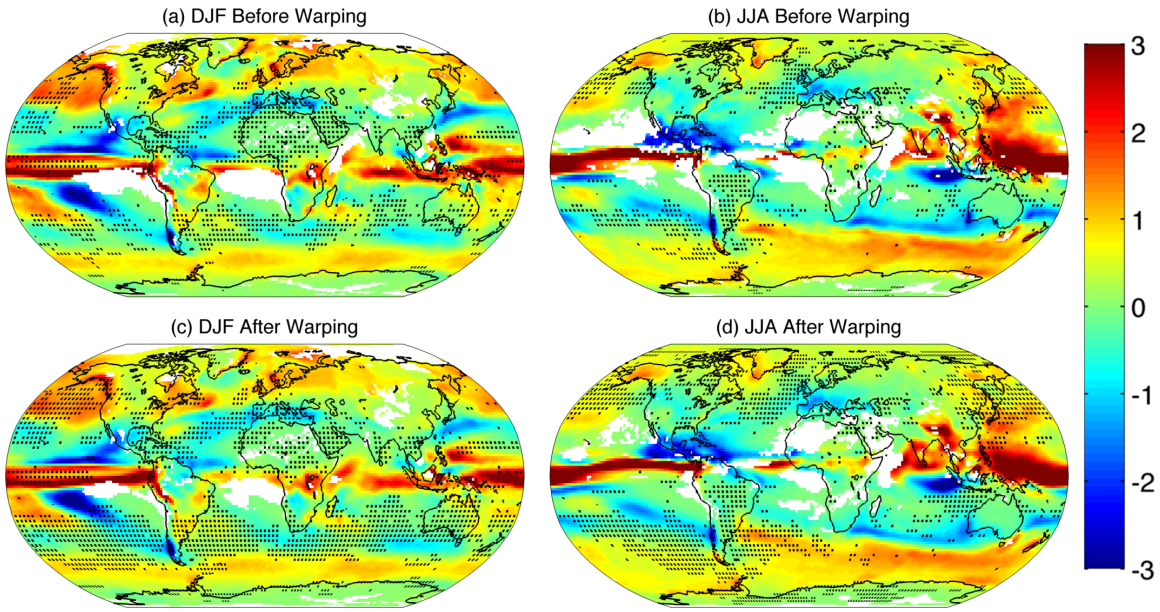

Figure 2.6: Projected change in precipitation seasonal climatology over 110 years under the 1% CO_2 CMIP5 experiment (mean of first thirty years, subtracted from mean of last thirty), displayed in mm/day. Areas with colour indicate an inter-model standard deviation smaller than the historical inter-annual variability at that location, and stippling denotes a standard deviation smaller than the estimated inter-decadal variability. Increases in colour and stippling after warping therefore illustrate greater agreement between GCMs. The panels show December-February (a and c) and June-August (b and d), both before (a and b) and after (c and d) warping.

regions it becomes less intense, as in the wetting seen in the Indian Ocean in DJF. As in Figure 2.5, the predominant signal is of increased precipitation (as one would expect from increased latent heat transport - *Allen and Ingram*, 2002), though there is clearer evidence of drying in some regions here, which were smoothed out in the zonal mean patterns.

Of most relevance to this study, though, is the increase in agreement (measured by standard deviation) between GCMs, as evidenced by the increase in stippling and colour in Figures 2.6c and 2.6d, relative to 2.6a and 2.6b. Though there are some regions where agreement deteriorates (such as Northern Africa in DJF), such regions are notably smaller than regions where agreement increases.

It is interesting to note that - though the clearest increases in agreement in Figure 2.5 took place in the tropics - the areas with the greatest increases in stippling are

mid-latitudes, indicating the impact of relatively small improvements in these regions. While many of the improvements take place over oceans (for example, the Southern Ocean in DJF), improvements are also seen over land, most notably Australia in DJF and East Asia in JJA. Such improvements over land imply the utility of location correction techniques for the detection and attribution of precipitation, where studies are often limited to land-based observations (e.g. *Polson et al.*, 2013a; *Zhang et al.*, 2007).

2.3 Conclusions

In this chapter, it has been demonstrated that techniques used in medical image registration can be adapted and applied to precipitation fields, in order to reduce location biases. Warping GCMs to reanalysis of the observations not only improves their ability to recreate the historical precipitation distribution, but also creates greater inter-model agreement on projected changes under a prescribed forcing (the CMIP5 1% CO_2 experiment). The latter result confirms that GCM errors of location do tend to persist under simulated climate change, and that such errors can mask agreement within GCMs' responses to forcings. In contrast, a simple local bias correction technique increased inter-model discrepancies, unless so restricted that it had little effect, illustrating the difficulties in applying such techniques when errors are location - rather than intensity - based.

Improvements are most notable over sea, particularly the North Pacific, Southern subtropical oceans and Southern Ocean, though Australia in DJF and East Asia in JJA also show improvements (Figure 2.6). The limited areas of improvement results partly from the relatively large value for λ that was required for physical and numerical constraints, and constrains warps to be smooth over larger scales. This limits warps to shifting larger features, such as the ITCZ, and prevents great improvements to smaller

scale features, such as those of orographic precipitation. Precipitation changes under global warming will include shifts of features, as well as changes in intensity (*Marvel and Bonfils*, 2013; *Scheff and Frierson*, 2012), and so transformations may be less applicable to future climatologies. The smoothness of the transformations used here means that such shifts are small relative to the length-scales of the transformations, and so warping will still reduce location errors.

While it is of interest that inter-model agreement on projected changes increases after warping, it is important to note that increased agreement does not necessarily imply increased accuracy. Here, a projection experiment was used in order to allow for a high signal to noise ratio, though this precluded validation against observations. It would therefore be of interest to study historical precipitation changes, in order to investigate whether warping brings simulated changes into better agreement with observations (see Chapter 5). Note also that there is less need to account for climatological shifts when investigating precipitation changes over the same time period used to derive the warp.

For this idealised study a reanalysis of observations (ERA-Interim) was used, as its climatology will be more similar to those of GCMs than a pure observational dataset. While reanalysis will be closer to observed precipitation than GCM simulations, by construction, ERA-Interim does not make use of precipitation observations in its reanalysis (*Dee et al.*, 2011). It will therefore be an important test of further work to evaluate any improvements relative to observations, as discussed in Chapters 4 and 5.

fnirt's inability to deal with the correct geometry of climatological fields not only limited the technique's ability to operate with periodic boundary conditions, but also meant that the areas of features could not be conserved when they are moved latitudinally (all grid squares are treated as having equal area). Both of these restrictions undermine the physical significance of the derived transformations. Further, it is not

clear whether it is more meaningful to conserve intensities or total precipitation upon warping, but *fnirt* currently limits our investigation to conservation of precipitation fluxes.

It would therefore be valuable to develop a registration tool designed specifically for this climatological problem. In Chapter 3 we describe the mathematical basis and implementation of a new technique that operates in the correct geometry of our problem (handling the spherical geometry of the Earth's surface, as well as the seasonal cyclicity), as well as the option to conserve either total integrated precipitation or precipitation fluxes. These changes will also allow us to find more physically meaningful and detailed transformations, and the hope is that this will allow for closer mapping to observations, particularly for features with smaller scales.

Chapter 3

Development and Implementation of New Tool[1]

While it is encouraging that GCM simulations of projected precipitation changes are brought into closer agreement using image warping techniques, the study described in Chapter 2 also highlights several shortcomings of using medical imaging software to study precipitation. As discussed, the registration of medical images *fnirt* requires markedly different considerations to the removal of location biases from GCM fields.

This Chapter therefore sets out to develop a new tool from the ground up, which is better suited to the requirements of warping precipitation. Crucially, the tool must operate in the correct geometry, such that longitudinal and seasonal periodicity are observed, and the lengths of displacements are treated correctly. Further, while in brain image registration intensities are simply translated, in precipitation this is equivalent to conserving fluxes upon warping. While this may be physically appropriate in some cases, as discussed in Chapter 2, this will lead to changes in the amounts of total global annual precipitation, breaking consistency with moisture budgets and tropospheric energy balance. As a result, the aim here is to include an

[1]This chapter is based upon (*Levy et al.*, 2014a), for which I had primary responsibility. Figure 3.6 is new here.

option in the new tool to warp either intensities, or integrated precipitation amount, upon warping. These options are equivalent to conserving precipitation fluxes, and mass, respectively.

This Chapter is organised as follows: Section 3.1 discusses various image transformation techniques developed in atmospheric sciences and medical imaging, and their relevance to the tool developed here; Section 3.2 outlines the mathematical framework of the new tool, and its implementation; and Section 3.3 investigates the properties of the derived transformations.

3.1 Existing Image Transformation Techniques

There exist several well-established disciplines that make use of image transformations for different purposes, each operating under different requirements (*Brown*, 1992). For example, in satellite imaging, registration can be used to tackle both 'mosaicking' (where multiple images are combined, as in a mosaic) and changes over time (*Falco et al.*, 2008). Registration techniques are also commonly used in the field of computer vision for a number of requirements, including stereo vision (*Lucas and Kanade*, 1981) and object modelling (*Chen and Medioni*, 1992).

Here, where the aim is to transform both spatially and seasonally, the registration technique must operate in either three or four dimensions (depending on how the spherical surface is characterised). While computer vision occasionally makes use of techniques that operate in more than two dimensions (e.g. *Papatheodorou and Rueckert*, 2004), medical image registration makes much more common use of non-2D approaches than any other field we are familiar with. Further, the availability of local expertise in this area (through supervision from Mark Jenkinson) made medical image registration the most convenient starting point for the development of this technique.

This section first discusses existing image transformation techniques that have been developed for use in transforming atmospheric fields, and then describes relevant approaches in medical image registration.

3.1.1 Transforming Atmospheric Fields

In numerical weather prediction (NWP) validation of a prediction is crucial to refining forecasts and, consequently, a number of validation metrics have been developed. These include several metrics tailored for the verification of spatial characteristics of the forecast field (*Gilleland et al.*, 2009, 2010a). These approaches include 'filtering' methods, which either smooth a predicted field until its skill reaches a certain threshold (*Yates et al.*, 2006), or decompose the field into its various spatial scales and measure the skill of each independently (*Briggs and Levine*, 1997). Such methods, though useful for evaluating the skill of a predicted field, are more appropriate for characterising than correcting errors, and so have limited relevance to to the applications of this Chapter. It should be noted, however, that reducing the dimensionality of climate fields (e.g. through EOF analysis or by taking zonal means) *is* common for many analyses, such as detection and attribution of an external forcing on a field (*Huntingford et al.*, 2006; *Zhang et al.*, 2007).

There exist several spatial verification techniques, categorised as 'displacement' techniques, which are more relevant to the technique developed here. One set of approaches is based on matching 'features' from the predicted field to the observed field. Although there are many different methods, which differ in their precise approach, they tend to rely on identifying distinct features as opposed to analysing and transforming the field as a continuous image (*Ebert and Gallus*, 2009; *Gilleland et al.*, 2009), and so tend to require somewhat arbitrary thresholds to define the features that are mapped.

Field deformation techniques, on the other hand, are a form of displacement tech-

nique that spatially transforms the forecast field as a whole. Many of these, however, still rely on the use of control points (*Gilleland et al.*, 2010b), rather than utilising the entire field in the derivation of the transformation.

Other NWP verification techniques make use of optical flow approaches to characterise location errors (*Keil and Craig*, 2007; *Marzban and Sandgathe*, 2010). Optical flow techniques characterise differences between images as fluid flow of intensities in the first image (*Horn and Schunck*, 1981). Although optical flow has great utility for applications such as video tracking, it tends to be a more restrictive technique, as it constrains intensities to remain constant as they 'flow'. Such techniques are therefore avoided here, as intensity modulation will be required in order to conserve total integrated precipitation (see Section 3.2.5).

Image mapping techniques are also used to assimilate data from various sources, in order to provide a single more accurate field. Several techniques demonstrate the usefulness of warping weather fields in different contexts, but these can also rely on explicit identification of features, making use of manually set 'tie points', for example (*Alexander et al.*, 1999). This can add a degree of subjectivity to an approach, which is undesirable for fields with as much complexity as the climate. While other approaches do not require the specific identification of features, they can be better suited to a smoother field (and hence transformation) than would be appropriate for this problem (*Hoffman and Grassotti*, 1996; *Nehrkorn et al.*, 2013). Other techniques make use of image warping to characterise the dynamics in weather fields, by warping from one time step to the next (*Aberg et al.*, 2005).

Although several of the approaches developed for the study of weather fields use image transformation techniques, therefore, there are some common shortcomings. These include restrictive transformations, and explicit identification of features. Most importantly, however, we know of no methods that take spherical geometry into account (a reasonable approximation for the small spatial scales of NWP), warp tem-

porally as well as spatially, or are capable of conserving total integrated precipitation upon warping.

3.1.1.1 Precipitation Warping on Climatological Scales

While image mapping techniques have already been applied to atmospheric fields, the majority of these applications have been for the analysis of weather systems. We are aware, however, of a small number of studies that have made use of such techniques for the study of climate-scale features.

Moise and Delage (2011) adapts the feature mapping technique developed in *Ebert and Gallus* (2009) for NWP verification, in order to apply it to a large scale climatological feature - the South Pacific Convergence Zone (SPCZ). This provides a useful metric for assessing GCM performance, though the adaptation consists solely of allowing uniform rotations, as well as translations. Thus, the limitations present in pattern matching techniques for NWP are again present here. Additionally, the simplicity of the possible transformations prevents this technique being applicable to global precipitation fields, or even large-scale features with substantially more complex structure than that of the SPCZ. In spite of these shortcomings, the technique was successfully used to analyse changes in the SPCZ for the CMIP3 projections, where it was used to derive changes in location of the SPCZ (*Brown et al.*, 2012b).

Through communication with the author, I have been made aware of a new, as yet unpublished, approach developed by F. Chauvin (*Chauvin et al.*, 2013, Working Group on Numerical Experimentation (WGNE) workshop poster). This technique is based upon optical flow, and has been developed in such a way that the transformation can vary if the precipitation climatology shifts over time. As discussed in Section 2.2.2, however, forced changes in the location of a GCM's features are generally expected to be small relative to location biases. Further, the method put forward in this Chapter takes into account the spherical geometry of the Earth, while allowing for

total integrated precipitation to be conserved, neither of which is taken into account in Chauvin's method. Indeed, as highlighted above, restrictions in optical flow techniques would prevent the intensity modulations required for this form of conservation (*Horn and Schunck*, 1981).

We are therefore not aware of any techniques that are capable of either correcting the spatial and seasonal location of climatological features, while respecting the correct geometry of the system, or providing the option for total integrated precipitation to be conserved upon warping.

3.1.2 Brain Image Registration Background

Brain image registration - introduced in Section 2.1 - is an example of medical image registration (*Sotiras et al.*, 2013), where transformations are derived to adjust an image of a subject's brain so that it more closely resembles a reference image (typically from another brain scan, or from a mean across many images). This can be carried out in order to align several images of the same brain, or to align the features between many different subjects (*Hajnal and Hill*, 2010). Intra-subject registration can be seen as a pure 'lining up' of the images, and so requires only affine transformations - i.e. linear transformations and translations (*Jenkinson and Smith*, 2001). Unless a climate model has a consistent seasonal or longitudinal bias, such transformations are of little application to the bias correction of GCM fields. While more complex transformations can be composed of several 'locally affine' transformations (as in *Feldmar and Ayache*, 1994), this Section focuses more generally on non-affine transformations.

Non-affine transformations are used in inter-subject registration, as anatomical features will have different sizes and locations in different subjects. Typically for inter-subject registration, the non-affine step is preceded by an affine step, which ensures the two images are approximately aligned before the non-linear warp is derived. As discussed above, such corrections are unlikely to be required here.

For the derived transformations to have physical significance for inter-subject registrations (as well as registrations of climatological fields) they must be invertible. This means that the warp is constrained to distort an image smoothly, avoiding the introduction of 'folds', 'tears', or the removal of features (*Beg et al.*, 2005). While it is possible that climate models may contain spurious features, or lack features present in the observed climate, correction of such errors would go beyond simple location corrections, and would undermine the physical significance - and therefore the applicability - of the derived transformation.

Transformation that avoid such folds and tears are termed diffeomorphisms and are 'one-to-one', in that every point in the input image will be mapped to a single point in the reference image upon warping (*Andersson et al.*, 2010). There are several approaches to constrain transformations to be diffeomorphic, though they can broadly be divided into two categories: the small and large deformation frameworks.

3.1.2.1 Small Deformation Framework

In the small deformation framework, intensities on a grid are translated by a single set of vectors, which are composed of a linear combination of basis functions. The coefficients of these basis functions are determined by minimising an objective function to find the optimal parameters (*Andersson et al.*, 2010; *Ashburner and Friston*, 1999; *Rueckert et al.*, 1999). If the objective function were based solely on the similarity of the warped input image to the reference image, the warp vectors would be unconstrained, and the resulting transformation would be unlikely to be physically meaningful.

The objective function generally includes two terms:

$$\mathrm{C}(\vec{W}(I), R) + \lambda \mathrm{O}(\vec{W}) \tag{3.1}$$

where I and R are the input and reference images respectively, W is the warp (composed of basis functions), C is a cost function, which quantifies the difference between the reference image and the warped input image, λ is a tuneable parameter, and O is a regularisation function.

The regularisation term is included in the objective function to penalise warps that vary less smoothly, and therefore enables the selection of warps that are more physically sensible. While this does not enforce diffeomorphisms, this mechanism can generate transformations that are usually sufficiently close to diffeomorphic for practical purposes. The exact form of the regularisation term is often based on the energy required to bend or distort a solid, as a penalty based upon such mechanical work enables more physically intuitive warps to be derived (*Andersson et al.*, 2010).

Note that this form of the objective function is the same as that described in Section 2.1 (see equation 2.1). Indeed, *fnirt* - the registration tool used in Chapter 2 - operates in the small deformation framework (*Andersson et al.*, 2010).

While the small deformation framework is generally computationally efficient, approximating a diffeomorphic transformation requires displacements of intensities to be constrained more than is necessary using the large deformation framework.

3.1.2.2 Large Deformation Framework

In the large deformation framework, intensities are shifted through many small steps by a velocity field, analogous to a viscous fluid flowing from one arrangement to another (*Bro-Nielsen and Gramkow*, 1996). Transformations composed of many diffeomorphic transformations are guaranteed to be diffeomorphic themselves, regardless of the transformation's complexity, and so all that is required is that each step is small enough to ensure no folds or tears are introduced (i.e. the determinant of the transformation's Jacobian is positive everywhere).

Traditionally, the velocity field varies as the transformation is derived, in an analo-

gous fashion to fluid flow (*Bro-Nielsen and Gramkow*, 1996; *Christensen et al.*, 1994). Such transformations are generally computationally expensive to derive, as they require many successive minimisations to find the velocity vectors at each step. Alternatively, *Ashburner* (2007) makes use of a constant velocity field, which allows more efficient computation of a transformation. Optimising the transformation is, however, more complex in this framework, as it requires the derivatives of the time-integrated velocities to be found.

In general, therefore, the large deformation framework can theoretically align images to a higher accuracy than the small deformation framework (*Andersson et al.*, 2010). However, the large deformation framework is both more difficult to implement, and more computationally intensive, as it requires minimisation of a succession of many velocity vectors. The technique developed here therefore uses the small deformation framework, on the understanding that if the warps derived are not sufficiently accurate, the technique can be adapted to a large deformation algorithm without substantial change to most of the underlying mathematics or code.

3.2 Mathematical Framework

As discussed above, in the small deformation framework of brain registration, the optimal warp is found by minimising an objective function, described in Equation 3.1. Here, in contrast with medical image registration or the warping of weather systems, the reference and input 'images' are global, monthly climate fields, and so they are spatially spherical, and seasonally periodic (to allow features to be moved between December and January).

Further, the aim here is to derive transformations that either conserve the intensity distribution, or the total integrated precipitation of features upon warping (discussed in Section 3.2.4). This provides an option to conserve either of two physical properties

of the climate upon warping: fluxes or total mass. We have not seen any analogue of the latter option in previous atmospheric or brain warping techniques, though it is desirable in that it allows consistency with the energy and moisture budgets. This section outlines the methods used to parameterise and minimise the warp for either of these options, while operating in our particular geometry.

3.2.1 Coordinate Systems and Parameterising Warp Vectors

While the most obvious coordinate system for the spatial component would be spherical polar coordinates, there are several limitations to this system. The most obvious are that close to the poles the polar and azimuth angles (θ and ϕ, respectively) will vary more rapidly with displacements, and that there is a discontinuity at the poles (*Tao et al.*, 2002). There is also a discontinuity in ϕ between 0 and values close to 2π, which can be accommodated, but can lead to complexities in implementation to avoid calculating spuriously large displacements. Further, as demonstrated in Section 3.2.5, a coordinate system based upon local Cartesian coordinates will allow a convenient formulation of the Jacobian (see Equation 3.13).

The location of points is therefore specified using Cartesian coordinates (i.e. x, y, z and t). A point, $\mathbf{X}$, is transformed by moving it by a vector, $\mathbf{U}$, which lies within $\mathbf{X}$'s tangent space. The periodicity of the temporal dimension is handled explicitly in the code, through taking the temporal component modulo 12. The transformed point, $\mathbf{Y}$, will not lie on the manifold unless it is translated purely seasonally, and so it is projected back to the nearest point on the surface, $\mathbf{P}$. Thus spherical polar coordinates are avoided. The tangent space at $\mathbf{X}$ is denoted by $\alpha(\mathbf{X})$, which is composed of the easterly, northerly and temporal directions, described by the unit vectors α_1, α_2 and α_3 respectively. The spatial component of this transformation is represented in Figure 3.1a.

(a) Translation of a point

(b) Translation of a shape

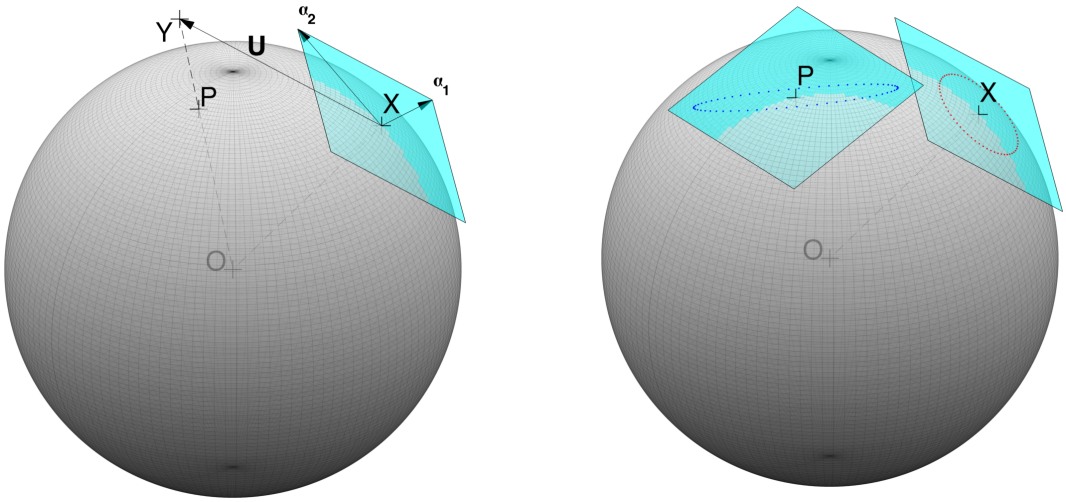

Figure 3.1: Figures demonstrating how points are translated on the spherical geometry. Figure (a) illustrates the mechanism through which a point, **X**, is warped spatially (i.e. seasonal component not shown). It is first translated by **U** along its tangent plane to **Y**, and then projected back onto the surface of the sphere to **P**. Note that α_1 and α_2 are not shown to scale, as they are in fact unit vectors, and so have the same length as the radius of the sphere (the sphere's radius is set to 1). Figure (b) shows a shape in X's tangent plane (red dots), which is warped to a different shape in P's tangent plane (blue dots). The coordinates in the two tangent planes of the red and blue dots are given by α and β, respectively, and the fractional change in area of the patterns are given by the determinant of the warp's Jacobian, $\left|\frac{\partial \beta_i}{\partial \alpha_j}\right|$.

To find the final Cartesian coordinate of the point, each step of:

$$\mathbf{X} \rightarrow \mathbf{Y} \rightarrow \mathbf{P} \tag{3.2}$$

must be calculated. The first step in this chain is given simply by:

$$\mathbf{Y} = \mathbf{X} + \mathbf{U} \tag{3.3}$$

While there are a number of spherical basis functions available (e.g. spherical harmonics and splines *Wang and Dahlen*, 1995), for simplicity each grid point's warp vector is

independently parameterised, unlike the spline-based parameterisation scheme used in *fnirt* (*Andersson et al.*, 2010)). The warp vector is some linear combination of the three unit vectors that define the tangent space (α_1, α_2 and α_3), and so can be written as:

$$\mathbf{U} = a\boldsymbol{\alpha_1} + b\boldsymbol{\alpha_2} + c\boldsymbol{\alpha_3} \tag{3.4}$$

where a, b and c are the easterly, northerly and temporal parameters, respectively. Thus, there are three parameters for each grid point. Although this means that the number of parameters is large, it simplifies the formulation of the warp, as each parameter affects only one set of translated coordinates.

Expanding $\mathbf{U}$ using the northward and eastward tangent vectors to the sphere gives:

$$\mathbf{U}(\mathbf{X}) = \begin{pmatrix} -\frac{ay+bxz}{r} \\ \frac{ax-byz}{r} \\ rb \\ c \end{pmatrix} \tag{3.5}$$

where (x, y, z, t) are the Cartesian coordinates of the point and r is the distance of the point from the z axis (i.e. $\sqrt{x^2+y^2}$). Note that the radius of the sphere, R (i.e. $\sqrt{x^2+y^2+z^2}$), has, without loss of generality, been set to 1.

$\mathbf{Y}$ is then projected onto the surface of the manifold, by normalising its spatial components, such that the total spatial distance from the origin is 1 (which has been set to be the radius of the sphere). This therefore gives:

$$\mathbf{P} = \begin{pmatrix} Y_x/\mathrm{d}_3(\mathbf{Y}) \\ Y_y/\mathrm{d}_3(\mathbf{Y}) \\ Y_z/\mathrm{d}_3(\mathbf{Y}) \\ Y_t \end{pmatrix} \tag{3.6}$$

where (Y_x, Y_y, Y_z, Y_t) are the Cartesian coordinates of the translated point, $\mathbf{Y}$, and $\mathrm{d}_3(\mathbf{Y}) = \sqrt{Y_x^2 + Y_y^2 + Y_z^2}$ (i.e. the spatial distance of $\mathbf{Y}$ from the origin).

3.2.2 Choice of minimisation technique

The minimisation of the objective function (Equation 3.1) with respect to the parameters can be described as finding a minimum of a function in a multi-dimensional space, where each dimension represents one of the parameters. Here, the conjugate gradient method is used, as it is widely utilised and requires only first order derivatives of the function (*Press et al.*, 2007).

3.2.3 Derivatives and Offset Grid

Calculating the derivative on the same grid that the vectors are defined requires a choice between centred and uncentered derivatives, unless complex interpolation techniques are used. The former was found to lead to grid splitting, whereas the latter distorted areas even for pure rotations. To avoid the problems associated with derivatives of these two forms, the objective function is evaluated on a grid that is offset from the grid on which the parameters are defined.

Although the mathematics for the registrations uses a Cartesian coordinate system, climate data (both of GCMs and observations) is generally provided on a grid that is based on longitude and latitude. The parameters are defined on this grid, which may be expressed as $l\Delta\phi$, $m\Delta\theta$ and $n\Delta\mathrm{t}$, where l, m and n are integers and $\Delta\phi$, $\Delta\theta$ and $\Delta\mathrm{t}$ are the longitudinal, latitudinal and temporal spacings, respectively. The location of the offset grid points is therefore given by $(l + \frac{1}{2})\Delta\phi$, $(m + \frac{1}{2})\Delta\theta$ and $(n + \frac{1}{2})\Delta\mathrm{t}$.

On the offset grid, each value for the intensity and the Cartesian displacement vector is taken as the average of the 8 grid points surrounding the offset grid point. Derivatives with respect to one of the three unit tangent vectors are taken as an

average of the four grid points 'ahead' of the offset grid point, minus the average of the four points 'behind', e.g.

$$\left.\frac{df}{d\alpha_1}\right|_{\phi_{1.5},\theta_{1.5},t_{1.5}} = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2}\left[\frac{f(\phi_2,\theta_i,t_j) - f(\phi_1,\theta_i,t_j)}{\Delta\alpha_1(\theta_i,t_j)}\right] \tag{3.7}$$

$\Delta\alpha_1$ is the easterly spacing between grid points, and would therefore be equal to $r\Delta\phi$, where r is the distance from the z axis, and ϕ is the azimuth angle.

3.2.4 Cost Function

A cost function is chosen that is based on the sum of squared differences between the warped input image, $W(I)$, and the reference image R:

$$C_{int}(W(I),R) = \sum_i A_i \left(W(I)_i - R_i\right)^2 \tag{3.8}$$

where the sum is over all grid points, i, and A_i is the area weighting of grid point i (for simplicity, the distance from the z axis, r, is used as the area weighting). Note that R_i denotes the intensity of the image R at point i. This form is commonly used as the cost function for the registration of MRI brain images (*Andersson et al.*, 2010; *Ashburner*, 2007; *Christensen et al.*, 1994), where the warp shifts intensities without changing them, as denoted by the subscript *int* (i.e. intensities are conserved). As mentioned above, it is also desirable to include an option to derive warps that conserve total integrated precipitation.

To illustrate how these mechanisms operate, consider a precipitation feature whose size is halved by warping. Using C_{int}, $W(I)$ would have the same intensity distribution as I, and so the total precipitation would halve. Alternatively, an approach that conserves integrated precipitation would, in this situation, double the intensities (i.e. local precipitation magnitudes) so that the total precipitation of the feature is kept

constant upon warping. These approaches are equivalent to conserving local precipitation magnitudes, or conserving integrated precipitation over large scale features, respectively.

Though it would be possible to correct intensities arbitrarily as part of the transformation (*Hoffman and Grassotti*, 1996), this technique focuses on correcting feature location errors. Either technique described here could be followed by a local bias correction to further correct local intensities. Further, local bias corrections would be smaller after warping, as grid point intensities would be brought closer together through correcting location biases.

To conserve integrated precipitation, the intensity of the warped input image must change to compensate compression or expansion of a region of the input image. The fractional change in area is provided by the determinant of the warp's Jacobian, J. When conserving in this way, therefore, the warped image is given by $W(I)_i|J(W_i)|$, which adjusts the sum of squares cost function to:

$$C_{cons}(W(I), R) = \sum_i A_i \left(W(I)_i|J(W_i)| - R_i\right)^2 \tag{3.9}$$

where $|J(W_i)|$ is the determinant of the warp's Jacobian at point i. Note that the quantity this aims to conserve (total integrated precipitation) is an integral over a continuous surface, whereas this algorithm is formulated by discrete points, and so one should not expect perfect conservation, due to discretisation errors.

The derivatives of these two forms of the cost function, with respect to a particular parameter a_j (a part of the representation of the warp - see Section 3.2.1), are given by:

$$\frac{\partial C_{int}}{\partial a_j} = 2\sum_i A_i \left[\left(W(I)_i - R_i\right)\frac{\partial W_i}{\partial a_j}\right] \tag{3.10}$$

and

$$\frac{\partial C_{cons}}{\partial a_j} = 2\sum_i A_i \left[(W(I)_i|J(W_i)| - R_i) \left(\frac{\partial W_i}{\partial a_j}|J_i| + W_i \frac{\partial |J_i|}{\partial a_j} \right) \right] \tag{3.11}$$

Note that differentiating with respect to a particular grid point's parameter will give zero for all terms in the sum other than the 8 offset grid points that surround that grid point.

In order to solve this registration problem, therefore, the form of the Jacobian and the warped intensities must be derived, and their derivatives must be calculated with respect to the parameters of the transformation.

3.2.5 Calculating the Jacobian

One convenient aspect of considering displacements in terms of local tangent surfaces is that these surfaces provide local Cartesian coordinate systems, locally equivalent to the manifold's surface. The Jacobian of a transformation is defined as:

$$J_{ij} = \frac{\partial y_i}{\partial x_j} \tag{3.12}$$

where $\mathbf{x}$ is the coordinate of a point before the transformation, and $\mathbf{y}$ is the coordinate after transformation. The Jacobian therefore measures how variation in points' coordinates after transformation depend on the coordinates before. Its determinant therefore quantifies the fractional change in area at a point.

For this problem, this can be most easily expressed in terms of the coordinates of a shape in the tangent planes before and after transformation (i.e. at $\mathbf{X}$ and $\mathbf{P}$). As in Section 3.2.3, α is used to denote the original local coordinates (as well as the unit tangent vectors). Here, the local coordinates β are also used, to express the location of the transformed points in $\mathbf{P}$'s tangent plane. In Figure 3.1b, the coordinates of the red dots (in X's tangent plane) are given by α and the blue dots' coordinates are

given by β. Using these local coordinate systems (both three-dimensional), gives:

$$J_{ij} = \frac{\partial \beta_i}{\partial \alpha_j} = \frac{\partial \beta_i}{\partial P_m} \frac{\partial P_m}{\partial \alpha_j} \tag{3.13}$$

with β being a point's transformed coordinate in $\mathbf{P}$'s tangent space, and α the point's original coordinate in the unwarped point's tangent space. The second equality follows by expanding using the chain rule, and using the Einstein summation convention.

The latter factor, $\frac{\partial P_m}{\partial \alpha_j}$, can be approximated as the derivative of $\mathbf{P}$ with respect to (normalized) longitude, latitude and time, calculated as described in Section 3.2.3.

The former factor, $\frac{\partial \beta_i}{\partial P_m}$, converts variation in the Cartesian coordinates of the transformed pattern (blue dots in Figure 3.1b) into the variation in the local coordinates of $\mathbf{P}$'s tangent surface. It is more straightforward to express how the Cartesian coordinates vary with respect to the local coordinates. As the same vectors are used to parameterise the warp vectors as for the coordinate system, this is equivalent to the derivative of $\mathbf{U}$ with respect to the parameters of our warp (see Equation 3.3, 3.4 and 3.5), which is given by:

$$\left.\frac{\partial P_i}{\partial \beta_j}\right|_{x,y,z,t} = \left.\frac{\partial u_i}{\partial a_j}\right|_{x,y,z,t} = \begin{pmatrix} -y/r & -xz/r & 0 \\ x/r & -yz/r & 0 \\ 0 & r & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.14}$$

and therefore, as an orthonormal set of basis vectors is used:

$$\frac{\partial \beta_i}{\partial P_j} = \left(\left.\frac{\partial u_i}{\partial a_j}\right|_{\mathbf{P}}\right)^T = \begin{pmatrix} -P_y/P_r & P_x/P_r & 0 & 0 \\ -P_x P_z/P_r & -P_y P_z/P_r & P_r & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.15}$$

where (P_x, P_y, P_z, P_t) are the Cartesian coordinates of the point, $\mathbf{P}$, and P_r is its distance from the z axis.

3.2.6 Jacobian derivatives

As discussed, the derivatives of the Jacobian's determinant (defined at each offset grid point) with respect to the warp's parameters (which define the vectors at each grid point) must be found. The determinant of a 3×3 matrix is given by:

$$|M_{ij}| = \epsilon_{ijk}\, M_{1i}\, M_{2j}\, M_{3k} \tag{3.16}$$

where ϵ_{ijk} is equal to $+1$ for even permutations of i, j and k, -1 for odd permutations, and 0 otherwise. The derivative of the Jacobian's determinant is therefore given by:

$$\frac{\partial |J_{ij}|}{\partial a_m} = \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\left[\epsilon_{ijk}\,\left(\frac{\partial J_{1i}}{\partial a_m}\,J_{2j}\,J_{3k} + J_{1i}\,\frac{\partial J_{2j}}{\partial a_m}\,J_{3k} + J_{1i}\,J_{2j}\,\frac{\partial J_{3k}}{\partial a_m}\right)\right] \tag{3.17}$$

with Equation 3.13 giving:

$$\frac{\partial J_{ij}}{\partial a_k} = \frac{\partial^2 \beta_i}{\partial P_m \partial a_k}\frac{\partial P_m}{\partial \alpha_j} + \frac{\partial \beta_i}{\partial P_m}\frac{\partial^2 P_m}{\partial \alpha_j \partial a_k} \tag{3.18}$$

This gives two new three dimensional arrays to be calculated. Although this has an apparently simple form, neither derivative in Equation 3.13 depends explicitly on the parameters of the warp, a_i. The second order derivatives can therefore be expanded into derivatives that can be found more simply. The first can therefore be expanded to:

$$\frac{\partial^2 \beta_i}{\partial P_j \partial a_k} = \frac{\partial^2 \beta_i}{\partial P_j \partial P_l}\frac{\partial P_l}{\partial Y_m}\frac{\partial Y_m}{\partial u_n^{\text{off}}}\frac{\partial u_n^{\text{off}}}{\partial u_q^{\text{on}}}\frac{\partial u_q^{\text{on}}}{\partial a_k} \tag{3.19}$$

where all symbols have their previous meanings, and u^{off} refers to the Cartesian displacement vector at the offset grid point where the Jacobian is defined, and u^{on}

refers to the Cartesian displacement vector at the grid point that a_k is associated with. The second three dimensional matrix in Equation 3.18 can be expanded to:

$$\frac{\partial^2 P_i}{\partial \alpha_j \partial a_k} = \frac{\partial^2 P_i}{\partial \alpha_j \partial P_l} \frac{\partial P_l}{\partial Y_m} \frac{\partial Y_m}{\partial u_n^{\text{off}}} \frac{\partial u_n^{\text{off}}}{\partial u_q^{\text{on}}} \frac{\partial u_q^{\text{on}}}{\partial a_k} \tag{3.20}$$

As can be seen, most terms that comprise the expansions of the three dimensional arrays in Equations 3.19 and 3.20 are shared. $\frac{\partial u_q^{\text{on}}}{\partial a_k}$ is given by Equation 3.14, and as U^{off} is simply an average of the surrounding 8 values of U^{on}, $\frac{\partial u_i^{\text{off}}}{\partial u_j^{\text{on}}}$ simply equals $\frac{1}{8}$. $\frac{\partial P_i}{\partial Y_j}$ can be simply derived from equation 3.6.

The remaining two factors can be computed from Equation 3.19 and the definition of a spatial derivative outlined in Equation 3.7. The first gives:

$$\begin{aligned}
\frac{\partial^2 \beta_i}{\partial P_j \partial P_1} &= \begin{pmatrix} \frac{P_x P_y}{P_r^3} & \frac{1}{P_r} - \frac{P_x^2}{P_r^3} & 0 & 0 \\ -\frac{P_z}{P_r} + \frac{P_x^2 P_z}{P_r^3} & \frac{P_x P_y P_z}{P_r^3} & \frac{P_x}{P_r} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\frac{\partial^2 \beta_i}{\partial P_j \partial P_2} &= \begin{pmatrix} \frac{P_y^2}{P_r^3} - \frac{1}{P_r} & -\frac{P_x P_y}{P_r^3} & 0 & 0 \\ \frac{P_x P_y P_z}{P_r^3} & -\frac{P_z}{P_r} + \frac{P_y^2 P_z}{P_r^3} & \frac{P_y}{P_r} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\frac{\partial^2 \beta_i}{\partial P_j \partial P_3} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ -\frac{P_x}{P_r} & -\frac{P_y}{P_r} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
\frac{\partial^2 \beta_i}{\partial P_j \partial P_4} &= \mathbf{0}
\end{aligned} \tag{3.21}$$

with the second giving:

$$\frac{\partial^2 P_i}{\partial \alpha_j \partial P_k} = \pm \frac{\delta_{ik}}{\Delta \alpha_j} \tag{3.22}$$

where the sign depends on whether the offset grid point is 'in front' or 'behind' the grid point, and $\Delta \alpha_j$ denoting the distance between grid points in the jth direction

(longitude, latitude or time). δ_{ik} denotes the Kronecker delta. Note that the subscript r on P_r denotes the distance of $\mathbf{P}$ from the z-axis, not its distance from the origin (which is always 1).

All the necessary components to compute the Jacobian and its derivatives have now been found.

3.2.7 Warped Intensities

A point's warp vector tells us where that point should sample its new intensity from. Put mathematically, this means the intensity of the unwarped image at $\mathbf{x} + \mathbf{W}(\mathbf{x})$ gives the intensity (before modulation, if this will be applied: Section 3.2.4) of the warped image at $\mathbf{x}$, or:

$$W(I(\mathbf{x})) = I\left(\mathbf{x} + \mathbf{W}(\mathbf{x})\right) \tag{3.23}$$

Since in general these translated positions will not lie precisely on the grid, intensity interpolation is necessary.

Trilinear interpolation is used, which depends only on the closest grid points to our point of interest (*Smith*, 1981). The eight points that enclose our translated point, $\mathbf{P}$, are found. The intensity at the grid points in this 'cuboid' are denoted with I_{ijk}, where the subscripts denote the corner the intensity is associated with:

$$\begin{aligned} I(\mathbf{P}) &= (1-dx)(1-dy)(1-dt)I_{000} + (1-dx)(1-dy)dt.I_{001} \\ &+ (1-dx)dy(1-dt)I_{010} + (1-dx)dy.dt.I_{011} \\ &+ dx(1-dy)(1-dt)I_{100} + dx(1-dy)dt.I_{101} \\ &+ dx.dy(1-dt)I_{110} + dx.dy.dt.I_{111} \end{aligned} \tag{3.24}$$

where:

$$\begin{aligned} dx &= \frac{P_\phi - \phi_{000}}{\phi_{100} - \phi_{000}} = \frac{P_\phi - \phi_{000}}{\Delta\phi} \\ dy &= \frac{P_z - z_{000}}{z_{100} - z_{000}} = \frac{P_z - z_{000}}{\Delta z} \\ dt &= P_t - t_{000} \end{aligned} \tag{3.25}$$

with P_ϕ denoting the azimuth angle of the point, and noting that the distance between time steps has been set to 1 (see Section 3.2.12 for discussion on dimensions). The cyclical nature of the time dimension is made explicit in the code that calculates dt.

3.2.8 Intensity derivatives

The derivatives of the intensity are again calculated via the chain rule:

$$\frac{\partial I}{\partial a_i} = \frac{\partial I}{\partial x_j}\frac{\partial x_j}{\partial P_k}\frac{\partial P_k}{\partial Y_l}\frac{\partial Y_l}{\partial u_m^{\text{off}}}\frac{\partial u_m^{\text{off}}}{\partial u_n^{\text{on}}}\frac{\partial u_n^{\text{on}}}{\partial a_i} \tag{3.26}$$

where all but the first two terms have been found in Section 3.2.6. Equation 3.24 gives:

$$\frac{\partial I}{\partial \mathbf{x}} = \begin{pmatrix} (1-dy)\left[(1-dt)(I_{100}-I_{000}) + dt(I_{101}-I_{001})\right] \\ \quad + dy\left[(1-dt)(I_{110}-I_{010}) + dt(I_{111}-I_{011})\right] \\ (1-dx)\left[(1-dt)(I_{010}-I_{000}) + dt(I_{011}-I_{001})\right] \\ \quad + dx\left[(1-dt)(I_{110}-I_{100}) + dt(I_{111}-I_{101})\right] \\ (1-dx)\left[(1-dy)(I_{001}-I_{000}) + dy(I_{011}-I_{010})\right] \\ \quad + dx\left[(1-dy)(I_{101}-I_{100}) + dy(I_{111}-I_{110})\right] \end{pmatrix} \tag{3.27}$$

and Equations 3.25 leads to:

$$\frac{\partial \mathbf{x}}{\partial \mathbf{P}} = \begin{pmatrix} -\frac{P_y}{\Delta\phi(P_x^2+P_y^2)} & \frac{P_x}{\Delta\phi(P_x^2+P_y^2)} & 0 & 0 \\ 0 & 0 & \frac{1}{\Delta z} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.28}$$

These two terms provide the additional information needed to calculate the derivatives of the grid point intensities with respect to the warp parameters.

3.2.9 Regularisation term

The regularisation term penalises warps that are less spatially smooth (i.e. warps whose vectors vary rapidly between neighbouring grid points), and in brain registration is generally based on the mechanical energy of a particular type of distortion, such as 'linear elastic', 'bending' or 'membrane' energy (*Andersson et al.*, 2010; *Ashburner*, 2007; *Christensen et al.*, 1994). Here, linear elastic energy is used, as it relies only on first order spatial derivatives, and has a well known form in both spherical and cylindrical polar coordinates (*Atanackovic and Guran*, 2000). Choices of regularisation are generally somewhat arbitrary (*Ashburner*, 2007), but a different choice could be used in the future without modification of the minimisation technique, or the form of the cost function and its gradient.

In general, elastic energy per unit volume is given by

$$U = \mu \left(\sum_{i=1}^{3} \epsilon_{ii} \right)^2 + 2\lambda \sum_{i=1}^{3} \sum_{j=1}^{3} \epsilon_{ij}^2 \tag{3.29}$$

where ϵ_{ij} is the strain tensor, and μ and λ are the Lamé parameters of the solid described (*Atanackovic and Guran*, 2000).

The first term primarily penalises changes in the volume of elements. The second term, on the other hand, encompasses the types of penalties that are desirable here,

by limiting scaling and shearing without penalising rotations (*Ashburner*, 2007). μ is therefore set to zero, which also makes the software simpler to tune. Note that although the regularisation term does not penalise pure rotations, one would not expect these to be obtained for any GCM remotely resembling observations.

Our regularisation term is:

$$K = 2 \sum_{p} A_p \sum_{i=1}^{3} \sum_{j=1}^{3} \lambda_{ij}^{\mathrm{rel}} \epsilon_{ij}^2 \tag{3.30}$$

where the sum over p denotes a sum over all offset grid points, A_p is the area weight of the point p, and $\lambda_{ij}^{\mathrm{rel}}$ is a tuneable, symmetric matrix, which sets the relative impact of each component of the strain tensor, and so the relative smoothness in each direction. The default values for all elements of this matrix are 1, but could be changed to impose a physical constraint on the warp, such as restricting transformations more in the latitudinal direction. The regularisation term is multiplied by a tuneable parameter $\bar{\lambda}$ in the objective function, which sets the overall smoothness of the warp, as in equation 3.1.

Our geometry is the surface of a sphere with time outwardly orthogonal to it, which incorporates aspects of both cylindrical and spherical polar coordinates. Combining the standard expressions for the strain matrix in these coordinates gives:

$$
\begin{aligned}
\varepsilon_{11} &= \frac{\partial w_1}{\partial \alpha_1} - \frac{z}{r} w_2 \\
\varepsilon_{22} &= \frac{\partial w_2}{\partial \alpha_2} \\
\varepsilon_{33} &= \frac{\partial w_3}{\partial \alpha_3} \\
\varepsilon_{12} &= \frac{1}{2}\left(\frac{\partial w_2}{\partial \alpha_1} + \frac{\partial w_1}{\partial \alpha_2} + \frac{z}{r} w_1 \right) \\
\varepsilon_{13} &= \frac{1}{2}\left(\frac{\partial w_1}{\partial \alpha_3} + \frac{\partial w_3}{\partial \alpha_1} \right) \\
\varepsilon_{23} &= \frac{1}{2}\left(\frac{\partial w_2}{\partial \alpha_3} + \frac{\partial w_3}{\partial \alpha_2} \right)
\end{aligned} \tag{3.31}
$$

where the three components of $\mathbf{w}(\mathbf{X})$ denote a point's final displacement in the longitudinal, latitudinal and temporal directions. Note that the strain tensor is symmetric, and so in three dimensions only contains 6 elements.

Now that a suitable expression has been calculated for the strain matrix in terms of derivatives with respect to the local coordinates, the overall displacement of a point (that is to say, $\mathbf{w}(\mathbf{X})$) must be expressed in terms of its spherical, not Cartesian, components:

$$
\begin{aligned}
w_1 &= r\Delta\phi = r\left[\tan^{-1}\left(\frac{P_y}{P_x} \right) - \tan^{-1}\left(\frac{X_y}{X_x} \right) \right] \\
w_2 &= \Delta\theta = \sin^{-1} P_z - \sin^{-1} X_z \\
w_3 &= c = P_t - X_t
\end{aligned} \tag{3.32}
$$

Recall that r denotes the distance to the z-axis.

With these definitions in place, a workable expression for the regularisation term has now been achieved. As with the other terms, the regularisation term's derivatives with respect to the warp's parameters must now be calculated.

3.2.10 Regularisation term derivatives

Equation 3.30 implies:

$$\frac{\partial K}{\partial a_m} = 4 \sum_p \sum_{i=1}^{3} \sum_{j=1}^{3} \lambda_{ij}^{\text{rel}} \epsilon_{ij} \frac{\partial \epsilon_{ij}}{\partial a_m} \tag{3.33}$$

As ϵ_{ij} depends both on the value and spatial derivatives of $\mathbf{w}$ (and hence the values of $\mathbf{w}$ on the regular and offset grids) expanding this derivative using the chain rule generates two distinct terms:

$$\begin{aligned} \frac{\partial \epsilon_{ij}}{\partial a_q} &= \frac{\partial \epsilon_{ij}}{\partial w_k^{\text{on}}} \frac{\partial w_k^{\text{on}}}{\partial P_l^{\text{on}}} \frac{\partial P_l^{\text{on}}}{\partial Y_m^{\text{on}}} \frac{\partial Y_m^{\text{on}}}{\partial u_n^{\text{on}}} \frac{\partial u_n^{\text{on}}}{\partial a_q} \\ &+ \frac{\partial \epsilon_{ij}}{\partial w_k^{\text{off}}} \frac{\partial w_k^{\text{off}}}{\partial P_l^{\text{off}}} \frac{\partial P_l^{\text{off}}}{\partial Y_m^{\text{off}}} \frac{\partial Y_m^{\text{off}}}{\partial u_n^{\text{off}}} \frac{\partial u_n^{\text{off}}}{\partial u_o^{\text{on}}} \frac{\partial u_o^{\text{on}}}{\partial a_q} \end{aligned} \tag{3.34}$$

where the superscript 'on' refers to the value of $\mathbf{w}$ at the grid point where a_q is defined, and 'off' refers to the value at the offset grid point where ϵ_{ij} is calculated.

This expansion introduces three new matrices. Equations 3.32 give:

$$\frac{\partial w_i}{\partial P_j} = \begin{pmatrix} -\frac{rP_y}{P_r^2} & \frac{rP_x}{P_r^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{1-P_z^2}} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3.35}$$

where r and P_r are the distances from the z axis of the original point and the transformed point, respectively. From 3.31 and our definition of spatial derivatives (Section 3.2.3):

$$\frac{\partial \epsilon_{ij}}{\partial w_k^{\text{on}}} = \pm \frac{1}{2} \left(\frac{\delta_{ik}}{\Delta \alpha_j} + \frac{\delta_{jk}}{\Delta \alpha_i} \right) \tag{3.36}$$

where the sign (as in equation 3.22) depends whether the offset point is 'in front of' or 'behind' the point at which the parameter is defined.

For the derivative with respect to $\mathbf{w}^{\text{off}}$:

$$\frac{\partial \epsilon_{ij}}{\partial w_k^{\text{off}}} = \begin{cases} -\frac{z}{r}, & \text{for } i,j,k = 1,1,2 \\ \frac{z}{2r}, & \text{for } i,j,k = 1,2,1 \text{ or } 2,1,1 \\ 0, & \text{otherwise} \end{cases} \tag{3.37}$$

3.2.11 Poles

Several of the forms calculated in this manuscript are undefined at the poles. The vectors in the tangent surface are defined to face eastwards and northwards, and hence a different local coordinate system would have to be used for polar grid-points. Further, a regularisation function was chosen that is not well defined at the poles. For simplicity, the polar grid points are neglected from the calculations here, which means that the derivatives of offset grid points neighbouring the poles cannot be calculated. With an input grid that has n latitudes, we have $n-1$ offset latitudes, but sum over only $n-3$, as the top and bottom rows are ignored.

Relatively little precipitation falls over the highest latitudes, and area weighting factors tend toward zero at the poles, and so this omission is unlikely to greatly affect the transformations derived, especially for high resolution models where the number of latitudes is large (typically around 100). Further, by using a different local coordinate system for the poles, and defining a form for the regularisation function for the poles, the algorithm could be adapted to include the polar grid-points.

Note that translated grid points can still pass over the poles.

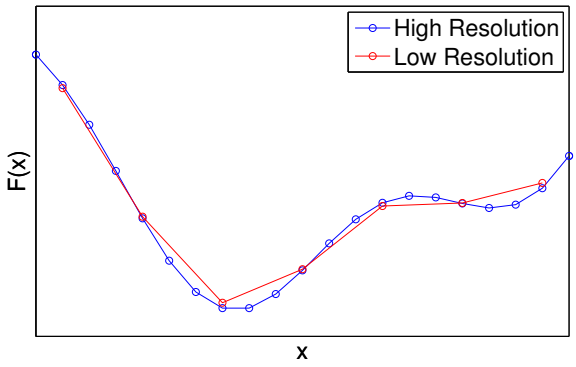

Figure 3.2: Figure illustrating a discrete function, F(x), at two resolutions. At the higher resolution (blue), the function contains two minima, one of which could 'trap' a minimisation technique, preventing it from finding the global minimum. At the lower resolution (red), the function only contains the global minimum, allowing minimisation techniques to find this solution more consistently. Deriving the warp on multiple coarser resolutions can therefore improve the quality of the minimum found, making it more robust to small perturbations to the input images, or warp parameters.

3.2.12 A note on dimensions

In all the calculations above, we set both the radius of the Earth, and the gap between months to be 1, for convenience. However, when spatial coordinates mix with the temporal coordinate, some equivalence between spatial and temporal distances must be chosen. A tuneable parameter, γ, is therefore included, which indicates how many degrees of latitude a temporal shift of one month is equivalent to in these calculations.

The quantity γ is used in elements of the strain tensor, ϵ_{ij}, that have spatial derivatives of temporal displacements, or vice versa. It is also used to normalise the temporal components of the gradient of the objective function, relative to the spatial components. For other quantities that involve both spatial and temporal units, such as the Jacobian, the units cancel, and so normalising is not required.

3.2.13 Transforming at multiple scales

As each grid point has its own independent parameters, minimising the objective function is more than an order of magnitude slower than *fnirt*, the medical registration

software used in *Levy et al.* (2013). To cut down on the time spent, the spatial resolution of the input and reference images is halved until each grid box is larger than 10^{o} in both the latitudinal and longitudinal directions. The optimal warp vectors are calculated at this scale, and are then up-sampled by a factor of two. These are then used as a 'first guess' to derive warp vectors for fields at this new spatial resolution (twice that of the previous). This process is repeated until a warp is derived at (approximately) the original resolution. Remapping and interpolation is performed with Climate Data Operators (CDO) (*Schulzweida and Kornblueh*, 2004), which make use of an implementation of *Jones* (1999). This remapping conserves total integrated precipitation upon regridding.

Registration on multiple scales has been shown to improve not only the speed, but also the robustness of transformations (*Bajcsy and Kovačič*, 1989; *Thevenaz et al.*, 1995), by considering a less detailed form of the image, which smooths out many local minima (see Figure 3.2). Similar techniques are used in atmospheric simulations (*Doyle and Shapiro*, 2000). The technique speeds up the derivation of a warp substantially, with the full resolution warp taking approximately one third the original number of iterations to converge, as it starts closer to the solution.

3.3 Properties of Warps

Having developed the mathematical framework presented in Section 3.2, the tool was first implemented in Matlab, so that each component of the warp could be visually debugged. It was then implemented in C++ for speed of execution. In this implementation, each warp was found to take approximately 3 hours to derive on a single 2.7GHz core with 2GB RAM, though speeds depend on the spatial resolution of the model, the form of conservation, and the value of $\bar{\lambda}$.

Warps are initially derived for the CMIP5 GCM, INM-CM4 (*Volodin et al.*, 2010),

(a) Percentage improvement (b) Warp variables

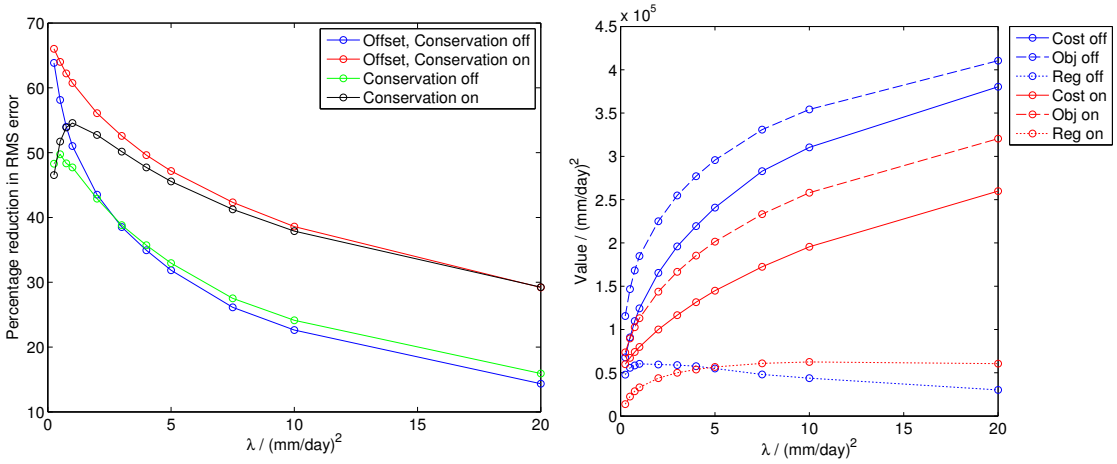

Figure 3.3: Figures showing statistics on the derived transformations for INM-CM4 as a function of the overall smoothness parameter, $\bar{\lambda}$. Figure (a) shows the percentage reduction in RMS error from GPCP both on the offset grid on which the warp is calculated (blue and red), and the GCM's original grid (green and black), both relative to their pre-warp values. The warps with total integrated precipitation conservation 'on' consistently fit better than those with it 'off' (i.e. those with intensity conservation), for the same value of $\bar{\lambda}$. In general, decreasing $\bar{\lambda}$ reduces RMS error, except at the lowest values, where decreasing $\bar{\lambda}$ increases the RMS error on the original grid. Figure (b) shows the objective function, cost function, and $\bar{\lambda}$ × regularisation term of the derived warps. The higher $\bar{\lambda}$, the more constrained the warps become, and so the higher the final cost and objective functions.

to investigate the general properties of the technique. This GCM was chosen, as it experienced the greatest percentage change in globally integrated precipitation when warped by *fnirt* in Chapter 2. The satellite/gauge based observational dataset, GPCP (*Huffman et al.*, 1997), is used as the reference image. This should provide a more realistic spatial distribution of precipitation than reanalyses of observations, such as ERA-Interim (*Dee et al.*, 2011), while still providing global, gridded coverage.

3.3.1 Parameter Sensitivity Studies

The most important tuneable parameter in the tool is the overall smoothness parameter, $\bar{\lambda}$. In preliminary tests, a value of 1 $(\text{mm/day})^2$ was found to generate warps

with few visible discontinuities, while still appearing to correct the location of medium scale features (as opposed to only correcting the location of very large features, such as the ITCZ). As some discontinuities were observed, $\bar{\lambda}$ was varied about this value from 0.25 to 20 $(\text{mm/day})^2$. All values of $\lambda_{ij}^{\text{rel}}$ were set to 1 (so that the warp was equally constrained in all directions), and the ratio of spatial to temporal distances, γ, was set to $10°$/month, which was found to be optimal in a parameter sensitivity study when using *fnirt* (Section 2.2).

Figure 3.3a shows how RMS error is reduced both on INM-CM4's original grid, and the offset grid on which the warp is derived. The most notable result of these curves is that, for the same value of $\bar{\lambda}$, the warps with integrated precipitation conservation remove more model error (i.e. reduce the cost function further) than the warps that conserve intensity on warping. This result implies that rather than acting as a constraint, adding the Jacobian term to the cost function allows an additional mechanism for intensities to be changed - not only can they be shifted, but they can also be stretched or compressed. Thus a better match can be achieved with the same strength of regularisation imposed. Figure 3.3a also shows that as $\bar{\lambda}$ is decreased, the RMS error is generally reduced, except at very low $\bar{\lambda}$ on the original grid, as the coherence between the offset and original grid breaks down when the warp's spatial scales of smoothness become very small.

The final values of the objective function, along with its two components (the cost function, and $\bar{\lambda}$ times regularisation) are shown in Figure 3.3b, and include results for warps derived both with integrated precipitation and intensity conservation. There are several characteristics that appear consistent between the two cases. As one would expect, when the smoothness of the warp is increased, both the objective function and cost function also increase. The third curve, $\bar{\lambda} \times$ regularisation, on the other hand, flattens off, as the regularisation is a decreasing function of $\bar{\lambda}$. Reassuringly, all the curves are smoothly varying, even at very low values of $\bar{\lambda}$, implying general

(a) Conservation Off

(b) Conservation On

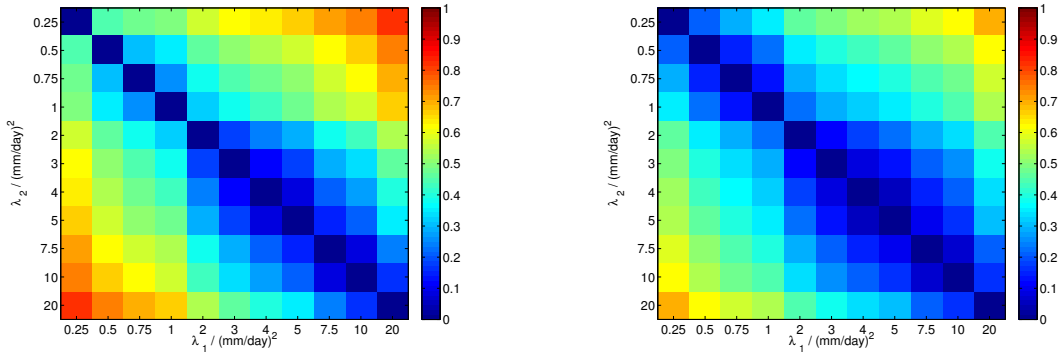

Figure 3.4: Plots showing RMS difference between warped versions of INM-CM4, for various values of $\bar{\lambda}$, in both the non-conserving (a) and conserving (b) cases. $\bar{\lambda}_1$ and $\bar{\lambda}_2$ denote the two values being compared. The RMS values are normalised by dividing by the RMS difference between unwarped INM-CM4, and the observation dataset, GPCP. Warps with similar values of $\bar{\lambda}$ generate more similar warped fields, implying stable derivation of transformations.

stability of derived warps.

This stability is confirmed by comparing different warped versions of INM-CM4, obtained with varying settings for $\bar{\lambda}$. 11 values of $\bar{\lambda}$ (between 0.25 and 20) were used, and the RMS difference between each pair of output images (warped versions of INM-CM4) was calculated. The results are normalised by the RMS differences between (unwarped) INM-CM4 and the observation dataset (GPCP). These normalised RMS differences, shown in Figure 3.4, demonstrate that the various warped images are consistently more similar to the warps with more similar $\bar{\lambda}$, and the warps with integrated precipitation conservation are more similar to each other than those without. Both these results are consistent with the changes in the cost function shown in Figure 3.3b, and confirm the overall stability to parameter perturbations of this technique. Any criteria for setting the smoothness of a warp will be somewhat subjective, but this result implies that the precise value selected for $\bar{\lambda}$ is not crucial, as warps with similar values behave in qualitatively similar ways.

An important characteristic of the warps is the extent to which they conserve total

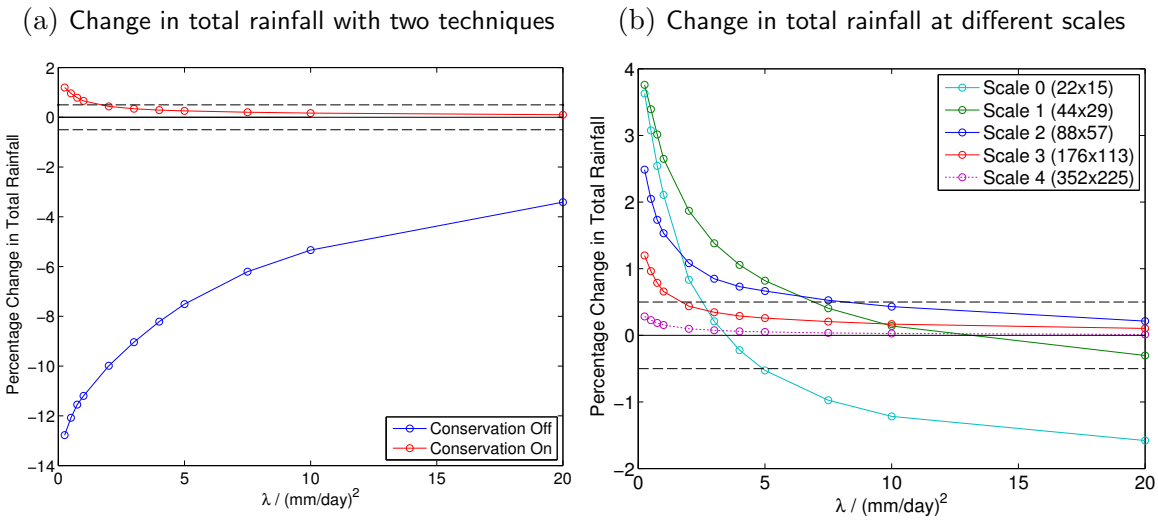

Figure 3.5: Plots showing how the warp's conservation of total precipitation varies as a function of $\bar{\lambda}$. Figure (a) shows the percentage change in total precipitation as the smoothness parameter changes, while Figure (b) shows how scale affects change in precipitation (with integrated precipitation conservation). As would be expected, they show that the percentage change in total precipitation is much smaller with total integrated precipitation conservation 'on' (as opposed to intensity conservation), and decreases as $\bar{\lambda}$ and grid resolution increase, implying that the departures from zero are discretisation errors. Scales 0 to 3 are generated from the multi-scale derivation of the warp, whereas scale 4 is generated by interpolating scale 3 to a higher resolution. The dashed lines denote an arbitrary threshold (0.5%) for the warp to be considered conservative.

globally, annually integrated precipitation. One would certainly not expect the warps derived using the intensity conservation algorithm to conserve total precipitation, but one would also expect small deviations to be present in those that used the integrated precipitation conservation algorithm, due to the discrete nature of this approach (Section 3.2.4). Figure 3.5a shows that deviation from perfect conservation of precipitation amount is indeed much lower when this conservation technique is applied. Further, both sets of warps deviate less from perfect conservation for higher values of $\bar{\lambda}$. This is to be expected, as the warps will get closer to a 'zero' warp as $\bar{\lambda}$ increases. For illustration, a threshold is included (chosen as an arbitrarily small value: 0.5%), under which the warps are considered to be sufficiently conservative.

To test the hypothesis that the discrete formulation of our algorithm causes the discrepancies from perfect conservation (when using the integrated precipitation conserving algorithm), we take advantage of the multi-scale nature of the technique. This provides warped fields at several different scales (four scales, in the case of INM-CM4). The change in total integrated precipitation for each of these scales is shown in Figure 3.5b (scales 0 to 3). At all scales, the total precipitation increases at low $\bar{\lambda}$, and seems to tend to a constant value for high $\bar{\lambda}$. The two lowest resolution scales (scales 0 and 1) tend toward a negative value (consistent with the low resolution biases in the Jacobian for zero transformations: not shown) whereas the higher resolution scales (2 and 3) appear to tend towards zero.

If the bias in the zero order transformation were removed (i.e. the curves for scales 0 and 1 were shifted up so that they tended towards 0%) we would see that higher resolution scales would reach the acceptable threshold faster. This implies that one could artificially interpolate the warp to a higher resolution, which should reduce the change in total precipitation. This is shown by the curve for scale 4, where the warp and GCM were interpolated to twice their original resolution (in both the longitudinal and latitudinal dimensions). As can be seen, this does indeed further reduce the change in total integrated precipitation, implying that an arbitrary level of conservation could be achieved, by interpolating the derived warps to higher resolutions.

Although $\bar{\lambda}$ is the most fundamental parameter, as it sets the overall smoothness of the derived warps, we should also investigate the other parameters introduced. $\lambda_{ij}^{\text{rel}}$ is a symmetric 3×3 matrix, and so introduces 6 additional parameters, besides the spacetime ratio, γ. Including $\bar{\lambda}$, therefore, we have an 7 independent parameters[2]. Clearly, an exhaustive exploration of this space would require a large amount of computational resources. The effect of perturbing any one of these parameters on the

[2]Note that changing all values in $\lambda_{ij}^{\text{rel}}$ by the same amount is equivalent to changing $\bar{\lambda}$, and so we lose one degree of freedom.

warped image is therefore investigated here.

For this investigation, one value of $\bar{\lambda}$ is selected for integrated precipitation conserving warps, and one for warps that conserve intensity. The $\bar{\lambda}$ with intensity conservation was chosen by maximizing the accuracy of fit, while requiring that the Jacobian determinant was between 0.1 and 10 for all grid points (very small or large values of the Jacobian determinant imply extreme distortions). The value for integrated precipitation conserving warps was chosen using the same criteria, while also requiring that total integrated precipitation changed by less than 0.5%. The chosen values of $\bar{\lambda}$ for INM-CM4 were 4 and 2 $(\text{mm/day})^2$, with intensity and integrated precipitation conservation, respectively. Although these criteria are simple, they balance goodness of fit with some constraint on the behavior of the warp. As can be seen from Figure 3.3a, these values of $\bar{\lambda}$ reduced RMS error on the original grid by 36% and 53%, respectively, when deriving warps with unperturbed parameters. Keeping $\bar{\lambda}$ fixed, each of the 7 other parameters were varied in turn by halving and doubling them relative to their initial value.

The results of this sensitivity study are shown in Figure 3.6. As one would expect, increasing any of the $\lambda_{ij}^{\text{rel}}$ parameters raised both the cost and objective functions (i.e. reduced the goodness of fit). The components of $\lambda_{ij}^{\text{rel}}$ in the latitudinal direction had the greatest effect, which is unsurprising as there is most structure in precipitation fields in the latitudinal dimension. The largest change found was for $\lambda_{22}^{\text{rel}}$, which changed the cost function by approximately 20% when increased with integrated precipitation conservation. Note, however, that this corresponds to only a 5% decrease in the RMS difference between the model and observations. The smallest changes were found to be in the temporal dimension. It is less intuitive what the effects of changing the size of the ratio, γ, would be. Here, it was found that for either form of conservation, halving or doubling this parameter altered the cost and objective functions by no more than 3%, with no consistent sign.

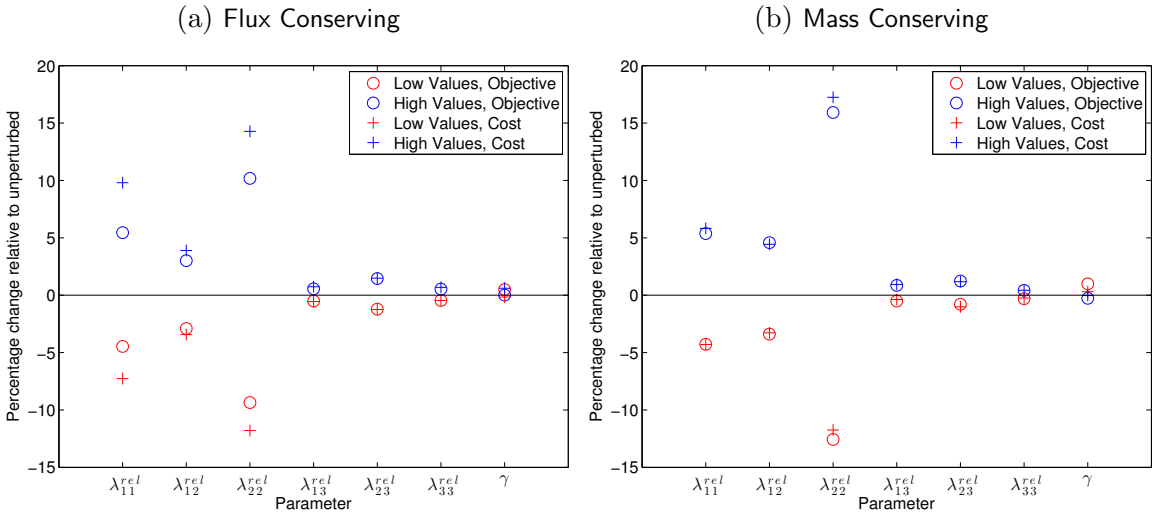

Figure 3.6: Plots showing percentage change in the objective and cost function that resulted from perturbing $\lambda_{ij}^{\text{rel}}$ and γ. The experiment was carried out with both the flux conserving (a) and the integrated precipitation conserving (b) techniques, and each parameter was halved (low value) and doubled (high value) while keeping all other parameters constant.

3.3.2 Warp Characteristics

In order to discuss some of the general characteristics of the warped fields, $\bar{\lambda}$ values are chosen from our range to generate an 'optimum' warp, for both forms of conservation. A $\bar{\lambda}$ value of 4 $(\text{mm/day})^2$ is chosen for both cases - the higher of the two values based on the criteria for the perturbed parameter example (see Section 3.3.1). Although in reality one would expect different values of $\bar{\lambda}$ would be chosen for the two forms of warp, using the same value allows for more direct comparisons to be made.

Figure 3.7 shows the spatial component of the warping process for one month (January), for both forms of conservation. Although both considerably decrease the error with respect to the reference image (GPCP) in a number of places, what is striking is that they do so in markedly different ways (see Figures 3.7c and 3.7d). For example, the intensity conserving warp reduces the error of one of the largest issues in the INM-CM4's field - the double ITCZ - by compressing it, to minimise its impact. In contrast, the integrated precipitation conserving warp shifts the feature,

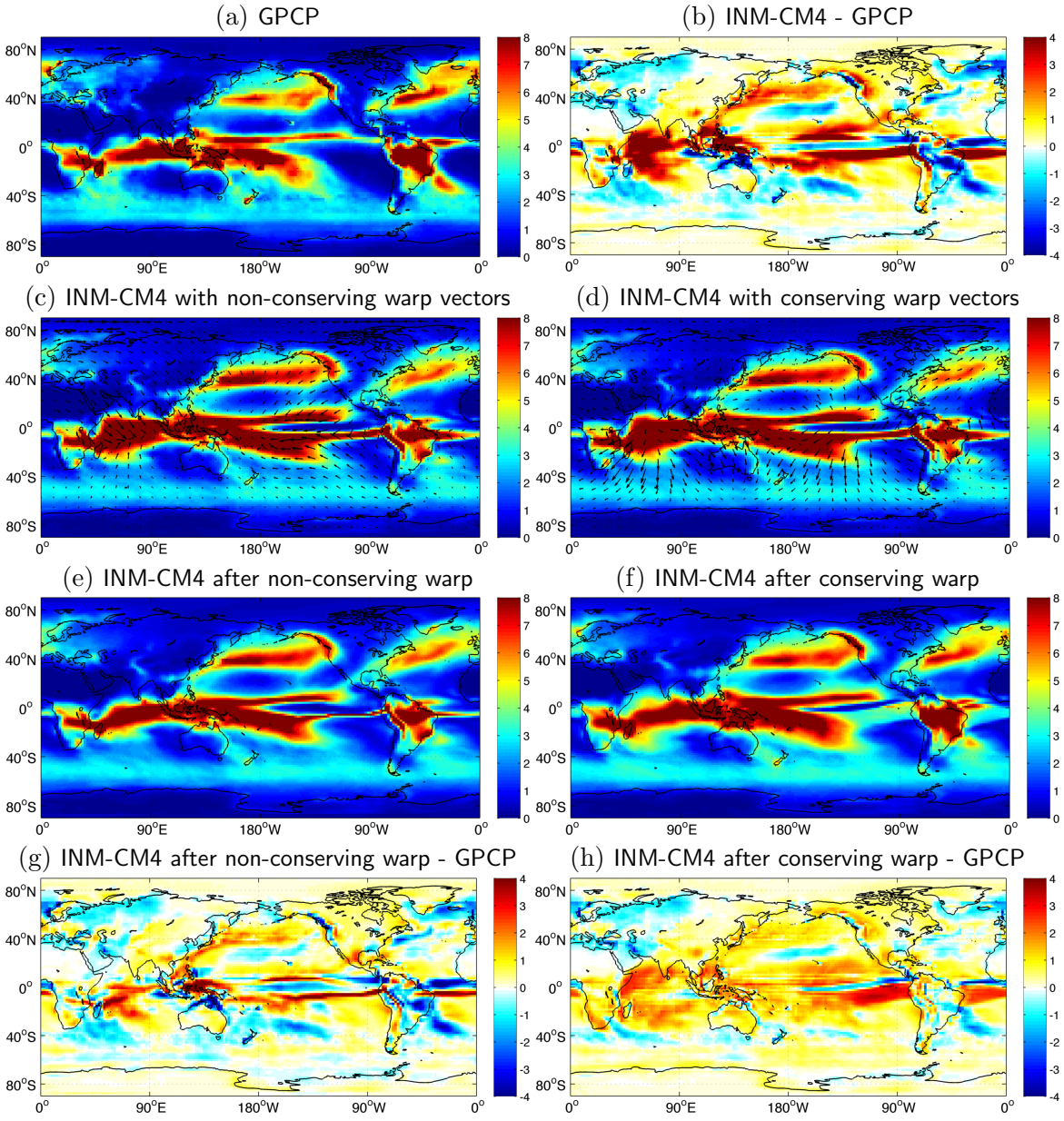

Figure 3.7: Plots illustrating the warping process with $\bar{\lambda}$ set to 4 (mm/day)^2, both with and without the total global rainfall conservation technique. Figures show GPCP (a), unwarped INM-CM4 minus GPCP (b), INM-CM4 with intensity conserving (c) and integrated precipitation conserving (d) warp vectors, warped INM-CM4 in the intensity conserving (e) and integrated precipitation conserving (f) cases, and INM-CM4 minus GPCP after warping with intensity conserving (g) and integrated precipitation conserving (h) warps. All units are mm/day, and all fields are for January. Vectors in (c) and (d) illustrate only the spatial components of the warp, with vectors scaled up by a factor of 5, and only a every 5th vector in latitude and longitude shown, for clarity.

which aligns it better with the actual ITCZ, while also expanding it to reduce its grid point intensities.

More generally, it can be seen that both warps are smoothly varying, while smaller scale errors are reduced more in the integrated precipitation conserving warp (consistent with the finding that warps with integrated precipitation conservation were able to reduce model error further). As a result, there are more regions of strong inconsistency remaining after warping in the intensity conserving case. With integrated precipitation conservation, these small areas of large error have been greatly reduced, but the GCM's global bias toward too much precipitation remains, as it cannot be removed due to the constraint on conserving total integrated precipitation.

Note that this GCM - INM-CM4 - is also used to exemplify the warping process with *fnirt* in Chapter 2, allowing for comparisons to be made between this technique's two forms of conservation and *fnirt* (see Figure 2.2). As one would expect, this Chapter's intensity conserving transformations operate in a qualitatively similar way to *fnirt*, which also conserves intensities upon warping. However, the intensity conserving technique in this Chapter is able to reduce errors to a greater extent, as can be seen, for example in the erroneous southern ITCZ (comparing Figures 3.7e and 2.2d).

The integrated precipitation conserving technique reduces error in a markedly different way to both *fnirt* and the intensity conserving technique of this Chapter (as described above). However, both versions of this Chapter's technique are able to reduce smaller scale errors than was possible with *fnirt*. This can be seen, for example, in the greater reduction of the bias on the West Coast of North America, highlighted in the difference plots. This is because the transformations derived here are restricted to being less smooth, and so are not constrained to correcting only the largest biases in the climate, which are typically in the tropics (both due to the difficulties of simulating convection, and the relatively large amount of tropical

precipitation).

To test the consistency of the results in this section, all experiments have been repeated with HadGEM2-ES (*Jones et al.*, 2011). This GCM was chosen in contrast to INM-CM4, as its total integrated precipitation changed by the smallest amount when warped by *fnirt*. All results discussed in this section are qualitatively similar to the results for HadGEM2-ES. Most notably, the warps that conserved total integrated precipitation again outperformed those that conserved grid point intensities, and the warp selected under the criteria above again reduced RMS error by more than 50%.

To compare the potential of this new technique to warps derived using *fnirt*, a range of warps is calculated for the same GCMs used in Chapter 2 (Table 2.1). Warps are selected using the same simple criteria used in that study - choosing the warps that reduce the most error, while requiring that the Jacobian determinant lies between 0.1 and 10 everywhere. Using this criteria it is found that the technique reduced RMS error of GCMs by 59% and 46% with integrated precipitation or intensity conservation technique, respectively. This is in contrast to the 35% reduction found using *fnirt*.

Qualitatively, this new technique allows us to apply a new conservation option, whereas it was only possible to conserve fluxes when using *fnirt*. While here it is found that the transformation corrects location errors in many regions of the climatology, *fnirt* only converged for very smooth warps, and so was only able to correct the dominant features of the climatology (such as the latitudinal width of the precipitation feature over the Indian Ocean, as illustrated by Figure 2.2). Further, *fnirt* had no periodic boundary conditions for either the longitudinal or seasonal dimensions, and so the climatology field was instead padded with replicas of itself. While this enabled an approximation of periodicity, it resulted in non-physical transformations, with vectors entering from *both* the easterly and westerly boundaries, for example. By operating in the correct geometry, the technique presented here is able to provide more physically meaningful location corrections than *fnirt* whilst also offering the

potential to correct more small scale features of the climatology.

3.4 Conclusions

In this chapter, a new technique has been introduced for the removal of spatial and seasonal biases from GCM precipitation fields. While previous work has demonstrated the potential of such techniques (*Brown et al.*, 2012b; *Levy et al.*, 2013), we are not aware of other techniques that deal with either the correct geometry of the problem (spherical spatial geometry, with periodic seasons), or have the capacity to approximately conserve total integrated precipitation upon warping.

This Chapter has investigated how several key aspects of the warp vary with respect to the warp's overall smoothness parameter, $\bar{\lambda}$, ensuring the stability of warps to perturbations. How errors in integrated precipitation conservation depend on this parameter, as well as the resolution of the grid on which the warps are calculated, is also analysed. Having introduced several new parameters not found in traditional medical image registration, this Chapter examines the warp's sensitivity to their variation, finding in particular that the latitudinal regularisation parameters have the greatest effect on the ability of the warps to match to observations.

It was found that warps that were derived to conserve total integrated precipitation were able to remove more error from GCM climatologies, due to the ability of these warps to alter intensities by translation, or through compression or expansion of a region. Using the same criteria for selecting optimal warps as used in Chapter 2, the average RMS error was reduced by 59% and 46%, with integrated precipitation and intensity conservation, respectively, both of which are substantially higher than the 35% achieved using *fnirt*. It is especially notable that conserving integrated precipitation allows the removal of more climatological error than possible with flux conservation as, rather than act to constrain the warps, it allows two mechanisms

for the correction of location errors (shifting features, and concentrating or diluting them).

While a simple criterion for selecting warps is used in this Chapter (similar to that used in Chapter 2), the choice of this parameter remains subjective to a degree. Different applications will require differing balances between accuracy of fit and smoothness of warp, and so a careful choice of criteria for each application should be made. However, warps with close values of $\bar{\lambda}$ are demonstrated to be similar - indicating the stability of the warping process. Further, using both forms of conservation in future work could further test the robustness of any results to the underlying technique.

It is found that the two conservation algorithms transform the GCM field in qualitatively different ways. This indicates that in certain situations, one approach may be more applicable, regardless of its effectiveness in reducing climatological error. It is noted that, due to the discrete nature of the algorithm used, integrated precipitation conservation does not perfectly conserve total global rainfall. Finding that this discrepancy depends on the resolution of the grid, it is demonstrated that it can be substantially reduced by increasing the resolution of a warp before applying it.

Further development of this technique could allow the overall smoothness parameter, $\bar{\lambda}$, to vary with location, potentially allowing for the derivation of warps that vary smoothly, yet prioritise the correction of precipitation in certain specific regions (over land, for example). It would also be possible to adapt the cost function so that it takes more than one pair of fields as its arguments, allowing warps to be derived that simultaneously correct related physical quantities (such as precipitation minus evaporation, and sea surface salinities - *Durack et al.* (2012)).

Chapter 4

Deriving and Analysing Transformations

In order to analyse either historical or projected simulated changes in precipitation, it is necessary to make use of an ensemble of different GCMs, so that a wide range of behaviour is explored. Uncertainty in changes can be inferred from the extent of disagreement on changes, though (as discussed in Chapter 1) some of this disagreement is caused by biases in the GCMs' mean climatologies, which we aim to correct here.

We therefore derive transformations for a range of GCMs, and develop criteria for the selection of 'optimal' transformations (Section 4.1). Such transformations should - to some degree - balance goodness of fit with the smoothness of the transformations. As the nature of location biases will vary between different GCMs, the optimal parameters should be selected for each model independently, based upon common criteria.

As well as removing location biases, the derived transformations also offer a unique opportunity to characterise the location biases of GCMs. In this Chapter, therefore, we also investigate properties of these transformations that are common across the ensemble of GCMs, and how physical properties of the models are related to differ-

ences in the derived transformations (Section 4.2). The results of this chapter are discussed in Section 4.3.

4.1 Deriving and Selecting Warps for an Ensemble

In order to remove spatial and seasonal biases from GCM simulated precipitation, we apply the tool described in Chapter 3. As detailed above, this technique allows either precipitation fluxes to be conserved, or total integrated precipitation. While conserving total integrated precipitation may seem the more appropriate option in that it maintains consistency with the energy and moisture budgets, as it changes precipitation flux values it is further from the basic intention of merely correcting errors of location. We therefore derive transformations for both algorithms to compare their behaviour.

As we aim to remove location biases from each GCM's mean climate, we use the GCM's mean monthly climatology as the input field, I. The satellite/gauge based observation dataset GPCP (*Huffman et al.*, 1997) is used as the reference field, R, as it provides global, gridded precipitation values. The GCM's climatology is calculated using its historical simulation, and only the years where the model and observations overlap (1979-2005) are used to form monthly means. Where more than one historical simulation is available for a particular GCM, we take a mean across these ensemble members.

As described in Chapter 3, besides the form of conservation used, there are a number of tuneable parameters included in our warping technique. These are the overall smoothness parameter, $\bar{\lambda}$, the relative smoothness parameter, $\lambda_{ij}^{\text{rel}}$, and the ratio of seasonal distances to spatial distances, γ. While future work could further refine the selection of transformations by varying all of these parameters (providing 7 degrees of freedom), here we keep $\lambda_{ij}^{\text{rel}}$ and γ fixed, for simplicity. All values in $\lambda_{ij}^{\text{rel}}$

are set to 1, such that translations in all directions are penalised equally, and γ is defined by the most fundamental climatological ratio between space and time in the climate: the mean rate of change of the sun's angle of latitudinal angle of incidence. The precise number is given by the axial tilt of the Earth (23.5° - *Williams*, 1993), divided by the number of months between an equinox and the preceding solstice (three).

We derive a range of warps from which to select our 'optimal' transformations, setting $\bar{\lambda}$ to a range of 12 values, varying from 0.25 to 100 (mm / day)^2, as we found that this range included warps that struck a balance between goodness of fit and smoothness (see Section 3.3.1). Warps are derived for both forms of conservation. As shown in Figure 3.3a, goodness of fit generally improves as $\bar{\lambda}$ decreases, except at very low values.

To investigate the nature of these improvements further, we calculate the change in RMS error relative to GPCP for 30° latitude bands for all GCMs. Again we find that goodness of fit generally improves as $\bar{\lambda}$ decreases, except at very low values. However, for many cases, this is not the case for the -90° to -60° band, as the lack of land in the Southern Ocean means that there are few identifiable precipitation features at these latitudes, and so the warps are relatively unconstrained here.

In selecting optimal transformations, however, it is important to consider more than simply the reduction in error relative to the observed precipitation climatology. Warps that are selected purely on their goodness of fit will have little applicability, as they will take no account of how physical the warp process is (*Rohlfing*, 2012). We therefore select warps based both on their goodness of fit, and a measure of how severely they distort the GCM's field.

To quantify deformations in the warp, we use its Jacobian determinants. The Jacobian determinant quantifies the extent to which volumes change upon warping, and so a very high or low value implies the introduction of folds or tears into the field.

(a) Flux Conserving (b) Mass Conserving

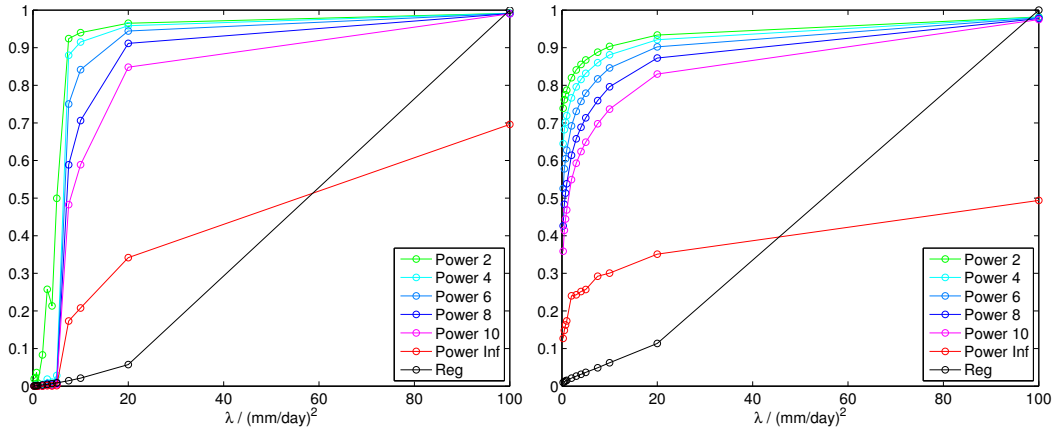

Figure 4.1: Figures showing how the quantity described by equation 4.1 varies with $\bar{\lambda}$ for HadGEM2-ES, and different values of n. 'Power Inf' (red lines) refers to taking the limit of n to infinity. Also shown is the value of the regularisation term (black lines). All values are unit-less, and values of the Jacobian powers close to 1 indicate a very smooth warp.

A Jacobian determinant of 1 implies no compression or expansion.

In Chapter 2, we selected warps to maximise the improvements seen, while also requiring that all the values of the Jacobian lay between 0.1 and 10. While this requirement prevents extreme distortions, it does not take into account the mean effect of the warp (only the extremal values). In order to allow a more complete diagnostic, we investigate the following quantity:

$$\left[\frac{\sum_i A_i \max\left(\frac{1}{J_i}, J_i\right)^n}{\sum_i A_i}\right]^{-\frac{1}{n}} \tag{4.1}$$

where the sum is over all grid points, A_i and J_i are the area weighting and Jacobian determinant of a gridpoint, i. This diagnostic evaluates the root mean nth power of the Jacobian, and the higher we choose n, the more emphasis placed on extremal values. The variation of this quantity (as well as the regularisation) with $\bar{\lambda}$ is shown in Figure 4.1 for several values of n.

As can be seen with the flux conserving algorithm, there is a sharp transition from very small values (implying sharply varying warps) to values close to 1 (implying smooth warps). This transition is considerably smoother for integrated precipitation conserving (mass conserving) warps, presumably as the Jacobian is included in the cost function (see equation 3.9). By assigning a minimum threshold for a particular power of the Jacobian, we constrain the warps to a particular smoothness. If we set our threshold to 0.1 and take the limit $n \to \infty$, this formula reduces to the requirement in previous work in Chapter 2. A threshold of 1 requires a transformation consisting solely of a uniform translation or rotation.

Here, we set n to 6 and the threshold to 0.67. This puts some emphasis on extreme areas of distortion, while requiring that the overall deviation from a zero transformation is not large and ensuring the selected flux conserving warps are on the smooth side of the transition. While any choice is somewhat subjective, we found that this choice of values ruled out transformations with visible discontinuities, or those where $\bar{\lambda}$ was low enough that transformations were unconstrained and goodness of fit was sacrificed. Further, we found that varying the limits did not qualitatively change the nature of the warps, as transformations of the same model with varying values of $\bar{\lambda}$ transformed in qualitatively similar ways (see Section 3.3).

We derive and select warps in this way for 21 GCMs from 13 different institutions (Table 4.1). All models' RMSs are reduced using either technique, as one would hope. For most GCMs, variation of the above quantity generates similar behaviour to that shown for HadGEM2-ES in Figure 4.1. However, for flux conserving warps of GISS-E2-R, the warps have regions of very strong distortion for all but the most constrained warp derived, which has a $\bar{\lambda}$ of 100 $(\text{mm/day})^2$. This seems to be the result of poorly differentiated precipitation features in this GCM - most notably a merged ITCZ and SPCZ (*Brown et al.*, 2013), which would provide incentive for the

Table 4.1: List of GCMs for which transformations are derived, with the number of historical ensemble members listed for each. Also shown are the values of $\bar{\lambda}$ selected for each GCM for the flux conserving ($\bar{\lambda}_f$) and mass conserving ($\bar{\lambda}_m$) cases, with units of $(\text{mm/day})^2$, and percentage reduction in RMS error after warping in the flux conserving (Δ_f^{rms}) and mass conserving (Δ_m^{rms}) cases.

Model Name(s)	Modelling Group	Runs	$\bar{\lambda}_f$	$\bar{\lambda}_m$	Δ_f^{rms}	Δ_m^{rms}
BCC	Program for Beijing Climate Center	6	7.5	2	45	57
CCSM4	University Corporation for Atmospheric Research	6	7.5	2	43	55
CNRM-CM5	Centre National de Recherches Meteorologiques	10	5	2	44	51
CSIRO-ACCESS 1.0	Commonwealth Scientific and Industrial Research Organisation	1	7.5	2	49	57
GFDL-CM3	NOAA Geophysical Fluid Dynamics Laboratory	5	5	3	50	54
GFDL-ESM2G	NOAA Geophysical Fluid Dynamics Laboratory	1	7.5	3	47	56
GFDL-ESM2M	NOAA Geophysical Fluid Dynamics Laboratory	1	7.5	3	43	53
GISS-E2-R 1.0	National Aeronautics and Space Administration	6	100	4	12	48
HadGEM2-AO	Met Office Hadley Centre	1	7.5	2	47	54
HadGEM2-CC	Met Office Hadley Centre	1	10	2	44	55
HadGEM2-ES	Met Office Hadley Centre	4	7.5	2	47	55
INM-CM4	Institute for Numerical Mathematics	1	4	3	48	50
IPSL-CM5A-LR	Institut Pierre-Simon Laplace	4	20	2	33	56
IPSL-CM5A-MR	Institut Pierre-Simon Laplace	1	10	2	42	56
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology	3	4	2	52	57
MIROC-ESM-CHEM	Japan Agency for Marine-Earth Science and Technology	1	7.5	2	45	57
MIROC5	Japan Agency for Marine-Earth Science and Technology	3	7.5	1	40	52
MPI-ESM-LR	Max Planck Institute for Meteorology	3	7.5	3	48	56
MPI-ESM-P	Max Planck Institute for Meteorology	2	7.5	3	49	57
MRI-CGCM3	Meteorological Research Institute	5	7.5	4	48	54
NorESM1-M	Norwegian Climate Centre	3	5	2	48	56

technique to introduce tears to the image. As a result of the high value of $\bar{\lambda}$, location biases are reduced by only 12% for GISS-E2-R when flux conservation is used.

On average, however, the warps selected reduced RMS error in the GCMs' climatologies by 44% and 55% in the intensity and integrated precipitation conserving cases, respectively. Despite GISS-E2-R being a clear outlier in the flux conserving case, we include it in our analyses for completeness. Taking an ensemble mean of all 21 GCMs before and after warping, we find that RMS error of the ensemble mean is reduced by 31% and 47%, respectively, relative to the GPCP climatology.

These results confirm the finding of Section 3.3, that conserving integrated precipitation enables the climatological error to be reduced further than conserving fluxes. We also note that the value of $\bar{\lambda}$ selected was smaller for integrated precipitation conserving warps for all GCMs (Table 4.1), indicating that these warps are generally smoother for a given value of $\bar{\lambda}$. This is to be expected, as the cost function in this case incorporates the warp's Jacobian determinants (equation 3.9), and so any areas of extreme distortion will effectively be penalised both by the cost function, and by the regularisation term.

We also found - as with our study of *fnirt* in Chapter 2 - that when conserving fluxes, the total amount of precipitation was reduced after warping, for all GCMs. While the majority of the GCMs were biased toward being too wet before warping, even the two GCMs that were too dry experienced reductions in total precipitation. As this occurs with two different techniques operating in two distinct geometries (the correct climatological geometry, and the Cartesian geometry in which *fnirt* operates), this strengthens the hypothesis that GCMs generally simulate precipitation features over too broad an area. One might expect this to be the case, as GCMs are generally restricted to simulating each grid box as uniform, limiting the lowest scale on which physical processes may be simulated, and hence leading to precipitation features that are larger than those in the observed climate.

Figure 4.2 shows the ensemble mean spatial component of the transformation process for one month (January). The warp vectors (shown in c and d) vary smoothly, and both transformations reduce the error in the ensemble mean effectively (as shown by the reduction in colour in g and h, relative to b). We see that prior to warping there is a propensity for the GCMs to simulate many features over too large an area on average, as described above. This can be seen in the case of the ITCZ, the North Pacific Storm Track, as well as precipitation over the Indian Ocean. Inspection of individual models demonstrates that these biases are common across the ensemble, though present to different extents, as has been noted elsewhere (e.g. *Li and Xie*, 2013). The integrated precipitation conserving transformations are able to reduce more of the small scale errors than the intensity conserving case, especially for superfluous features, such as the spurious southern Inter-Tropical Convergence Zone (ITCZ), which the integrated precipitation conserving transformations are able to dilute through expansion.

To explore how sampling uncertainty could affect the warps derived, we examine differences in the derived warps when different ensemble members of the same GCM

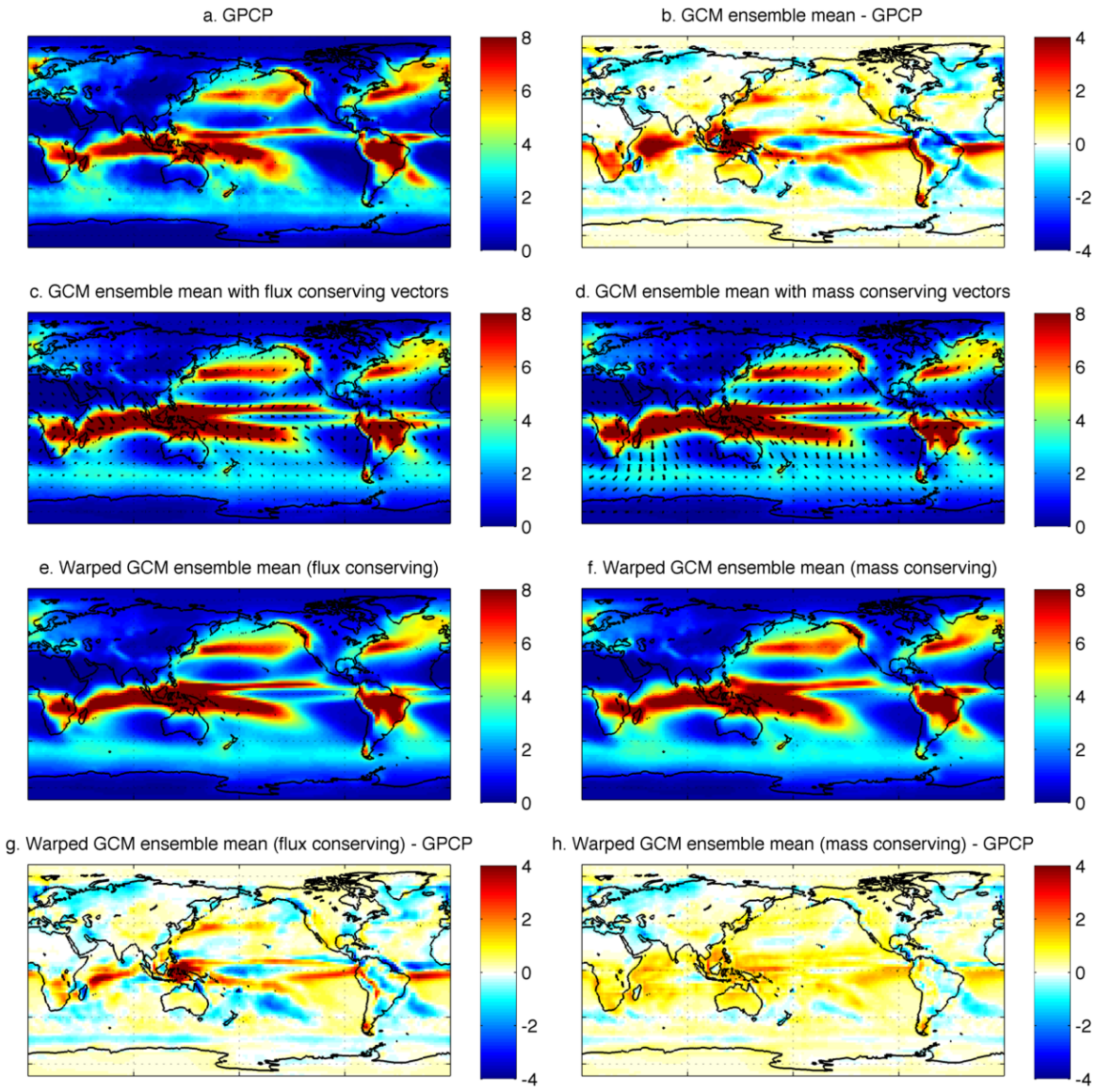

Figure 4.2: Figures showing mean historical precipitation climatology for January. They show the precipitation patterns for GPCP (a) and for the GCM ensemble means before (c and d) and after (e and f) warping. To highlight changes, difference plots are shown before (b) and after (g and h) warping. For each GCM, we derive and select (see equation 4.1) transformations to remove spatial and seasonal biases relative to the GPCP climatology. The ensemble mean vectors (c and d) are displayed above for both forms of warp. All units are in mm day^{-1}, and ensemble means are across 21 GCMs. Historical means are taken over the period of overlap of the CMIP5 historical simulations and GPCP (1979 - 2005).

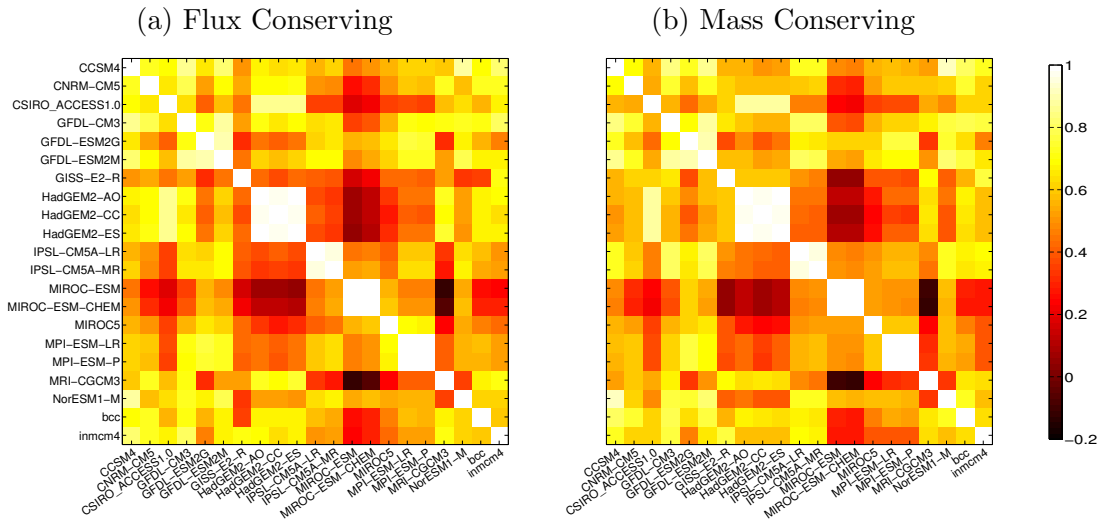

Figure 4.3: Figures showing the correlation between the warps for all different GCMs listed in Table 4.1. Correlations are calculated for both the flux conserving (a) and total integrated precipitation conserving (b) cases. The temporal component of the warp vectors are multiplied by the spacetime ratio, γ (see Section 3.2.12), before the correlations are calculated.

(IPSL-CM5A-LR) are used as input fields. These different ensemble members constitute different samplings of the GCM's variability. This allows us to estimate the sensitivity of the technique both to initial conditions (i.e. the input fields) and to limitations with regards to optimisation, which result from a finite number of iterations and becoming 'stuck' in local minima rather than the global minimum. We find the mean lengths of the three vector components (longitude, latitude and season) and the mean standard deviation in these three directions. For the flux conserving warps, these are $(0.22 \pm 0.02, 0.49 \pm 0.03, (7.9 \pm 0.8) \times 10^{-3})$, and for integrated precipitation conserving warps we find $(1.4 \pm 0.1, 1.2 \pm 0.1, (4.6 \pm 0.5) \times 10^{-2})$, with units of the first two dimensions in degrees and the third in months. The large size of the vector lengths relative to the uncertainty demonstrates the robustness of the derived warps to sampling uncertainty both in the input fields and the algorithm.

4.2 Characterising Location Errors

There have been several efforts previously to characterise the similarity between various GCMs, either through direct comparisons of their simulated mean climate (*Knutti et al.*, 2013; *Masson and Knutti*, 2011) or through comparison of spatial distribution errors in their simulation of present-day climate (*Pennell and Reichler*, 2010). Our derived transformations provide another mechanism for comparing GCM climatologies, through characterisations of location biases. We therefore resample all transformations onto a common grid, and calculate the correlation coefficients between GCMs, for both flux conserving and integrated precipitation conserving warps.

Figure 4.3 shows the correlations in these two cases. As can be seen, the correlations between GCMs are similar for either form of conservation, implying that although the two techniques behave in qualitatively different ways, there is consistency in how they vary between models. The one exception to this is GISS-E2-R, which has weaker correlations in the flux conserving case due to the high constraints placed on the warp (discussed above). It is interesting to note that in all but two cases (between MRI-CGCM3 and the two versions of MIROC-ESM), the correlations are positive, with an average correlation of 0.5 in both cases. This demonstrates that there are strong similarities in location biases across almost all 21 CMIP5 GCMs. This reaffirms that certain large scale biases in precipitation distributions are common across many GCMs, as in the case of the double ITCZs (*Li and Xie*, 2013), for example.

Some GCMs have much more similar warps than others, however. We can see that various versions of the same model tend to have very similar location biases, as demonstrated by the high correlations among the different variants of HadGEM2, IPSL-CM5A, MIROC-ESM and MPI-ESM. While the GFDL models are all similar, their correlations are comparable to some calculated between GCMs of different modelling groups. Indeed, it is not uncommon for different modelling groups to share code

(*Knutti et al.*, 2013) and some GCMs, although developed by different groups, have very similar location errors. For example, the correlations between the HadGEM2 models and CSIRO_ACCESS1.0 are very high, which is unsurprising as the latter is based upon the atmosphere used in the former (*Dix et al.*, 2013).

In general, inspection of the correlations between different GCMs implies similar relationships to those depicted in the 'family tree' of *Knutti et al.* (2013)[1]. This demonstrates that the metrics used there (taking into account seasonal cycle, the inter-annual variations, and the spatial correlation of precipitation and temperature) are closely linked to location errors of precipitation. This is as expected, as location errors will take seasonal and spatial variations into account. Nevertheless, our findings here reinforce the notion that GCMs generally have similar biases, with the closeness of certain GCMs reflecting a commonality of institution or parameterisation. Indeed, accounting for such similarities, *Pennell and Reichler* (2010) estimated that the 'effective number' of models in a 24 member ensemble lies between 7.5 and 9.

We next investigate how the derived warps depend on underlying properties of the GCMs. As discussed in Chapter 1, simulation of precipitation remains difficult for GCMs primarily because many of the relevant physical processes take place on scales much smaller than that of an average grid box. As a result, such sub-grid processes must be parameterised based on grid scale properties. *Kendon et al.* (2012), for example, demonstrated that realism of precipitation can be dramatically improved with increased resolution. GCMs with sufficiently high resolutions to simulate the relevant convective processes directly, however, would require substantially more computational resources than are currently available with modern supercomputers.

In Figure 4.4, we investigate whether resolution affects climatological error in GCMs by plotting RMS error (relative to GPCP's climatology) against the resolution

[1]Note that *Pennell and Reichler* (2010) includes few GCMs from the CMIP5 ensemble, and so we do not compare Figure 4.3 to their findings.

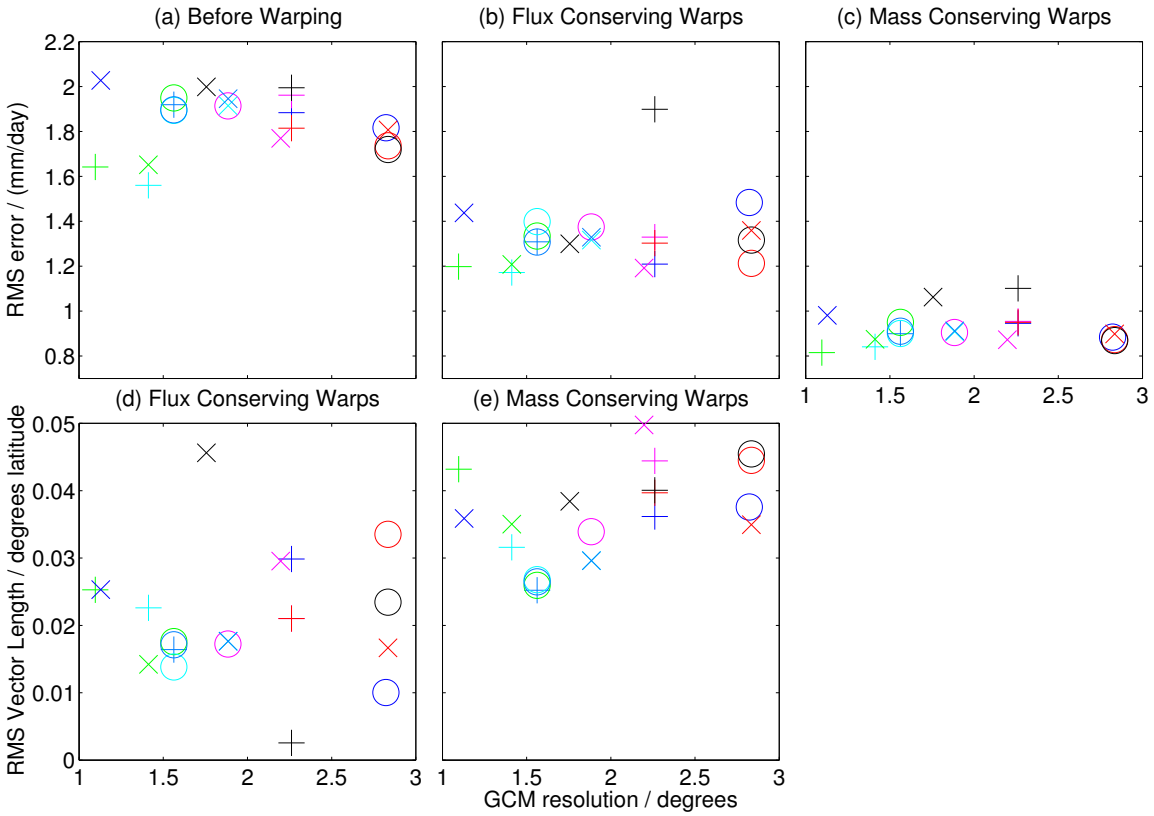

Figure 4.4: Figures showing how the RMS error varies with resolution for each of the 21 GCMs shown in Table 4.1. RMSs are relative to GPCP, and are shown before (a) and after warping, with either flux conservation (b) or total integrated precipitation conservation (c). Also shown are RMS vector lengths (in degrees latitude) against GCM resolution in the flux (d) and mass (e) conserving cases. GCM resolution is calculated as the mean of the longitudinal and latitudinal resolutions of the GCM. Marker shape and colour are consistent across all subplots.

of the GCM. While it can be seen that before warping the three GCMs with the lowest RMS error have relatively high resolution, there is no clear relationship between grid spacing and the accuracy of precipitation simulations. We also investigate whether there is a correlation after warping, as location biases may be less dependent on resolution than intensity biases, and so the former may mask a relationship in the latter. Again we find that - after warping with either form of conservation - there is little correlation between climatological error and resolution, implying that at the range of resolutions used by modern GCMs, changes in resolution have little impact

on the realism of local precipitation fluxes.

However, RMS error is only one approach to characterise GCM biases. The warping technique also gives us a tool to characterise the location biases, through the RMS lengths of warp vectors (note that these are calculated by normalising the seasonal vector component with the spacetime ratio, γ). Figures 4.4d and 4.4e show how the mean length of warp vectors vary with the resolution of the GCM. We find that although there is no significant correlation in the flux conserving case, there is a correlation of 0.5 in the integrated precipitation conserving case (p-value of 0.03). The reason this correlation is apparent in the latter, but not the former case, may be because the range of values of $\bar{\lambda}$ selected for flux conserving warps ranges from 4 to 100 $(\text{mm/day})^2$, whereas $\bar{\lambda}$ only varies from 1 to 5 $(\text{mm/day})^2$ in the integrated precipitation conserving case. As $\bar{\lambda}$ constrains the warp, this means there may be more consistency between GCMs on the physical significance of warp vectors' lengths in the integrated precipitation conserving case.

This demonstrates that although resolution seems to have little impact on the realism of grid-point precipitation fluxes, it may affect a GCM's ability to simulate a more realistic precipitation climatology. However, it is important to note that this correlation does not explain all variation in GCM realism. Indeed, while increasing resolution may have many benefits, the resolutions of CMIP5 ensemble are much lower than those used in 'high resolution' models, such as that described in *Kendon et al.* (2012). Therefore, increases in resolution will have limited impact on the need for parameterisation of physical processes for the range of resolutions explored here, and the realism of the parameterisation schemes used will remain crucial to the quality of the GCMs' simulated precipitation.

It is also of interest to evaluate how errors in the GCMs' simulations of precipitation relate to their realism in simulating clouds and water vapour. The distributions of water vapour at various heights, as well as the formation of clouds, are clearly two

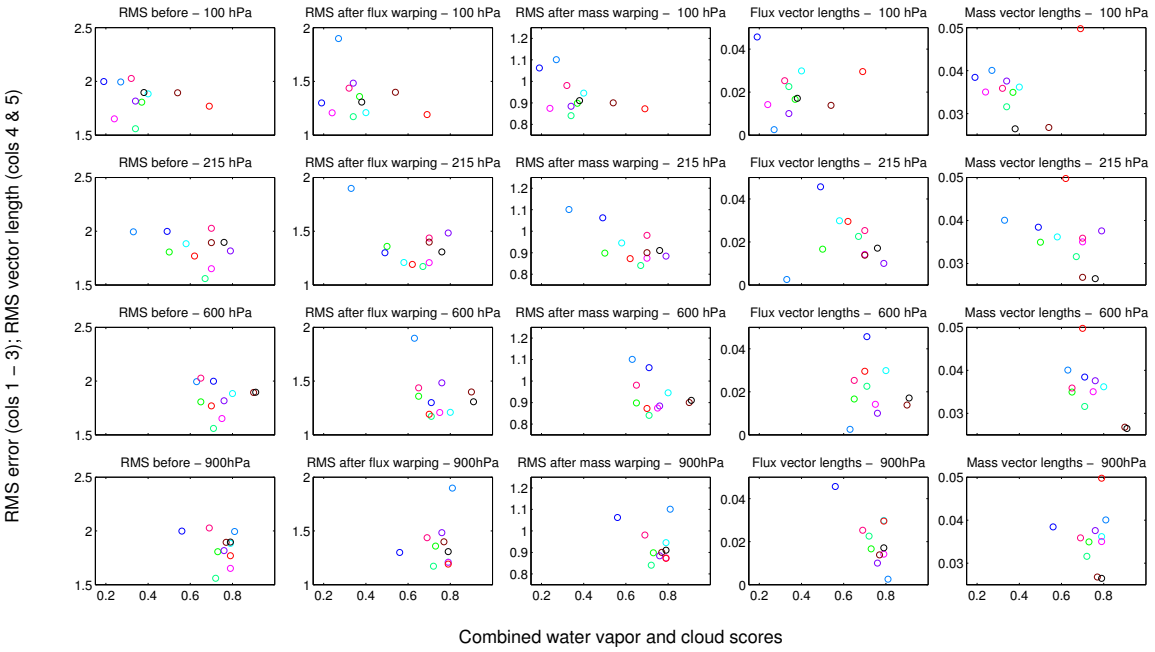

Figure 4.5: Figures showing how the GCM precipitation errors depend on the water vapour and cloud scores (x-axes), as evaluated by *Jiang et al.* (2012). Plots include all GCMs for which warps were derived and scores are available, and are shown for a range of heights (rows). y-axes show the RMS error before and after warping (columns 1 - 3) in mm/day, and RMS vector lengths in latitudinal degrees (columns 4 and 5). The water vapour and cloud scores are unit-less. Note that marker colour and x-axes ranges are consistent across all subplots, whereas y-axes vary in each column.

of the atmospheric variables most linked to precipitation. Water vapour and clouds are also two of the most important feedbacks, and so their realism has further implications in estimating the sensitivity of the climate to increases in greenhouse gases (*Held and Soden*, 2000).

To see how precipitation varies with these properties of climate models, we use the combined scores derived in *Jiang et al.* (2012), which use the A-Train dataset (*L'Ecuyer and Jiang*, 2010) to evaluate model realism in simulating the spatial mean, variance and correlation of the water vapour mixing ratio, cloud ice water content and cloud liquid water content. We plot RMS precipitation error relative to GPCP against these scores at each of the four heights at which they are evaluated, as shown in the first three columns of Figure 4.5. Also shown in the fourth and fifth columns

are RMS vector lengths relative to model score.

We find, however, that there are no consistent patterns of correlation here, though the majority of correlations are found to be negative, implying that the higher a GCM's score, the lower the error, as one would expect. As we only find three cases in which the p-value is less than 0.05, it is not possible to determine robust information on the relationship between precipitation errors and cloud and water vapour realism.

The reason for these weak relationships may in part be the sample size - there are only 12 GCMs featured in *Jiang et al.* (2012) for which we have derived transformations. This will potentially obscure any relationship, and may lead to spurious correlations that are caused by outliers. For example, we find that in the case of *RMS vector lengths with flux conservation - 900 hPa* a correlation of -0.72 and a p-value of 0.01, though this seems to be caused primarily by the dark blue datapoint. Indeed, the algorithm we use to calculate p-values assumes large sample sizes, which we lack here. These problems are exacerbated further in the flux conserving cases, as the value of $\bar{\lambda}$ varies greatly between models, as discussed above.

4.3 Conclusions

In this chapter, we have derived and selected warps for 21 CMIP5 GCMs. Here, we keep two of the parameters, $\lambda_{ij}^{\text{rel}}$ and γ, fixed for simplicity, varying only $\bar{\lambda}$. To select warps based upon more than goodness of fit, we use the root mean sixth power of the warp's Jacobian, which emphasises regions with greater distortion, while still taking into account the entire warp field, unlike the criterion used in Chapter 2. Note that the requirement for the transformation to be physically sensible is stronger here where we wish to remove location biases in order to better characterise changes. In cases where we simply wish to characterise the location biases of a GCM (as in Chapter 6), this requirement can be somewhat relaxed.

We demonstrate that the derived warps reduce climatological error in the GCMs by an average of 44% and 55% in the flux conserving and total integrated precipitation conserving cases, respectively. As found with INM-CM4 in Chapter 3, we see that the improvements in the inter-model ensemble mean are qualitatively different for the two types of transformation. Again, the flux conserving technique is limited to changing the size of features, whereas the total integrated precipitation conserving warps are able to correct features by concentrating or diluting them.

In order to investigate whether the derived warps are robust to uncertainty in the input fields, we derive transformations for different ensemble members of the same GCM, and demonstrate that standard deviations in the vectors are small relative to mean vector lengths. We also note that GCMs from the same institutions have similar precipitation fields, as suggested by *Knutti et al.* (2013), and that these lead to similar transformations. Further, GCMs from other modelling groups generally have similar location biases, in some cases as close as GCMs from the same institution. Our results are in general agreement with *Knutti et al.* (2013) on the 'genealogy' of CMIP5 GCMs. Although this lowers the 'effective number' of GCMs in our ensemble (*Pennell and Reichler*, 2010), throughout this thesis we treat all GCMs listed in Table 4.1 as independent samplings for simplicity. Future work could, however, use weightings when taking ensemble means, in order to reflect similarities between certain GCMs.

To investigate how GCM errors in precipitation simulation depend on other properties of the GCM, we investigate model resolution, as well as water vapour and cloud realism. By investigating both RMS error (before or after warping), and RMS vector length, we can investigate both grid-point and location errors' relationship to these GCM properties.

It is interesting to note that there is little evidence of correlation between grid resolution and precipitation grid box intensity errors, although there is reasonable correlation between warp vector size and resolution, in the integrated precipitation

conserving case. This implies that increasing resolution primarily improves the spatial and seasonal distribution of the precipitation climatology, rather than the realism of precipitation fluxes. Resolution cannot explain all variation in GCM accuracy, however, as the parameterisation schemes used remain crucial for GCM realism. Further, as discussed in Section 1.3, new approaches for designing GCMs may break down relationships between resolution and realism if smaller scale processes could be simulated directly, by relaxing the requirement that they are simulated with perfect precision (*Palmer et al.*, 2014).

We also investigate how precipitation error relative to GPCP depends on the realism of water vapour and cloud properties at various different heights. We find that there there are no clear relationships between either RMS error or RMS vector length, and the scores given in *Jiang et al.* (2012). This is likely due to the small sample size, and so this relationship could be better investigated in the future by deriving transformations for other models that feature in *Jiang et al.* (2012), which includes several CMIP3 models.

While one might expect that scores evaluating the spatial distribution of water vapour and clouds would correlate more strongly with measures of precipitation error, it is important to note that - especially in the flux conserving case - $\bar{\lambda}$ will vary between GCMs, introducing another influence on the strength of the warps, beyond pure GCM error. We note that it would be possible to extend the scoring mechanism of *Jiang et al.* (2012) to the realism of precipitation, though - of course - this could not be applied at multiple levels. Through calculating precipitation scores in this way, the nature of the improvements seen through warping could be better evaluated, through study of which aspect of the score correlates best with RMS changes in precipitation.

Further work could look to investigate how similarities between the derived warps depend on the underlying parameterisation schemes of the GCMs, and compare to other metrics of model similarity (such as *Knutti et al.*, 2013). Further, the precise

role of certain aspects of the climate system in biasing precipitation locations could be determined by investigating differing versions of GCMs. Most conveniently, warps could be derived for atmosphere only implementations of climate models, and differences between transformations could be used to quantify the effect of sea surface temperature biases on precipitation.

To validate whether the corrections derived are physically meaningful, however, it is necessary to apply these transformations to simulated changes in precipitation. While we have already done this for idealised projections, using *fnirt* (see Chapter 2), investigation of future precipitation only allows evaluation of GCM agreement. In other words, it does not demonstrate whether improved GCM agreement indicates improved realism. As a result, in the following Chapter we investigate the effect of applying our derived transformations to historical precipitation changes, where we can compare to observed changes.

Chapter 5

Location Biases in Historical Precipitation[1]

5.1 Introduction

While it is encouraging that correcting location biases can reduce inter-model disagreement on projected precipitation changes, as shown in Chapter 2, investigation of projections does not offer an opportunity for validation of results. In other words, increased model agreement does not necessarily imply that the GCM simulated projections are closer to describing reality. It is therefore of great interest to investigate historical precipitation, not only as it allows for such validation, but removing location biases could also enable new physical insight into the nature of historical forcings on precipitation.

As discussed in Section 1.2.1, understanding of historical precipitation changes remains challenging for a number of reasons. Limited observations that are either limited to land (*Fuchs et al.*, 2007) or have global coverage, but only since the satellite era (*Huffman et al.*, 2009), restrict the ways in which precipitation can be investigated.

[1]This chapter is based upon (*Levy et al.*, 2014b), for which I had primary responsibility. Figures 5.1 and 5.4 are new here.

These limitations are compounded by precipitation's large time scales of internal variability (*Dai*, 2012), and the relatively small length scales of precipitation changes, which are tied to the underlying climatology (*Chou et al.*, 2009).

Indeed, it seems that errors in GCMs' climatologies may mask agreement on changes (as discussed in Chapter 2, and *Allen and Ingram*, 2002). Studies of historical precipitation therefore tend to employ techniques to circumnavigate these discrepancies, for example by averaging over zonal bands (*Polson et al.*, 2013a; *Zhang et al.*, 2007) or by characterising changes relative to each GCM's climatology (*Allan and Soden*, 2007; *Marvel and Bonfils*, 2013; *Polson et al.*, 2014).

In this Chapter, the impact of removing location biases from GCMs is investigated, using the technique introduced in Chapter 3, and the transformations derived in Chapter 4. By correcting features' locations in this way, the aim is to compare GCM simulated changes that are relative to a climatology that resembles the observed mean climate more closely. Any improvements resulting from such corrections are investigated relative to global precipitation observations (GPCP), which have previously been challenging for such studies, both due to the relatively short time period of their coverage, as well as differences between the simulated and observed climatologies. This approach is similar to techniques that assess changes relative to climatology, in that differences in the mean climate are accommodated, though it allows an assessment of how precipitation changes in particular regions.

In this Chapter, the effects of these transformations on errors in simulated historical precipitation changes are investigated first (Section 5.2). We then investigate whether such corrections can affect the extent to which an external forcing can be detected on global historical precipitation patterns (Section 5.3). Finally, the implications of this investigation, as well as ways in which it could be extended, are discussed (Section 5.4).

5.2 Errors in Historical Precipitation Changes

The usefulness of the derived transformations for understanding historical precipitation rests on the assumption that precipitation changes are tied to the underlying climatology, so that correcting the seasonal and geographical location of climatological precipitation features will also correct the location of precipitation *changes* (see Section 1.3.1). This would be expected, as some climatological precipitation changes can be characterised as an intensification of the mean distribution (*Chou et al.*, 2009; *Held and Soden*, 2006). Although shifts are also expected to take place over time (*Yin*, 2005), provided these are small relative to location biases, and the transformation is smooth, the derived warps should still serve to reduce errors. To evaluate the benefit of correcting mean feature location, the transformations derived in Chapter 4 are applied to statistical measures that evaluate historical changes (see Table 4.1 for list of GCMs).

The simplest such field derivable from each GCM's time series is that of anomalies, where the mean climatology is subtracted from each year in the time series. As the warps improve the accuracy of simulated climatology, by construction, subtracting this from the model fields allows a simple approach to assess how warping affects the accuracy of simulated precipitation change. This process is repeated for the GPCP observational dataset, and only years where there is data for both the GCMs and the observations (1979 - 2005) are used. Applying conserving and non-conserving warps to each year of the anomalies, it is found that the RMS error of the GCMs relative to GPCP is reduced by an average of 10.9% in the integrated precipitation conserving case and 10.7% in the flux conserving case.

This reduction in RMS error demonstrates that the warps derived using the new technique have potential to correct simulated precipitation *changes*, rather than just correcting the climatology. While anomalies are the simplest field to derive that lacks the climatology, they are not easily physically interpretable. The effect of warps on

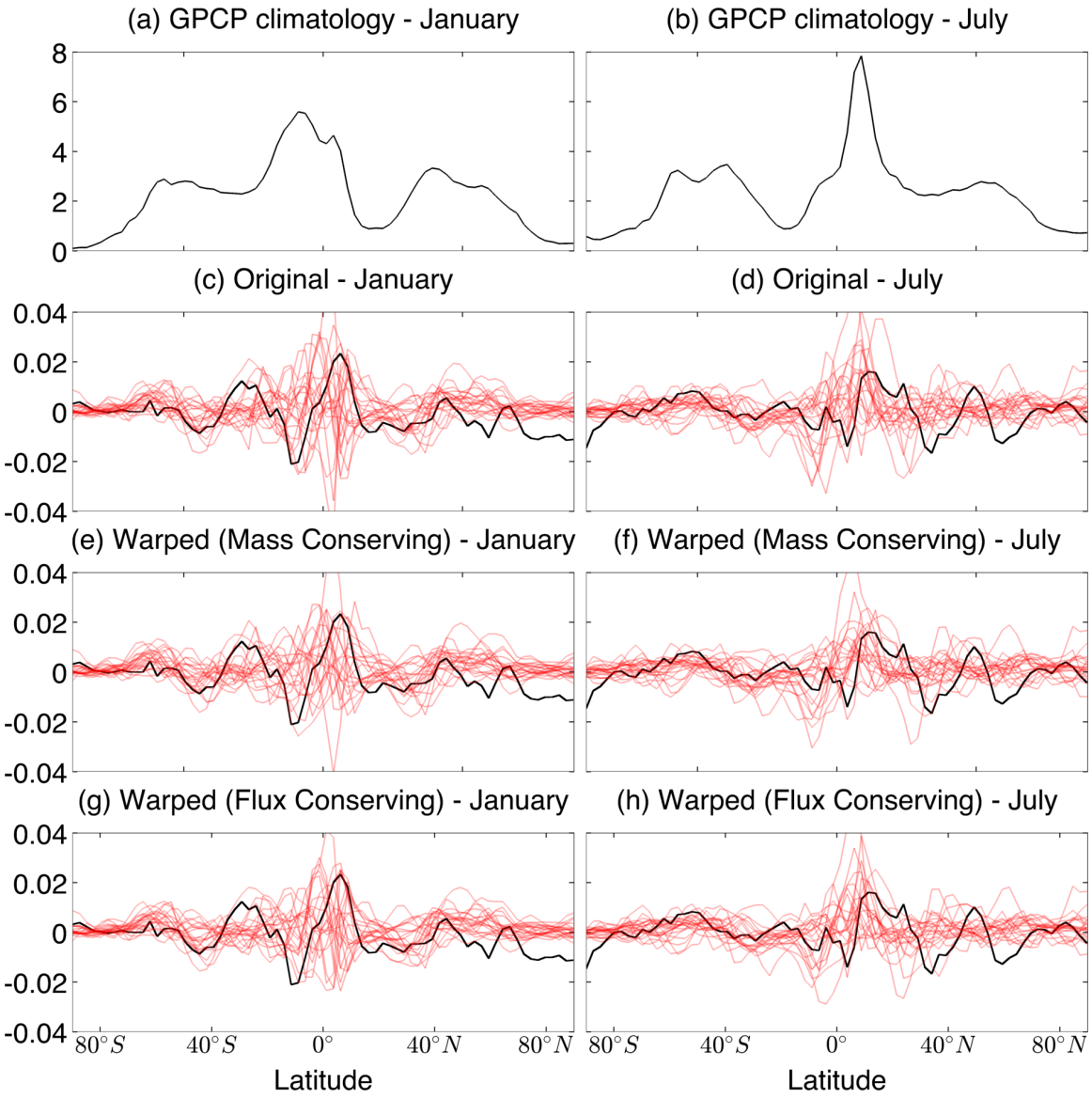

Figure 5.1: Zonal mean climatologies (a and b) and trends (c to h) from 1979-2005. Climatologies are shown for observations, GPCP (black), and trends are shown for 21 GCMs (red lines) and for GPCP (heavy black lines). GCM trends are shown before (c and d) and after warping, with the integrated precipitation conserving (e and f) and flux conserving (g and h) techniques. Units for (a) and (b) are mm/day, and all other units are in (mm/day)/year. The central months of boreal winter (January, left) and winter (July, right) are chosen for their extremal nature.

trend patterns of the 21 GCMs is therefore also investigated.

Trend patterns are created here by taking the trend of each grid point, for each month of the year independently. These trend patterns are created for all GCMs, and for the observational dataset, GPCP, for years with data from both the obser-

vations and the GCMs' historical simulations (1979 - 2005). The effect of warping is then investigated by applying the intensity and integrated precipitation conserving transformations to each GCM's trend pattern.

Zonal means of these trend patterns are shown in Figure 5.1 for January and July. Also shown, for reference, are the zonal means of the GPCP climatology (a and b). The challenge of detecting and attributing a signal in precipitation changes - even in zonal mean patterns - is immediately apparent. The trends in the observations (solid black lines, c - h) vary over small scales relative to variations in the climatology patterns (a and b). While physical insights predict that precipitation changes will involve both an increased contrast between wet and dry regions (*Chou et al.*, 2009; *Held and Soden*, 2006), and poleward shifts due to changing circulation patterns (*Yin*, 2005), it is difficult to clearly infer these processes here. This may be due to complicating factors, such as limited moisture availability over land (*Greve et al.*, 2014), discussed in Section 1.2. Though there is evident drying of subtropical regions, it is not clear whether changes to the ITCZ and storm tracks constitute shifts or intensifications. There exist techniques to decompose, and therefore characterise, these changes, though these involve further reducing the dimensionality of the field (*Marvel and Bonfils*, 2013; *Scheff and Frierson*, 2012).

Not only is there large variability in the observed field, but there is also large disagreement between different GCMs, and with GCMs and the observations. Both before and after warping, there is much disagreement both on the magnitude and sign of precipitation change. Warping using either technique does reduce the deviation of the GCMs from observations, especially in regions with a stronger signal, such as the tropics.

However, even with warping it is notable that there remain regions where the observed trend lies outside the range of responses seen by the models. Most notably, this includes the drying trend in the Northern storm tracks during boreal summer

(shown between 55 and 65° in d, f and h), which none of the GCM simulations capture. Note that this pattern also exists in Southern storm tracks in boreal winter, though it is captured by the range of GCM simulations in that case. It is possible, however, that this discrepancy between GPCP and the GCMs is caused by an artefact in the observation dataset. Land-based datasets display a mixed response in this region over the same latitudes averaged over the Boreal summer, over the longer time period of the 1950s to the late 20th or early 21st Century (*Polson et al.*, 2013a, Figure 2). GPCPs response in this region may therefore be the result of a bias in this dataset, or the internal variability of the climate system.

To calculate the extent of the improvement on trends, the RMS error between the GCMs and observations is evaluated before and after warping (using the whole field, not zonal means). It is found that RMS error is reduced for all GCMs, using either form of transformation. On average, the error is reduced by 11.1% and 11.2% for the integrated precipitation and flux conserving methods, respectively.

While it is clear, therefore, that some disagreement on trends is the result of location biases within GCMs, it is also apparent that there is substantial disagreement due to both the variability of the system, and due to biases simulating the relevant physical processes, and their responses to forcings.

5.3 Detection and attribution case study

Detection and attribution studies aim to identify whether an external effect on a climatological field can be detected, and if this effect can be attributed to a particular cause (e.g. anthropogenic or natural). This is achieved by projecting observed patterns onto expected patterns (fingerprints). Generally, these fingerprints are derived

using GCMs. This process takes the form (*Allen and Stott*, 2003):

$$\boldsymbol{y} = \sum_i \left[\beta_i \left(\boldsymbol{X}_i + v_i\right)\right] + v_0 \tag{5.1}$$

where $\boldsymbol{y}$ is the observed pattern and $\boldsymbol{X}_i$ is the ith fingerprint. v_0 and v_i represent estimates of the variability of the system, and of the fingerprints respectively. β_i is the ith regression coefficient, and so indicates a potential link between observations and the fingerprints. Characterising the noise allows confidence limits to be estimated, and a signal is generally considered to be detected if zero lies below the 5-95% limits for an individual β_i. When more than one fingerprint is used, detectability for some combination of these fingerprints enables changes to be attributed to the physical mechanisms described by those fingerprints.

In many studies, the noise of the fingerprint is low, and so v_i can be neglected (*Stott et al.*, 2001). As the ensemble mean is taken over 21 GCMs, each of which averaged over the available runs (see Table 4.1), the model contribution to internal variability will be small compared to that of the observations. While there will be a substantial contribution to v_i from differences between models, the effect of reducing this noise through warping is precisely what is investigated here. Fingerprint noise is therefore neglected for simplicity. Note that this may bias the best estimate and upper bounds of β toward zero, though *Tett et al.* (1999) found that it did not affect the lower bounds.

5.3.1 Input Fields

To improve detectability, the component of internal variability due to ENSO can be removed before performing the detection study (e.g. *Polson et al.*, 2013b). ENSO variability is removed from GPCP by performing a multivariate regression onto the anthropogenic and natural forcings, as well as the Multivariate ENSO Index (MEI)

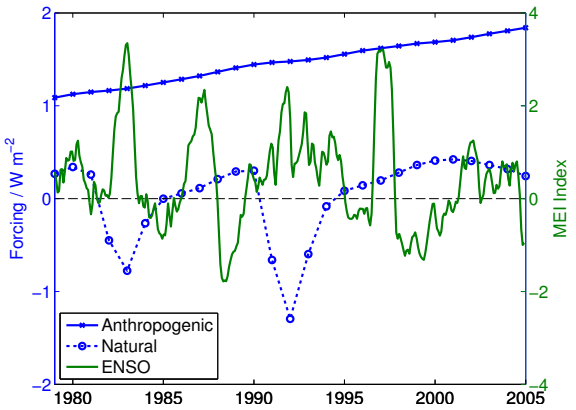

Figure 5.2: Annual mean, globally averaged radiative forcings, estimated for anthropogenic and natural components (blue). Also shown is the monthly MEI Index, which can be used to characterise ENSO (*Wolter and Timlin*, 2011). The forcing time series are used to decompose GCM simulations into anthropogenic and natural components, and the MEI index is used to remove ENSO variability from observations (GPCP).

(*Wolter and Timlin*, 2011), shown in Figure 5.2. The outer product of the MEI time series and the MEI component of GPCP is then subtracted from GPCP's historical simulation, except when the result would imply negative precipitation (fewer than 1% of values). A multivariate (as opposed to single) regression is used in an attempt to reduce the impact of the ENSO removal on the natural and anthropogenic components. The resulting field is referred to as the 'ENSO-free' GPCP precipitation field. While ENSO variability exists within the models, its impact will be greatly reduced in the ensemble mean.

While trend patterns are commonly used to characterise fingerprints, when there is more than one forcing acting on the global climate, linear trends provide limited insight into how these changes are linked to different forcing patterns. This is the case for the historical period studied here, where the external forcings acting on the climate can broadly be categorised into anthropogenic (dominated by greenhouse gases) and natural (dominated by the cooling impact of Sulphate aerosols from volcano eruptions). While the anthropogenic radiative forcing for this period was indeed close to linear, the natural forcing has a pattern that is both non-linear with time,

and asymmetrical over the time period considered here (see Figure 5.2). As a result, one would expect both forcings (natural and anthropogenic) to influence the trend patterns derived, and so separating these influences will require further analysis.

The GCM fields are therefore decomposed into two patterns - representing projections onto the natural and anthropogenic forcings. To achieve this, a multiple regression (*Crooks and Gray*, 2005) is performed of each GCM onto the two global annual mean radiative forcings (*Meinshausen et al.*, 2011), shown in Figure 5.2. As with the trend, this is repeated for each grid point, for each month of the year, independently. Though radiative forcings vary with both latitude and season, the responses to them are spread out by heat transport and storage by atmosphere and ocean, and so depend more on the physical processes of the climate system than the distribution of the forcing (*Boer and Yu*, 2003). Precipitation also depends on the surface temperature response to forcings (the slow response - see Section 1.1), which will further reduce the temporal and spatial specificity of precipitation responses (*Lambert et al.*, 2011). Annual global mean values for these forcings are therefore used.

The two types of conserving warps can again be applied to these patterns. The ensemble means are displayed in Figure 5.3 for both the anthropogenic and natural patterns. Note that these ensemble means are taken across all GCMs, each of which is itself a mean over that model's available historical simulations. As can be seen, the GCMs' time variation projects much more strongly onto the anthropogenic forcing pattern (effectively a trend pattern). For both components, and for both types of transformation, warping reduces the small scale features in the ensemble mean, while maintaining the broad structure of the original patterns. This is especially clear for the ITCZ and the South Pacific Convergence Zone (SPCZ), where the anthropogenic pattern is much smoother after warping.

Further, regardless of warping, one can see that many regions of the natural pattern show the negative response to that shown in the anthropogenic pattern. This is

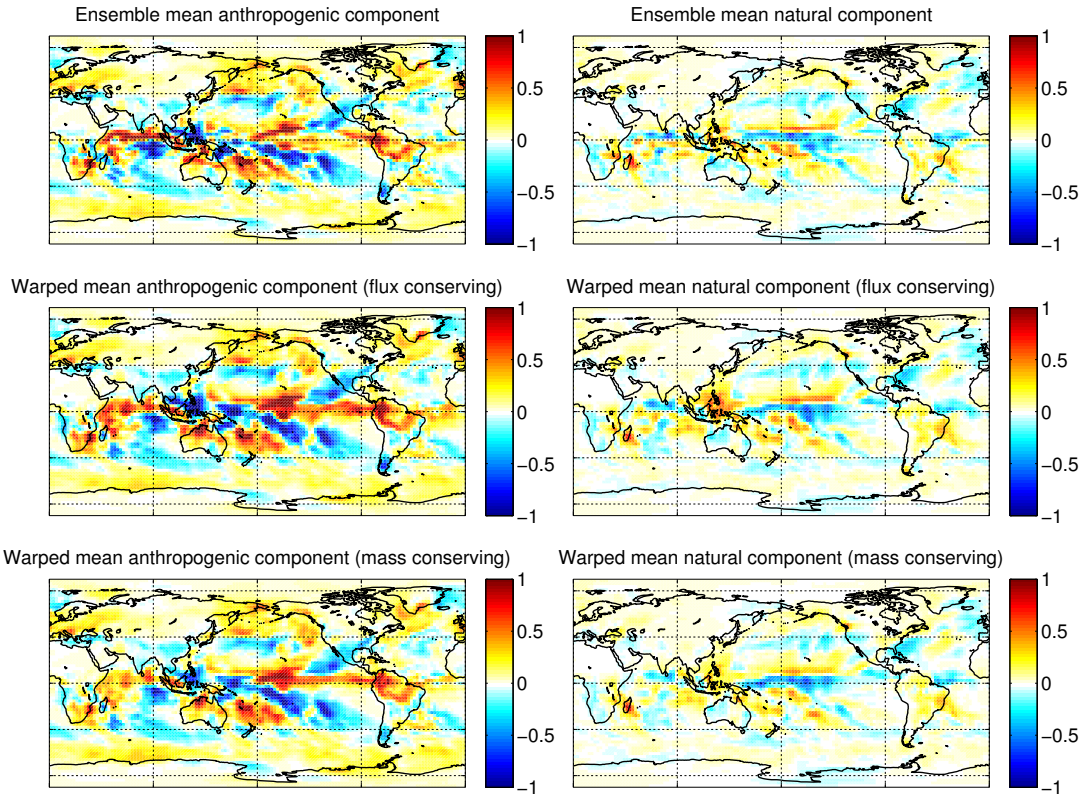

Figure 5.3: Plots showing ensemble mean regressed patterns for January, with anthropogenic (left) and natural (right) components. These patterns are derived by performing a multiple ordinary least squares regression of each GCM's historical run (for 1979 - 2005) onto global annual mean radiative forcings, independently for each grid point, for each month of the year. The ensemble means are taken of the 21 GCMs before, and after applying intensity or integrated precipitation conserving warps (labeled flux and mass above, respectively). Units are in mm day^{-1}(W m^{-2})$^{-1}$.

not an artefact of this particular month, or the form of warping, as the correlation coefficient between the two patterns is -0.2 either with or without warping, taking all months into account. Note that this indicates differing precipitation response per unit forcing, and so does not result from the opposing signs of the natural and anthropogenic forcings. This supports the notion that although an increase in temperature will lead to an atmosphere that holds more water, understanding the form of this heating is crucial to understanding the precipitation response. The nature of the forcing must be taken into account as evaporation and precipitation serve to trans-

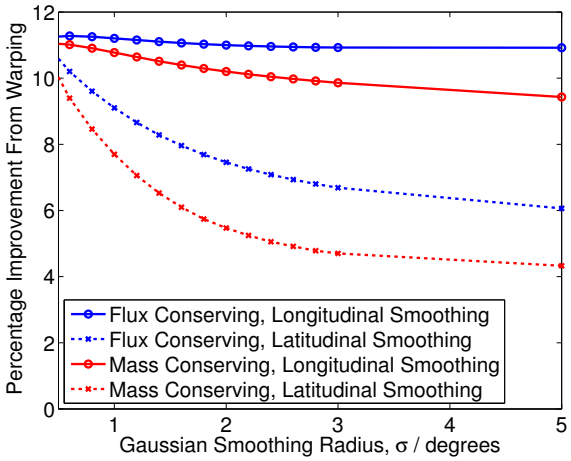

Figure 5.4: Mean percentage change in RMS between the trend patterns of the GCMs, and that of GPCP. Improvements shown are for flux conserving and integrated precipitation conserving (i.e. mass conserving) warps, and are calculated while smoothing in either the latitudinal or longitudinal dimensions by the specified angle. Smoothing in either direction reduces the improvement seen through warping, though latitudinal smoothing has a much larger effect than longitudinal smoothing, for either form of conservation.

port heat from the lower to upper troposphere, and so a forcings will have different impacts on precipitation if they heat the lower or upper troposphere, as discussed in Section 1.1.

Previously, internal-variability and GCM biases have meant that studies that aim to detect and attribute an external signal in precipitation changes have been required to greatly reduce the dimensionality of the field. Approaches include taking annual or seasonal means, latitude bands (*Allen and Ingram*, 2002; *Polson et al.*, 2013a; *Zhang et al.*, 2007), and characterising changes in terms of shifts and intensifications of extremal zonal values (*Marvel and Bonfils*, 2013). While studies that attempt to attribute changes of a particular nature (e.g. intensification of the contrast between wet and dry regions) have been able to make use of shorter satellite/gauge based observational datasets (*Marvel and Bonfils*, 2013; *Polson et al.*, 2013b), studies that attempt to identify a signal in the entire field generally require longer time series, and so have been limited to precipitation over land.

Here, we attempt to detect and attribute a signal using the entire year, without limiting the investigation to land. It remains, however, desirable to reduce the dimensionality of the field partially, to reduce the impact of internal variability. As the aim is to determine the potential impact of warping techniques on such studies, we first investigate whether reducing the dimensionality in each dimension impacts on the improvements we see from warping. Naturally, if a global, annual mean was taken, the effect of the warp would be reduced to near zero, as there can be no location biases in such a field.

Figure 5.4 shows how smoothing in the latitudinal and longitudinal directions affects the improvement resulting in warping. As can be seen, smoothing in the east-west direction does not substantially sacrifice improvements seen from warping, though north-south smoothing greatly reduces the effect of the warp. It is also found that taking annual means of the fields reduces the improvements seen through warping by more than a third for both forms of conservation. This is to be expected, as the circulations that drive large scale precipitation patterns vary much more strongly with latitude and season than with longitude. Although there is of course variation in orography in the longitudinal direction, location biases arising from these forms of precipitation are not dominant in the fields. Zonal means are therefore taken for this study, without reducing dimension in the seasonal or latitudinal dimensions, except to match the GCMs' latitudinal resolution to that of GPCP (2.5° spacing).

For the observed climate pattern, therefore, we simply take the zonally averaged ENSO-free GPCP time-series, with the monthly mean subtracted, so that it forms a time-series of monthly anomalies. This is projected onto anthropogenic and natural fingerprints, which are composed by taking the outer product of the GCM ensemble mean coefficient patterns (described above, and shown in Figure 5.3) with the relevant forcing time-series (Figure 5.2). The fingerprints are therefore a simulated estimation of the natural and anthropogenic components of the historical time-series.

Performing an ordinary least squares regression requires an estimate of the external noise. Here, 90% confidence intervals are derived by replacing the GPCP field with sections of pre-industrial control runs of the same length. The control runs used are from the same 21 GCMs used to simulate historical precipitation, which provide us with a distribution of 463 correlation coefficients. This large sample size forms an approximately Gaussian distribution, which allows us to derive confidence bands based upon the standard deviation of these coefficients. Note that by using climate model control simulations in computing confidence intervals, the covariance structure of internal climate variability will be taken into account.

No attempt is made to remove ENSO from these control runs, since El-Niño variability will be uncorrelated between different GCMs, and any attempt to derive a suitable index for each one would necessarily be both onerous and subjective. Thus the model simulations that provide an estimate of natural variability will contain extra noise, which will tend to make tests conservative. In order to test the impact of correcting location biases, this experiment is repeated with warped historical runs, and with both historical and control runs warped, with both forms of conserving warp. By applying the warps to simulated internal variability as well as the fingerprint signals, the aim is to address model-model differences in climate variability.

5.3.2 Results

The results of these regressions are shown in Figure 5.5. As can be seen, with or without warping, no signal is detected when projecting onto the natural component. Indeed, in all cases the best guess is a slightly negative correlation coefficient. For the anthropogenic component, on the other hand, there is consistency with one in all cases. Before warping, there is marginal consistency with zero, indicating that a signal cannot be detected in this case. With either intensity or integrated precipitation conserving warps (with or without warped control runs), however, the 90% error bars

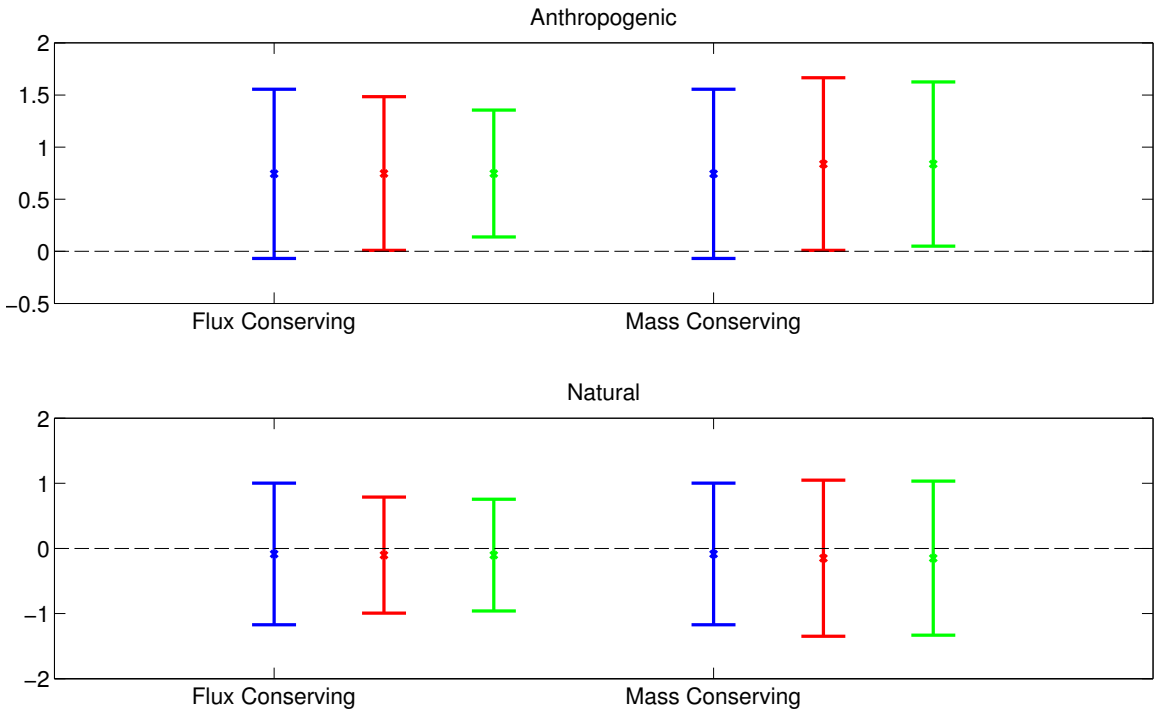

Figure 5.5: Regression coefficients ('β') from regressing GPCP anomalies onto anthropogenic (top) and natural (bottom) components of CMIP5 ensemble mean. The observed patterns are taken from GPCP, with ENSO variability removed by regression as described in the text (Section 5.3.1). The error bars represent the 90% confidence range, derived using GCM control runs to capture the variability. The best guess values for the regression coefficients are shown by crosses. Shown are results without warping (blue), with warped historical runs (red) and with warped historical and control runs (green). Detection is marginally achieved in all anthropogenic cases where warping is applied, but in none of the natural cases (where the best guess is slightly negative). The GCMs were warped using intensity conserving (left) and integrated precipitation conserving (right) warps.

no longer cross the x-axis, indicating that the anthropogenic signal is detected at the 5% level by this particular analysis.

While 95% confidence bands are the standard used, they reflect an arbitrary choice. It is therefore more useful to consider the *change* in confidence of the detection, rather than the binary of whether or not a signal has been detected under a particular criterion. Here, an anthropogenic signal was detected with 93% confidence before warping. After warping both historical and control runs, the confidence in

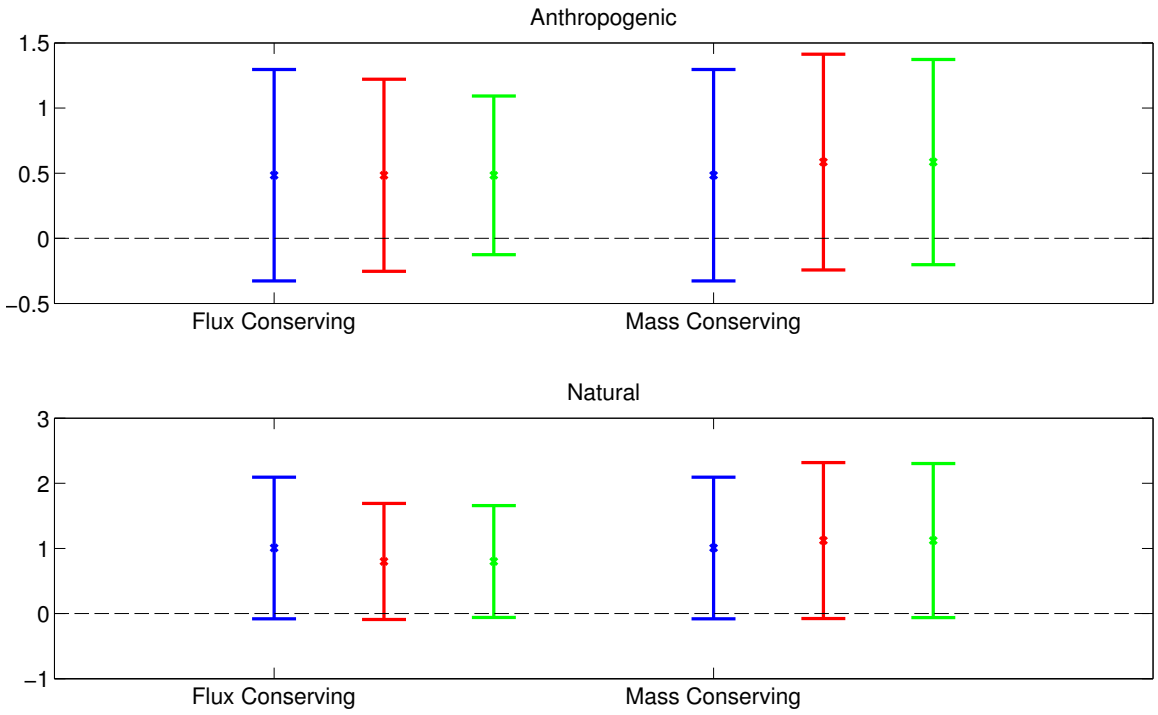

Figure 5.6: As with Figure 5.5, but without ENSO variability removed from GPCP. Detection at the 95% level is no longer achieved in the anthropogenic case, though warping the models still results in improved detectability. Improvements in the natural component are negligible.

the result is increased to 98% and 96%, in the intensity and integrated precipitation conserving cases, respectively.

To verify that these improvements are not an artefact of the removal of ENSO from GPCP, results where ENSO is not removed from GPCP are also considered. Figure 5.6 shows the regression coefficients that are found when using the same procedure described above, without removal of ENSO variability from GPCP. While detection is not achieved with 95% confidence, it can be seen that warping again improves the detectability of the anthropogenic component. Here, confidence is increased from 84% before warping to 90% and 89% in the intensity and integrated precipitation conserving cases, respectively. Again, warping does not substantially affect the detectability of the natural component, although not removing ENSO from GPCP provides better consistency with 1 for this component while bringing it closer to detection at the 95%

confidence level.

5.4 Discussion and Conclusion

The results of this detection and attribution study demonstrate that our ability to detect a human influence on historical precipitation changes can be enhanced by correcting the location of features in GCMs. In the case presented here, these changes are attributable only where location correction techniques are used. While this attribution relies on an arbitrary threshold, it has have also demonstrated that the likelihood of a non-positive regression coefficient is reduced from 7% to 2% with intensity conservation, or to 4% with integrated precipitation conservation.

There are several reasons why correcting the location of features enhances the detectability of the signal. By transforming the GCMs' historical runs, features are lined up better before taking ensemble means, so that the signal is not smoothed out before the fields are merged. More importantly, the ensemble mean changes are then better aligned with the observed patterns, increasing our chances of detecting a signal. Further, correcting feature location in control runs allows a better estimate of the variability of the system to be made, which is not artificially inflated by varied biases in different GCMs' mean climates. The improvements in detectability support the expectation that both inter-model agreement, and agreement between models and observations on precipitation changes are partially masked by the various location biases in the GCMs' underlying climatologies.

This does not apply for the natural signal, where the best guess correlation coefficient is found to be small and negative in all cases. This is a result of the ENSO removal technique, as when this study is repeated without removing ENSO variability, the best guess - indicated by the crosses in Figure 5.6 - is found to be close to one for both natural and anthropogenic coefficients (detection with a 95% confidence is lost

for the anthropogenic correlation coefficient, however). Further, as shown in Figure 5.6, the result that warping improves detectability of the anthropogenic component is resilient to whether or not ENSO is removed from the observations.

This negative effect on the natural component due to ENSO removal is due to the known negative correlation in recent decades between the natural forcing and ENSO variability, though there is no clear evidence that this relationship is causal (e.g. *Self et al.*, 1997). For the time period used here (1979-2005), the correlation coefficient between these two time series is -0.4 (see Figure 5.2). As a result, in spite of removing the ENSO signal using a multiple regression that incorporates both the anthropogenic and natural forcings, this has reduced the extent to which a component of observed changes can be attributed to natural forcing. While future work could attempt to better decouple the ENSO and natural signals, the focus here is on investigating the effect of warping on the anthropogenic signal.

It is important to note the sensitivity of the method presented here. The detection of the anthropogenic signal with 95% confidence is only marginal, as suggested by the loss of detectability when ENSO variability is not removed from GPCP. Although leaving ENSO variability in the control runs will have increased the size of the confidence bars, other factors may suggest that the error bars calculated underestimate the uncertainty. Most importantly, model simulated variability may underestimate the true natural variability of precipitation in the climate system. Further, by using an ordinary least squares rather than a total least squares regression, the noise internal to the climate models has been neglected. Although taking an ensemble mean across 21 GCMs (most of which are comprised of several ensemble members) reduces the size of this noise, there remains model noise through differences in model climatologies and underlying physics, though this will primarily bias only the central and upper estimates (*Tett et al.*, 1999).

Nevertheless, the increases in confidence (with or without ENSO removal) demon-

strate that correcting feature location reduces model error in simulated historical changes, which in turn improves the possibility of detecting and attributing changes to external forcings. It is found that in spite of substantial differences in the behaviour of the transformations, both the intensity and integrated precipitation conserving warps enhance detectability. The reduction of RMS error with the observed trend patterns - using either technique - also indicates that more than a tenth of the error in simulated historical precipitation changes can be attributed to errors in location, which in previous studies have gone uncorrected. This is not only encouraging for investigation of historical precipitation changes, but adds confidence that increased inter-model agreement on projections upon warping (see Chapter 2) implies GCMs are closer to the 'true' climatological changes after warping. Further, it reinforces previous results, which have been able to detect external influence on global precipitation patterns by taking climatological differences into account (*Marvel and Bonfils*, 2013; *Polson et al.*, 2014).

In spite of the ability of the integrated precipitation conserving technique to better reduce errors in the climatology, both conservation approaches removed a similar amount of RMS error from GCM trend patterns. This indicates that the applicability of a particular transformation cannot be exclusively determined based on its success at removing errors from a GCM's mean historical climate. The flux conserving technique improves the consistency of GCM simulated precipitation changes with observations, in spite of introducing inconsistency with simulated moisture budgets and tropospheric energy balance. This provides evidence that bias correction of feature location is a valuable tool regardless of its precise form, though one form may be more appropriate than others for a particular application. The uncertainty in the warping technique could be explored further in future work by, for example, using a different form for the regularisation term (R in equation 3.1), or criteria for selecting warps (see equation 4.1).

Note that, in this study, errors in the observation dataset, GPCP, have not been considered. Climatological biases in the observations will tend to be due to errors in precipitation intensity, not location (*Huffman et al.*, 2009). However, land-sea contrast may be affected by differences in satellite retrieval algorithms for these two contexts, which may introduce biases into the warps derived for GCMs . Further, introductions of new satellites may introduce biases in the GPCP trend (*Polson et al.*, 2013b). Any source of error that is uncorrelated between observations and GCMs will tend to obscure any common signal that is present, and so serves to reduce the detectability of a signal, with or without warping. Future work could therefore seek to quantify these sources of uncertainty by using other precipitation observations, both in deriving warps and in comparing observations to fingerprints.

Similarly, this study does not explicitly consider the effect of aerosols on precipitation, which are resolved to varying extents in modern GCMs. Errors in the mean response to aerosols are one of the sources of location biases that the warping technique hopes to reduce, and so are dealt with implicitly. However, there will also be a precipitation response to aerosol forcing over the time period studied, though there remains substantial uncertainty both in the observed (*Stocker et al.*, 2013) and simulated (*Forster et al.*, 2013) radiative forcing. Aerosol microphysics may of course have further influence on precipitation than from their radiative effect, and GCMs' varying abilities to resolve these processes will further reduce one's ability to detect a signal. The techniques used in this Chapter, however, could help quantify such effects by separating climatological and trend biases that result from limitations in the simulation of aerosols.

There have been several efforts previously to detect and attribute precipitation changes by increasing the signal to noise in the observations and fingerprints being studied. This may be achieved, for example, by reducing the dimensionality (*Marvel and Bonfils*, 2013), or by using a longer, land-limited observational dataset (*Polson*

et al., 2013a). By correcting spatial and seasonal biases in GCMs, this Chapter has demonstrated that the agreement between observed and fingerprint patterns can be improved, further enhancing our ability to attribute observed changes to external forcings. This allows more robust understanding of changes in observed precipitation records, and builds confidence in model simulated projections.

Chapter 6

Location Biases in the West African Monsoon

We have demonstrated that removing location biases from GCMs using techniques based on medical image registration can improve agreement on simulated changes, both between GCMs (Chapter 2) and between GCMs and observations (Chapter 5). However, such techniques also have great potential for diagnosing biases in precipitation (*Brown et al.*, 2013), which could aid the development of GCMs, improving fidelity and reducing the need for such bias corrections in the future. We therefore analyse location biases in a particular climatological feature in a GCM, and investigate how these biases vary with physical properties of the model.

6.1 Background

In order to explore underlying causes of location biases in some depth, we investigate a particular climatological phenomenon, rather than investigating the entire climate, as in previous chapters. Due to the small scales of the convective processes involved (relative to grid scales), challenges simulating tropical precipitation are particularly pronounced. We choose to investigate the West African Monsoon, which is charac-

terised by heavy rains over the Sudan and Sahel regions, resulting from the Northward migration of the ITCZ onto land in boreal summer.

The West African monsoon has been extensively studied, in part due to the severe drought experienced in this region during the 1970s and 1980s (*Held et al.*, 2005). Projected changes in the monsoon are therefore of great importance, as the lives and livelihoods of several million people are dependent on it (*Marsham et al.*, 2013). However, there is much disagreement between projections (including on the sign of change), and several GCMs fail to capture the movement of the ITCZ onto land, thus failing to simulate the monsoonal system (*Cook and Vizy*, 2006). Further, there has been little evidence of improved realism or agreement in projections from the CMIP3 to the CMIP5 ensemble (*Roehrig et al.*, 2013). As a result of biases in the region, *Lott et al.* (2014) stress that caution should be taken in interpreting simulations of the Sahel region, and note that simulations of precipitation are generally less reliable than those of temperature. It is therefore of interest to investigate here, as understanding the origin of GCM biases in this feature may assist in its future simulation, and hence improve understanding of how it will vary under climate change.

The onset and location of the monsoon have been shown to depend on sea surface temperatures (SSTs) in the Gulf of Guinea (*Joly and Voldoire*, 2010), as meridional temperature contrast between the Atlantic Cold Tongue and land serve to set up pressure gradients that enhance northward winds, thus moving the ITCZ further North (*Caniaux et al.*, 2011; *Okumura and Xie*, 2004). GCM biases in simulating SSTs have been linked to poor simulation of the monsoon, with warm SST biases preventing sufficient migration of the ITCZ onto land. Indeed, across GCMs the mean latitudinal position of the ITCZ correlates with the meridional temperature gradient between the Gulf of Guinea and the Sahara Desert (*Roehrig et al.*, 2013). By prescribing SSTs, it is possible to successfully simulate the main features of the monsoon, though large errors remain (*Hourdin et al.*, 2010; *Roehrig et al.*, 2013).

As West African Monsoonal rainfall is made up of convective precipitation, the scales of many key processes will be substantially smaller than grid scales for even state-of-the-art GCMs. Some work has therefore questioned the extent to which GCMs can capture key properties of the monsoon without explicitly simulating convection. For example, mesoscale convective systems - responsible for the great majority of West African Monsoonal precipitation (*Mathon et al.*, 2002) - are poorly simulated by convective parameterisation schemes (*Davis et al.*, 2003), as their scales (order 100km) are comparable to current grid scales of GCMs.

Increases in computational resources have only recently enabled convection-resolving simulation of the West African Monsoonal region over many days. By comparing two versions of the same model where one explicitly simulates convection, while the other uses a traditional parameterisation scheme, *Marsham et al.* (2013) show that resolving convection leads to more realistic simulation of both diurnal timing and the northward location of the monsoon. Indeed, it is common for precipitation to occur too early in GCMs (*Stratton and Stirling*, 2012), which may bias the location of the monsoon by introducing biases into the diurnal cycle of heating, and hence boundary layer winds. It is therefore of interest to investigate the extent to which a GCM is capable of correctly simulating the spatial distribution of the monsoon, in spite of requiring convection to be parameterised.

6.2 Experiment Setup

In order to investigate location biases in West African Monsoonal rainfall, we make use of a perturbed physics ensemble (PPE). In PPEs, parameters of interest are varied within the parameter scheme of a single GCM. In contrast, in multi-model ensembles (MMEs, which are used in Chapters 2, 4 and 5) a single 'optimum' set of parameters is generally set for each GCM variant, allowing uncertainty associated with general

model structure and parameterisation schemes to be investigated. While MMEs offer great utility due to the differing approximations used by different modelling groups, and so are widely used (e.g. *Stocker et al.*, 2013), it can often be difficult to unpick the physical cause of differences between simulations. Further, it is not possible to define the 'space' of models that one wishes to explore, or to generate truly independent samplings of that space (*Pennell and Reichler*, 2010).

While PPEs keep the structure of a GCM fixed, they allow parameter uncertainty to be assessed. The 'space' to be explored (the parameter space) is more easily defined than in the case of MMEs (*Murphy et al.*, 2004), and the direct relationship between different ensemble members allows a clearer understanding of the physical cause of disagreements to be assessed. However, each GCM can have hundreds of tuneable parameters (*Irvine et al.*, 2013), and so thorough investigations of the parameter space of interest can prove challenging due to computational constraints.

In particular, for state-of-the-art GCMs, which require several months of supercomputer time to perform a typical experiment, it is generally not possible to investigate perturbed parameter ensembles larger than ten or so members. As a result, parameters are often varied independently, and are only set to one of several discrete values. This limits investigation into the role of different parameters, especially as some parameters interact non-linearly.

6.2.1 climate*prediction*.net Experiment

A more thorough exploration of parameter space is possible, however, by making use of distributed computing (*Allen and Stainforth*, 2002). In climate*prediction*.net (CPDN), a relatively low resolution GCM is compiled so that it can be run on standard home computers. This allows many thousands of simulations to be run, by making use of volunteers' computers during idle processor time, thus enabling a detailed evaluation of the role of the parameterisations in specifying important aspects of the

GCM's climatology (e.g. *Sanderson et al.*, 2008b).

Several CPDN ensembles have been generated, which explore the parameter space using different methods. Here, we make use of the experiment outlined in *Yamazaki et al.* (2013), where an emulator is used to select parameter combinations that generate energetically balanced variants of the model, that do not require radiative flux corrections. This allows us to explore a broad range of parameters, that generate simulations with some level of physical consistency. This experiment is made up of 20,000 variations of HadCM3 (a coupled ocean-atmosphere GCM - *Gordon et al.*, 2000), generated with the intention of estimating the uncertainty of the climate's sensitivity to carbon-dioxide. The parameters are co-varied within this ensemble. By using a coupled model, we hope to investigate not only the role of atmospheric parameters on the West African Monsoon, but also the effect of SST biases.

It should be noted that none of the ensemble members have all parameters within the ranges suggested in previous work (*Murphy et al.*, 2004), though this is not obstructive for our study, as we aim to investigate how properties of the GCM variants' climatologies vary with different parameters, even if those parameters may take non-physical values. The ensemble of *Yamazaki et al.* (2013) has been used to simulate pre-industrial climate, as well as the 'historical' climate (1880-2000). Here, we wish to investigate location biases relative to observations, and so make use of the historical simulations, as they allow for the most direct comparison to be made.

6.2.1.1 HadCM3 Parameterisations

HadCM3 has been used for the last three IPCC reports (*Houghton et al.*, 2001; *Solomon et al.*, 2007; *Stocker et al.*, 2013), as a result of being an early GCM that did not require flux corrections in its default configuration, while performing relatively well compared to other GCMs from the same 'generation' (*Covey et al.*, 2003). The resolution of the atmosphere in this GCM is 2.5° in latitude by 3.75° in longitude

Table 6.1: List of parameters perturbed to generate the ensemble described in *Yamazaki et al.* (2013) (reproduced from that paper).

Description	Component	Name
Atmosphere Parameters		
Rate at which cloud liquid water is converted to precipitation	Cloud	*ct*
Threshold cloud liquid water content over sea	Cloud	*cw_sea*
Threshold cloud liquid water content over land	Cloud	*cw_land*
Empirically adjusted cloud fraction	Cloud	*eacf*
Critical relative humidity for cloud formation	Cloud	*rhcrit*
Ice fall speed	Cloud	*vf1*
Entrainment rate coefficient	Convection	*entcoef*
Albedo at melting point of sea ice	Radiation	*alpham*
Temperature range over which ice albedo varies	Radiation	*dice*
Ice particle size	Radiation	*ice_size*
Horizontal diffusion coefficient for temperature and wind speed	Dynamics	*diff_coeff*
Horizontal diffusion coefficient for water vapor	Dynamics	*diff_coeff_q*
Order of horizontal diffusion for temperature and wind speed	Dynamics	*diff_exp*
Order of horizontal diffusion of water vapor	Dynamics	*diff_exp_q*
Surface gravity wave drag: typical wavelength	Dynamics	*kay_gwave*
Surface gravity wave trapped lee wave constant	Dynamics	*kay_lee_gave*
Lowest model level for gravity wave drag	Dynamics	*start_level_gwdrag*
Number of soil levels from which water can be extracted	Land surface	*r_layers*
Vertical distance over which air parcels travel before mixing with their surroundings	Boundary layer	*asym_lambda*
Constant in Charnock formula for calculating roughness length for momentum transport over sea	Boundary layer	*charnock*
Used in calculation of stability function for heat, moisture, and momentum transport	Boundary layer	*g0*
Roughness length for free heat and moisture transport over the sea	Boundary layer	*z0fsea*
Ocean Parameters		
Isopycnal diffusion of tracer at surface	Dynamics	*isopyc*
Background vertical diffusion of tracer at surface	Dynamics	*vdiffsurf*
Increase of background diffusion of tracer with depth	Dynamics	*vdiffdepth*
Background vertical diffusion of momentum (viscosity)	Dynamics	*vertvisc*
Decay of wind mixing energy with depth	Mixed layer	*model*
Wind mixing energy scaling factor	Mixed layer	*mllam*
Forcing Parameters		
Scaling factor for emission from anthropogenic sulphate aerosols	Chemistry	*anthsca*
Sulphate mass scavenging parameter L0	Chemistry	*l0*
Sulphate mass scavenging parameter L1	Chemistry	*l1*
Model level for SO2 (high level) emissions	Chemistry	*so2_high_level*
Scaling factor for emission from natural (volcanic) emissions	Chemistry	*volsca*

with 19 vertical layers, while the ocean resolution is 1.25° in both dimensions with 20 vertical layers. In total, 33 parameters are covaried in the ensemble studied here, which are comprised of 22 parameters from the atmospheric model, six from the ocean model, and five associated with aerosol forcings (see Table 6.1).

Of primary interest to the West African Monsoonal rainfall, of course, is the parameterisation scheme used for convection (*Gregory and Rowntree*, 1990). In this

scheme, every grid box that is evaluated as being convectively unstable contains an ensemble of convective plumes. These plumes may mix with the surrounding air through the processes of entrainment and detrainment, at each vertical level. The entrainment coefficient, *entcoef*, sets the rate of these processes, such that the higher the value of the parameter, the greater the rate of mixing between the convective plumes and the surrounding air, all other things being equal. The realism of spatial and temporal variations in tropical precipitation have been shown to depend strongly on the parameterisation scheme used to simulate entrainment (*Hirota et al.*, 2014). It is therefore of interest how GCM location biases vary when the entrainment scheme is fixed, but the entrainment rate is varied.

It should be noted that the GCM used in this experiment has been investigated in a number of studies of projected changes in the West African Monsoon. Although *Sylla et al.* (2010) emphasises HadCM3's realism in simulating the vertical zonal wind profile of the Sahelian atmosphere, other studies point out deficiencies in this model. *Lau et al.* (2006), for example, uses a drought monitoring index (*McKee et al.*, 1993) to evaluate CMIP3 GCMs' realism in simulating the Sahel drought, finding that HadCM3's indices correlate negatively with observations. In *Cook and Vizy* (2006), HadCM3 is eliminated from the study as, in spite of simulating the broad spatial structure of the monsoon well, it failed to simulate sufficient precipitation over land. This is, of course, especially of interest here, as it implies that HadCM3 is capable of simulating the correct monsoon distribution, but that it does so with biases in location.

6.2.2 Deriving Warps for CPDN

We derive warps, using GPCP precipitation (*Huffman et al.*, 2009) as our reference field (as in Chapter 4). Fields for CPDN ensemble members are saved as decadal, seasonal means, and so we convert GPCP to seasonal means, and use only decades

for which both GPCP and the ensemble have complete coverage. The historical period used for deriving warps is therefore 1980-1999. We use the entire global field to derive the warps, though only a specific region is used to select and characterise the transformations (see below). As discussed above, there are many thousand variations of HadCM3 included in the ensemble. However, deriving transformations for each model variant takes several hours, and so producing transformations for every ensemble member is not practical, even with the use of a supercomputer. We therefore select 500 ensemble members at random, though more transformations could of course be derived subsequently, if required.

Here, we wish to characterise location biases, rather than correct them, and so deriving transformations with both forms of conservation is unnecessary. In this study, we make use only of the flux conserving warps, as these transform precipitation fields only through the translation of intensities, as opposed to integrated precipitation conserving transformations, which can also concentrate and dilute precipitation features. This can make it difficult to diagnose the nature of location biases in the integrated precipitation conserving case, as compressions and expansions will simultaneously affect intensity and location. While it may be of interest to investigate such intensity changes in future work, here we use flux conserving transformations, as they allow direct characterisation of location biases alone.

Flux conserving warps are derived as in Chapter 4, with no modifications to the range of $\bar{\lambda}$ or algorithm, other than to allow the technique to transform a climatology with four time steps instead of twelve. When investigating the effect of removing location biases on simulated changes in precipitation (Chapter 5), we selected values of $\bar{\lambda}$ that were most appropriate for each individual GCM. Here, it is important that the same smoothness parameter is used for all versions of the model, so that any difference between transformations is a clear characterisation of that variant's fields, and does not result from differing smoothness parameters.

We therefore initially apply the criteria described in Section 4.1 to each GCM variant, using only the region of interest for West African Monsoonal rainfall. We take this region to be an expansion of the 5°-20°N, 18°W-18°E region (taken from *Fontaine et al.*, 2003), so that displacement vectors in the vicinity are also taken into account. We therefore apply the criteria to 20S°-40°N, 30°W-30°E. Although the monsoon generally takes place from July to September (*García-Serrano et al.*, 2013), the seasonal means available for this ensemble precludes us from using this precise period, and so we use the June to August seasonal mean, for simplicity. Note that the larger region described above is used in plotting Figure 6.1, whereas the smaller Sahel region is used in plotting Figure 6.3.

By applying the criteria to the appropriate region of all ensemble members, we obtain a range of values of $\bar{\lambda}$. We note that the standard deviation of $\bar{\lambda}$ across the ensemble is 2.2 $(\text{mm/day})^2$, indicating that there is modest variation in warp smoothness across the ensemble. Here, we simply take the median value, accepting that there may be some regions of high compression or expansion for some ensemble members. Using this approach, we select a $\bar{\lambda}$ of 4 $(\text{mm/day})^2$. Having derived and selected flux conserving warps, we can now analyse properties of the transformations over the West African Monsoonal region (as defined in *Fontaine et al.*, 2003).

6.3 Properties of Derived Transformations

We first investigate the properties of the mean warp, in order to understand the mean biases across the ensemble (Figure 6.1). Although the general structure of the ensemble mean precipitation field is correct, several location biases are clearly apparent. In particular, the ITCZ is too broad, and too far south, leading to too much precipitation over sea, and too little over land.

These relatively simple location biases are corrected well by the warping technique,

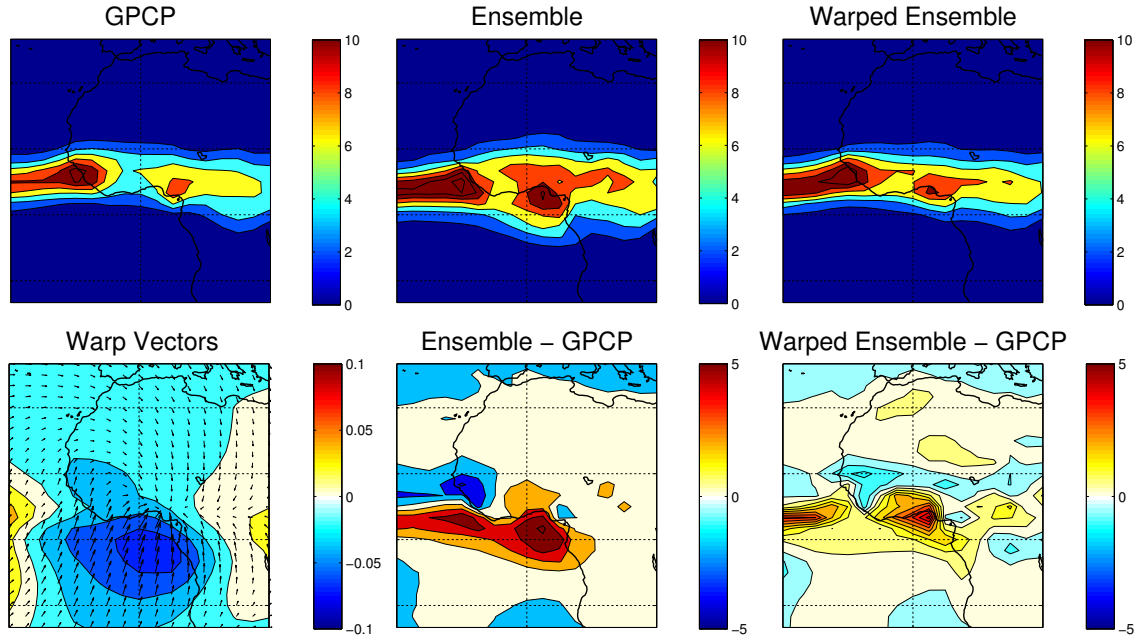

Figure 6.1: Plot of the mean warping process across all 500 versions of HadCM3. All units are in mm/day, except for the warp vectors, where the contours depict the seasonal component of the warp, with units of months. Warps are derived for all GCM variants independently, and for the entire global, annual field, although we only investigate the region relative to the West African Monsoon. Note that these Figures plot over the expanded Sahel region (see Section 6.2.2).

and the warped field closely resembles the observed precipitation pattern. The warp generally shifts the precipitation field northwards, while compressing it in the latitudinal direction, and moving the precipitation pattern slightly earlier. While there is still a slight southerly location bias, it is greatly reduced relative to the unwarped mean field.

By compressing the precipitation band in the ensemble mean, the flux conserving transformation reduces total rainfall across the region. We find here that mean daily precipitation across the West African Monsoon region is reduced in the ensemble from 5.4 mm/day to 4.8 mm/day, bringing it closer to the GPCP average of 4.4 mm/day (note that all averages are for June to August). This exemplifies the ensemble, where the majority (79%) of the GCM variants were originally wetter than GPCP before warping, and warping reduced the mean precipitation of 75% of the ensemble members

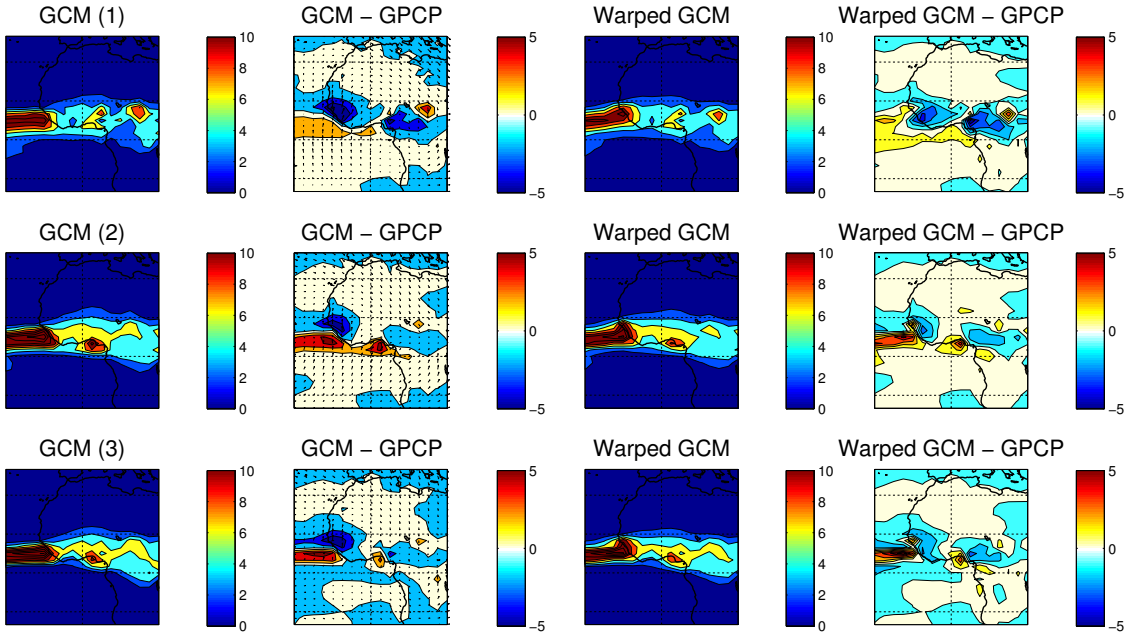

Figure 6.2: Plot of the precipitation fields of the best three GCM variants, as evaluated by RMS error before (1) and after (2) warping, and by RMS warp vector length (3). Each row shows that version of the GCM's precipitation field, and error relative to GPCP, both before and after warping. All units are in mm/day, and spatial vectors are overlaid on the error field before warping.

(taking only the West African Monsoon region into account).

We next evaluate how realistically variants of HadCM3 can simulate the monsoonal precipitation distribution. We evaluate the performance of the simulations using three different metrics: RMS error before warping, RMS error after warping, and RMS warp vector lengths. The fields for the three 'best' versions of the GCM, according to these metrics, are shown in Figure 6.2. All three demonstrate the potential of HadCM3 to simulate the spatial distribution of the West African Monsoon realistically, with certain parameters. Nevertheless, all three variants simulate precipitation slightly too far south before warping, which is well removed (and hence characterised) by the warp vectors.

It is interesting to note that all three criteria select a different version of the GCM as 'best', though they have broadly similar spatial distributions both before and after

warping. GCM versions 1 and 2 are different, as a realistic simulated precipitation distribution does not necessarily mean that the errors in this distribution are easily corrected. Which is more valuable - a GCM with fairly realistic precipitation; or a GCM with a very good distribution in the wrong location - depends on applications, and the applicability of location corrections. It is also interesting to note that the GCM version with the shortest RMS vector length is more markedly different to the other two versions, displaying the largest error both before and after warping. This reinforces the idea that a simulated climatology may have small location biases, but that biases may exist in other aspects of the field, such as the diurnal cycle, or (as we see here) mean precipitation intensity.

From studying other warps with notable properties (such as those with the largest RMS error and vector lengths - not shown), we discern that there is great variation in this ensemble, as one would expect in such a large perturbed physics ensemble in which parameters are not restricted to their recommended ranges. This includes versions of HadCM3 that simulate an ITCZ that covers too broad or too narrow a latitudinal area, and versions that fail to simulate any intense precipitation at all over land. Indeed, there are more diverse behaviours than we have seen in multi-model ensemble studies (*Cook and Vizy*, 2006; *Lau et al.*, 2006; *Sylla et al.*, 2010), implying that this study of a PPE may also capture much of the variation observed in MMEs.

We wish to investigate how location errors in precipitation vary with other properties of the model. We find that the warping technique can generally reduce errors well, implying that location biases are reasonably well characterised by the warping process. RMS error in the West African Monsoon region is reduced by an average of 40%, with a standard deviation of 5%. As noted above, we would not expect complete removal of RMS error, as some error is the result of intensity or distribution biases, rather than incorrect location or size of the precipitation feature.

To check that the warp is characterising location biases in a physically intuitive

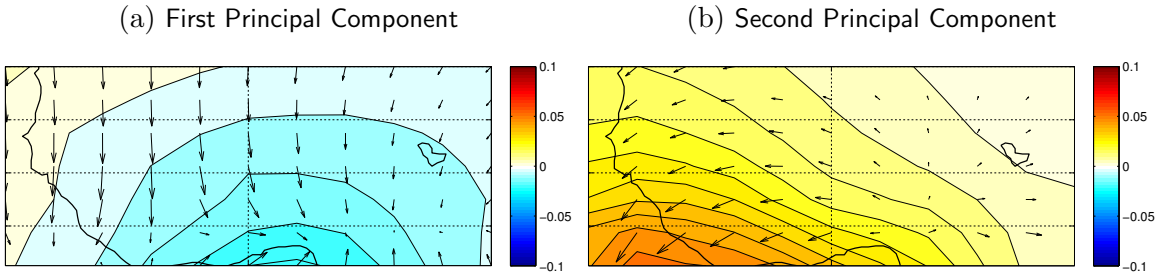

Figure 6.3: Plot of the first (a) and second (b) principal component of the warp vectors (a southerly vector, indicates a bias too far North, for example) over the West African Monsoon region, calculated across the 500 versions of HadCM3. These component capture 64% and 11% of the variability described in the warps, respectively. Colour indicates the variation in the seasonal component of the warp vectors. Note that these Figures plot over the standard Sahel region (described in Section 6.2.2). Also note that mean magnitude of the seasonal component of the first principal component is 16 times smaller than the mean magnitude of the latitudinal component, and that principal components are dimensionless.

way, we investigate the correlation coefficients of the mean latitudinal displacement with mean precipitation latitude, and mean Jacobian determinant with latitudinal precipitation standard deviation. We find the correlations to be 0.7 and 0.3, respectively, both with p-values $\ll 0.05$ (calculated assuming normal distributions). It is unsurprising that the Jacobian determinant has a relatively small correlation with latitudinal standard deviation, as errors in the longitudinal and seasonal dimensions will also contribute to the warps' Jacobians. Nevertheless, it is encouraging that these correlations are robust, indicating that the warp is behaving in a physically intuitive way.

To investigate how location biases vary with properties of the GCM, we first perform a principal component analysis on the derived warps, in order to characterise warp variation. In order to analyse all three dimensions of the transformation, we normalise the seasonal component using the spacetime ratio of the warp (see Section 3.2.12) before performing the analysis. The first two principal components of the warp vectors are shown in Figure 6.3. It can be seen that both modes of variability characterise land-sea shifts to an extent, though the first is primarily latitudinal, and

the second predominantly longitudinal. The first principal component appears to be composed of both latitudinal displacements, with some latitudinal compression also taking place, and accounts for the majority of the variability (64%), while, the second component captures only 11% of the variability of the location biases. We therefore focus our analyses on the primarily latitudinal first component.

6.3.1 Role of Parameters

We now investigate how location biases in the GCM vary with the underlying physical parameters. As mentioned above, parameters are simultaneously varied across the ensemble, and fully decomposing the effects of 33 different parameters is an onerous task. In practice, however, many parameters will have little or no effect on the precipitation distribution, and our sample size is large enough that looking at each parameter independently should highlight those parameters that do impact precipitation, in spite of some amount of spread. We therefore plot scores for the first principal component against each parameter, and investigate whether any relationship is apparent.

Figure 6.4 shows these plots for all 33 parameters that are varied in this ensemble. Most parameters show only a very weak correlation, even with a sample size of 500. There are, however, several parameters that have a clear relationship with the first principal component score. Most striking is the entrainment coefficient, which has a correlation 0.81 with the scores, indicating that location biases in the ensemble depend very strongly on this coefficient. As the relationship between these scores and *entcoef* is clearly non-linear, we also investigate $\log(entcoef)$ and the reciprocal of *entcoef*. We found that the correlation with $\log(entcoef)$ was greatest (0.88), and the relationship is shown in the bottom right plot of Figure 6.4.

There are several reasons to expect that *entcoef* would strongly dictate the location biases of the Sahel Monsoon. Not only does this parameter directly affect convective processes (*Gregory and Rowntree*, 1990), but changes to *entcoef* have been

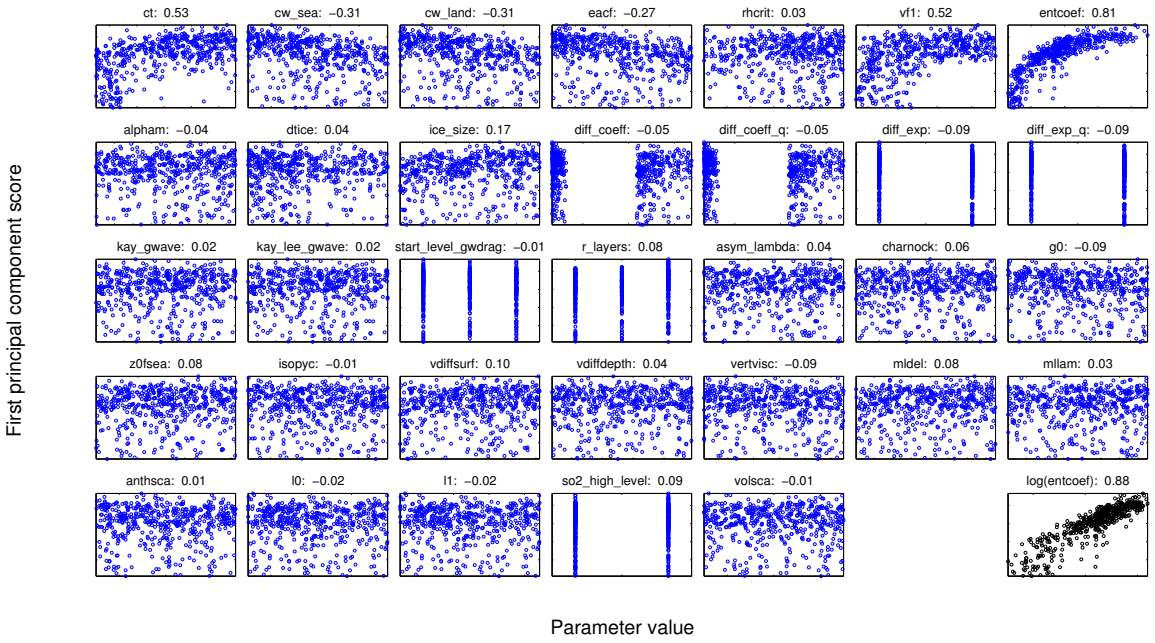

Figure 6.4: Plots showing how the scores of the first principal component vary with parameter value for all 33 parameters, and across all 500 versions of HadCM3. Titles show the name given to the parameter, and the correlation coefficients. For all multi-dimensional parameters, the first dimension is used that captures the variation of that parameter. Note that certain parameters may only take discrete values. Also shown is the score against $\log(entcoef)$ (bottom right). Although p-values are not shown, we find that they are small (less than 2%) for all correlations stronger than 0.1.

found to be primarily responsible for controlling the sensitivity of the atmosphere to increases in carbon-dioxide (*Joshi et al.*, 2010; *Sanderson et al.*, 2008a,b). *Sanderson et al.* (2008a) shows that *entcoef* affects the tropical relative humidity distribution - with low values leading to lower relative humidity in the lower troposphere, and vice versa - and hence the water vapour feedback.

Although the first principal component score also correlates with several other parameters, it is possible that these correlations do not imply a causal relationship. Rather, the entrainment coefficient is so central to the dynamics of the model, that it may dictate to a certain extent the values the other parameters may take in order to avoid the need for radiative flux corrections. Therefore correlation between the location bias of the West African Monsoon and a particular parameter may simply result from both depending on *entcoef*. We therefore perform a multiple linear regression of

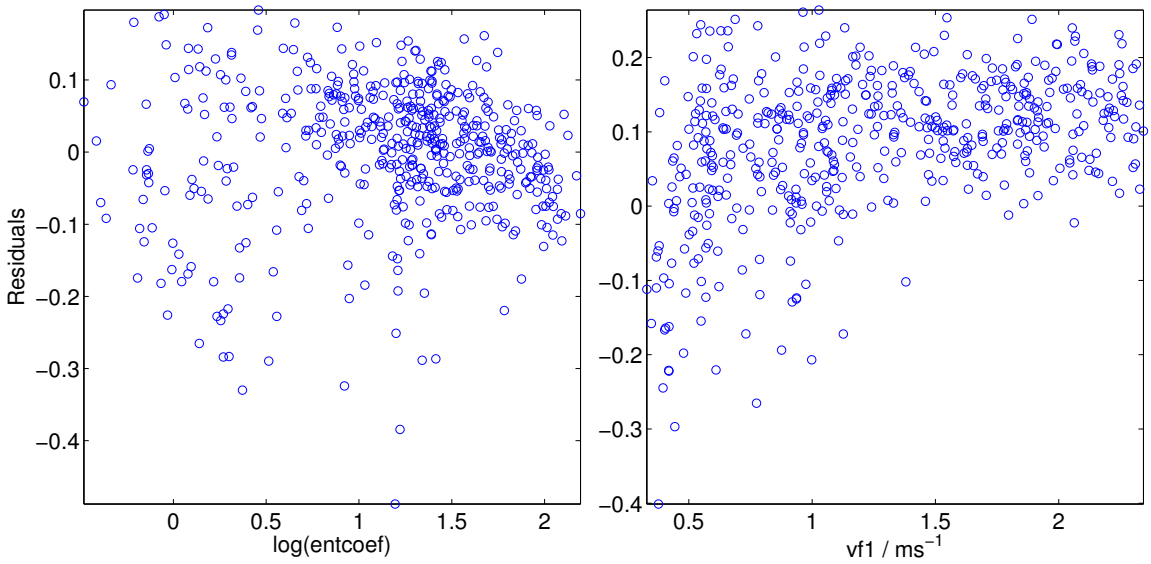

Figure 6.5: Plots showing the residuals after the first principal component scores are regressed onto patterns. The plot on the left shows the residuals after the scores are regressed onto log(*entcoef*), as a function of log(*entcoef*). The right hand plot shows the scores after they are multiply regressed onto *vf1* and log(*entcoef*) and the component explained by the latter is subtracted, as a function of *vf1*.

scores onto each parameter in turn, alongside log(*entcoef*). This multiple regression allows the robustness of any correlation to be determined independently of a linear relationship with log(*entcoef*). We only consider parameters where the regression coefficient is robust with 95% confidence, and which have a correlation magnitude greater than 0.2 with the first principal component score. This criterion indicates that only the ice fall speed (*vf1*) has a robust and independent linear relationship with the location biases in the monsoon.

Figure 6.5 shows the residuals from log(*entcoef*) when regressing the scores onto either log(*entcoef*) alone, or both log(*entcoef*) and *vf1* simultaneously, as a function of log(*entcoef*) and *vf1*, respectively. There is no clear pattern in the left hand plot, indicating that a linear dependence of the scores on log(*entcoef*) explains the sizeable majority of the relationship between first principal component score and this parameter. Note that - by construction - the correlation in this plot is zero. There remains some dependency of the scores on *vf1*, however, indicating that this parameter is

linked to monsoonal location beyond its linear dependency on log(*entcoef*). In spite of *ct*'s apparent relationship with principal component scores, this test indicates that this is a result of its dependence on *entcoef*.

The importance of the ice fall speed in dictating the position of the West African Monsoon is to be expected, as it has also been shown to be of fundamental importance to the simulated climatology of HadCM3 (*Sanderson et al.*, 2008a). Increases to this parameter, and hence the speed at which ice particles fall within both ice and mixed-phase clouds, create larger particles and hence increased precipitation (*Gregory and Morris*, 1996). Indeed, *Wu* (2002) shows using a cloud resolving coupled cloud-ocean model that increasing the ice fall speed strengthens convection and precipitation.

Despite the importance of *vf1*, however, we note that *entcoef* has by far the strongest relationship with the first principal component scores. Analysis of the first principal component pattern (Figure 6.3a), combined with inspection of versions of HadCM3 with extremal values of this parameter (not shown), demonstrates that the greater the value of *entcoef*, the further North the GCM tends to simulate precipitation, as well as simulating it as a latitudinally broader feature. As *entcoef* sets the rate of mixing of a convective plume with the surrounding air, the latitudinal broadening of the precipitation band is unsurprising. It is not apparent, however, why the entrainment rate would set the position of the convergence zone.

We first investigate whether changes in the water vapour distribution of the troposphere could be driving a 'local greenhouse' effect (*Stephens and Greenwald*, 1991), that could in turn drive variation in the location of the monsoon between different variants of HadCM3, by altering land-sea temperature contrast. While full 3D humidity fields are not available for this ensemble, the 'local greenhouse' effect can be estimated by calculating the ratio of clear-sky top of atmosphere longwave radiation to the irradiance of the surface (calculated using the Stefan-Boltzmann law). We find, however, that there is no discernible relationship between *entcoef* and this quantity's

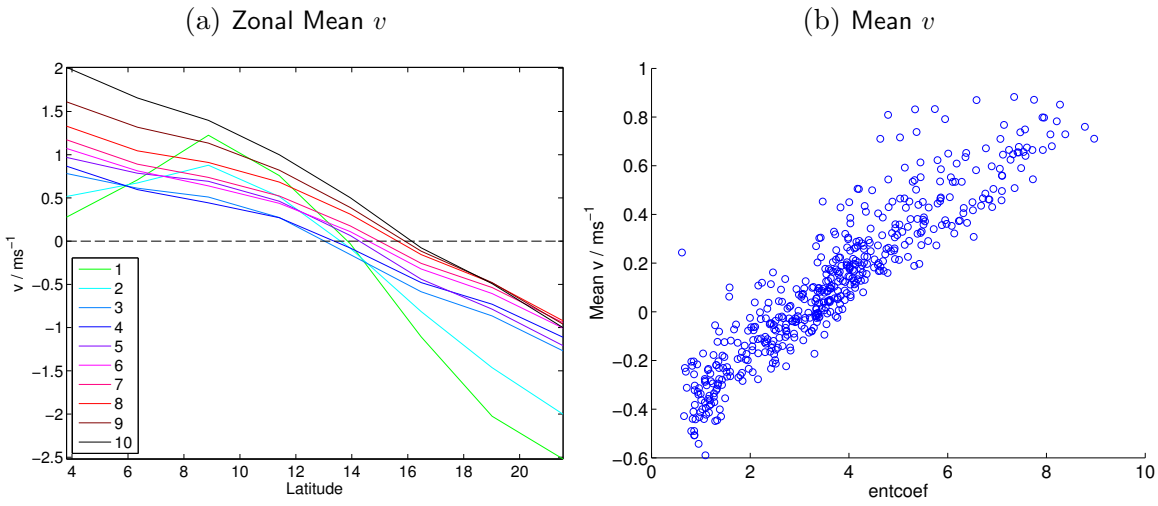

Figure 6.6: Figures showing how the 850hPa meridional winds (v) vary with *entcoef* across our ensemble. Figure 6.6a shows zonal means across the West African Monsoon region, for sub-ensembles ordered by *entcoef*, such that the first group ('1' in the legend) contains the 50 model variants with the lowest values of *entcoef*. Figure 6.6b shows how mean v across the region varies with *entcoef*. The correlation between mean v and *entcoef* is found to be 0.92, with a p-value $\ll 0.05$.

distribution over the West African Monsoon region. It is therefore unlikely that total column moisture - for which the 'local greenhouse' effect is a proxy - is responsible for the West African Monsoon's latitudinal dependence on the rate of entrainment.

In order to investigate the nature of the latitudinal convergence, we next investigate how meridional winds (v) vary with *entcoef* (Figure 6.6). Figure 6.6a shows that the meridional component of the winds decreases with latitude, consistent with convergence of moisture. Further, the latitude at which the mean meridional wind speed is zero increases with the value of *entcoef*. While this may be caused by the varying location of the ITCZ, not the other way round, we note the consistently higher values of meridional wind speed with *entcoef* at almost all latitudes, except for the two groups of model variants with the lowest values of *entcoef*.

We also look at how the value of the meridional wind speed, averaged across the entire West African Monsoon region, varies with *entcoef* (Figure 6.6b). Here, the dependence of meridional wind speed on *entcoef* is clearer still, with a correlation

of 0.92 - higher than the correlation between the first principal component score and *entcoef* (or its logarithm). It is possible, therefore, that increases in *entcoef* strengthen the Northward winds, thus shifting the West African Monsoon further North.

HadCM3 makes use of the Hadley Centre Atmospheric Model version 3 (HadAM3 - *Pope et al.*, 2000), which simulates the effect of convection on momentum through the convective momentum transport (CMT) scheme developed by *Gregory et al.* (1997). CMT simulates the vertical transport of horizontal momentum, and the inclusion of this scheme is associated with changes in both zonal and meridional wind profiles (*Gregory et al.*, 1997; *Pope et al.*, 2000). In particular, CMT increases the strength of near surface winds, as it allows strong winds at the top of the boundary layer to be more easily mixed with those near the surface (*Pope et al.*, 2000, Section 6.1.2). By increasing the rate at which convective plumes mix with their environment, higher values of *entcoef* may enhance this effect, by serving to strengthen low level northward winds, and thus shifting the ITCZ further North.

Noting that location biases depend on the two parameters *entcoef* and *vf1*, it is of interest to compare the values for these parameters for our 'best' three versions of the GCM (Figure 6.2) to the default parameter values for HadCM3 (taken from *Murphy et al.*, 2004). For *vf1*, the values for the three versions of HadCM3 are 0.57, 1.1 and 1.2 ms^{-1}, with a mean of 0.96 ms^{-1}. While this parameter therefore shows variation by a factor of more than two between these three versions of the GCM, its mean value is close to the default: 1 ms^{-1}.

In contrast, all three GCM variants have approximately the same value of *entcoef*: 1.8. It is unsurprising that the value of entrainment is similar for these three versions of the model, as entrainment plays such a crucial role in the location biases of the monsoon. However, it is notable that the default value of *entcoef* in HadCM3 is 3; 67% higher than the value that we find to be optimal for the simulation of the West

African Monsoon.

6.3.2 Role of Sea Surface Temperatures

We next investigate the role of SSTs in the Gulf of Guinea in location biases of the West African Monsoon. As the location biases are characterised purely using the precipitation field, any relationships between warp vectors and other fields will imply a link between precipitation processes and other physical quantities. We evaluate the SSTs as a mean over 10°S-10°N, 15°W-10°E (taken from *Joly and Voldoire*, 2010), excluding land grid points.

The proposed mechanisms linking SSTs in the Gulf of Guinea to the location of the West African Monsoon stress the temperature of the 'cold tongue' in the Gulf relative to the relatively warm surrounding SSTs (*Caniaux et al.*, 2011), and the land-sea temperature contrast (*Joly and Voldoire*, 2010; *Roehrig et al.*, 2013). We therefore investigate how the mean latitudinal displacement vectors depend on SSTs - first in terms of the SST anomaly to the tropics, and then in terms of the temperature gradient between the Sahara and the Gulf.

We evaluate mean latitudinal displacement across the broader West African Monsoon region (20S°-40°N, 30°W-30°E), described in Section 6.2.2. We use the same expanded region used to select our warps, as the first principal component of the warp (Figure 6.3a) demonstrates that there is variation both in terms of latitudinal location and width of the monsoon. Using a smaller domain could allow these corrections to become conflated. For example, if the monsoon is near the Northern edge of the domain and requires compression, averaging latitudinal vectors over the domain will imply the monsoon is too far South, even if no overall latitudinal displacement is required.

We find that the mean latitudinal displacement has a correlation of -0.72 with the Gulf of Guinea SST anomaly relative to the mean SST of the Tropics (Figure 6.7a),

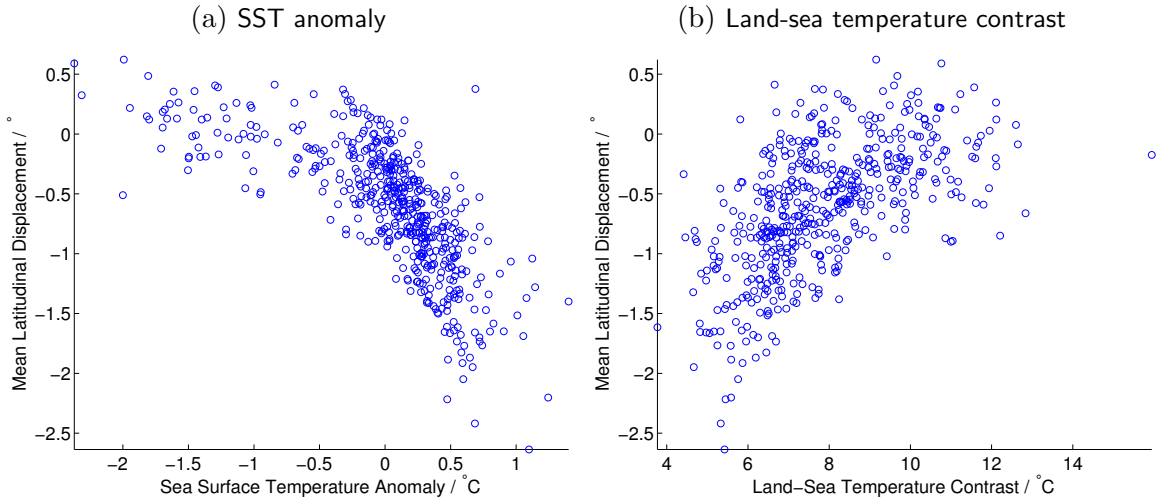

Figure 6.7: Plots showing how the mean latitudinal displacement across the expanded Sahel region depends on SSTs in the Gulf of Guinea. Figure 6.7a shows the relationship between the displacement and SST anomaly of the Gulf of Guinea relative to the tropics, which have a correlation of -0.72. Figure 6.7b shows the relationship between latitudinal displacement and land-sea temperature contrast between the Sahara and Gulf of Guinea regions. Correlation in this case is found to be 0.55.

demonstrating the importance of the 'cold tongue' in correctly simulating the location of the monsoon. Similarly, correlation between latitudinal displacement and the temperature contrast between the Sahara and the Gulf of Guinea is found to be 0.55 (Figure 6.7b), confirming that land-sea temperature gradients contribute to location biases in the monsoon. Interestingly, the correlation between the mean latitudinal displacement and the raw Gulf of Guinea SSTs is 0.12, which is not only substantially smaller than the correlations for relative SSTs, but also implies higher SSTs shift the monsoon further North. This demonstrates that it is the *relative* temperature of the Gulf of Guinea that contributes to the shift of the ITCZ onto land, as one would expect.

6.4 Conclusion

In this Chapter, we have shown that image transformation techniques are not only powerful tools for the removal of location biases from GCMs (as demonstrated in Chapters 2 and 5), but that they also have great potential in characterising, and hence understanding, such biases. By applying the technique developed in Chapter 3 to a large PPE, we are able to robustly investigate the physical origin of location biases in one of the most extensively studied precipitation phenomena - the West African Monsoon.

Several studies have linked the SSTs in the Gulf of Guinea to the location of the monsoon - both in terms of model biases (*Roehrig et al.*, 2013) and variability (*Caniaux et al.*, 2011). We are able to confirm that there is a clear relationship between SSTs and the latitudinal location of the monsoon, but only when we consider SST anomalies, or land-sea temperature contrast. This emphasises that it is not the 'raw' SSTs that influence the location of the monsoon, but rather the relative temperature of the SSTs in the Gulf of Guinea relative to the Tropics as a whole (*Caniaux et al.*, 2011), or the Sahara (*Roehrig et al.*, 2013). This may contribute to the large range of projections for West African Monsoonal rainfall (*Roehrig et al.*, 2013) as - despite GCMs generally experiencing increased SSTs - biases in the relative temperature of the Gulf of Guinea will be likely to persist.

Besides the potential for studying many model variants, one of the greatest benefits of working with a PPE is that it allows variation in GCM behaviour to be robustly linked to the physical parameterisations of the GCM. While there have been extensive studies to investigate variation in large scale properties of the climate, such as climate sensitivity (*Sanderson et al.*, 2008b) and the Atlantic meridional overturning circulation (*Yamazaki et al.*, 2013), our warping technique allows us the unique opportunity to investigate the effect of parameter variations on precipitation location biases.

We are able to demonstrate that the principal component of variation in mon-

soonal location consists primarily of latitudinal displacement and extent, accounting for 64% of variation. We determine that the parameter that sets the rate of entrainment, *entcoef*, correlates very strongly with the model scores of the first principal component. As entrainment sets the rate at which a convecting plume mixes with the surrounding air, it is intuitive that a large value of *entcoef* may broaden the latitudinal width of the monsoon. However, it is less apparent how this parameter could set the seasonal mean latitude of the convergence zone.

We investigate the effect of entrainment rate on meridional winds, and note that *entcoef* has a greater correlation with meridional wind strength than it does with the location of the monsoon. We propose that entrainment may serve to strengthen winds, which could be possible through the convective momentum transport implementation included in HadAM3 (*Gregory et al.*, 1997; *Pope et al.*, 2000).

This hypothesis would suggest that the link between *entcoef* and the location of the convergence band is a local effect, rather than a result of changes in the global circulation. This hypothesis could be tested in future work, with the use of a regional climate model (RCM). In such an experiment, the boundary conditions for the RCM would be provided by a GCM with a fixed value of *entcoef*. The value of *entcoef* in the RCM could then be perturbed, and the effect on the wind fields and precipitation distribution could be investigated. Through such a study, the local effects of *entcoef* on the region could be separated from any global circulation changes resulting from perturbing the parameter. Further, it could be of interest to investigate variation in the location of the monsoon in an ensemble with prescribed SSTs, as this could further clarify the extent to which the relationship with *entcoef* is dependent on convective momentum transport.

With an understanding that *entcoef* is the most important parameter in dictating the location biases of the monsoon, we investigate how it varies across three of the variants of HadCM3 that are best at simulating the distribution of monsoonal rainfall

(as evaluated by RMS before and after warping, and RMS warp vector length - see Figure 6.2). Interestingly, these three ensemble members all perform markedly better in terms of precipitation distribution in this region than the default configuration of HadCM3 (*Cook and Vizy*, 2006). We also find that they all have an entrainment coefficient of 1.8; 40% lower than the value in the default configuration (*Yamazaki et al.*, 2013). This demonstrates that there could be some utility in using perturbed parameter ensembles' simulations of historical precipitation climatology to constrain the values of parameters. It must be stressed, however, that improvements in the precipitation realism of the monsoon may be counteracted by degenerations elsewhere in the model, and so any attempt to constrain parameter values must take a more holistic approach.

There are limitations to the physical mechanisms that this PPE allows us to study, however. Notably, meridional soil moisture gradients may affect the precipitation of the West African Monsoon, through the formation of the African easterly jet (*Cook*, 1999). In this study, the only perturbed parameter linked to soil moisture is the number of soil levels from which moisture can be extracted (*r_layers*), limiting the extent to which this physical mechanism can be investigated. We find no relationship between *r_layers* and location biases of the ITCZ, however, and so our study does not provide any evidence to support the importance of soil properties.

Nevertheless, we have demonstrated here the great utility of image registration techniques and perturbed physics ensembles on diagnosing location biases in precipitation. This study could be extended to also investigate the ways in which integrated precipitation conserving transformations characterise location biases in the region, though results from such transformations may be harder to interpret (as discussed in Section 6.2.2). It would also be of interest to investigate the physical origins of location biases in other aspects of precipitation climatologies. In particular, the spatial and seasonal distribution of the West African Monsoon is relatively simple, and

thus the location biases are easily characterised in terms of latitudinal breadth and location. The techniques we introduce in this thesis, however, may have the ability to correct and characterise more complex location biases, and so investigation of precipitation features with more complex biases (such as the double ITCZ - *Li and Xie*, 2013) could be especially of interest.

Chapter 7

Conclusions

7.1 Discussion of Research

Precipitation is one of the most important climate fields in terms of its impact on human lives, and so understanding how it changes under a climatological forcing is crucial to unpicking the true impacts of that forcing. This has proved difficult not only because of the relatively large internal variability and short length scales of the field, but also because of large differences between simulations from different GCMs. Such differences exist not only in simulated changes, but also in GCMs' mean climatologies. As a result, it is possible that differences in GCMs' climatologies are to some extent masking agreement on the physical process of precipitation change (e.g. *Allen and Ingram*, 2002).

In this thesis we have therefore investigated whether removing location biases from GCM simulated precipitation fields could reduce disagreement on changes. Initially, we investigated projected precipitation changes under an idealised forcing scenario (Chapter 2), as this would provide a strong signal-to-noise ratio, driven by a single forcing agent. By making use of an existing medical image registration tool, *fnirt*, we demonstrated that inter-model disagreement was reduced, confirming that uncer-

tainty in projections arises to some extent from climatological differences between models. This work has been published in *Levy et al.* (2013).

Encouraged by this result, we have developed a tool that is tailored to removing location biases from GCM precipitation fields (Chapter 3). Not only does this tool respect the correct geometry of the problem, but it also allows us to conserve either precipitation fluxes, or total integrated precipitation amount, upon warping. This allows locations to be corrected in two qualitatively different ways, both of which are shown to be able to remove smaller scale errors than was possible with the use of *fnirt*. Further, by respecting the boundary conditions and geometry of climatological fields, this technique is able to provide more physically meaningful transformations than could be derived with tools developed for other applications. We demonstrate that the derived transformations are qualitatively similar for similar values of the smoothness parameter, $\bar{\lambda}$, and that conserving total integrated precipitation does not constrain, but rather increases the goodness of fit we are able to achieve. As well as being discussed in this thesis, the framework for this tool and descriptions of its behaviour have been published in *Levy et al.* (2014a).

Using this new technique, we then go about deriving and selecting transformations across an ensemble of 21 CMIP5 GCMs (Chapter 4). We note that transformations for different models from the same institutions are generally similar, and that other similarities are consistent with the findings of previous work on the CMIP5 ensemble (*Knutti et al.*, 2013). We also find that when fluxes are conserved, total integrated precipitation is reduced upon warping for all GCMs. This reinforces the same result found with *fnirt*, implying that GCMs have a propensity to simulate precipitation features over too great an area. We also find that although there is no discernible relationship between the resolution of a GCM's grid and the realism of grid box intensities, the RMS vector length of integrated precipitation conserving warps does correlate robustly with resolution. This implies that while precipitation fluxes are

not necessarily made more realistic by increasing resolution, the realism of the precipitation climatology may be enhanced through reducing the spacing of the grid.

Although inter-model disagreement was reduced when investigating projections using *fnirt*, this study did not allow the opportunity for validations of the results. We therefore set out to investigate the effects of correcting feature location in historical precipitation changes (Chapter 5). We initially investigate the effect of removing location biases on historical precipitation anomalies and trends, finding that the error in both is reduced relative to observations. In order to explore the potential impact of these reductions in error, we perform a simple detection and attribution experiment. We demonstrate that - for the assumptions made - an anthropogenic influence on precipitation is only detected when we remove location biases.

While 'detectability' is dependent both on the exact methodology and the confidence required, this nevertheless demonstrates the potential benefits of correcting location biases in GCMs when studying historical precipitation, especially in cases where a signal was previously undetectable. This work on historical precipitation confirms that the physical realism of GCM simulations of precipitation changes may be obscured by climatological biases. We have published an investigation into the effects of location biases on historical precipitation and detection and attribution in *Levy et al.* (2014b).

We have therefore demonstrated that correcting feature location in GCMs can provide insights into the source of GCM errors in their simulations of precipitation change. However, it is also possible to use image registration tools of this nature to investigate the origin of location errors themselves (Chapter 6). To this end, we study causes of location biases in simulating the West African Monsoon - a convective precipitation system that many GCMs simulate poorly (both in terms of its location, and many other properties - *Roehrig et al.*, 2013).

We therefore warp 500 versions of a coupled ocean-atmosphere model, HadCM3,

which all use the same parameterisation schemes, with differing parameter values. We find that the location of the monsoon depends strongly on SSTs in the Gulf of Guinea, as has been found in other work (e.g *Roehrig et al.*, 2013), though only when considered as anomalies, or in terms of land-sea contrast. Further, we show that the value of the entrainment parameter, *entcoef*, is a powerful indicator of both the latitudinal location and extent of the West African Monsoon. As entrainment serves to mix moist air from a convecting plume with the surrounding environment, it is intuitive that it could serve to broaden the monsoonal feature. However, it is less apparent how the rate of entrainment would effect the location of the monsoon.

We find that the Northward strength of meridional winds correlates strongly with entrainment, and propose that greater entrainment may strengthen low level convective momentum transport. This could serve to increase low level winds by allowing greater mixing across the boundary layer (as found in *Pope et al.*, 2000), thus shifting the convergence zone further North.

We also note that there exist members in the ensemble that perform markedly better than the default configuration of HadCM3 in their simulation of the location of the monsoon, though we find that their value for *entcoef* is 40% lower than the default. This implies that there could be scope to constrain parameter values based - in part - on location biases in their precipitation climatologies, though of course other properties of the climate should also be taken into account.

7.2 Areas for Future Work

The work presented in this thesis could be expanded in a number of ways. Firstly, it would be possible to study both historical and projected precipitation changes in more depth than we have in this work. An investigation of the effect of warping on more realistic forcing scenarios - such as the Representative Concentration Pathways (*van*

Vuuren et al., 2011) - could help constrain possible precipitation responses over the 21st Century. Note that when applying transformations to projected precipitation, care must be taken that the derived warps are not over-fitted to observations, such that they will no longer be applicable due to climatological shifts (see below for one way that the smoothness of warps could be constrained for investigating projected precipitation).

Further insight into the effects of external influence on historical precipitation could be gleaned through investigating other detection and attribution questions, particularly where it has previously been impossible to detect a signal. This work could be expanded to consider the effects of different observational datasets, both for the observed pattern used to detect an external forcing (see Equation 5.1), and for the reference field that is used to derive the GCM transformations. In order to warp to land-based datasets, however, some minor adjustments to the algorithm would be needed, in order to allow it to operate with missing data.

While this thesis has investigated GCM location biases relative to observations, the tool we have developed could also be used to consider other examples of differences in location. For example, by warping a GCM's past climatology onto its future climatology, shifts associated with changes in circulation could be removed. Not only would this allow these shifts to be characterised, but it would also allow intensifications in the hydrological cycle to be disentangled, and so allow for a more direct evaluation of the 'wet get wetter' mechanism (*Scheff and Frierson*, 2012). Further, these transformations would provide length scales of such climatological shifts, which would help to constrain the smoothness of warps designed to remove location biases from projected precipitation fields.

It would also be possible to use the technique to investigate differences between datasets. For example, different observational datasets could be warped onto each other, allowing discrepancies in both intensity and location to be better characterised.

A similar approach could also be carried out to investigate more precisely the ways in which different versions of a GCM differ. This would be especially interesting, as our work in this thesis has implied that location biases tend to be very consistent across different versions of the same GCM (Section 4.2).

As discussed in Chapter 6, the tool developed here provides a convenient way of characterising location biases in particular precipitation features. Although here we have studied the West African Monsoon, location biases in more complex precipitation features could be investigated. By adapting the technique so that it could study regional climate models (i.e. so that it could warp based on a restricted domain), comparisons between models with parameterised and explicit convection simulation could be made. This could allow the physical origins of location biases in GCMs to be explored more directly than has previously been possible. Further, applying warps to an ensemble could help to provide additional metrics for assessing the realism of GCMs, or (as in Chapter 6) GCM variants.

Although we have focussed on near-term historical and projected precipitation in this study, due to their relevance to modern civilisation, it could also be possible to study precipitation in palaeoclimate. Any proxies for precipitation observations for these time periods will of course be too sparse to derive location bias corrections, though careful application of transformations derived for recent historical climatologies could still provide precipitation fields with improved realism (in a similar way to applying such transformations to projected precipitation).

Finally, we hope that the techniques presented in this thesis could be used for the study of fields beyond precipitation. Of special interest would be other hydrological fields, due to their relationship with precipitation. For example, sea surface salinities can serve as a proxy for evaporation minus precipitation, as discussed in Section 1.2.2 (*Durack et al.*, 2012). Further, a simple modification of the cost function of the technique outlined in Chapter 3 would allow transformations to be derived based

on several variables simultaneously. This could not only allow the simultaneous corrections of these variables, but also increase the robustness of any transformations derived. Expanding our technique so that it can transform fields vertically (as well as longitudinally, latitudinally and seasonally) could expand the number of fields that it could be applied to, allowing investigation of water vapour distributions, for example.

In this thesis, we have demonstrated the potential of location correction techniques, based upon approaches used in medical image registration. It is our hope that such techniques continue to be improved upon, and applied to new problems. In so doing, we hope that location corrections can continue to refine our understanding of the climate, and help to inform decisions on how best to mitigate and adapt to global warming.

Bibliography

Aberg, S., F. Lindgren, A. Malmberg, J. Holst, and U. Holst (2005), An image warping approach to spatio-temporal modelling, *Environmetrics*, *16*(8), 833–848, doi:10.1002/env.741.

Adler, R. F., G. Gu, J.-J. Wang, G. J. Huffman, S. Curtis, and D. Bolvin (2008), Relationships between global precipitation and surface temperature on interannual and longer timescales (1979–2006), *Journal of Geophysical Research: Atmospheres*, *113*(D22), doi:10.1029/2008JD010536.

Alexander, G. D., J. A. Weinman, V. M. Karyampudi, W. S. Olson, and A. C. L. Lee (1999), The effect of assimilating rain rates derived from satellites and lightning on forecasts of the 1993 superstorm, *Monthly Weather Review*, *127*(7), 1433–1457, doi:10.1175/1520-0493(1999)127<1433:TEOARR>2.0.CO;2.

Allan, R. P., and B. J. Soden (2007), Large discrepancy between observed and simulated precipitation trends in the ascending and descending branches of the tropical circulation, *Geophysical Research Letters*, *34*(18), doi:10.1029/2007GL031460, l18705.

Allen, M., and P. Stott (2003), Estimating signal amplitudes in optimal fingerprinting, part I: theory, *Climate Dynamics*, *21*(5-6), 477–491, doi:10.1007/s00382-003-0313-9.

Allen, M. R., and W. J. Ingram (2002), Constraints on future changes in climate and the hydrologic cycle, *Nature*, *419*(6903), 224–232.

Allen, M. R., and D. A. Stainforth (2002), Towards objective probabalistic climate forecasting, *Nature*, *419*(6903), 228–228.

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

Andersson, J., M. Jenkinson, and S. Smith (2010), Non-linear registration, aka spatial normalisation. FMRIB technical report, available from http://fsl.fmrib.ox.ac.uk/analysis/techrep/tr07ja1/.

Andrews, T., P. M. Forster, and J. M. Gregory (2009), A surface energy perspective on climate change, *Journal of Climate*, *22*(10), 2557–2570, doi:10.1175/2008JCLI2759.1.

Andrews, T., P. M. Forster, O. Boucher, N. Bellouin, and A. Jones (2010), Precipitation, radiative forcing and global temperature change, *Geophysical Research Letters*, *37*(14), doi:10.1029/2010GL043991.

Ashburner, J. (2007), A fast diffeomorphic image registration algorithm, *NeuroImage*, *38*(1), 95–113.

Ashburner, J., and K. J. Friston (1999), Nonlinear spatial normalization using basis functions, *Human Brain Mapping*, *7*, 254–266.

Ashfaq, M., L. C. Bowling, K. Cherkauer, J. S. Pal, and N. S. Diffenbaugh (2010), Influence of climate model biases and daily-scale temperature and precipitation events on hydrological impacts assessment: A case study of the United States, *Journal of Geophysical Research: Atmospheres*, *115*(D14), D14,116, doi:10.1029/2009JD012965.

Atanackovic, T., and A. Guran (2000), *Theory of elasticity for scientists and engineers*, Birkhäuser.

Bajcsy, R., and S. Kovačič (1989), Multiresolution elastic matching, *Computer Vision, Graphics, and Image Processing*, *46*(1), 1–21.

Beg, M. F., M. I. Miller, A. Trouvé, and L. Younes (2005), Computing large deformation metric mappings via geodesic flows of diffeomorphisms, *International Journal of Computer Vision*, *61*(2), 139–157, doi:10.1023/B:VISI.0000043755.93987.aa.

Boer, G., and B. Yu (2003), Climate sensitivity and response, *Climate Dynamics*, *20*(4), 415–429, doi:10.1007/s00382-002-0283-3.

Briggs, W. M., and R. A. Levine (1997), Wavelets and field forecast verification, *Monthly Weather Review*, *125*(6), 1329–1341, doi:10.1175/1520-0493(1997)125<1329:WAFFV>2.0.CO;2.

Bro-Nielsen, M., and C. Gramkow (1996), Fast fluid registration of medical images, in *Visualization in Biomedical Computing*, pp. 265–276, Springer.

Brown, B. G., E. Gilleland, and E. E. Ebert (2012a), Forecast of spatial fields, in *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, second ed., chap. 6, John Wiley & Sons, Ltd.

Brown, J., A. Moise, and R. Colman (2013), The South Pacific Convergence Zone in CMIP5 simulations of historical and future climate, *Climate Dynamics*, *41*(7-8), 2179–2197, doi:10.1007/s00382-012-1591-x.

Brown, J. R., A. F. Moise, and F. P. Delage (2012b), Changes in the South Pacific Convergence Zone in IPCC AR4 future climate projections, *Climate Dynamics*, *39*(1-2), 1–19, doi:10.1007/s00382-011-1192-0.

Brown, L. G. (1992), A survey of image registration techniques, *ACM Computing Surveys*, *24*(4), 325–376, doi:10.1145/146370.146374.

Caniaux, G., H. Giordani, J.-L. Redelsperger, F. Guichard, E. Key, and M. Wade (2011), Coupling between the Atlantic cold tongue and the West African monsoon in boreal spring and summer, *Journal of Geophysical Research: Oceans*, *116*(C4), doi:10.1029/2010JC006570.

Chadwick, R., I. Boutle, and G. Martin (2012), Spatial patterns of precipitation change in CMIP5: Why the rich do not get richer in the tropics, *Journal of Climate*, *26*(11), 3803–3822, doi:10.1175/JCLI-D-12-00543.1.

Chauvin, F., P. Arbogast, and H. Douville (2013), A new method of bias correction and its impact on multi-model climate projections, http://cclics.rcec.sinica.edu.tw/xms/read_attach.php?id=1339.

Chen, Y., and G. Medioni (1992), Object modelling by registration of multiple range images, *Image and Vision Computing*, *10*(3), 145–155, doi: http://dx.doi.org/10.1016/0262-8856(92)90066-C.

Chou, C., J. D. Neelin, C.-A. Chen, and J.-Y. Tu (2009), Evaluating the "rich-get-richer"mechanism in tropical precipitation change under global warming, *Journal of Climate*, *22*(8), 1982–2005, doi:10.1175/2008JCLI2471.1.

Christensen, G. E., R. D. Rabbitt, and M. I. Miller (1994), 3D brain mapping using a deformable neuroanatomy, *Physics in Medicine and Biology*, *39*(3), 609.

Cook, K. H. (1999), Generation of the African easterly jet and its role in determining west African precipitation, *Journal of Climate*, *12*(5), 1165–1184, doi: 10.1175/1520-0442(1999)012<1165:GOTAEJ>2.0.CO;2.

Cook, K. H., and E. K. Vizy (2006), Coupled model simulations of the West African Monsoon system: Twentieth- and twenty-first-century simulations, *Journal of Climate*, *19*(15), 3681–3703, doi:10.1175/JCLI3814.1.

Covey, C., K. M. AchutaRao, U. Cubasch, P. Jones, S. J. Lambert, M. E. Mann, T. J. Phillips, and K. E. Taylor (2003), An overview of results from the coupled model intercomparison project, *Global and Planetary Change*, *37*(1–2), 103 – 133, doi:http://dx.doi.org/10.1016/S0921-8181(02)00193-5, evaluation, Intercomparison and Application of Global Climate Models.

Crooks, S. A., and L. J. Gray (2005), Characterization of the 11-year solar signal using a multiple regression analysis of the ERA-40 dataset, *Journal of Climate*, *18*(7), 996–1015, doi:10.1175/JCLI-3308.1.

Dai, A. (2006), Precipitation characteristics in eighteen coupled climate models, *Journal of Climate*, *19*(18), 4605–4630, doi:10.1175/JCLI3884.1.

Dai, A. (2012), The influence of the inter-decadal Pacific oscillation on US precipitation during 1923–2010, *Climate Dynamics*, pp. 1–14, doi:10.1007/s00382-012-1446-5.

Dai, A., and T. M. L. Wigley (2000), Global patterns of ENSO-induced precipitation, *Geophysical Research Letters*, *27*(9), 1283–1286, doi:10.1029/1999GL011140.

Davis, C. A., K. W. Manning, R. E. Carbone, S. B. Trier, and J. D. Tuttle (2003), Coherence of warm-season continental rainfall in numerical weather prediction models, *Monthly Weather Review*, *131*(11), 2667–2679, doi:10.1175/1520-0493(2003)131<2667:COWCRI>2.0.CO;2.

Dee, D. P., S. M. Uppala, A. J. Simmons, P. Berrisford, P. Poli, S. Kobayashi, U. Andrae, M. A. Balmaseda, G. Balsamo, P. Bauer, P. Bechtold, A. C. M. Beljaars,

L. van de Berg, J. Bidlot, N. Bormann, C. Delsol, R. Dragani, M. Fuentes, A. J. Geer, L. Haimberger, S. B. Healy, H. Hersbach, E. V. Hólm, L. Isaksen, P. Kållberg, M. Köhler, M. Matricardi, A. P. McNally, B. M. Monge-Sanz, J. J. Morcrette, B. K. Park, C. Peubey, P. de Rosnay, C. Tavolato, J. N. Thépaut, and F. Vitart (2011), The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553–597.

Dix, M., P. Vohralik, D. Bi, H. Rashid, S. Marsland, S. O'Farrell, P. Uotila, A. Hirst, E. Kowalcyzk, A. Sullivan, et al. (2013), The ACCESS coupled model: Documentation of core CMIP5 simulations and initial results, *Australian Meteorological and Oceanographic Journal*, *63*(1), 83–99.

Doyle, J. D., and M. A. Shapiro (2000), A multi-scale simulation of an extreme downslope windstorm over complex topography, *Meteorology and Atmospheric Physics*, *74*, 83–101, doi:10.1007/s007030070027.

Durack, P. J., S. E. Wijffels, and R. J. Matear (2012), Ocean salinities reveal strong global water cycle intensification during 1950 to 2000, *Science*, *336*(6080), 455–458.

Ebert, E. E., and W. A. Gallus (2009), Toward better understanding of the contiguous rain area (CRA) method for spatial forecast verification, *Weather and Forecasting*, *24*(5), 1401–1415, doi:10.1175/2009WAF2222252.1.

Ermert, V., A. Fink, A. Morse, and H. Paeth (2012), The impact of regional climate change on malaria risk due to greenhouse forcing and land-use changes in tropical africa., *Environmental Health Perspectives*, *120*(1), 77–84.

Falco, I. D., A. D. Cioppa, D. Maisto, and E. Tarantino (2008), Differential evolution as a viable tool for satellite image registration, *Applied Soft Computing*, *8*(4), 1453 – 1462, doi:http://dx.doi.org/10.1016/j.asoc.2007.10.013.

Feldmar, J., and N. Ayache (1994), Locally affine registration of free-form surfaces, in *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pp. 496–501, doi: 10.1109/CVPR.1994.323872.

Filippini, N., B. J. MacIntosh, M. G. Hough, G. M. Goodwin, G. B. Frisoni, S. M. Smith, P. M. Matthews, C. F. Beckmann, and C. E. Mackay (2009), Distinct patterns of brain activity in young carriers of the APOE-ϵ4 allele, *Proceedings of the National Academy of Sciences*, *106*(17), 7209–7214, doi:10.1073/pnas.0811879106.

Fischer, E. M., J. Sedláček, E. Hawkins, and R. Knutti (2014), Models agree on forced response pattern of precipitation and temperature extremes, *Geophysical Research Letters*, doi:10.1002/2014GL062018.

Fontaine, B., P. Roucou, and S. Trzaska (2003), Atmospheric water cycle and moisture fluxes in the West African monsoon: mean annual cycles and relationship using NCEP/NCAR reanalysis, *Geophysical Research Letters*, *30*(3), doi: 10.1029/2002GL015834.

Forster, P. M., T. Andrews, P. Good, J. M. Gregory, L. S. Jackson, and M. Zelinka (2013), Evaluating adjusted forcing and model spread for historical and future scenarios in the CMIP5 generation of climate models, *Journal of Geophysical Research: Atmospheres*, *118*(3), 1139–1150, doi:10.1002/jgrd.50174.

Fuchs, T., U. Schneider, and B. Rudolf (2007), Global precipitation analysis products of the GPCC, *Global Precipitation Climatology Centre (GPCC). Deutscher Wetterdienst, Offenbach.*

García-Serrano, J., F. J. Doblas-Reyes, R. J. Haarsma, and I. Polo (2013), Decadal prediction of the dominant West African monsoon rainfall modes, *Journal of Geophysical Research: Atmospheres*, *118*(11), 5260–5279, doi:10.1002/jgrd.50465.

Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert (2009), Intercomparison of spatial forecast verification methods, *Weather and Forecasting*, *24*(5), 1416–1430, doi:10.1175/2009WAF2222269.1.

Gilleland, E., D. A. Ahijevych, B. G. Brown, and E. E. Ebert (2010a), Verifying forecasts spatially, *Bulletin of the American Meteorological Society*, *91*(10), 1365–1373, doi:10.1175/2010BAMS2819.1.

Gilleland, E., J. Lindström, and F. Lindgren (2010b), Analyzing the Image Warp Forecast Verification Method on Precipitation Fields from the ICP, *Weather and Forecasting*, *25*(4), 1249–1262, doi:10.1175/2010WAF2222365.1.

Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments, *Climate Dynamics*, *16*(2-3), 147–168, doi:10.1007/s003820050010.

Gregory, D., and D. Morris (1996), The sensitivity of climate simulations to the specification of mixed phase clouds, *Climate Dynamics*, *12*(9), 641–651, doi:10.1007/BF00216271.

Gregory, D., and P. R. Rowntree (1990), A mass flux convection scheme with representation of cloud ensemble characteristics and stability-dependent closure, *Monthly Weather Review*, *118*(7), 1483–1506, doi:10.1175/1520-0493(1990)118<1483:AMFCSW>2.0.CO;2.

Gregory, D., R. Kershaw, and P. M. Inness (1997), Parametrization of momentum transport by convection. II: Tests in single-column and general circulation models, *Quarterly Journal of the Royal Meteorological Society*, *123*(541), 1153–1183, doi:10.1002/qj.49712354103.

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, *Nature Geoscience*, *7*(10), 716–721.

Haerter, J. O., S. Hagemann, C. Moseley, and C. Piani (2011), Climate model bias correction and the role of timescales, *Hydrology and Earth System Sciences*, *15*(3), 1065–1079, doi:10.5194/hess-15-1065-2011.

Hajnal, J. V., and D. L. Hill (2010), *Medical image registration*, CRC press.

Hargreaves, J. C., and J. D. Annan (2014), Can we trust climate models?, *Wiley Interdisciplinary Reviews: Climate Change*, *5*(4), 435–440, doi:10.1002/wcc.288.

Hartmann, D., A. K. Tank, M. Rusticucci, L. Alexander, S. Brönnimann, Y. Charabi, F. Dentener, E. Dlugokencky, D. Easterling, A. Kaplan, B. Soden, P. Thorne, M. Wild, and P. Zhai (2013), Observations: Atmosphere and surface, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Cambridge University Press.

Hasselmann, K. (1979), On the signal-to-noise problem in atmospheric response studies, *Meteorology of tropical oceans*, pp. 251–259.

Hawkins, E., M. Joshi, and D. Frame (2014), Wetter then drier in some tropical areas, *Nature Climate Change*, *4*(8), 646–647.

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

Hegerl, G. C., E. Black, R. P. Allan, W. J. Ingram, D. Polson, K. E. Trenberth, R. S. Chadwick, P. A. Arkin, B. B. Sarojini, A. Becker, A. Dai, P. J. Durack, D. Easterling, H. J. Fowler, E. J. Kendon, G. J. Huffman, C. Liu, R. Marsh, M. New, T. J. Osborn, N. Skliris, P. A. Stott, P.-L. Vidale, S. E. Wijffels, L. J. Wilcox, K. M. Willett, and X. Zhang (2014), Challenges in quantifying changes in the global water cycle, *Bulletin of the American Meteorological Society*, doi: 10.1175/BAMS-D-13-00212.1.

Held, I. M., and B. J. Soden (2000), Water vapor feedback and global warming, *Annual Review of Energy and the Environment*, *25*(1), 441–475, doi: 10.1146/annurev.energy.25.1.441.

Held, I. M., and B. J. Soden (2006), Robust responses of the hydrological cycle to global warming, *Journal of Climate*, *19*(21), 5686–5699, doi:10.1175/JCLI3990.1.

Held, I. M., T. L. Delworth, J. Lu, K. L. Findell, and T. R. Knutson (2005), Simulation of Sahel drought in the 20th and 21st centuries, *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 17,891–17,896, doi: 10.1073/pnas.0509057102.

Hirota, N., Y. N. Takayabu, M. Watanabe, M. Kimoto, and M. Chikira (2014), Role of convective entrainment in spatial distributions of and temporal variations in precipitation over tropical oceans, *Journal of Climate*, *27*(23), 8707–8723, doi: 10.1175/JCLI-D-13-00701.1.

Hoffman, R. N., and C. Grassotti (1996), A Technique for Assimilating SSM/I Observations of Marine Atmospheric Storms: Tests with ECMWF Analyses, *Journal of Applied Meteorology*, *35*(8), 1177–1188, doi:10.1175/1520-0450(1996)035<1177:ATFASO>2.0.CO;2.

Horn, B. K., and B. G. Schunck (1981), Determining optical flow, *Proc. SPIE*, *0281*, 319–331, doi:10.1117/12.965761.

Houghton, J., Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, and K. Maskell (2001), *Climate Change 2001: The Scientific Basis*, Cambridge University Press.

Hourdin, F., I. Musat, J.-Y. Grandpeix, J. Polcher, F. Guichard, F. Favot, P. Marquet, A. Boone, J.-P. Lafore, J.-L. Redelsperger, P. M. Ruti, A. Dell'aquila, M.-A. Filiberti, M. Pham, T. L. Doval, A. K. Traore, and H. Gallée (2010), AMMA-Model Intercomparison Project, *Bulletin of the American Meteorological Society*, *91*(1), 95–104, doi:10.1175/2009BAMS2791.1.

Huffman, G. J., R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, and U. Schneider (1997), The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset., *Bulletin of the American Meteorological Society*, *78*, 5–20, doi:http://dx.doi.org/10.1175/1520-0477(1997)078<0005:TGPCPG>2.0.CO;2.

Huffman, G. J., R. F. Adler, D. T. Bolvin, and G. Gu (2009), Improving the global precipitation record: GPCP Version 2.1, *Geophysical Research Letters*, *36*(17), doi:10.1029/2009GL040000.

Huntingford, C., P. A. Stott, M. R. Allen, and F. H. Lambert (2006), Incorporating model uncertainty into attribution of observed temperature change, *Geophysical Research Letters*, *33*(5), doi:10.1029/2005GL024831.

Irvine, P. J., L. J. Gregoire, D. J. Lunt, and P. J. Valdes (2013), An efficient method to generate a perturbed parameter ensemble of a fully coupled AOGCM without flux-adjustment, *Geoscientific Model Development*, *6*(5), 1447–1462, doi:10.5194/gmd-6-1447-2013.

Jenkinson, M., and S. Smith (2001), A global optimisation method for robust affine registration of brain images, *Medical Image Analysis*, *5*(2), 143 – 156, doi: http://dx.doi.org/10.1016/S1361-8415(01)00036-6.

Jiang, J. H., H. Su, C. Zhai, V. S. Perun, A. Del Genio, L. S. Nazarenko, L. J. Donner, L. Horowitz, C. Seman, J. Cole, A. Gettelman, M. A. Ringer, L. Rotstayn, S. Jeffrey, T. Wu, F. Brient, J.-L. Dufresne, H. Kawai, T. Koshiro, M. Watanabe, T. S. LÉcuyer, E. M. Volodin, T. Iversen, H. Drange, M. D. S. Mesquita, W. G. Read, J. W. Waters, B. Tian, J. Teixeira, and G. L. Stephens (2012), Evaluation of cloud and water vapor simulations in CMIP5 climate models using NASA "A-Train" satellite observations, *Journal of Geophysical Research: Atmospheres*, *117*(D14), doi:10.1029/2011JD017237.

Joly, M., and A. Voldoire (2010), Role of the Gulf of Guinea in the inter-annual variability of the West African monsoon: what do we learn from CMIP3 coupled simulations?, *International Journal of Climatology*, *30*(12), 1843–1856, doi: 10.1002/joc.2026.

Jones, C. D., J. K. Hughes, N. Bellouin, S. C. Hardiman, G. S. Jones, J. Knight, S. Liddicoat, F. M. O'Connor, R. J. Andres, C. Bell, K.-O. Boo, A. Bozzo, N. Butchart, P. Cadule, K. D. Corbin, M. Doutriaux-Boucher, P. Friedlingstein, J. Gornall, L. Gray, P. R. Halloran, G. Hurtt, W. J. Ingram, J.-F. Lamarque, R. M. Law, M. Meinshausen, S. Osprey, E. J. Palin, L. Parsons Chini, T. Raddatz, M. G. Sanderson, A. A. Sellar, A. Schurer, P. Valdes, N. Wood, S. Woodward, M. Yoshioka, and M. Zerroukat (2011), The HadGEM2-ES implementation of CMIP5 centennial simulations, *Geoscientific Model Development*, *4*(3), 543–570, doi:10.5194/gmd-4-543-2011.

Jones, P. W. (1999), First- and second-order conservative remapping schemes for

grids in spherical coordinates, *Monthly Weather Review*, *127*(9), 2204–2210, doi: 10.1175/1520-0493(1999)127<2204:FASOCR>2.0.CO;2.

Joshi, M. M., M. J. Webb, A. C. Maycock, and M. Collins (2010), Stratospheric water vapour and high climate sensitivity in a version of the HadSM3 climate model, *Atmospheric Chemistry and Physics*, *10*(15), 7161–7167, doi:10.5194/acp-10-7161-2010.

Keil, C., and G. C. Craig (2007), A displacement-based error measure applied in a regional ensemble forecasting system, *Monthly Weather Review*, *135*(9), 3248–3259, doi:10.1175/MWR3457.1.

Kendon, E. J., D. P. Rowell, R. G. Jones, and E. Buonomo (2008), Robustness of future changes in local precipitation extremes, *Journal of Climate*, *21*(17), 4280–4297, doi:10.1175/2008JCLI2082.1.

Kendon, E. J., N. M. Roberts, C. A. Senior, and M. J. Roberts (2012), Realism of rainfall in a very high-resolution regional climate model, *Journal of Climate*, *25*(17), 5791–5806, doi:10.1175/JCLI-D-11-00562.1.

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. Schär, R. Sutton, G. J. van Oldenborgh, G. Vecchi, and H. J. Wang (2013), Near-term climate change: Projections and predictability, in *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. Stocker, D. Qin, G.-K. Plattner, M. Tignor, S. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. Midgley, Cambridge University Press.

Knutti, R., and J. Sedlacek (2013), Robustness and uncertainties in the new CMIP5 climate model projections, *Nature Climate Change*, *3*(4), 369–373.

Knutti, R., D. Masson, and A. Gettelman (2013), Climate model genealogy: Generation CMIP5 and how we got there, *Geophysical Research Letters*, *40*, 1194–1199, doi:10.1002/grl.50256.

Lambert, F. H., and N. E. Faull (2007), Tropospheric adjustment: The response of two general circulation models to a change in insolation, *Geophysical Research Letters*, *34*(3), doi:10.1029/2006GL028124.

Lambert, F. H., and M. J. Webb (2008), Dependency of global mean precipitation on surface temperature, *Geophysical Research Letters*, *35*(16), doi:10.1029/2008GL034838.

Lambert, F. H., M. J. Webb, and M. M. Joshi (2011), The relationship between land–ocean surface temperature contrast and radiative forcing, *Journal of Climate*, *24*(13), 3239–3256, doi:10.1175/2011JCLI3893.1.

Lau, K. M., S. S. P. Shen, K.-M. Kim, and H. Wang (2006), A multimodel study of the twentieth-century simulations of Sahel drought from the 1970s to 1990s, *Journal of Geophysical Research: Atmospheres*, *111*(D7), doi:10.1029/2005JD006281.

L'Ecuyer, T. S., and J. H. Jiang (2010), Touring the atmosphere aboard the A-Train, *Physics Today*, *63*(7), 36–41.

Levy, A. A. L., W. Ingram, M. Jenkinson, C. Huntingford, F. Hugo Lambert, and M. Allen (2013), Can correcting feature location in simulated mean climate improve agreement on projected changes?, *Geophysical Research Letters*, doi:10.1029/2012GL053964.

Levy, A. A. L., M. Jenkinson, W. Ingram, and M. Allen (2014a), Correcting precipitation feature location in general circulation models, *Journal of Geophysical Research: Atmospheres*, doi:10.1002/2014JD022357.

Levy, A. A. L., M. Jenkinson, W. Ingram, F. H. Lambert, C. Huntingford, and M. Allen (2014b), Increasing the detectability of external influence on precipitation by correcting feature location in GCMs, *Journal of Geophysical Research: Atmospheres*, *119*(22), 12,466–12,478, doi:10.1002/2014JD022358.

Li, G., and S.-P. Xie (2013), Tropical biases in CMIP5 multi-model ensemble: The excessive equatorial Pacific cold tongue and double ITCZ problems, *Journal of Climate*, doi:10.1175/JCLI-D-13-00337.1.

Lott, F. C., M. Gordon, R. J. Graham, A. A. Scaife, and M. Vellinga (2014), Reliability of African climate prediction and attribution across timescales, *Environmental Research Letters*, *9*(10), 104,017.

Lucas, B. D., and T. Kanade (1981), An iterative image registration technique with an application to stereo vision, in *International Joint Conferences on Artificial Intelligence*, pp. 674–679.

Marsham, J. H., N. S. Dixon, L. Garcia-Carreras, G. M. S. Lister, D. J. Parker, P. Knippertz, and C. E. Birch (2013), The role of moist convection in the West African monsoon system: Insights from continental-scale convection-permitting simulations, *Geophysical Research Letters*, *40*(9), 1843–1849, doi:10.1002/grl.50347.

Marvel, K., and C. Bonfils (2013), Identifying external influences on global precipitation, *Proceedings of the National Academy of Sciences*, doi:10.1073/pnas.1314382110.

Marzban, C., and S. Sandgathe (2010), Optical flow for verification, *Weather and Forecasting*, *25*(5), 1479–1494, doi:10.1175/2010WAF2222351.1.

Masson, D., and R. Knutti (2011), Climate model genealogy, *Geophysical Research Letters*, *38*(8), doi:10.1029/2011GL046864.

Mathon, V., H. Laurent, and T. Lebel (2002), Mesoscale convective system rainfall in the Sahel, *Journal of Applied Meteorology*, *41*(11), 1081–1092, doi:10.1175/1520-0450(2002)041<1081:MCSRIT>2.0.CO;2.

McKee, T. B., N. J. Doesken, and J. Kleist (1993), The relationship of drought frequency and duration to time scales, in *Proceedings of the 8th Conference on Applied Climatology*, 22, pp. 179–183, American Meteorological Society Boston, MA.

McSweeney, C., and R. Jones (2013), No consensus on consensus: the challenge of finding a universal approach to measuring and mapping ensemble consistency in GCM projections, *Climatic Change*, *119*(3-4), 617–629, doi:10.1007/s10584-013-0781-9.

Meehl, G. A., T. F. Stocker, W. D. Collins, P. Friedlingstein, A. T. Gaye, J. M. Gregory, A. Kitoh, R. Knutti, J. M. Murphy, A. Noda, et al. (2007), Global climate projections, *Climate change*, *283*.

Meinshausen, M., S. Smith, K. Calvin, J. Daniel, M. Kainuma, J.-F. Lamarque, K. Matsumoto, S. Montzka, S. Raper, K. Riahi, A. Thomson, G. Velders, and D. Vuuren (2011), The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, *Climatic Change*, *109*(1-2), 213–241, doi:10.1007/s10584-011-0156-z.

Min, S.-K., X. Zhang, F. W. Zwiers, and G. C. Hegerl (2011), Human contribution to more-intense precipitation extremes, *Nature*, *470*(7334), 378–381.

Moise, A. F., and F. P. Delage (2011), New climate model metrics based on object-orientated pattern matching of rainfall, *Journal of Geophysical Research: Atmospheres*, *116*(D12), doi:10.1029/2010JD015318.

Muller, C. J., and P. A. O'Gorman (2011), An energetic perspective on the regional response of precipitation to climate change, *Nature Climate Change*, *1*(5), 266–271.

Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, *430*(7001), 768–772.

Nehrkorn, T., B. Woods, T. Auligné, and R. N. Hoffman (2013), Application of feature calibration and alignment to high-resolution analysis: Examples using observations sensitive to cloud and water vapor, *Monthly Weather Review*, *142*(2), 686–702, doi: 10.1175/MWR-D-13-00164.1.

O'Gorman, P. A., R. P. Allan, M. P. Byrne, and M. Previdi (2012), Energetic constraints on precipitation under climate change, *Surveys in Geophysics*, *33*(3-4), 585–608.

Okumura, Y., and S.-P. Xie (2004), Interaction of the Atlantic equatorial cold tongue and the African monsoon, *Journal of Climate*, *17*(18), 3589–3602, doi: 10.1175/1520-0442(2004)017<3589:IOTAEC>2.0.CO;2.

Pall, P., M. Allen, and D. Stone (2007), Testing the Clausius–Clapeyron constraint on changes in extreme precipitation under CO2 warming, *Climate Dynamics*, *28*(4), 351–363, doi:10.1007/s00382-006-0180-2.

Palmer, T., P. Düben, and H. McNamara (2014), Stochastic modelling and energy-efficient computing for weather and climate prediction, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *372*(2018), doi:10.1098/rsta.2014.0118.

Papatheodorou, T., and D. Rueckert (2004), Evaluation of automatic 4D face recognition using surface and texture registration, in *Automatic Face and Gesture Recog-*

nition, 2004. Proceedings. Sixth IEEE International Conference on, pp. 321–326, doi:10.1109/AFGR.2004.1301551.

Pennell, C., and T. Reichler (2010), On the effective number of climate models, *Journal of Climate*, *24*(9), 2358–2367, doi:10.1175/2010JCLI3814.1.

Philander, S. G. H., D. Gu, G. Lambert, T. Li, D. Halpern, N.-C. Lau, and R. C. Pacanowski (1996), Why the ITCZ is mostly North of the equator, *Journal of Climate*, *9*(12), 2958–2972, doi:10.1175/1520-0442(1996)009<2958:WTIIMN>2.0.CO;2.

Piani, C., J. Haerter, and E. Coppola (2010), Statistical bias correction for daily precipitation in regional climate models over Europe, *Theoretical and Applied Climatology*, *99*, 187–192, 10.1007/s00704-009-0134-9.

Polson, D., G. C. Hegerl, X. Zhang, and T. J. Osborn (2013a), Causes of robust seasonal land precipitation changes, *Journal of Climate*, doi:10.1175/JCLI-D-12-00474.1.

Polson, D., G. C. Hegerl, R. P. Allan, and B. B. Sarojini (2013b), Have greenhouse gases intensified the contrast between wet and dry regions?, *Geophysical Research Letters*, doi:10.1002/grl.50923.

Polson, D., M. Bollasina, G. C. Hegerl, and L. J. Wilcox (2014), Decreased monsoon precipitation in the Northern Hemisphere due to anthropogenic aerosols, *Geophysical Research Letters*, *41*(16), 6023–6029, doi:10.1002/2014GL060811.

Pope, V. D., M. L. Gallani, P. R. Rowntree, and R. A. Stratton (2000), The impact of new physical parametrizations in the Hadley Centre climate model: HadAM3, *Climate Dynamics*, *16*(2-3), 123–146, doi:10.1007/s003820050009.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery (2007), *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski (2003), Breaking the cloud parameterization deadlock, *Bulletin of the American Meteorological Society*, *84*(11), 1547–1564, doi:10.1175/BAMS-84-11-1547.

Roe, G. H. (2005), Orographic precipitation, *Annual Review of Earth and Planetary Sciences*, *33*(1), 645–671, doi:10.1146/annurev.earth.33.092203.122541.

Roehrig, R., D. Bouniol, F. Guichard, F. Hourdin, and J.-L. Redelsperger (2013), The present and future of the West African Monsoon: A process-oriented assessment of CMIP5 simulations along the AMMA transect, *Journal of Climate*, *26*(17), 6471–6505, doi:10.1175/JCLI-D-12-00505.1.

Rohlfing, T. (2012), Image similarity and tissue overlaps as surrogates for image registration accuracy: Widely used but unreliable, *Medical Imaging, IEEE Transactions on*, *31*(2), 153–163, doi:10.1109/TMI.2011.2163944.

Rueckert, D., L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes (1999), Nonrigid registration using free-form deformations: application to breast MR images., *IEEE transactions on medical imaging*, *18*(8), 712–721.

Sanderson, B., C. Piani, W. Ingram, D. Stone, and M. Allen (2008a), Towards constraining climate sensitivity by linear analysis of feedback patterns in thousands of perturbed-physics GCM simulations, *Climate Dynamics*, *30*(2-3), 175–190, doi:10.1007/s00382-007-0280-7.

Sanderson, B. M., R. Knutti, T. Aina, C. Christensen, N. Faull, D. J. Frame, W. J. Ingram, C. Piani, D. A. Stainforth, D. A. Stone, and M. R. Allen (2008b), Constraints on model response to greenhouse gas forcing and the role of subgrid-scale processes, *Journal of Climate*, *21*(11), 2384–2400, doi:10.1175/2008JCLI1869.1.

Scaife, A., C. Folland, and J. Mitchell (2007), A model approach to climate change, *Physics World*.

Scheff, J., and D. Frierson (2012), Twenty-first-century multimodel subtropical precipitation declines are mostly midlatitude shifts, *Journal of Climate*, *25*(12), 4330–4347, doi:10.1175/JCLI-D-11-00393.1.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, and B. Rudolf (2014), GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle, *Theoretical and Applied Climatology*, *115*(1-2), 15–40, doi:10.1007/s00704-013-0860-x.

Schulzweida, U., and L. Kornblueh (2004), Climate data operators, *Max-Planck-Institute for Meteorology*.

Seager, R., N. Naik, and G. A. Vecchi (2010), Thermodynamic and dynamic mechanisms for large-scale changes in the hydrological cycle in response to global warming, *Journal of Climate*, *23*(17), 4651–4668, doi:10.1175/2010JCLI3655.1.

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, *Environmental Research Letters*, *9*(4), 044,008.

Self, S., M. R. Rampino, J. Zhao, and M. G. Katz (1997), Volcanic aerosol perturbations and strong El Niño events: No general correlation, *Geophysical Research Letters*, *24*(10), 1247–1250, doi:10.1029/97GL01127.

Smith, P. (1981), Bilinear interpolation of digital images, *Ultramicroscopy*, *6*(2), 201 – 204, doi:10.1016/0304-3991(81)90061-9.

Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor, and H. L. Miller (2007), *Climate change 2007: the Physical Science Basis.*

Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Summary for Policymakers., Cambridge University Press, Geneva.

Sotiras, A., C. Davatzikos, and N. Paragios (2013), Deformable medical image registration: A survey, *Medical Imaging, IEEE Transactions on*, *32*(7), 1153–1190, doi:10.1109/TMI.2013.2265603.

Sperber, K., H. Annamalai, I.-S. Kang, A. Kitoh, A. Moise, A. Turner, B. Wang, and T. Zhou (2013), The Asian summer monsoon: an intercomparison of CMIP5 vs. CMIP3 simulations of the late 20th century, *Climate Dynamics*, *41*(9-10), 2711–2744, doi:10.1007/s00382-012-1607-6.

Steiner, M., and J. A. Smith (1998), Convective versus stratiform rainfall: An ice-microphysical and kinematic conceptual model, *Atmospheric Research*, *47–48*(0), 317 – 326, doi:http://dx.doi.org/10.1016/S0169-8095(97)00086-0.

Stephens, G. L., and T. J. Greenwald (1991), The earth's radiation budget and its relation to atmospheric hydrology: 1. observations of the clear sky greenhouse effect, *Journal of Geophysical Research: Atmospheres*, *96*(D8), 15,311–15,324, doi:10.1029/91JD00973.

Stocker, T. F., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (2013), Climate change 2013: The physical science basis, *Intergovernmental Panel on Climate Change Fifth Assessment Report, Working Group I*.

Stott, P. A., S. F. B. Tett, G. S. Jones, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell (2001), Attribution of twentieth century temperature change to natural and anthropogenic causes, *Climate Dynamics*, *17*(1), 1–21, doi:10.1007/PL00007924.

Stratton, R. A., and A. J. Stirling (2012), Improving the diurnal cycle of convection in GCMs, *Quarterly Journal of the Royal Meteorological Society*, *138*(666), 1121–1134, doi:10.1002/qj.991.

Sylla, M. B., A. T. Gaye, G. S. Jenkins, J. S. Pal, and F. Giorgi (2010), Consistency of projected drought over the Sahel with changes in the monsoon circulation and extremes in a regional climate model projections, *Journal of Geophysical Research: Atmospheres*, *115*(D16), doi:10.1029/2009JD012983.

Tao, X., J. Prince, and C. Davatzikos (2002), Using a statistical shape model to extract sulcal curves on the outer cortex of the human brain, *Medical Imaging, IEEE Transactions on*, *21*(5), 513 –524, doi:10.1109/TMI.2002.1009387.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2011), An Overview of CMIP5 and the Experiment Design, *Bulletin of the American Meteorological Society*, *93*(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tebaldi, C., J. M. Arblaster, and R. Knutti (2011), Mapping model agreement on future climate projections, *Geophysical Research Letters*, *38*(23), L23,701, doi:10.1029/2011GL049863.

Tett, S. F. B., P. A. Stott, M. R. Allen, W. J. Ingram, and J. F. B. Mitchell (1999), Causes of twentieth-century temperature change near the Earth's surface, *Nature*, *399*(6736), 569–572.

Thevenaz, P., U. E. Ruttimann, and M. Unser (1995), Iterative multi-scale registration without landmarks, in *Image Processing, 1995. Proceedings., International Conference on*, vol. 3, pp. 228–231 vol.3, doi:10.1109/ICIP.1995.537622.

Trenberth, K. E., and A. Dai (2007), Effects of Mount Pinatubo volcanic eruption on the hydrological cycle as an analog of geoengineering, *Geophysical Research Letters*, *34*(15), doi:10.1029/2007GL030524.

van Pelt, S. C., P. Kabat, H. W. ter Maat, B. J. J. M. van den Hurk, and A. H. Weerts (2009), Discharge simulations performed with a hydrological model using bias corrected regional climate model input, *Hydrology and Earth System Sciences*, *13*(12), 2387–2397, doi:10.5194/hess-13-2387-2009.

van Vuuren, D., J. Edmonds, M. Kainuma, K. Riahi, A. Thomson, K. Hibbard, G. Hurtt, T. Kram, V. Krey, J.-F. Lamarque, T. Masui, M. Meinshausen, N. Nakicenovic, S. Smith, and S. Rose (2011), The representative concentration pathways: an overview, *Climatic Change*, *109*(1-2), 5–31, doi:10.1007/s10584-011-0148-z.

Volodin, E., N. Dianskii, and A. Gusev (2010), Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, *Izvestiya Atmospheric and Oceanic Physics*, *46*, 414–431, 10.1134/S000143381004002X.

Wang, Z., and F. A. Dahlen (1995), Spherical-spline parameterization of three-dimensional Earth models, *Geophysical Research Letters*, *22*(22), 3099–3102, doi:10.1029/95GL03080.

Westlye, L. T., K. B. Walhovd, A. M. Dale, A. Bjørnerud, P. Due-Tønnessen, A. Engvig, H. Grydeland, C. K. Tamnes, Y. Østby, and A. M. Fjell (2010), Life-Span Changes of the Human Brain White Matter: Diffusion Tensor Imaging (DTI) and Volumetry, *Cerebral Cortex*, *20*(9), 2055–2068, doi:10.1093/cercor/bhp280.

White, R. H., and R. Toumi (2013), The limitations of bias correcting regional climate model inputs, *Geophysical Research Letters*, *40*(12), 2907–2912, doi:10.1002/grl.50612.

Williams, G. E. (1993), History of the Earth's obliquity, *Earth-Science Reviews*, *34*(1), 1 – 45, doi:http://dx.doi.org/10.1016/0012-8252(93)90004-Q.

Wolter, K., and M. S. Timlin (2011), El Niño/Southern Oscillation behaviour since 1871 as diagnosed in an extended multivariate ENSO index (MEI.ext), *International Journal of Climatology*, *31*(7), 1074–1087, doi:10.1002/joc.2336.

Woolrich, M., S. Jbabdi, B. Patenaude, M. Chappell, S. Makni, C. B. T. Behrens, M. Jenkinson, and S. Smith (2009), Bayesian analysis of neuroimaging data in FSL, *NeuroImage*, *45*(1, Supplement 1), S173–S186.

Wu, X. (2002), Effects of ice microphysics on tropical radiative–convective–oceanic quasi-equilibrium states, *Journal of the Atmospheric Sciences*, *59*(11), 1885–1897, doi:10.1175/1520-0469(2002)059<1885:EOIMOT>2.0.CO;2.

Xu, Z., and Z.-L. Yang (2012), An improved dynamical downscaling method with GCM bias corrections and its validation with 30 years of climate simulations, *Journal of Climate*, *25*(18), 6271–6286, doi:10.1175/JCLI-D-12-00005.1.

Yamazaki, K., D. J. Rowlands, T. Aina, A. T. Blaker, A. Bowery, N. Massey, J. Miller, C. Rye, S. F. B. Tett, D. Williamson, Y. H. Yamazaki, and M. R. Allen (2013), Obtaining diverse behaviors in a climate model without the use of flux adjustments, *Journal of Geophysical Research: Atmospheres*, *118*(7), 2781–2793, doi:10.1002/jgrd.50304.

Yates, E., S. Anquetin, V. Ducrocq, J.-D. Creutin, D. Ricard, and K. Chancibault (2006), Point and areal validation of forecast precipitation fields, *Meteorological Applications*, *13*(1), 1–20, doi:10.1017/S1350482705001921.

Yin, J. H. (2005), A consistent poleward shift of the storm tracks in simulations of 21st century climate, *Geophysical Research Letters*, *32*(18), doi:10.1029/2005GL023684.

Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa (2007), Detection of human influence on twentieth-century precipitation trends, *Nature*, *448*(7152), 461–465.