UNIVERSITY OF BIRMINGHAM

Development of a correction approach for future precipitation changes simulated by General Circulation Models

Jonathan Michael Eden

A thesis submitted to the University of Birmingham for the degree of Doctor of Philosophy.

School of Geography, Earth and Environmental Sciences

College of Life and Environmental Sciences

University of Birmingham

March 2011

Abstract

Producing reliable estimates of changes in precipitation at local- and regional-scales remains an important challenge in climate change science. Statistical downscaling methods are often utilised to bridge the gap between the coarse resolution of General Circulation Models (GCMs) and the higher-resolutions at which information is required by the majority of end users. However, the skill of GCM precipitation, particularly in simulating temporal variability, is not fully understood and statistical downscaling typically adopts a 'Perfect-Prog' (short for perfect prognosis) approach in which the derivation of high-resolution precipitation projections is based on real world statistical relationships between large-scale atmospheric 'predictors' and local-scale precipitation.

Here, a 'nudged' simulation of the ECHAM5 GCM is conducted in which the large-scale climatic state is forced towards historical observations of large-scale circulation and temperature for the period 1958-2001. By comparing simulated and observed precipitation it is possible to, for the first time, quantify GCM skill in simulating temporal variability of precipitation. Correlation between simulated and observed monthly mean precipitation is shown to be as strong as 0.8-0.9 in many parts of Europe, North America and Australia.

A nudged simulation permits the development of an alternative approach to statistical downscaling, known as Model Output Statistics (MOS), to correct precipitation as simulated by ECHAM5. It is also shown that MOS correction offers greater skill than Perfect-Prog methods when estimating local-scale monthly mean precipitation. The strongest-performing MOS models are applied to ECHAM5 climate change simulations and are shown to produce high-resolution precipitation projections that support those of RCM simulations. The potential for extending the MOS approach to daily precipitation is also assessed, with recommendations made for further research and application.

Acknowledgements

I wish to thank my main supervisor Martin Widmann for his vision and enthusiasm throughout this project. I am grateful for the initial and continued belief shown in me, and for the advice and encouragement offered over the last three years. I very much hope to continue what has become fine working relationship during the coming years.

I also thank my three co-supervisors for their valued support and advice during the undertaking of this project. I thank Xiaoming Cai, particularly for his faith shown in me following my initial application and his assistance in setting up the ECHAM5 simulation on BlueBEAR. I thank Chris Kidd, particularly for the advice and assistance afforded in the acquisition of precipitation datasets and for including me in the Fortran tutorial classes. I also thank David Grawe for his assistance in solving many, many technical difficulties.

I am indebted to the University of Birmingham's BlueBEAR supercomputer and to the facility's management and support team, including Paul Hatton, Aslam Ghumra and Jonathan Hunt. I am also very grateful to the Linux and Windows support persons at the School of Geography, Earth and Environmental Sciences, namely Amarjit Sathi and Steve Swoffer.

At the Max Planck Institute for Meteorology, Hamburg, Germany, I thank Sebastian Rast for his continued interest and support throughout this project, and particularly for the provision of the ECHAM5 nudging files and the assistance offered in the nudging set-up. I also owe many thanks to Irina Fast at the University of Hamburg, Germany for her assistance in setting up ECHAM5 to run on BlueBEAR.

I thank Tim Osborn (University of East Anglia) and David Hannah (University of Birmingham) for their examination of this thesis and suggested amendments. I am also grateful to Andrew Eden and Matthew Dowell for their review of this work and subsequent valued comments.

Finally, I would also like formally thank the School of Geography, Earth and Environmental Sciences at the University of Birmingham for funding this three-year PhD project.

Glossary

CCA: Canonical Correlation Analysis.

CCA-PCR: Nomenclature for the downscaling method based on PC-MLR, equivalent to one-dimensional CCA.

CMAP: Climate Prediction Center Merged Analysis of Precipitation (Xie and Arkin, 1997).

CORDEX: Coordinated Regional Climate Downscaling Experiment.

E-OBS: European land-only daily precipitation gridded ($0.25° \times 0.25°$) dataset (1950-2006) used in the ENSEMBLES EU-FP6 project (Haylock et al., 2008).

ECHAM5: Atmosphere GCM developed by ECMWF at the Max Planck Institute for Meteorology in Hamburg.

ECMWF: European Centre for Medium-Range Weather Forecasting.

ENSEMBLES: EU project in which ensembles of general circulation models and regional climate models were developed for Europe and North Africa.

ERA-40: ECMWF 40-year reanalysis (Uppala et al., 2005).

GCM: General Circulation Model.

GPCP: Global Precipitation Climatology Project.

GPCC: Global Precipitation Climatology Center.

IPCC: Intergovernmental Panel on Climate Change.

IPCC AR4: IPCC Fourth Assessment Report, published in 2007.

MCA: Maximum Covariance Analysis.

MOS: Model Output Statistics.

MPIOM: Max Planck Institute Ocean Model.

NARCCAP: North American Regional Climate Change Assessment Program.

NCAR: National Center for Atmospheric Research.

NCEP: National Centers for Environmental Prediction.

NWP: Numerical Weather Prediction.

PCA: Principal Component Analysis.

PC-MLR: Principal Component Multiple Linear Regression (alternatively PCR).

PP: Perfect-Prog.

PRUDENCE: Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects.

RCM: Regional Climate Model.

SRES: IPCC Special Report of Emissions Scenarios.

SVD: Singular Value Decomposition.

SVD-RM: Nomenclature for the downscaling method based on orthogonal projection of predictor anomalies onto a regression map; equivalent to SVD of the covariance matrix (in turn equivalent to one-dimensional MCA).

UKCP09: UK Climate Projections (published in 2009).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Modelling precipitation in a changing climate

Advances in modelling and an improved understanding of the climate system have led to the widespread availability of increasingly reliable climate change projections. These suggest that future climatic changes, brought about by increases in the concentration of greenhouse gases, will have a wide range of impacts at varying spatial scales. In addition to a perceived increase in global temperature are changes in global precipitation patterns. Precipitation is the most important input variable for many natural systems and the most recent Intergovernmental Panel on Climate Change (IPCC) findings suggest that precipitation changes are likely to occur in all parts of the world. Such changes exist on a range of time scales, from shifts in long-term monthly or seasonal climatologies to changes in the frequency and intensity of extreme daily precipitation events. The IPCC Fourth Assessment Report (AR4) (IPCC, 2007) provides an in-depth account of the anticipated changes in patterns of precipitation across the globe and, as a result of improved modelling approaches and agreement between simulations over larger regions, there is now a greater certainty of what these changes will be.

General Circulation Models (GCMs) are the most important tool in estimating future climate. The projections of large-scale climate change made by GCMs are considered to be relatively skilful, at least on planetary and continental scales. Whilst circulation variables are computed on the basis of fundamental physical equations, the processes which govern precipitation variability, such as convection and cloud formation, often occur at a finer resolution than is resolved by GCMs. Additionally, precipitation simulation is hindered by the coarse representation of the physiographic features that exert influence on the formation and distribution of precipitation in regions of complex topography. It is therefore challenging to quantify the various sources of error in GCM-simulated precipitation, and in particular how well temporal variability is simulated.

Crucially, it is at the local scale that impacts are felt by people and the environment and consequently such scales hold the most interest to the climate change community. In simulating precipitation, GCMs rely on parameterisation schemes to estimate the smaller-scale processes that are not resolved by the GCM. Dynamical downscaling uses a higher-resolution numerical model which is nested into a GCM over a fixed limited area, in order to represent physical approaches at a greater grid cell resolution. Whilst based on model physics to the greatest possible extent, dynamical downscaling is computationally intensive and usually bound by an upper limit in terms of what can be physically resolved. Statistical downscaling is often favoured by impact studies due to its wide applicability to a range of climates and a general performance comparable with dynamical methodologies. The aim of this approach is to derive empirical relationships between large-scale 'predictor' variables and a local-scale 'predictand'. As the skill of GCM-simulated precipitation, particularly in representing temporal variability, is not fully understood, statistical downscaling typically adopts a so-called 'Perfect-Prog' approach in which a predictor-predictand statistical link is defined by a relationship between large-scale atmospheric variables (typically geopotential height, temperature or humidity) and local-scale precipitation as observed in the historical record. In estimating a future value of precipitation, the relationship is applied to values of the predictor(s) simulated by the GCM at some future time.

1.2 An alternative statistical downscaling approach

Previously, assessment of GCM skill has been limited to a comparison of simulated and observed long-term precipitation means or distributions. As the sequence of day-to-day 'weather' in a freely-evolving GCM is stochastic, extending this assessment to a comparison of simulated and observed time series is not possible. In order to fully assess GCM skill in representing temporal variability, a hindcast GCM simulation is required in which the temporal evolution of the large-scale circulation and temperature fields are forced to match those of the observed record. An 'in-phase' simulation would permit an assessment of GCM skill in simulating precipitation given a realistic large-scale climatic state. Widmann and Bretherton (2000) and Widmann et al. (2003) previously followed this approach to assess reanalysis precipitation (Widmann and Bretherton, 2000; Widmann et al., 2003) but this has not yet been extended to GCMs used for climate change simulations.

It is possible to formulate a statistical correction for GCM-simulated precipitation based on the relationship between local-scale precipitation in the observed record and precipitation as simulated by an 'in-phase' GCM. This is known as Model Output Statistics (MOS) and can be considered as an alternative to the Perfect-Prog statistical downscaling approach. In a conceptual sense, a MOS approach to downscaling precipitation offers a number of benefits over a Perfect-Prog approach. First of all, arguably

the most obvious advantage is the inclusion of the GCM in not only the implementation of a downscaling relationship but also the development of that relationship. Whereas in Perfect-Prog downscaling a simulation of some future climatic state is considered to be a 'perfect' error-free representation, MOS is able to accept the existence of and account for model-inherent error and bias. Secondly, there is little consensus in previous studies as to what constitutes the most powerful set of predictor variables. The parameterisation of GCM precipitation is based on large-scale circulation and temperature and so can be considered to contain the predictive information of all relevant climate variables. In basing a downscaling relationship solely on simulated precipitation, the subjective decision of the choice of predictor is removed from the downscaling process. Finally, in any downscaling study it is necessary to consider the usefulness of the downscaled information produced. Fowler et al. (2007) called for a greater transparency of downscaling methods and explanation of how results should be interpreted. Users of downscaled information vary in expertise and may not be fully aware of a particular method's limitations or the uncertainty of the results (Maraun et al., 2010). The MOS approach uses model physics to the fullest extent and yet the physical nature of statistical correction (linking large- to small-scale precipitation) is more simple to interpret and understand.

1.3 Aim and objectives

The over-arching aim of this work is thus: to develop a MOS downscaling correction approach for estimating regional precipitation changes for the latest generation of climate change simulations. To achieve this aim, a number of research objectives are outlined. Each objective is hereby discussed in terms of its importance, with the methods of analysis used to achieve each objective also summarised.

1. **To conduct a hindcast simulation using the ECHAM5 GCM in which the large-scale circulation and temperature variables are forced to corresponding fields in the ERA-40 reanalysis.** As a GCM is designed to be freely-evolving and unconstrained by the assimilation of observations, an apparent 'mismatch' exists between simulated and observed day-to-day 'weather'. Thus, the extent to which a GCM is able to simulate temporal precipitation variability remains troublesome to quantify. Here, a solution is proposed in the form of a 'nudged' simulation of the ECHAM5 GCM in which the circulation and temperature fields are forced towards corresponding fields in the ECMWF reanalysis (ERA-40) (Uppala et al., 2005). Precipitation is not explicitly nudged and remains an independent parameterisation.

2. **To assess and quantify the skill of ECHAM5 to simulate precipitation given large-scale climatic conditions that are in temporal phase with real world observations.**

In forcing the ECHAM5 prognostic variables into temporal phase with the observed record, it is possible to ask the question, "How well does ECHAM5 simulate precipitation given a realistic large-scale circulation?". Model to real world comparison can be undertaken on long-term time series, and an assessment of the skill of ECHAM5-simulated precipitation can be extended to a spatial quantification of the representation of temporal variability.

3. **To develop and validate several statistical downscaling techniques following a MOS approach, in which local-scale monthly mean precipitation is estimated from precipitation as simulated by ECHAM5.** In regions where skill of simulated precipitation is strong, there may exist potential for simulated precipitation to be subjected to a post-simulation statistical correction. Such a correction is a form of MOS and has been used previously to correct the bias of Regional Climate Models (RCMs). It may also be possible to develop a MOS bias correction to directly downscale raw GCM output, which previously, although conceptually possible, could not consider the skill of GCM-simulated precipitation to represent temporal variability. In addition to a MOS downscaling bias correction, the temporally co-varying nature of the precipitation field of the nudged simulation with the observed record permits the development of further MOS downscaling methodologies based on regression. In total, three MOS downscaling methods (local bias correction and two regression-based techniques) are developed for global application to ECHAM5-simulated precipitation.

4. **To develop and validate Perfect-Prog statistical downscaling techniques that have been used successfully in the literature.** As mentioned, traditional precipitation downscaling methods follow a Perfect-Prog approach and seek to derive a statistical link between simultaneous observations of some large-scale climate variable(s) and local-scale precipitation. Typically, statistical downscaling work is region-specific and focuses on the predictive power of a small number of variables. Here, an alternative systematic approach is adopted in the global development of numerous Perfect-Prog models based on individual and combinations of predictor variables at different atmospheric levels. Whilst such an approach is computationally-intensive and offers limited flexibility to fine-tune the calibration of the most suitable downscaling model during a given season in a particular region, producing a selection of models for global application provides a comprehensive platform on which to base an inter-model comparison.

5. **To make a comparison between the relative merits of MOS and Perfect-Prog approaches to statistical downscaling.** A key component in assessing the benefits of a MOS downscaling approach is a comparison with existing methods used in statistical downscaling. As the spatial applicability of a MOS downscaling correction approach is currently unknown and likely

to be highly variable, a holistic understanding of the relative benefits of MOS necessitates a comparison with Perfect-Prog statistical downscaling models.

6. **To assess the applicability of a MOS downscaling correction in future climates.** There are a number of issues that may impinge upon the effectiveness of any statistical downscaling methodology when applied to climate change applications. Arguably, the most serious of these is the assumption that the empirical predictor-predictand relationship established in the historical observed record is stationary over time. That is, that the relationship remains stable under perturbed climatic conditions. In estimating future climate, GCMs are relied upon to simulate a physically consistent climate change signal and the chosen predictor variable(s) should also capture this signal. Here, MOS downscaling methods are assessed in terms of the ability of each to reproduce precipitation characteristics likely to be associated with a changing climate.

7. **To apply successful MOS corrections to the latest ECHAM5 climate change simulations used in the IPCC AR4 and to make estimates for future changes in local-scale precipitation across selected regions of interest.** In regions where MOS downscaling is shown to be successful, corrections are applied to simulated precipitation in the most recent ECHAM5 climate change simulations. Absolute difference and percentage change in downscaled seasonal precipitation totals during the late twentieth and twenty-first centuries are analysed in detail and compared with raw ECHAM5 projections.

8. **To evaluate the potential for a MOS downscaling correction of GCM-simulated daily precipitation distributions across Europe.** Up until this point, all MOS downscaling corrections within this work have dealt with monthly and seasonal precipitation. Here, a MOS correction for daily precipitation distribution is developed and assessed. Similarly to the corrections developed for monthly precipitation, corrections of daily distributions are both local and non-local in nature. This objective should be considered a 'feasibility study' of the potential for a MOS correction, and as such is limited to European precipitation.

1.4 Structure of thesis

A general review of relevant existing literature is given in Chapter 2. The content of Chapter 2 includes a summary of the key themes in understanding and estimating future precipitation changes, and most importantly introduces the concept of MOS in the context of statistical downscaling. The following five chapters (Chapters 3-7) constitute the empirical component of the thesis, although each contains an introductory section where specific themes are discussed in more detail if necessary. Chapter 3 focuses on the

assessment and quantification of ECHAM5-simulated precipitation and the potential for a development of a downscaling correction for the simulated precipitation field. In Chapter 4, MOS and Perfect-Prog downscaling models are developed and validated in terms of their respective skill in correctly estimating local-scale precipitation. In an extension of the validation process, Chapter 5 assesses the ability of the MOS downscaling models to behave realistically in a perturbed climate and thus their transferability to future simulations. In Chapter 6, successful MOS downscaling models are applied to climate change projections in regions where model skill (as ascertained in Chapters 4 and 5) is strong. Chapter 7 explores the potential for extending the MOS downscaling process to daily time-scales. In Chapter 8, all findings are summarised and discussed in relation to the overall objectives defined in section 1.3. Chapter 7 concludes with a discussion of the limitations of this work and the scope for its further development.

Chapter 2

Literature review

2.1 Outline

The following review of existing literature provides an introduction to statistical downscaling in the context of estimating future precipitation changes. Application of MOS as a correction for climate change projections is a relatively new concept and does not feature greatly in the literature. Therefore, a key role of this review is to introduce the MOS approach and distinguish it from traditional statistical downscaling, so-called Perfect-Prog, approaches. The content of this chapter is relevant for all analysis and discussion that follows in Chapters 3-7. The introduction in each following chapter briefly summarises the chapter's aims and references are made to appropriate material discussed in this literature review.

The remainder of this chapter consists of five main sections and a summary. Section 2.2 outlines how global precipitation is expected to change in future climates and summarises the current projections of future precipitation in different parts of the world. The limitations of the current generation of global climate models are also highlighted, followed by an introduction to the value of downscaling. Section 2.3 forms a discussion on dynamical downscaling and its current limitations. In section 2.4, focus is given to statistical downscaling and an overview of its methods. A comparison between dynamical and statistical downscaling methods is also given in this section. The distinction between the Perfect-Prog and MOS approaches to statistical downscaling is made in section 2.5. With the differences between these approaches in mind, section 2.6 details the steps taken in the development of a downscaling methodology and the considerations necessary at each stage. In the summary provided in section 2.7, focus is given to uncertainty in climate modelling and the challenges that persist in making reliable estimates of future precipitation.

2.2 Projected changes in global precipitation

Climate change discussion often centres around the 'global warming' phenomenon meaning that the greatest perceived changes are associated with temperature. Temperature changes are arguably more obvious and easier to measure, not to mention more highly publicised than alterations in any other climate variables. Additionally, confidence in projections of future temperature changes made by GCMs is greater than for other variables, particularly precipitation. But changes in temperature are inherently linked to changes in atmospheric moisture, precipitation and atmospheric circulation (Trenberth et al., 2007). Evaporation is directly affected by surface heating and increased temperatures are responsible for a greater moisture-holding capacity of the atmosphere, estimated to be a rate of 7% per °C (Trenberth et al., 2007).

The characteristics of precipitation, which, in addition to amount and type, include frequency, duration and intensity, may be altered greatly in a changing climate (Trenberth et al., 2003; Groisman et al., 2005). Even on occasions when the total amount of precipitation may not change much, greater intensity of precipitation may be associated with a reduction in event duration and/or frequency (Osborn et al., 2000; Trenberth et al., 2007). Indeed, the IPCC AR4 (Trenberth et al., 2007) accounts widely-observed occurrences of heavy precipitation events in regions where the total amount of precipitation has actually decreased. The gradual change in precipitation type from snowfall to rainfall in the higher latitudes of the northern hemisphere has also been noted (Trenberth et al., 2007).

2.2.1 Regional precipitation variation

The variation in global climate from region to region is a consequence of a number of factors. Christensen et al. (2007) discuss the unequal distribution of solar heating along with the responses of the atmosphere, oceans and land surface and the interactions between them as the key cause of regional variation. The physical characteristics of each region are also important, and arguably most influential in dictating precipitation distribution.

The projected changes in global climate appear to be partially dependent upon latitude, at least where increasing temperatures are concerned. Precipitation changes are more complex but still exhibit latitude-dependent features (Christensen et al., 2007). For example, current projections suggest a precipitation increase in polar regions, whereas many sub-tropical regions are expected to become drier. Seemingly much more crucial to future precipitation change on a regional scale are the topographical features of the local area. The presence of a dominant mountain range, proximity to an ocean and the direction of the prevailing wind flow are all important factors in precipitation response.

Projects such as the IPCC reports seek to comprehensively evaluate climate projections in all regions

of the world. The IPCC AR4 (Christensen et al., 2007) uses the same continental-scale regions that appear in previous reports and are divided into a number of sub-continental areas that are small enough to be reasonably analysed. The boundaries of these sub-regions closely follow those developed by Giorgi and Francisco (2000) and Ruosteenoja et al. (2003). A common set of sub-continental regions is important in developing robust statements about climate change that can be quantitatively digested by the scientific and policy-making communities.

2.2.2 Limitations of GCMs

GCMs are the most fundamental means by which climate change can be studied. GCMs are highly complex dynamical models capable of representing the processes involved with the earth's atmosphere, oceans and land surface and the interactions between them. The primitive equations, which comprise a number of well-established physical laws, are the fundamental basis upon which the dynamical climate system is simulated, with real-world observations used to define the parameterisation schemes used in the simulation of precipitation and other diagnostic variables. The most up-to-date models are able to provide simulations of long-term climate means that are, on the planetary scale at least, considered to be realistic. At this largest scale, good model skill is largely attributed to realistic simulation of important features of global climate, such as jet streams, atmospheric circulation cells and location of intertropical convergence zones (Zorita and von Storch, 1999). The accuracy of GCM runs simulating past climate have led to good confidence in prediction of future climate conditions under different atmospheric forcings.

Although GCMs are generally able to simulate climate features at planetary scales, simulations at finer spatial resolutions are associated with smaller skill and consequently less confidence. This is particularly true at the regional-scale where smaller scale processes that govern spatial variability in temperature and precipitation are not well captured. This difference in skill is possibly attributable to the scale-dependent response of climate to different factors. For instance, global climate is largely a response to solar forcing, earth rotation, the large-scale structure of the earth's physiographic features, whereas regional climate is a function of global climate and the regional environment (Zorita and von Storch, 1999). It may be considered that although planetary scale climate may be realistically simulated, a GCM may not be able to replicate the regional scale features that are so important in driving climate at this scale (Zorita and von Storch, 1999). One reason for the poor skill at regional scales is the inadequacy of GCM spatial resolution in simulating the detailed structure of the earth's surface (Zorita and von Storch, 1999). A set of 200 km grid-cells, for example, would not accurately detail land-sea boundaries and the 'smoothing' out of topographic features may reduce mountain ranges to stepped-sections or broad flat hills (Zorita and von Storch, 1999).

Also, the limited resolution of GCMs means that many climate system processes, which interact at

a variety of temporal and spatial scales, are not resolved (Randall et al., 2007). Such processes occur at a 'sub-grid scale' and include cloud formation, hydrological processes, boundary-layer conditions, and radiation and chemical processes. Although these processes occur at small scales, they are dependent upon, and in turn affect, those larger-scale processes that are well-resolved by the GCM (Kalnay, 2003). In order for this interaction to be resolved, sub-grid scale processes must be 'parameterised'. Important to precipitation are evaporation, convection and cloud processes, the parameterisations of which aim to describe easily-interpretable statistics such as mean cloud cover and area-averaged precipitation (Randall et al., 2007). One of the main reasons why climate model results differ is the differences observed between these parameterisations (Randall et al., 2007). With increasing resolution of climate models it is becoming possible to explicitly represent many more of these sub-grid scale processes. It is likely, however, that future generations of climate models will not be able to fully resolve those processes that act at the smallest scales (Zorita and von Storch, 1999). Crucially, it is at these scales that the anthropological and ecological environments experience the greatest impacts. In order for the power of GCMs to be utilised in impact studies there is a need to produce accurate simulations local climates.

2.2.3 A need to downscale

Producing climate information at resolutions smaller than that of GCMs is known as downscaling. Downscaling techniques can generally be categorised into two broad approaches: dynamical and statistical. The vast majority of literature focuses on just one approach, or occasionally on an amalgamation of the two, but most studies have, if only briefly, defined both approaches and provided a description of the differences between them. In this work, research is firmly focussed on statistical downscaling methods but it is necessary to discuss aspects of both approaches, the overall success of each in previous work and also what is involved in deciding upon which approach is appropriate for a particular application. It is fortunate that the fundamental differences between each approach mean that both can be easily defined and compared.

2.3 Dynamical downscaling

Dynamical downscaling typically uses RCMs to generate high-resolution climate information that is consistent with large-scale GCM output (Mearns et al., 2003). These high-resolution models are 'nested', in that they use GCM output as boundary conditions in order to better represent finer-scale physical processes (Christensen et al., 2007). Mearns et al. (2003) highlighted an issue concerning feedback from small- to large-scales. The real atmosphere experiences interaction between feedback processes in different regions, whilst any model-simulated feedback is limited to the fixed boundaries of the area of interest.

To avoid a high-resolution model being run under the guise of a closed system, it is recommended that an area of 'sufficient minimal resolution' is retained.

High-resolution atmosphere-only GCMs are run as an alternative to atmosphere-ocean (coupled) models, using either observed or simulated sea surface temperature (SST) and sea ice information as a lower boundary condition. Compared to those in the ocean, atmosphere and land surface components are associated with far shorter time scales meaning that a particular period of interest can be identified using a coupled model and then analysed under a higher resolution atmosphere-only model (AGCM) (Mearns et al., 2003; Christensen et al., 2007). These so-called 'time-slices' can then be examined in greater spatial detail. Mearns et al. (2003) consider this approach to be based on two assumptions: that large-scale circulation patterns are consistent in both the coarse and high resolution GCMs; and that the atmospheric state is considered to be in equilibrium with its oceanic boundary conditions.

Both nested RCMs and high-resolution atmosphere-only GCMs are able to produce highly resolved information derived from models based upon true physical processes but the fact that this is usually at great computational expense is a well-documented drawback and a hindrance to running multiple scenarios (Mearns et al., 2003).

2.3.1 Progress in RCM applications

The development and application of RCMs has featured heavily in downscaling literature during the last two decades. A large number of simulations have been conducted for different regions, with different model formulations and boundary forcings (Maraun et al., 2010). In recent years, international collaborations have established exchange and inter-comparison of multiple RCMs with the ultimate goal of making quantitative statements about uncertainty in regional projections of climate change. In Europe, such projects include ENSEMBLES (van der Linden and Mitchell, 2009) and PRUDENCE (Prediction of Regional Scenarios and Uncertainties for Defining European Climate Change Risks and Effects) (Christensen and Christensen, 2007), whilst the NARCCAP project (Mearns et al., 2009) is a recent intercomparison of RCMs for North America. Most recently, the Coordinated Regional Climate Downscaling Experiment (CORDEX) has been established with the dual-purpose of evaluating the multiple-model performance and establishing a framework through which to make refined regional climate projections (Giorgi et al., 2009).

As mentioned in section 2.2.3, skill in the representation of precipitation in GCM simulations is weaker than that of other climate variables. Precipitation is highly dependent on a model's dynamical formulation and the subsequent parameterisation of a number of relevant processes. Although precipitation in an RCM is parameterised in the same fundamental way, greater spatial and temporal resolution allows for the simulation of far more useful precipitation information (Durman et al., 2001; Frei et al., 2006; Buonomo

et al., 2007). Compared with real world observations, the amplitude of variance in GCM precipitation may be considered to be inhibited, or 'dampened', by a smoothing of topography and small-scale physical processes. The relative intensification of precipitation in an RCM improves the representation of the daily precipitation distribution, which is crucial in the analysis of extreme events (Durman et al., 2001; Christensen and Christensen, 2007; Maraun et al., 2010). Furthermore, RCMs provide an improved representation of spatial variability over topographically complex regions (Frei et al., 2006).

Frei et al. (2003) and Buonomo (2010) have shown that it is possible to make a clean assessment of the skill of the RCM-simulated precipitation when an atmospheric reanalysis is implemented as the driving GCM. The boundary conditions expressed by the reanalysis can be considered 'quasi-observations' and, thus, the subsequent RCM simulation is in temporal coherence with the real world (Maraun et al., 2010). That is, the day-to-day variability in the observed record is reflected in the simulation. The removal of random weather variability from the RCM allows for comparison with precipitation observations and a more comprehensive understanding of model skill. These concepts are discussed in more detail in Chapter 3.

2.3.2 Limitations of RCMs

Whilst RCM simulations have been widely shown to improve the representation of local-scale precipitation, there are still some important sources of error to consider. Many of these are inherent to the RCM itself (as with any dynamical model), including parameterisation schemes and representation of topography. However, the most obvious theoretical limitation is the reliance of large-scale fields as simulated by the driving GCM. All systematic errors in the GCM are inherited by the RCM, and so biases in GCM-simulated precipitation are likely to remain in the equivalent RCM field (Noguer et al., 1998; Durman et al., 2001). This limitation is, of course, common to any application of downscaling (whether dynamical or statistical) to climate change simulations, which by definition requires output from a GCM, but it is important to note that RCM output contains components of error from two dynamical models.

There is a further issue when considering the importance of the driving GCM, namely the lack of two-way interaction between the RCM and the GCM (Jones et al., 1995). Theoretically, local-scale variability would exert an influence on large-scale climate but at present there have been no successful attempts to establish a two-way system, where output from the RCM would theoretically provide feedback to the GCM. However, Mearns et al. (2003) note that whilst RCMs are not designed to perturb the large-scale climatic state simulated by a driving GCM, simulation over large spatial domains may induce a significant modification of precipitation-bearing flows. The authors further suggest these influences may in fact improve the representation of such large-scale phenomena, which is important to the continued development of high-resolution GCMs.

The extent to which RCM skill increases in unison with model resolution is an important component in understanding the benefits of RCM applications. Model parameterisation schemes are sensitive to grid cell resolution (Giorgi and Marinucci, 1996), and a scheme that may be sufficient in resolving physical and dynamical processes on a 50 km grid may not be automatically transferable to a 25 km grid. This issue is particularly important in resolving small-scale convective processes, where the parameterisation schemes involved may have been originally developed for tropical regions in coarser resolution models (Hohenegger et al., 2008). Maraun et al. (2010) note that such schemes may be inappropriate when applied in extra-tropical regions during the summer months, when the majority of precipitation events are likely to be convective. Other factors, including seasonal and regional dependences and observational station density, are known to contribute to the potential of high-resolution model skill.

2.4 Statistical downscaling

The rationale for statistical downscaling is based on the assumption that regional climate is a function of the large-scale climatic state and the physical features of the local environment such as topography, land-sea distribution and land-use (von Storch, 1995, 1999; Wilby et al., 2004). It is possible to ascertain relationships between the large- and regional-scale climate, where large-scale climate parameters are chosen as 'predictors' for local variable 'predictands'. The relationships between large- and regional-scale climate parameters for a particular region and/or period of interest are derived from observations. GCM-simulated predictor variables are then fed into the statistical model to derive estimates of equivalent local characteristics (Wilby et al., 2004).

Unlike dynamical downscaling, statistical methodologies have the advantage of being computationally inexpensive to undertake (Wilby et al., 2004) and it is far easier to derive multiple regional climate scenarios than under a dynamical approach (Christensen et al., 2007). Consequently, statistical techniques are popular with site-specific impact studies (Wilby et al., 2004). However, statistical downscaling is associated with a number of disadvantages. Crucially, the approach assumes that the empirical relationships remain constant in the future. It is not possible to verify whether or not a statistical model developed from historical or present day climate will continue to be true under (unknown) future climate forcings (Wilby et al., 2004; Christensen et al., 2007). Statistical downscaling is also limited by the inherent requirement of sufficient observational data from which to derive relationships (Christensen et al., 2007). Daily surface and/or upper air data is not available at the optimal resolution in all regions of the world, and is often lacking in the parts of the world most at risk from climate change impacts.

Additionally, the statistical approach is usually dependent on the output of a GCM. In a Perfect-Prog approach (section 2.5.1), the GCM is assumed to be ideal, and thus large-scale predictors are considered

a 'perfect' representation of the true global climate. However, GCMs inevitably contain biases and it is usually a substantial challenge to account for these biases when formulating statistical relationships.

2.4.1 Classification of statistical downscaling methods

Many attempts have been made in the categorisation of statistical downscaling techniques (e.g. Wilby and Wigley, 1997; Fowler et al., 2007; Maraun et al., 2010). In general, techniques can broken down into three broad categories: regression models, weather generators and weather classification schemes. This categorization is used in the IPCC AR4 (Christensen et al., 2007) after recommendation in the supporting guidance document (Wilby et al., 2004). Similar classifications of statistical downscaling methodologies are also used in the IPCC Third Assessment Report (TAR) (Giorgi et al., 2001) and a number of review papers (Wilby and Wigley, 1997; Wilby et al., 1998; Fowler et al., 2007).

In a recent review paper, which focused specifically on precipitation downscaling under climate change, Maraun et al. (2010) based their categorisation on a distinction between Perfect-Prog and MOS approaches. Whilst such a categorisation may be relevant here, the discussion of MOS downscaling by Maraun et al. (2010) is dedicated almost entirely to bias correction of RCM output rather than a direct downscaling correction for GCMs. In this research, MOS is proposed as an alternative approach to Perfect-Prog downscaling but may potentially utilise the same techniques developed for Perfect-Prog methodologies over recent years, and these techniques are discussed with reference to their implementation in previous work. The categorisation used in the IPCC AR4 (Wilby et al., 2004; Christensen et al., 2007) remains a convenient manner in which to outline and discuss different techniques and related studies. It is important to note that, in practice, many approaches can take on attributes of more than one of these techniques (Wilby and Wigley, 1997).

2.4.1.1 Weather classification

Weather classification schemes, also known as weather typing, attempt to relate local climate to a set of synoptically defined weather types or 'states' (Wilby et al., 2004; Fowler et al., 2007). Instead of estimating the predictand based on a single continuous predictor field, each day within a period of analysis is assigned a particular categorical weather type, either through subjective or, more commonly, objective means (Maraun et al., 2010). Usually, cluster analysis is applied to atmospheric fields (Corte-Real et al., 1999; Huth, 2000; Kidson, 2000; Hewitson and Crane, 2002; Fowler et al., 2000, 2005) but other means of classification have used sea level pressure indices (Conway et al., 1996) and empirical orthogonal functions (EOFs) from pressure data (Goodess and Palutikof, 1998).

Subjective classification techniques can be traced back to the Lamb Weather Types (LWTs) of the British Isles (Lamb, 1972) and more recent schemes (Bardossy and Caspary, 1990; Jones et al., 1993;

Wilby, 1994; Conway et al., 1996). Wigley and Jones (1987) related LWTs to spatial and temporal variations in precipitation. Conway et al. (1996) extended the attempts of Jones et al. (1993) to develop an objective version of Lamb's Weather Types using three continuous indices of geostrophic air flow, namely total shear vorticity, strength of the resultant flow and angular direction. In a regression and resampling methodology, vorticity was shown to be most useful in estimating precipitation.

Once a classification scheme is in place, local surface variables, often precipitation, are conditioned on the corresponding daily weather patterns. Typically, this is achieved by deriving conditional probability distributions for observed data (Wilby and Wigley, 1997; Fowler et al., 2007). The change in frequency of GCM-simulated weather classes can be evaluated to estimate future changes in climate, although such a method is limited in its assumption that the characteristics of each weather class will remain constant (Fowler et al., 2007).

Weather classification schemes are favoured for their versatility and their usefulness in the analysis of extremes (Wilby et al., 2004). Conversely, the additional task of classification that is inherent to this downscaling approach is often time-consuming and its use is not as common in the literature as, say, regression techniques.

2.4.1.2 Weather generators

Weather generators can be described as statistical models that generate random sequences of weather which reproduce the observed statistical attributes of a given climate variable, such as the mean and variance, but not variability in the observed sequence of events (Wilks and Wilby, 1999; Wilby et al., 2004; Maraun et al., 2010). In their simplest form, model calibration is performed on local-scale observations and so is not strictly a downscaling approach (Maraun et al., 2010). Modern weather generator applications are more sophisticated, with parameters conditioned on large-scale weather in some cases (Wilks and Wilby, 1999). In the context of downscaling, this may include atmospheric predictors, climatic states, or the properties of precipitation (e.g. Wilks, 1992; Katz, 1996; Palutikof et al., 2002; Feddersen and Andersen, 2005). In application for climate change scenarios, it is possible to adjust the model parameters by applying 'change factors' diagnosed from GCM or RCM output (e.g. Kilsby et al., 2007; Maraun et al., 2010). Change factor conditioned weather generators were used in the UKCP09 regional climate projections (Jones et al., 2009).

The most obvious advantage offered is the synthesis of a random weather sequence for a theoretically unlimited period, which is particularly useful when observational records are unavailable or incomplete (Yang et al., 2005). Furthermore, weather generators permit the production of large ensembles and thus prove attractive for analysis of uncertainty (Wilby et al., 2004). Additionally, it is possible to generate sub-daily information (Fowler et al., 2000).

One weakness of weather generators in statistical downscaling is that they are calibrated on local relationships that may not be constant in a future, perturbed climate, and the extent of such a limitation has not yet been fully investigated (Fowler et al., 2007). Weather generators have also been shown to underestimate inter-annual variability and the frequency of extreme events, largely because of an insufficient representation of the longer term variability of key climatic processes that drive local weather (Katz and Parlange, 1998; Fowler et al., 2007; Maraun et al., 2010). Hybrid techniques that link weather generators with regression models have been shown to improve this (Wilby et al., 2002). An additional limitation is the representation of spatial consistency in many weather generators (Jones et al., 2009). In the context of precipitation downscaling, Maraun et al. (2010) make a distinction between single- and multi-station weather generators, and it is the latter that involves additional challenges. Multi-station generation must consider the distribution of precipitation at a number of (usually gridded) locations in unison in order to reflect the spatial coherence that would exist between neighbouring stations in the real world, and by definition necessitates multi-variate techniques.

2.4.1.3 Regression methods

Arguably the most widely applied statistical downscaling approach is the use of regression models to directly quantify linear or non-linear relationships between predictors and predictands. In their most fundamental form, a simple linear regression model may be constructed using a climate variable as a predictor for surface values of, typically, temperature and precipitation (Fowler et al., 2007). In practice, the estimate for the predictand is usually based upon a one or more predictor fields (i.e. quantities of predictor over a spatial domain), which necessitates a multiple regression (e.g. Huth, 1999; Murphy, 1999; Spak et al., 2007).

More advanced techniques have taken alternative predictors, often based on a transformation of the original predictor field(s), such as principal components extracted from values of, say, pressure and geopotential height (e.g. Kidson and Thompson, 1998; Hanssen-Bauer et al., 2003). More sophisticated approaches have attempted to identify coupled patterns in the variability of spatially-corresponding predictor and predictand fields. Common techniques adopted include maximum covariance analysis (MCA) (e.g. Huth, 1999; Widmann et al., 2003; Tippett et al., 2008) and, particularly, canonical correlation analysis (CCA) (e.g. Busuioc et al., 2001; Zorita and von Storch, 1999; Shongwe et al., 2006), both are discussed in more detail in section 2.6.2. The application of non-linear regression models is less frequent although the most common technique, artificial neural networks (ANNs) (e.g. Hewitson and Crane, 1996; Zorita and von Storch, 1999; Schoof and Pryor, 2001), has been used to downscale precipitation.

An important issue raised by von Storch (1999) and in a number of subsequent review papers (e.g. Wilby et al., 2004; Maraun et al., 2010) is the under-prediction of local-scale variance in regression

downscaling approaches; in other words, the chosen predictors are only able to account for a portion of the variability in the predictand. It is problematic to implement a scaling (or 'inflation') (e.g. Karl et al., 1990; Burger, 2002) of the simulated predictor to better reflect quantities of the predictand as local-scale variation contains an element of noise that, fundamentally, cannot be explained in a deterministic model (von Storch, 1999; Maraun et al., 2010).

2.4.2 Inter-comparison of statistical downscaling methods

In recent years, much work has focused on the comparison of different statistical downscaling methodologies (e.g. Wilby and Wigley, 1997; Wilby et al., 1998; Zorita and von Storch, 1999; Wilby and Wigley, 2000; Cavazos and Hewitson, 2005; Goodess et al., 2010) and such studies offer an important basis when attempting to quantify the relative value of each method. In an early example, Wilby et al. (1998) compared the skill of two weather generator techniques, two vorticity-based models and two ANN techniques in downscaling precipitation from the HadCM2 GCM, finding marked differences between the methods. The smallest differences between observed and simulated precipitation were found with the weather generators, although these methods are shown to poorly simulate low-frequency variability.

In downscaling daily and monthly winter rainfall in the Iberian Peninsula, Zorita and von Storch (1999) compared an analog method with more sophisticated techniques considered to be representative of each of the broad groupings in statistical downscaling. These included CCA, an ANN and a classification method. The analog method compared favourably with all methods and actually shows greater skill than the ANN, surprising given the far superior complexity of the ANN. CCA (representative of linear methods) is found to offer the clearest physical interpretation of rainfall patterns and is considered most suitable for estimates of future climate.

Amongst the six statistical downscaling methods evaluated by Schmidli et al. (2007) are regression methods, weather typing and a conditional weather generator. Additionally, a bias correction approach, based on the local scaling technique developed by Widmann et al. (2003), and a two-step analog method are also used. Although all methods are shown to strongly underestimate interannual variations, the weather classification and analog approaches do show the highest skill. The choice of study area, the European Alps, is a region of varied topography, and previous studies have shown statistical downscaling methods to lack skill in such regions. One contrasting example is Widmann et al. (2003) in which dynamically-corrected local scaling improved precipitation estimates in the rain shadow of the Cascade mountains of the north-west United States.

The majority of comparison studies have focused on statistical downscaling of monthly means, although recently more attention has been given to precipitation on a daily time scale, with specific focus on extreme events. As part of the European Union funded STARDEX (STAtistical and Regional Down-

scaling of EXtremes for European regions) project, twenty-two different statistical downscaling models were developed and systematically evaluated on their ability to reproduce indices of extreme temperature and precipitation. Many of the STARDEX findings are summarised by Goodess et al. (2010) although a breadth of other publications exists (e.g Busuioc et al., 2006; Kostopoulou et al., 2007; Schmidli et al., 2007). The methods used encompass each category detailed in section 2.4.1 and a range of predictor variables. Much of the STARDEX work experiences difficulty when attempting to identify the most suitable method or combination of predictors. Goodess et al. (2010) also noted that those downscaling models exhibiting skilful representation of temporal variability are not always those with lowest biases. An important recommendation suggested by Goodess et al. (2010) is that, given the lack of a consistently superior downscaling method, estimation of future precipitation scenarios should be based on a range of (at least the most skilful) methods. This is conceptually similar to the practice of basing projections of global (regional) climate on an ensemble of GCM (RCM) simulations (Mearns et al., 2003).

2.4.3 Comparing dynamical and statistical downscaling

Focus has also been given to the relative merits of dynamical and statistical downscaling methodologies (e.g. Kidson and Thompson, 1998; Murphy, 1999; Hanssen-Bauer et al., 2001). In such a comparison study, Kidson and Thompson (1998) found a nested climate model and a regression-based statistical method to give similar representations of present day observations of temperature and precipitation in New Zealand. The authors pointed out that, in this case, the extra expense and effort involved in producing a suitable RCM does not result in an improvement over statistical techniques. Murphy (1999) carried out a similar comparison of methods for downscaling historical climate in Europe. A regression model was produced alongside a nested RCM with both techniques showing similar overall skill, with estimates for temperature showing a higher degree of skill than those for precipitation. Murphy (2000) took this work a step further, comparing downscaled predictions of future climate (2080-2100) across the same European region. Whereas statistical and dynamical estimations of historical local climate were similar, predictions of future climate scenarios differ significantly. In particular, there is a large difference in estimates of precipitation.

Other work has found notable differences in the skill of each method. Mearns et al. (1999) examined precipitation scenarios for Nebraska, USA offered by a weather typing technique and a RCM, finding marked differences between the two. Similarly, Hellstrom et al. (2001) found that statistically downscaled scenarios of precipitation change across Sweden were associated with a greater temporal and spatial variability than the corresponding dynamical approaches. A number of studies have noted an apparent weakness of statistical methods in estimating summer precipitation, relative to winter precipitation, possibly due to a poor representation of an intensified hydrological cycle (Hanssen-Bauer et al., 2001;

Haylock et al., 2006; Schmidli et al., 2007). Hanssen-Bauer et al. (2001) concluded that dynamically downscaled estimates of summer precipitation are likely to be more realistic. The authors suggested that using a humidity measure as an additional predictor would yield improved results, although the inclusion of such variables has not always been associated with more skilful models (e.g. Hellstrom et al., 2001).

Schmidli et al. (2007) attempted to make a more thorough comparison of different downscaling methods for estimating precipitation over the Alps. Several models in each category were evaluated, including three RCMs and six statistical models. Additionally, the statistical techniques chosen encompass a range of fundamentally different methodologies. In general, both sets of models were shown to have similar biases, but the statistical techniques were found to underestimate interannual variations. The authors concluded that the means of deriving regional climate information via either dynamical or statistical downscaling contributes to great uncertainty. This appears particularly true in summer when stochastic processes occurring at the mesoscale are of greater importance.

Goodess et al. (2010) suggest that an additional benefit of statistical downscaling is the potential to provide information on the skill of climate models in reproducing large-scale predictor variables and, equally, their relationships with predictands. A necessary step in statistical downscaling is quantifying predictor-predictand relationships, which is potentially useful in identifying sources of error and bias in GCMs (or RCMs) (e.g. Osborn et al., 1999).

In general, statistical methods tend to be more favourable to researchers, a notion largely down to their ease of application to many studies. Even though dynamical methods perhaps have greater theoretical potential to produce accurate downscaled estimates, carrying them out effectively is often impractical. There tend to be relatively few RCM simulations available and these may not cover the period for which estimates are required. Furthermore, most are relatively short (e.g. 30 years), making it difficult to assess the representation of multi-decadal variability (Goodess et al., 2010). The substantially greater costs and computational power required to carry out such methods is not necessarily matched by a worthwhile improvement of local-scale estimates over those produced via statistical means. There remains a challenge, however, to further improve the statistical downscaling process and subsequently the estimates of future local climate conditions.

2.5 Perfect-Prog and MOS approaches to statistical downscaling

As discussed in section 2.4, statistical downscaling encompasses a range of methods and techniques of which there are many examples in the literature. In the context of this research, it is necessary to examine the two fundamental approaches that any statistical downscaling methodology can follow, namely Perfect-Prog and MOS.

The evolution of these approaches is rooted in weather forecasting and it is important to consider the development in statistical methods used in Numerical Weather Prediction (NWP) over the course of the last few decades. Statistics has long been used in weather forecasting as a means to correct for errors and develop transfer functions (Zwiers and Von Storch, 2004). NWP systems are able to forecast weather at large-scales with greater success than statistical methods but lack the skill to forecast at local-scales. NWP models can be considered coarse in nature and are unable to simulate weather variation on a local-scale. Statistical methods can be used as the sole means of forecasting but are more often used in conjunction with the output of NWP systems. This approach involves 'post-processing' the output of NWP models in order to make more accurate forecasts at local-scales.

Statistical reinterpretation of the output of NWP models is theoretically analogous to downscaling GCM output. In this section, the focus is initially on outlining the principles of the Perfect-Prog and MOS approaches and their use in weather forecasting, but it is permissible to consider the application of each approach to statistical downscaling to be conceptually similar.

2.5.1 Perfect-Prog

The Perfect-Prog approach to statistical downscaling was developed by Klein et al. (1959) for application in weather forecasting and typically involves deriving a statistical relationship between simultaneous observations of the predictor(s) and predictand (Zwiers and Von Storch, 2004). That is, the relationship derived is based entirely upon historical observations. In this sense, Wilks (2006) suggest that the formulation of Perfect-Prog regression equations is analogous to that of the 'classical' regression equations that formed the basis of pre-NWP forecasting. Such equations quantify links between present observations of a predictor(s) to future observations of the predictand. Thus, one may be able to use today's value of, say, sea-level pressure to predict tomorrow's value of, say, temperature. Perfect-Prog equations differ primarily from classical forecast equations in the absence of this time lag. Instead, the observations of predictors and predictands used to derive a relationship are simultaneous. In this sense, tomorrow's predictor variable is used to determine tomorrow's predictand.

Wilks (2006) goes on to point out that in a classical weather forecasting undertaking of this approach, the predictor variable would obviously not be known until tomorrow. Perfect-Prog forecasting, however, uses tomorrow's predictors as forecasted by NWP models. When applying the derived relationship to the output of a NWP model, the model itself is assumed to be 'perfect' in accurately simulating the required predictor variable. If the NWP model is indeed perfect, and consequently provides excellent estimates of predictors, one can expect very accurate forecasts. However, as the approach does not consider any possible errors or biases in the NWP model and takes forecasts of atmospheric variables at 'face value', the accuracy of any resultant forecast is entirely dependent on the skill of the initial dynamical model

(Wilks, 2006).

The Perfect-Prog approach is heavily reliant on an extensive historical dataset but is especially useful when forecasting for an individual location or season with a long-term observational record. Although this approach has been replaced by MOS in weather forecasting, downscaling output from GCMs still largely follows the Perfect-Prog approach.

2.5.2 Model Output Statistics (MOS)

MOS is an alternative to the Perfect-Prog approach and was initially developed for NWP application by Glahn and Lowry (1972) and Klein and Glahn (1974). MOS "consists of determining a statistical relationship between a predictand and variables forecast by a numerical model at some projection time" (Glahn and Lowry, 1972). MOS is the preferred method of improving NWP forecasts as, unlike the Perfect-Prog approach, simulated predictor variables are used to both develop and implement the statistical equations (Wilks, 2006). The MOS approach thus recognises that NWP models are not perfect and allows for statistical relationships that are able to account for model errors and biases.

Similarly to Perfect-Prog and classical forecasting, the MOS approach is based upon a relationship between a predictor and predictand. Wilks (2006) expresses the distinctions between the three in a series of equations. Following the classical approach, a relationship is derived between today's observations of the predictor(s) and tomorrow's observation of the predictand:

$$y_t = f(x_0) \tag{2.1}$$

where the y_t is the predictand at time t, f is the relationship (i.e. regression function), and x_0 is the predictor variable(s) at time 0 (i.e. current values, earlier than time t).

The Perfect-Prog approach differs in that it uses observations in the development of a forecast equation but applies NWP-simulated predictors in its implementation. Thus in development,

$$y_0 = f_{PP}(x_0) \tag{2.2}$$

where f_{PP} is the Perfect-Prog function and the concurrent values of x_0 and y_0 are taken from the observed record. In implementation, a future value of the predictand, y_t, is estimated thus:

$$y_t = f_{PP}(x_t^*) \tag{2.3}$$

where f_{PP} is the same Perfect-Prog function used in Equation 2.2 and x_t^* is simulated at time t. Conversely, the MOS approach includes model-simulated predictors in both the development and imple-

mentation of the function. In development,

$$y_0 = f_{MOS}(x_0^*) \tag{2.4}$$

where f_{PP} is the MOS function and x_0^* is a hindcast simulation of the predictor x at time 0. In implementation, the future value of the predictand is estimated in the same way as in Equation 2.3:

$$y_t = f_{MOS}(x_t^*) \tag{2.5}$$

where f_{MOS} is the same MOS function used in Equation 2.4 and x_t^* is the predictor x simulated at time t. Unlike the Perfect-Prog approach, in which the same statistical model can be used for any location and any forecast projection, the MOS approach requires separate forecast equations for each numerical model. The error characteristics of any NWP model differ depending on the forecast projection and one should expect different statistical relationships between variables at, say, 24 hours and 48 hours into the future (Wilks, 2006).

Glahn and Lowry (1972) are credited with the first MOS approach which they applied to make improved predictions of surface wind, precipitation, maximum temperature and cloud cover in the United States. The MOS approach was found to be very successful, particularly when forecasting probabilities and the authors also provided evidence of application of their new technique. Subsequently, Klein and Glahn (1974) compared local weather observations with NWP model output, showing that MOS allows for both the inaccuracy of the model and the local climatology to be included within the forecast system. Additionally, more detailed applications of the method are discussed, specifically in the prediction of thunderstorms and tornadoes.

A number of studies have followed the early development of MOS in numerical weather forecasting. McCutchan (1978) used MOS to successfully predict synoptic weather types in southern California. Karl (1979) focused on forecasting concentrations of ozone in Missouri, using MOS to formulate statistical equations. This study represents the first successful attempt using MOS to predict a weather-dependent variable, as opposed to simply a weather variable. Godfrey (1982) used an MOS application to forecast the Levante wind regime in southern Spain, showing successful application of the approach in the prediction of a locally-occurring dynamic phenomenon. Carter et al. (1989) discussed the merits of MOS and Perfect-Prog for operational forecasts in the USA, with MOS shown to be superior. MOS has also been applied to numerical models used in seasonal forecasting applications (e.g. Landman and Goddard, 2002; Feddersen and Andersen, 2005; Shongwe et al., 2006) but as yet has not been fully utilised in GCMs used for climate change simulations.

2.5.3 Application of MOS to downscaling of climate change simulations

After its initial development and implementation (Klein et al., 1959), Perfect-Prog made way for the more 'efficient' MOS in NWP (Zwiers and Von Storch, 2004). Perfect-Prog was still occasionally utilised for forecasting on occasions where local factors were especially important (Klein and Glahn, 1974). The Perfect-Prog approach, however, is now commonly associated with downscaling GCM climate simulations but recent developments suggest there remains potential for MOS to play a role.

As discussed in section 2.3, precipitation simulated in RCMs, whilst more informative than that simulated in GCMs, is still representative of an area mean at model resolution rather than a local-scale value (Maraun et al., 2010). Recent developments in RCM application have focused on an additional step in the downscaling process in which namely a statistical link is established between simulated precipitation and observed precipitation (e.g. Schmidli et al., 2006; Leander and Buishand, 2007; Graham et al., 2007; Themessl et al., 2011). This statistical 'bias correction' of RCM output is a simple form of MOS, although usually these corrections operate at the same resolution and thus are not explicitly a downscaling function.

A direct MOS downscaling correction (i.e. without the intermediate RCM) of GCM precipitation using a similar 'scaling' concept has been proposed and shown to exhibit potential (e.g. Widmann and Bretherton, 2000; Widmann et al., 2003; Schmidli et al., 2006). In the case of Widmann and Bretherton (2000) and later Widmann et al. (2003), this was demonstrated through the local scaling of precipitation from an atmospheric reanalysis, the parameterisation of which can be considered similar to that of GCM precipitation. Conceptually, there are some clear benefits of such an approach. As mentioned, MOS allows for a GCM's errors and biases to be accounted for and it is possible that increased computing resources along with a wider availability of GCMs to a greater number of researchers has meant that focus can be given to a particular model's characteristics. Additionally, as MOS uses simulated variables, this approach is able to include many predictors, such as three-dimensional air trajectories and boundary layer potential temperature, which are not typically observed and thus cannot be included in a Perfect-Prog approach (Klein and Glahn, 1974).

A obvious limitation of the MOS approach is the requirement of a long series of hindcasts using the same model which is subject to the downscaling (Feddersen and Andersen, 2005). GCMs are in constant development and any upgrade may potentially alter systematic errors of the model physics and dynamics. Thus, any large-to-regional scale relationship must be derived using historical simulations of the latest model version. It is not always practical to attain updated historical GCM simulations from which MOS downscaling models can be formulated, and a Perfect-Prog approach (in which relationships are based upon true observations) has become the favoured downscaling method over the last fifteen years or so (Zwiers and Von Storch, 2004).

However, the most crucial component in the application of MOS for GCM precipitation is the initial hindcast simulation upon which downscaling models are calibrated. A standard GCM simulation (or an RCM driven by a standard GCM simulation) is freely-evolving and although long-term simulated climate fields may agree with observations, day-to-day 'weather' is not in temporal alignment with the real world. Thus, sequences of monthly and seasonal means are not in agreement with reality so only MOS based on long-term precipitation means or distributions is possible. MOS correction of RCM-simulated precipitation permits an alternative approach. As explained in section 2.3.1, a comparison of simulated and observed time series can be achieved by conducting an RCM simulation driven by reanalysis data, which forces the temporal evolution of individual variables to match that in the observed record. Subsequently, it is possible to fit more sophisticated MOS models (i.e. using regression) and fully quantify the skill of each.

There is potential to extend the concept of MOS correction for the purpose of downscaling GCM climate change simulations. In addition to performing local scaling corrections, Widmann et al. (2003) developed statistical models to directly downscale precipitation from the NCEP-NCAR reanalysis, in which the day-to-day and year-to-year sequence of the large-scale climatic state corresponds to observations. Although a reanalysis is designed to assimilate real world observations, its background forecast model can be considered similar to that of a GCM used for climate change simulations. Widmann et al. (2003) found that a downscaling correction of reanalysis precipitation, based either on local scaling or singular value decomposition, produced more skilful estimates of local-scale precipitation in the North-West United States than Perfect-Prog techniques based on other large-scale predictors.

Extending the approach of Widmann et al. (2003) to a climate change GCM has not yet been undertaken, largely due to the additional challenge of producing a hindcast GCM simulation which captures the temporal variability of real world observations. These concepts are discussed further in section 2.7 and in greater detail in Chapter 3.

2.6 Necessary considerations in the development of a statistical downscaling scheme

The majority of downscaling work has focused on precipitation as the predictand variable and a large number of methods have been developed to estimate small-scale precipitation changes. There have been various attempts to classify these methods in the literature; Wilby and Wigley (1997) describe three categories, namely regression models, weather generators and weather-type approaches, whereas Rummukainen (1997) and recently Maraun et al. (2010) have categorised methods according to the conceptual differences between Perfect-Prog and MOS approaches. To understand the full evolution of

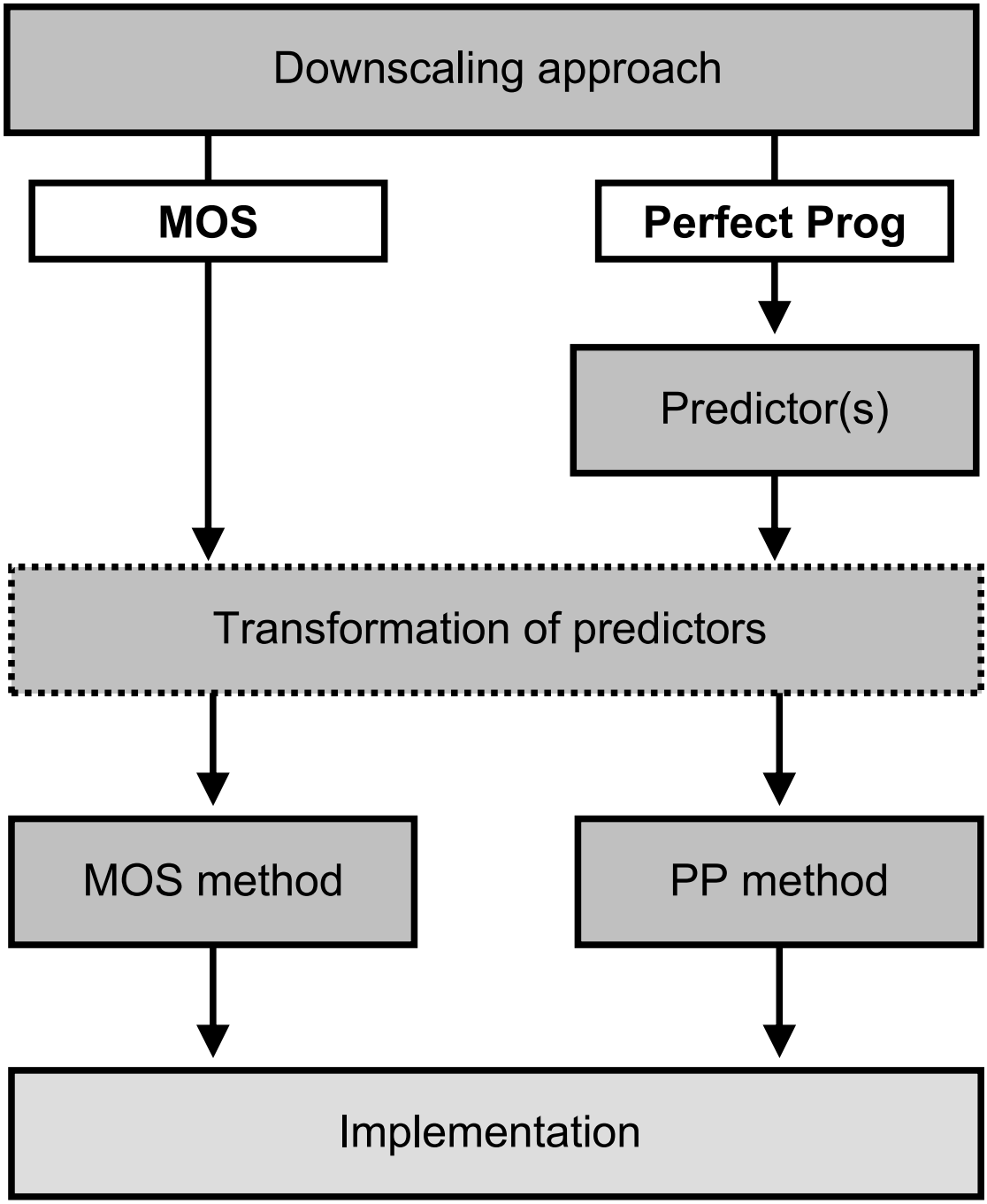

Figure 2.1: The conceptual development of a downscaling scheme. In the context of the themes discussed here, this begins with the approach to be used: MOS or Perfect-Prog. The latter requires an additional step, the choice of predictor(s). 'Transformation of predictors' can be considered an optional step.

a downscaling methodology, it is important to consider the process as a number of steps, as shown in Figure 2.1. It is first necessary to define the downscaling approach as either MOS or Perfect-Prog. With the latter, an additional choice is made about which observational variables are to be used as predictors. Under a MOS framework in the context of this work, it is assumed that the sole predictor variable will be simulated precipitation itself. The choice of method is approach-dependent, with MOS offering a directly empirical correction of the simulated precipitation field. Other methods, including linear regression, can be implemented under each framework. It may also be necessary to transform or reduce predictor data into modes of variability. The steps identified in Figure 2.1 are now discussed in more detail.

2.6.1 Downscaling approach and choice of predictors

2.6.1.1 Perfect-Prog predictors

Certainly, the majority of methods seeking to estimate changes in monthly mean precipitation have been developed under the classical Perfect-Prog approach, in which a statistical link is derived between observations of a large-scale predictor(s) and some local-scale predictand. The development of reanalysis datasets (e.g. Kalnay et al., 1996; Uppala et al., 2005) has meant that a far greater number of atmospheric variables can be considered as potential predictors. Reanalysis products are especially useful when the study area is large or involves regions of poor station data quality or density (Maraun et al., 2010). The initial choice of the predictor variable(s) is a key component in the development of a downscaling model. Wilby et al. (2004) and Goodess et al. (2010) specify that an ideal predictor variable for statistical downscaling should fulfil all the following criteria.

- To exhibit a strong correlation with the predictand (target variable).
- To have a realistic physical link to the predictand.
- To be able to explain low-frequency variability and trends.
- To be well represented by the GCM.
- To reproduce realistic interannual variability.
- To capture signals of climate change.

Although the characteristics of a strong predictor are well-defined and widely agreed upon, there is no clear consensus within the existing literature as to which predictor variables are most useful in downscaling precipitation. Table 2.1 summarises the predictor variables used in a number of downscaling studies (extended from Wilby and Wigley, 2000). Commonly used variables include those representing atmospheric circulation or temperature (e.g. Karl et al., 1990; Conway et al., 1996; Wilby et al., 1998), but

Table 2.1: A summary of predictor variables and methods used to downscale precipitation in recent studies (adapted and extended from Wilby and Wigley, 2000).

Author(s)	Predictor variable(s)	Method(s)
Bardossy and Plate (1992)	500-hPa geopotential height	Weather classification
Busuioc et al. (2001)	Sea-level pressure	Canonical correlation analysis
Cavazos (1999)	Sea-level pressure	Artificial neural network
Conway et al. (1996)	Vorticity	Semi-stochastic, regression
Crane and Hewitson (1998)	Geopotential heights, specific humidity	Artificial neural network
Goodess and Jones (2002)	Sea-level pressure, North Atlantic Oscillation (NAO) index, circulation types	Regression
Goodess and Palutikof (1998)	Mean sea-level pressure, airflow indices	Weather classification
Hanssen-Bauer et al. (2003)	Temperature and sea-level pressure	Regression
Haylock et al. (2006)		Artificial neural network
Hay et al. (1992)	Wind direction, cloud cover	Weather classification
Hellstrom and Chen (2003)	Large-scale circulation indices and 850 hPa humidity	Regression
Hellstrom et al. (2001)	u, v, vorticity, humidity at 850 hPa	Regression
Karl et al. (1990)	Geopotential heights, thickness, sea-level pressure, relative humidity	Principal components analysis and canonical correlation analysis
Katz (1996)	Sea-level pressure	Stochastic
Kilsby et al. (1998)	Vorticty, sea-level pressure, airflow strength and direction, altitude, distance from coast, grid reference	Regression
Landman and Goddard (2002)	850 hPa geopotential height	Canonical correlation analysis
Matyasovszky and Bogardi (1996)	500 and 700 hPa geopotential heights	Weather classification
Perica and FoufoulaGeorgiou (1996)	Thermodynamic parameters	Stochastic-dynamic
Shongwe et al. (2006)	850 hPa geopotential height	Canonical correlation analysis
Schoof and Pryor (2001)	Circulation predictors	Regression, artificial neural network, principal components analysis and cluster analysis
Tolika et al. (2007)	500 and 700 hPa geopotential height, 1000-500 hPa thickness, specific humidity	Artificial neural network
Trigo and Palutikof (2001)	Sea-level pressure	Artificial neural network
Wilby et al. (1998)	Vorticity, sea surface temperature, North Atlantic Oscillation (NAO) index	Semi-stochastic, regression
Woodhouse (1997)	Teleconnection indices	Rotated principal components analysis

recent research has emphasised the value of using a measure of atmospheric moisture content (e.g. Crane and Hewitson, 1998; Charles et al., 1999; Wilby and Wigley, 2000). The use of circulation predictors alone is considered insufficient in capturing important precipitation formation processes based on moisture and thermodynamics (Fowler et al., 2007). Humidity or otherwise some measure of atmospheric moisture is increasingly used in conjunction with circulation variables as part of a multiple-predictor method (e.g. Karl et al., 1990; Wilby and Wigley, 1997; Murphy, 2000; Hellstrom and Chen, 2003). It is important to consider the relative success of including many predictor variables in downscaling schemes given the spatial and temporal variation of the explanatory power of any predictor (Huth, 1999; Wilby and Wigley, 2000).

Special considerations need to be made when downscaling in the tropics where the climate system is often characterised by different complexity. The role of the ocean is more prominent in determining atmospheric conditions in such regions and oceanic predictor variables may provide critical information

in the estimation of local variables (Wilby et al., 2004). Furthermore, predictor-predictand relationships are more susceptible to within-year variation and downscaling models that represent a solitary month or season may be more useful (Jimoh and Webster, 1999).

2.6.1.2 MOS predictors

In the context of this research, a MOS correction is proposed as an approach to downscaling the GCM-simulated precipitation field. Under this approach, GCM-simulated precipitation acts as the sole predictor and an alternative to the common suite of circulation- and moisture-based variables typically used to downscale precipitation under a Perfect-Prog framework. In assessing the capability of the simulated precipitation field to adhere to the criteria defined by Wilby et al. (2004), it is useful consider how precipitation is resolved in a GCM. Widmann et al. (2003) (and later Schmidli et al., 2006) suggested that as GCM-simulated precipitation is a diagnostic parameterisation based on the dynamically-derived prognostic variables, it integrates all the significant large-scale predictors. That is, the precipitation field can be considered to include all relevant 'predictive' information contained in these variables. Furthermore, all relevant information pertaining to the climate change signal is, conceptually at least, contained in the precipitation field.

Following the efforts of Widmann et al. (2003) in demonstrating the potential of GCM precipitation as a predictor for downscaling precipitation, some subsequent studies have adopted a similar approach. Salathe (2003) attempted to simulate streamflow in a rainshadow river basin, again within the northwestern United States. Schmidli et al. (2006) extended the temporal resolution of Widmann et al. (2003) from a monthly to daily timescale across a region of varied topography, the European Alps.

Furthermore, in accepting simulated precipitation as the sole predictor, an obvious advantage of MOS over Perfect-Prog is that the (usually subjective) choice of predictor(s) is removed from the downscaling process. In principle, MOS may be based on an alternative simulated predictor (e.g. Themessl et al., 2011), although this is rare. There is suggestion that combining precipitation data with circulation data as predictors would provide information on the thermodynamic and fluid dynamic controls on precipitation (Salathe, 2003). Tolika et al. (2007) found that including simulated precipitation within a downscaling model for Greece alongside 500-hPa geopotential height and specific humidity gave improved results. Additionally, in the context of seasonal forecasting, Landman and Goddard (2002) used a MOS approach to produce regional rainfall forecasts for Southern Africa. A screening procedure, 'forward selection' (Wilks, 2006), is used to determine that 850 hPa geopotential height is the best predictor variable. These findings were supported by Shongwe et al. (2006) in a subsequent study of the same region which also showed geopotential height at the other pressure levels to have some skill in determining regional rainfall. However, in the majority of applications of MOS, particularly those using RCMs, a link is generally

derived between simulated and observed precipitation (Maraun et al., 2010). With this in mind, the 'predictor selection' step is removed in the MOS process, as detailed in Figure 2.1.

2.6.2 Transformation of predictors

Before fitting a downscaling model on predictor information, it may be desirable to subject the predictor data to a transformation. In applications in which a predictor-predictand relationship is established between time series at individual grid points (including MOS bias correction/local scaling), this process is not required. Thus, in the context of the development of a downscaling scheme, predictor transformation can be considered an 'optional' step, as indicated in Figure 2.1. However, when predictor data is in the form of a large grid based array, it is often necessary to reduce the number of variables, or dimensions, and extract modes of variability (Maraun et al., 2010). Such a transformation requires analysis of the structures that exist within a dataset, which in its most simplistic form may involve correlation between pre-defined indices (Bretherton et al., 1992). For example, the North Atlantic Oscillation (NAO) index is an indicator of variability in the North Atlantic pressure field, and pressure fields have been used to define indices describing air flow direction, strength and vorticity in precipitation downscaling applications (Conway and Jones, 1998; Maraun et al., 2010).

In a recent review paper, Maraun et al. (2010) considered downscaling with weather types (see section 2.4.1) to be an implicit form of predictor transformation. In this instance, the new predictors are the weather types themselves, defined by cluster analysis or some other form of classification and based on synoptic meteorological situations. Whilst it is relatively straight-forward to conduct predictor transformation that is physically consistent, there is invariably an element of subjectivity in the classification process.

More complex analysis using matrix operations provide a more objective measure of predictor-predictand relationships and are also able to account for the co-dependence of information at grid points in close proximity to one another. Principal component analysis (PCA), or empirical orthogonal function (EOF), analysis is widely used in reducing dimensionality of a single data field. The new variables created are linear combinations of the original variables, and capture the largest possible portion of variability that exists within the original data (Wilks, 2006). These new variables are a set of orthogonal vectors known as principal components (or alternatively as EOFs) (Hannachi et al., 2007).

When establishing a predictor-predictand relationship it is necessary to identify coupled modes of variability between two time series (Bretherton et al., 1992). PCA is usually only applied to a single dataset and therefore limited to a decomposition of the predictor field without consideration of variability in the predictand field (Tippett et al., 2008). As such, PCA is unable to optimise the predictor-predictand relationship and downscaling studies often use alternative methods to quantify the association between

two datasets. CCA is an example of such a method that has been widely implemented for downscaling applications (e.g. Karl et al., 1990; Bretherton et al., 1992; Huth, 1999; Landman and Goddard, 2002; Widmann, 2005; Tippett et al., 2008). MCA is conceptually similar to CCA but is known to yield different results (Widmann, 2005; Tippett et al., 2008). The majority of such methods have been developed for large-scale observed predictors as part of a Perfect-Prog approach, but there is scope for extending these techniques for a 'non-local' MOS correction. All appropriate methods are discussed in greater detail in Chapter 4.

2.6.3 Choice of downscaling method

Once a decision is made on which approach to follow (and, if following a Perfect-Prog approach, a decision on which predictor(s) to use) a statistical downscaling model can be developed. Key review papers in this area have been inclined to group all statistical downscaling methodologies into several categories, which are outlined in Chapter 2 (section 2.4.1). Different methods have been applied in many areas of the world (also summarised in Table 2.1) and the choice of method is undoubtedly region-specific.

The most common MOS method is bias correction or, in the context of downscaling GCM output, a local scaling correction (e.g. Widmann et al., 2003). Other methods, originally developed for Perfect-Prog applications, may also be implemented for a MOS approach. Arguably, the most common of the methods described in section 2.4.1, particularly when downscaling monthly precipitation means, is linear regression (e.g. Kilsby et al., 1998; Hellstrom et al., 2001; Goodess and Jones, 2002). Regression methods are especially useful when the overall goal is to derive a predictor-predictand relationship that can be easily implemented to estimate a future value of the predictand. Regression methods are also easily applicable to both MOS and Perfect-Prog downscaling approaches, which permits a comparison of each approach and of different predictor variables. The different options available for MOS and Perfect-Prog downscaling are discussed in greater detail in Chapter 4, with reference to the particular downscaling methods implemented in this research.

2.7 Summary

2.7.1 Modelling and uncertainty

A review of existing literature has shown the downscaling of GCM output to be an important and wide-ranging aspect of climate change research. Ultimately, it is at local-scales that impacts are felt most by human beings and the environment and improving upon projections of future conditions at these scales is crucial. Precipitation is particularly challenging to estimate given its greater variability compared to

many other climate variables. GCMs produce output at coarse resolutions and are currently unable to resolve the processes responsible for precipitation formation and variability. Until GCMs are developed sufficiently to account for this issue, which is likely to be decades away, producing good estimates of regional precipitation remains a major challenge.

There is still great uncertainty about anticipated changes in global precipitation patterns as highlighted by the most recent IPCC report. Downscaling tools have been developed to reduce this uncertainty and improve projections in specific regions. The merits of both dynamical and statistical downscaling methods have been discussed, with particular attention given to statistical methods. Although downscaling work is shown to be a vital part of making realistic precipitation estimates, it could be argued that the downscaling process is responsible for quantifying additional uncertainty. For instance, across all the downscaling work reviewed there seems to be little agreement as to which techniques are the most reliable and robust. Similarly, the choice of the most appropriate predictor variable(s) is contested throughout the literature.

It is important to note at this stage that the downscaling models developed in the following chapters can all be considered deterministic. As such, random variability, or 'noise', in the expected predictand that cannot be explained by the predictor(s) is disregarded. Uncertainty due to natural variability is a fundamental limitation of predictability and being able to quantify this uncertainty is an important challenge in producing probabilistic climate projections

2.7.2 Challenges and open questions

There is potential for using GCM-simulated precipitation as a predictor variable as part of a MOS approach (e.g. Widmann et al., 2003). Precipitation from atmospheric reanalyses, which is derived in a similar way to precipitation in a GCM, has been shown to have some skill in estimating local precipitation. GCM precipitation generally has a poor reputation and its use as a predictor variable is rare in the downscaling literature. But as it is parameterised directly from circulation variables simulated by the GCM, the simulated precipitation field inherently includes a number of relevant predictors. The derivation of GCM precipitation can be considered conceptually similar to that of precipitation in an atmospheric reanalysis (this is discussed further in Chapter 3). The use of reanalysed precipitation in the literature has been summarised with some success noted in the potential of its predictive power (e.g. Widmann and Bretherton, 2000).

The most prominent issue with directly including simulated precipitation in a downscaling model is that the day-to-day sequence of weather in a GCM simulation does not match that of the observed record; GCM and observed precipitation are thus uncorrelated. A possible solution to this problem is to nudge a GCM run towards some form of historical observation or reanalysis with respect to large-scale weather

states. There are few examples in the literature of the application of nudging techniques to GCMs but there exists enough evidence to support such a methodology.

Whilst much downscaling work has been previously undertaken, especially with respect to precipitation, it is the approach and methodology of the current research that is unprecedented. The MOS approach has been successful in weather forecasting but has received little attention in downscaling research. This study represents the first attempt to develop MOS models for estimating precipitation changes from the latest generation of climate change simulations.

Chapter 3

Skill and simple correction of ECHAM5-simulated precipitation

3.1 Introduction

In analysing the skill of model-simulated precipitation over an historical period, it is possible to make direct comparisons with real world observations. Groisman et al. (2005) have shown that the multi-model mean global precipitation trend patterns during the twentieth century do not agree very well with observed trends. It should be noted however that the differences might be partly related to random atmospheric variability rather than to problems in the models. When making inferences about the simulation of future climates, for which observational data are obviously not available, inter-model consensus is an accepted indicator of skill. Multi-model mean trends for the twenty-first century were calculated in the IPCC AR4, accompanied by an analysis of the areas over which the models agree with each other with respect to the sign of the change (Randall et al., 2007). Again, it is partly unclear which of these differences are due to genuine differences in the model's response to greenhouse gas forcing, and which are due to random, unpredictable differences in atmospheric variability. Detection of anthropogenic influence on global precipitation is made troublesome by the averaging of varying regional response, which may serve to weaken the overall global signal (Allen and Ingram, 2002; Held and Soden, 2006) and successful quantification of anthropogenic forcing must usually consider changes in particular regions or latitudinal bands (e.g. Zhang et al., 2007). Moreover, it is problematic to associate areas of good model agreement with a high level of confidence in predictions because simulations in such areas may still be wrong for a common reason.

Uncertainty about the skill of the simulated temporal variability means that, in estimating small-scale

precipitation changes, simulated precipitation is often disregarded in favour of the alternative statistical downscaling approaches. However, the response of hydrological processes, including moisture transport and evaporation in addition to precipitation, to large-scale warming is robust in many of the GCM simulations included in the IPCC AR4 (Held and Soden, 2006). One may conclude that, conceptually at least, the simulated precipitation captures the climate change signal in temperature and other dynamically-resolved variables, and to discard this field is to possibly exclude critical information describing future precipitation changes.

It is the purpose of this chapter to, for the first time, establish a quantification of the skill of a GCM in simulating temporal precipitation variability and to assess the potential for the application of a simple statistical downscaling correction. A hindcast (1958-2001) simulation of the ECHAM5 GCM is subjected to a nudging technique, used to force the prognostic circulation and temperature fields to corresponding values from ERA-40. Consequently, the temporal variability of the large-scale atmospheric state is captured by the simulation, and subsequent derivation of precipitation is based upon a realistic circulation. The skill of the simulated precipitation, which of course is likely to vary geographically, is then quantified on a seasonal basis. Focus is also given to the implementation of a MOS downscaling approach. The second part of the analysis evaluates the potential for a simple scaling correction of the simulated precipitation field to produce skilful estimates of regional precipitation.

The remainder of the chapter is structured as follows. In section 3.2, a distinction is made between three sources of model error and it is demonstrated that, by isolating the error relating to parameterisation deficiencies, it is possible to quantify GCM skill in reproducing temporal variability of precipitation. Section 3.3 describes the formulation of the nudged simulation and the observational precipitation data sets used for validation purposes. The skill of ECHAM5 precipitation is assessed and discussed in section 3.4 and the potential for a downscaling correction is evaluated in section 3.5. A summary of the main conclusions is given in section 3.6.

3.2 Quantification of model skill

3.2.1 Three sources of error

A strict assessment of model skill requires the distinction between three sources of errors in simulated precipitation (Figure 3.1). If the simulated large-scale atmospheric conditions differ from reality the simulated precipitation will be different from observations even if the convection and precipitation parameterisations in the model are perfect. Two causes are identified for this difference. Firstly, a particular model may have a partly unrealistic mean state and/or large-scale response to climate forcings, which here is termed a 'type 1' error. Secondly, internally-generated variability is unpredictable and will be

different from the real world. This is termed a 'type 2' error but it is noted that this 'mismatch' between simulated and observed variability is not strictly a model deficiency but a consequence of a freely evolving GCM. Standard GCM simulations for historical periods assess the climatic response to forcings such as solar variability and changes in atmospheric concentrations of greenhouse gases, and of anthropogenic and volcanic aerosols. These simulations are not constrained by the historic meteorological observations and thus, due to the chaotic nature of the climate system, the circulation and temperature fields differ from reality with respect to the random, internally-generated component of variability. This random component dominates daily to interannual time scales and is still a major component in decadal time scales.

While the type 2 error can be ameliorated to an extent by temporal averaging, and is usually taken into account by initial value ensemble simulations, understanding the causes of and quantifying the type 1 error is highly important and at the core of many model validation studies (Jansen et al., 2007; Randall et al., 2007). However, it is also necessary to consider a further source of error, which here is termed a 'type 3' error, relating to the deficiencies in precipitation parameterisation as well as differences between the real orography and the 'model world', which will lead to precipitation errors even if the large-scale atmospheric conditions are in agreement with reality. When the task is to assess the skill of simulated precipitation variability, it is useful to focus on the type 3 error and to ask the question 'How well is precipitation simulated for given large-scale conditions?'. Conceptually, this is consistent with the validation and application of numerical and statistical downscaling methods, which also assume that the large-scale atmospheric states simulated by a GCM can be considered an adequate sample of the true distribution, and thus correct. More precisely, the large-scale states are on what the estimated regional precipitation is conditioned.

It is possible to evaluate the skill of simulated precipitation associated with particular synoptic situations for a given region (e.g. Osborn et al., 1999) but a global quantification of skill is troublesome to achieve. While until now it has not been possible to isolate the type 3 error for GCMs used for climate change simulations, Widmann and Bretherton (2000) have show that this error can be quantified for the GCMs that are used in atmospheric reanalyses (Kalnay et al., 1996; Kistler et al., 2001; Uppala et al., 2005). This approach is possible because due to the assimilation of meteorological measurements such as pressure, wind speeds, temperature and humidity, the large-scale states in a reanalysis are the best estimates for the state of the real atmosphere. They are consistent with both the assimilated observations and the physical laws that govern the atmosphere, as represented in the model. However, no precipitation observations are assimilated and precipitation is simulated using the parameterisations and the large-scale atmospheric states derived in the assimilation process (Fig. 3.1b). Several studies have shown that reanalyses capture observed temporal precipitation variability well (Gutowski et al., 1997;

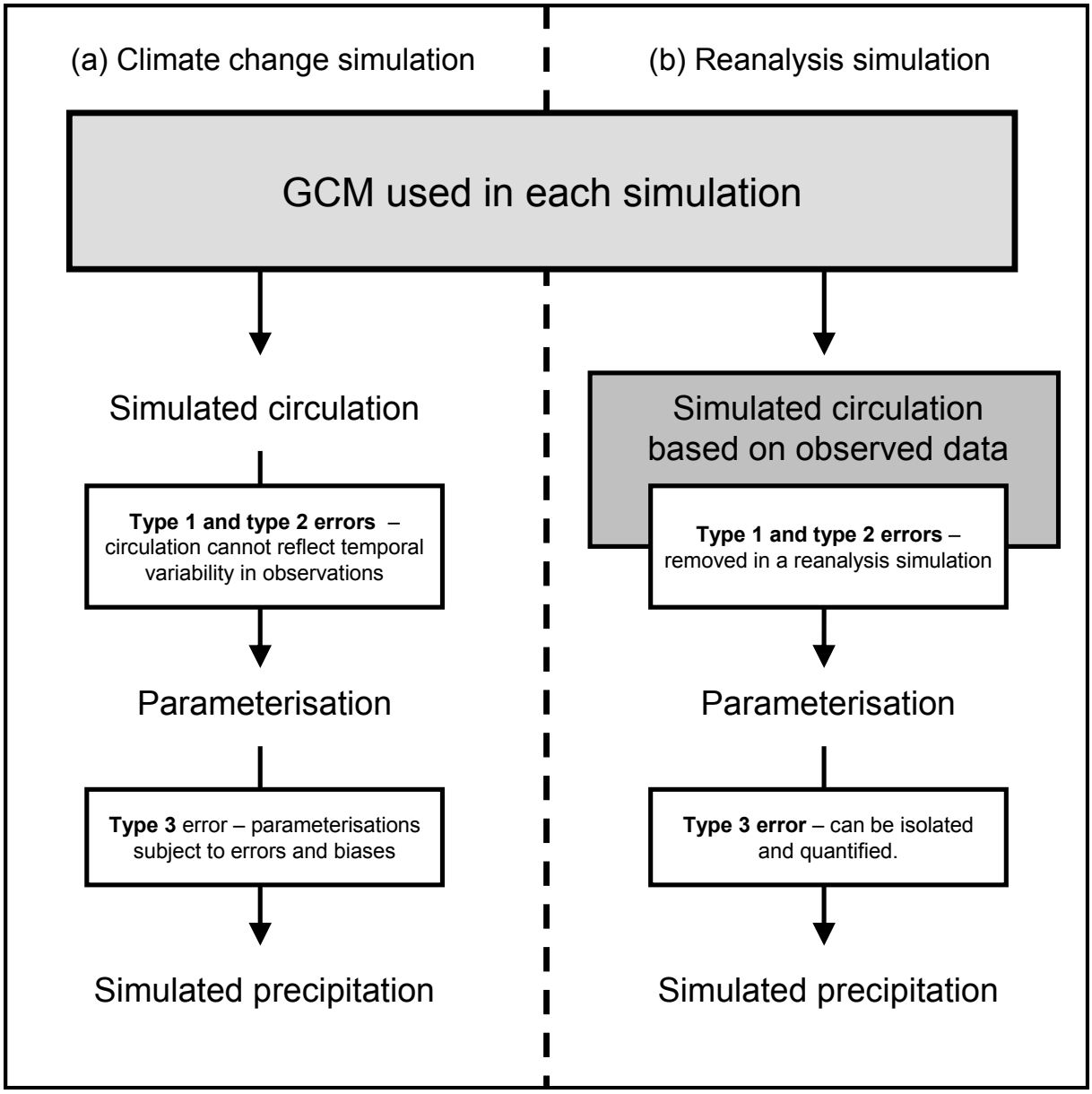

Figure 3.1: Climate change (a) and reanalysis (b) GCM simulations of precipitation. Large-scale circulation fields simulated in the reanalysis are forced to real world observations. The 'type 3' error relating to parameterisation error still exists but can now be fully evaluated.

Janowiak et al., 1998; Widmann and Bretherton, 2000; Widmann et al., 2003; Bosilovich et al., 2008). Whilst regional differences are apparent in many parts of the world, reanalysed precipitation is at least able to account for the interannual variability seen in observations as its derivation comes directly from large-scale variables which are based upon a real world historical record (Gutowski et al., 1997; Widmann and Bretherton, 2000). The atmospheric reanalysis may therefore be considered an 'ideal' GCM, in which the large-scale circulation is in good agreement with reality (Widmann and Bretherton, 2000). Although the GCMs used for reanalyses are similar to those used for climate simulations, they differ in terms of resolution and parameterisations, and thus the reanalysis-based results can not be directly transferred to

other GCMs.

The present study extends the reanalysis-based studies and quantifies how well precipitation variability is simulated in a GCM (ECHAM5) used for climate prediction in the IPCC AR4 (Randall et al., 2007; Trenberth et al., 2007) if the error related to large-scale atmospheric states is approximately removed. Here, a simulation of ECHAM5 is conducted in which the prognostic variables are forced toward reanalysis values for an historical period. The subsequently parameterised precipitation field is thus expected to represent observed temporal variability, at least for regions where reanalysis fields are skilful.

3.2.2 Skill of reanalysis precipitation

The validity of precipitation from reanalyses has been evaluated in several studies, although these did not focus on downscaling. A series of papers followed the establishment of the NCEP-NCAR reanalysis (Kalnay et al., 1996). Mo and Higgins (1996) evaluated the usefulness of reanalysis for hydrological research finding that, in general, large-scale precipitation features are in good agreement with an alternative reanalysis (NASA Data Assimilation Office reanalysis) and satellite observations but that there are significant regional differences. In a similar study, Higgins et al. (1996) again compared the NCEP and NASA reanalysis, this time over central United States during May for the period 1985-89. In general, the NCEP reanalysis exhibited a higher correlation with observed precipitation than the NASA equivalent, but the precipitation fields in each analysis are more closely related to one another than to the observations.

Gutowski et al. (1997) assessed atmospheric water transport in the NCEP reanalysis by comparing convergence in the Upper Mississippi and Ohio-Tennessee basins with observed river discharge. The authors's premise is that, over several years, atmospheric water input should roughly equate to output in streamflow but the study finds a difference of around 40%. The work does conclude that the temporal variability of the reanalysis is in good agreement with observations.

Janowiak et al. (1998) compared precipitation fields from the NCEP reanalysis and the Global Precipitation Climatology Project (GPCP) for the period 1988-95. The GPCP has monthly estimates of precipitation that incorporates both gauge observations and estimates from satellite measurements. As with Mo and Higgins (1996), large-scale features were well resolved by the reanalysis but there are significant regional differences. Poor agreement between the two datasets was generally observed over oceans, equatorial land regions and the Pacific ITCZ. Wu and Xie (2003) reiterated the need to exercise caution when using NCEP-NCAR reanalysis output for the tropical Pacific.

Other studies have focused on specific areas. Poccard et al. (2000) evaluated the accuracy of rainfall data from the NCEP-NCAR reanalysis compared to observations over tropical Africa between 1958 and 1997. The reanalysis was shown to generally underestimate amounts during the rainy season and also

correlated poorly with observations in terms of interannual variability. However, the reanalysis was able to replicate fairly well the main teleconnections between ENSO and African rainfall variations. Reid et al. (2001) compared NCEP-NCAR reanalysis output with observations in Central and Eastern England and Italy. Precipitation is not as well simulated by the reanalysis as other variables, particularly in regions and periods that are dominated by convective precipitation. Enough examples exist in the literature to suggest that there are stronger relationships between reanalysed and observed precipitation during winter than in summer (e.g. Mo et al., 2005).

A great benefit of a reanalysis is completeness of its spatial and temporal coverage (Reid et al., 2001). However, the reanalysis output is usually in a gridded, uniform format which cause difficulty when attempting to compare with point-scale, station observations. Some attempts have been made to compensate for this. Widmann and Bretherton (2000) produced a 50-year precipitation dataset on a 50 x 50 km grid for the north-west United States. The grid values were interpolated from station data and, because of the varying physical relief of the region, corrected for slope aspect and elevation. This dataset was compared with precipitation from the NCEP-NCAR reanalysis with the overall aim of understanding how useful precipitation from such a model is in estimating local temporal variability. On scales of about 500 km, long-term variability was well-captured by the reanalysis, although a poor representation of topography meant that there were systematic biases at the individual grid cell scale. A similar approach was undertaken by Maurer et al. (2002) to derive a dataset of land surface states and fluxes. Widmann and Bretherton (2000) acknowledged that the results are region- and model-specific given that previous work has already established that correlation between reanalysed and observed precipitation is regionally dependent. The good skill of the reanalysis in this case is partly attributed to the inclusion of assimilated humidity values in the calculation of precipitation. The authors suggested that GCMs with similar precipitation parameterisations may also produce decent results in other regions where good links between reanalysed and observed precipitation have been found.

The limitations of reanalysed precipitation should still be noted. Bosilovich et al. (2008) evaluated the precipitation fields of five reanalyses, including the NCEP-NCAR and ERA-40 systems, and showed each to be in reasonable agreement with satellite-rain gauge observations over large areas. However, all reanalyses showed a tendency to over-estimate precipitation in tropical regions. In general, the literature suggests that the accuracy of reanalysed precipitation varies greatly across the world. Whilst other reanalysed variables show greater agreement with observations, precipitation is often shown to have substantial biases, or in some cases is subject to a pre-defined correction (e.g. Sheffield et al., 2004).

3.3 Methods of validation and analysis

All hindcast simulations are performed using the ECHAM5 GCM (Roeckner et al., 2006). ECHAM5 forms the atmosphere component of the ECHAM5/MPI-OM atmosphere-ocean coupled GCM, simulations from which contributed to the IPCC AR4. ECHAM5 is run with a time-dependent post-industrial radiation scheme, which includes prescribed quantities of atmospheric constituents. Long-term changes in water vapour, cloud water content and cloud cover are resolved by the model. The simulations performed here are relatively short thus do not involve the ocean component (MPI-OM), and are instead forced at the atmosphere-ocean boundary by monthly sea-surface temperature fields from ERA-40. The nudging procedure is wholly atmospheric and does not include ocean fields.

3.3.1 Performing a nudged simulation

A nudging technique (also known as Newtonian relaxation) (Krishnamurti et al., 1991; Jeuken et al., 1996; Timmreck et al., 1999; Timmreck and Schulz, 2004) is used to force the ECHAM5 simulated divergence, vorticity, temperature and surface pressure fields to corresponding fields from ERA-40 at all atmospheric levels. The simulation is conducted between September 1957 and August 2002, the same period for which ERA-40 data is available. In all subsequent analysis, only model output for full calendar years is considered (1958-2001). Model output is on a T63 Gaussian grid, which equates to $1.875° \times 1.875°$ or roughly 200km latitude $\times$ 150km longitude at 45°N.

The nudging procedure does not simply replace these prognostic variables with the reanalysis values, but merely guides the GCM towards reality without substantially compromising model physics and dynamics. Early examples of nudging in the literature are usually concerned with NWP (e.g. Lyne et al., 1982; Ramamurthy and Carr, 1987) although there are some studies that have applied such methods to GCMs (e.g. Murphy, 1999; Timmreck et al., 1999; Timmreck and Schulz, 2004). This method has been used successfully in previous studies to force the circulation in earlier ECHAM versions towards a reanalysis for short periods (Jeuken et al., 1996; Timmreck et al., 1999; Timmreck and Schulz, 2004), but this is the first time it has been applied to ECHAM5 and for the entire reanalysis period.

An additional term, $N(X_t) \cdot (X^{obs} - X)$, is added to the tendency equations for each variable, which take the form,

$$\frac{\partial X}{\partial t} = F_m(X_t) + N(X_t) \cdot (X^{obs} - X) \tag{3.1}$$

as described in Krishnamurti et al. (1991). Here F_m represents the model tendency (the dynamical and physical processes that determine the temporal evolution of X) (Wilks, 2006). N is the relaxation coefficient and X^{obs} represents a observed value to which the nudging is aimed. In practice, the full

integration is a two-step process, with the normal tendencies first carried out,

$$X^*_{t+\triangle t} = X_t + F_m(X_t) \cdot \triangle t \tag{3.2}$$

where $X^*_{t+\triangle t}$ represents a pre-nudging predicted value of X at time $t + \triangle t$. The nudging procedure is applied in a second step,

$$X_{t+\triangle t} = X^*_{t+\triangle t} + N(X^{obs}_{t+\triangle t} - X^*_{t+\triangle t}) \cdot \triangle t \tag{3.3}$$

resulting in a new value, $X_{t+\triangle t}$. The relaxation e-folding time, τ, is equivalent to $1/N$. The choice of τ is variable specific and based on the results of previous work (Jeuken et al., 1996; Kaas et al., 2000) (Table 3.1). In each case, τ should be of sufficient length that the observed fields exert the required influence but not so much that the relaxation term is dominant over the model tendencies (Kaas et al., 2000).

Table 3.1: Relaxation coefficients used in the nudging procedure. Based on simple nudging assimilations used at the Danish Meteorological Institute (DMI) (Kaas et al., 2000).

Prognostic variable	τ (hours)	$N(s^{-1})$
Vorticity	6.00	0.0000463
Temperature	24.00	0.0000116
log(surface pressure)	24.00	0.0000116
Divergence	48.00	0.0000058

The nudging process requires a number of important considerations to be made. Deciding upon a suitable value of the relaxation coefficient is a crucial step in the nudging procedure (Stauffer and Seaman, 1990; Jeuken et al., 1996). If too large a relaxation coefficient is used, the relaxation itself could compromise the model physics. Conversely, a small value will make little difference to the simulated variable. Jeuken et al. (1996) varied the relaxation coefficient in a number of experiments in order to find the most suitable value. These results have been adopted by subsequent works (e.g. Timmreck et al., 1999). In most applications, the relaxation coefficient for a given variable is constant in space and time. Hoke and Anthes (1976) suggested, however, that the value of the relaxation coefficient should reflect the accuracy of the observations. As the confidence in reanalysed data may vary on both temporal and spatial scales, it is perhaps conceivable that a relaxation coefficient is flexible enough to account for this.

As pointed out by Jeuken et al. (1996), there are also different approaches in the choice of the simulated variables to be adjusted. Timmreck et al. (1999) use a nudging technique in their study of Mt. Pinatubo volcanic cloud with the ECHAM4 GCM. Prognostic variables, including temperature, vorticity and divergence, are relaxed towards observations in order to simulate the true circulation. Timmreck and

Schulz (2004) applied relaxation coefficients to temperature, vorticity, divergence and the logarithm of surface pressure ranging from $1.16 \times 10^{-5} s^{-1}$ to $9.29 \times 10^{-4} s^{-1}$. In most cases, the chosen coefficient value is based upon work from previous studies. There is scope for the optimal value of a relaxation coefficient to be derived mathematically (Zou et al., 1992). However, whilst this is feasible for regional-scale studies (e.g. Genthon et al., 2002), the computing power required to make such derivations for a global model makes this additional step impractical.

It is important to identify an additional source of error that may be induced by the nudging procedure. In a standard (non-nudged) simulation, the balance between the dynamically-resolved and parameterised fields is physically consistent and the addition of a nudging term in (3.1) inhibits the parameterised quantities to reach this balanced state (Jeuken et al., 1996). An implication is a modification of diabatic heating in the GCM and a possible violation of energy conservation which in turn may result in spurious precipitation. Jeuken et al. (1996) state that such errors will be greatest where the nudging term is large relative to the other terms in each tendency equation and that the relaxation coefficient chosen for temperature must be small enough to prevent a distortion in the diabatic heating quantities. This nudging-induced error, which is termed 'type 4' in accordance with the definitions used thus far, contributes along with parameterisation deficiencies (type 3 error) to form the total precipitation error. Whilst the set of coefficients in (Table 3.1) have been previously shown to be sufficient in suppressing the dominance of the nudging term in the model tendency equations, the magnitude of a potential type 4 error is expected to be associated with considerable geographical variability.

3.3.2 Observational datasets

The output from the nudged simulation is compared to that of a standard (normal; non-nudged) free-running ECHAM5 simulation at the same resolution for the period 1958-2001. All analysis is conducted using seasonal mean precipitation which is consistent with other climate change projections and also allows for a broad understanding of the seasonal dependence of model skill and thus the potential of a proposed downscaling correction. Seasonal means from each simulation are compared to observations from the Global Precipitation Climatology Project (GPCP) dataset (Huffman et al., 1997; Adler et al., 2003). Version 2 of this dataset, described by Adler et al. (2003), provides gridded ($2.5^\circ \times 2.5^\circ$) monthly means based on satellite and rain gauge observations for the period 1979-2001. Consideration was also given to a similar merged dataset, the Climate Prediction Center Merged Analysis of Precipitation (CMAP) (Xie and Arkin, 1997) that also provides monthly precipitation at the same spatial resolution. Gruber et al. (2000) compared CMAP and a previous version of the GPCP dataset (Huffman et al., 1997), generally finding good agreement between the datasets but also some acute differences. CMAP was shown to treat rain gauge measurements differently to the GPCP dataset, such as in the inclusion of atoll gauge data

to calibrate oceanic satellite estimates. The GPCP dataset has been used in other GCM and reanalysis validation work (e.g. Janowiak et al., 1998; Trenberth and Shea, 2005) and is considered most suitable for this analysis.

For the development of a high-resolution scaling correction, a comparison is made with the Global Precipitation Climatology Center (GPCC) dataset (Rudolf et al., 1994; Beck et al., 2005; Rudolf and Schneider, 2005; Schneider et al., 2008), which is based only on interpolated rain gauge observations and covers only land areas but with a higher resolution of $0.5° \times 0.5°$ and for the period 1958-2001. Precipitation from the nudged simulation was linearly interpolated to the same $0.5° \times 0.5°$ grid in which the GPCC data is available.

3.4 Validation of GCM precipitation

3.4.1 Representation of the large-scale circulation in the nudged simulation

The nudged ECHAM5 simulation was evaluated alongside equivalent output from the free-running (non-nudged) simulation for the period 1958-2001. It was first of all necessary to compare circulation and temperature fields from the nudged simulation with observations so as to ensure confidence in the nudging procedure. Figure 3.2 shows global correlation coefficients between annual and monthly mean values of example circulation variables from ERA-40 and corresponding fields from the standard and nudged ECHAM5 simulations. As expected, standard simulated geopotential height and temperature at the 850-hPa and 500-hPa levels show no significant relationship with ERA-40, with correlation coefficients fluctuating around zero in all case (Figure 3.2a,c,e,g). The strongest correlation is +/-0.3 which reflects that the ability of the normal simulation to reproduce interannual variability on a month-by-month (or indeed, day-by-day) basis is non-existent.

In contrast, correlation between large-scale fields in the nudged simulation and ERA-40 are strong (Figure 3.2b,d,f,h). Correlation coefficients are consistently greater than 0.9 across the majority of the globe. Tropical regions exhibit weaker correlations than extra-tropical regions. Geopotential height at 850-hPa shows noticably weak correlation with ERA-40 across much of northern South America (Figure 3.2b). This may be due to a number of factors, including the poor representation of real-world upper air observations over this region in ERA-40, and the simulation of the intense convective processes that prevail over the Amazon basin for much of the year. Correlations for geopotential height at 500-hPa are not noticeably different in the tropics, and this northern South America discrepancy is not apparent (Figure 3.2d).

The representation of 850-hPa geopotential height is investigated further by examining correlation coefficients over each month of the year (Figure 3.3). A marked tropical-extratropical difference is evident

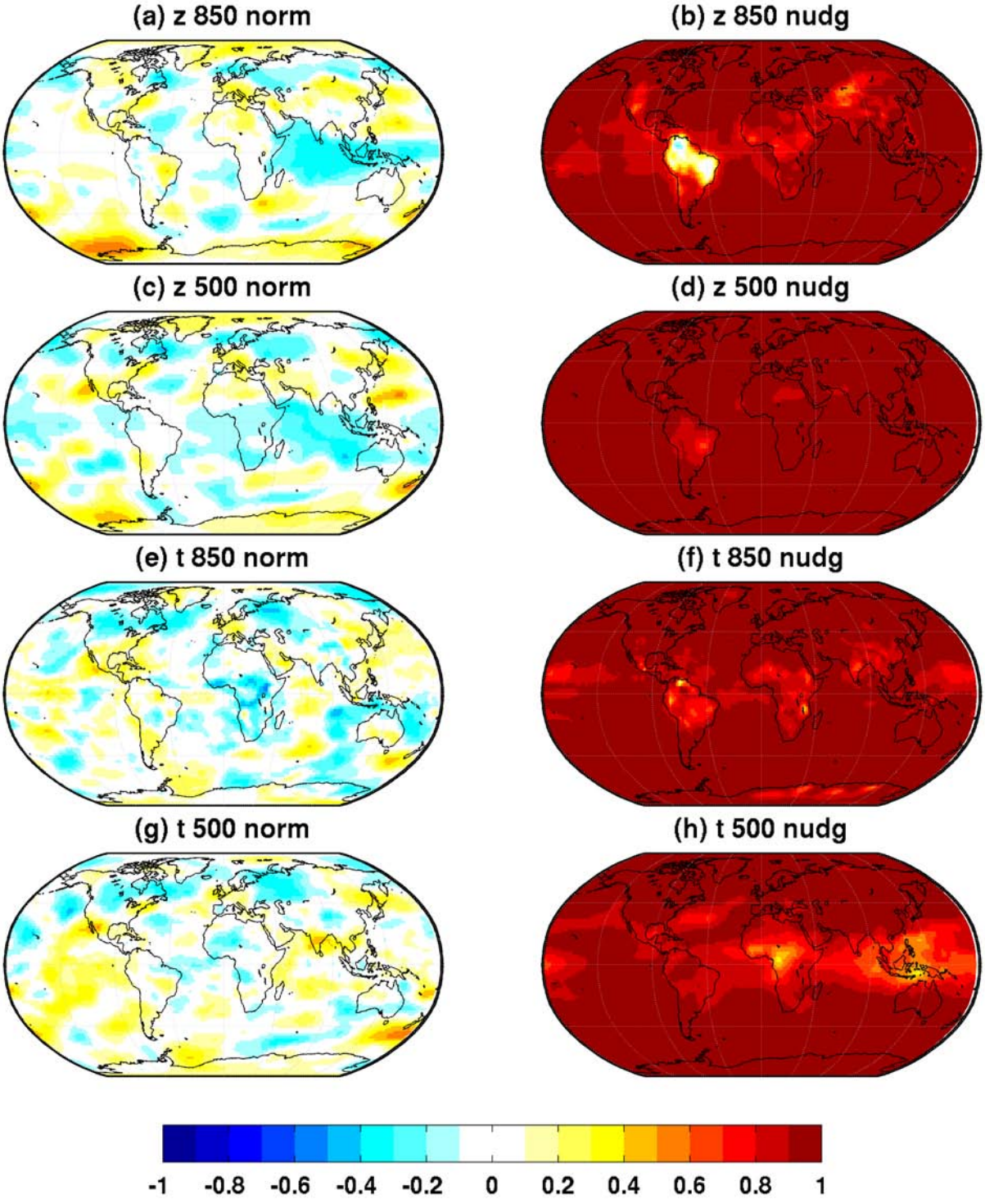

Figure 3.2: Correlations of annual ERA-40 temperature (t) and geopotential height (z) at 850-hPa and 500-hPa with corresponding fields from the standard ('norm'; left panels) and nudged ('nudg'; right panels) ECHAM5 simulations (1958-2001).

in all months, although is most pronounced between April and September when there is a clear band across the tropics in which correlations are consistently around 0.7 over the ocean. Whilst this tropical-extratropical difference is less obvious during the rest of the year, the northern tip of South America and parts of north-eastern Brazil show noticeably weak correlations (0.2-0.3) during all months. It is interesting that correlations of this magnitude do not exist in other tropical continental areas, which suggests that poor upper air representation of the northern South America region in ERA-40 is more likely the cause of poor reanalysis-GCM correlation. Monthly mean correlations of observed and simulated geopotential height at 500-hPa appear slightly weaker in the tropics between June and October (around 0.7) (Figure 3.4), but the tropical-extratropical difference is much smaller than at 850-hPa.

Simulated temperature is also more weakly correlated with ERA-40 in the tropics, with 500-hPa temperature showing a particular tropical-extratropical difference which is most pronounced in the eastern hemisphere (Figure 3.2f,h). Weak correlations ($<$0.3) of simulated 500-hPa temperature with ERA-40 are most prominent in central Africa in the later half of the year (July to November) and in South-East Asia from September to November (Figure 3.6). In contrast, simulated temperature at 850-hPa shows weakest correlation with ERA-40 over the Amazon basin during November and December (Figure 3.5). Root mean square errors (RMSEs) were calculated to quantify model skill in reproducing realistic temperature values. RMSEs of 850-hPa temperature are greater over land, where variation in surface temperature is greater than over the ocean (Figure 3.7). RMSEs are especially high over the Amazon basin between August and November, which is consistent with other measures of model skill over this region. Errors in 500-hPa temperature are not especially localised, but greater on a global scale between November and April (Figure 3.8).

3.4.2 Interannual precipitation variability

Simulated precipitation from both the nudged and standard (non-nudged) simulations is in reasonable agreement with the long-term observed climatology in terms of the spatial distribution of large-scale precipitation patterns. However, some key regional differences between the simulations are noted (Figure 3.9; right panels), particularly lower precipitation in the western Pacific in the nudged simulation compared to the standard simulation, which points towards the relevance of internal variability even for multi-decadal averages. Other parts of the tropics also exhibit large differences, particularly south-east Asia and the Amazon basin.

As the standard simulation is unable to represent observed interannual variability in seasonal mean precipitation the correlation coefficients fluctuate randomly around zero (Figure 3.10a). In contrast, seasonal mean precipitation from the nudged simulation exhibits in many areas high correlations with observations (Figure 3.10b-f), including the extra-tropics (30° to 60°), especially over large parts of

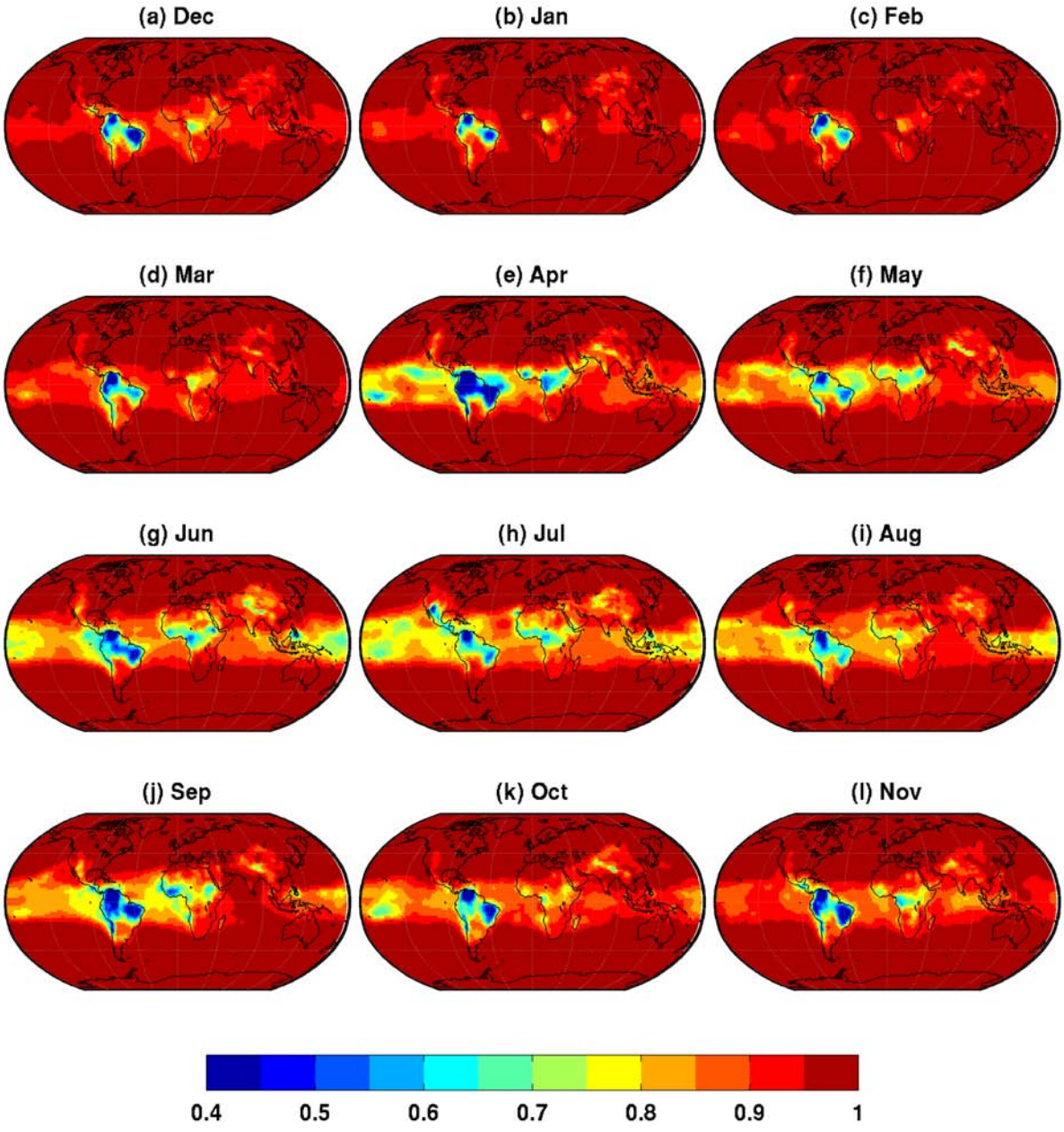

Figure 3.3: Correlations of monthly ERA-40 geopotential height at 850-hPa with corresponding fields from the nudged ECHAM5 simulation (1958-2001).

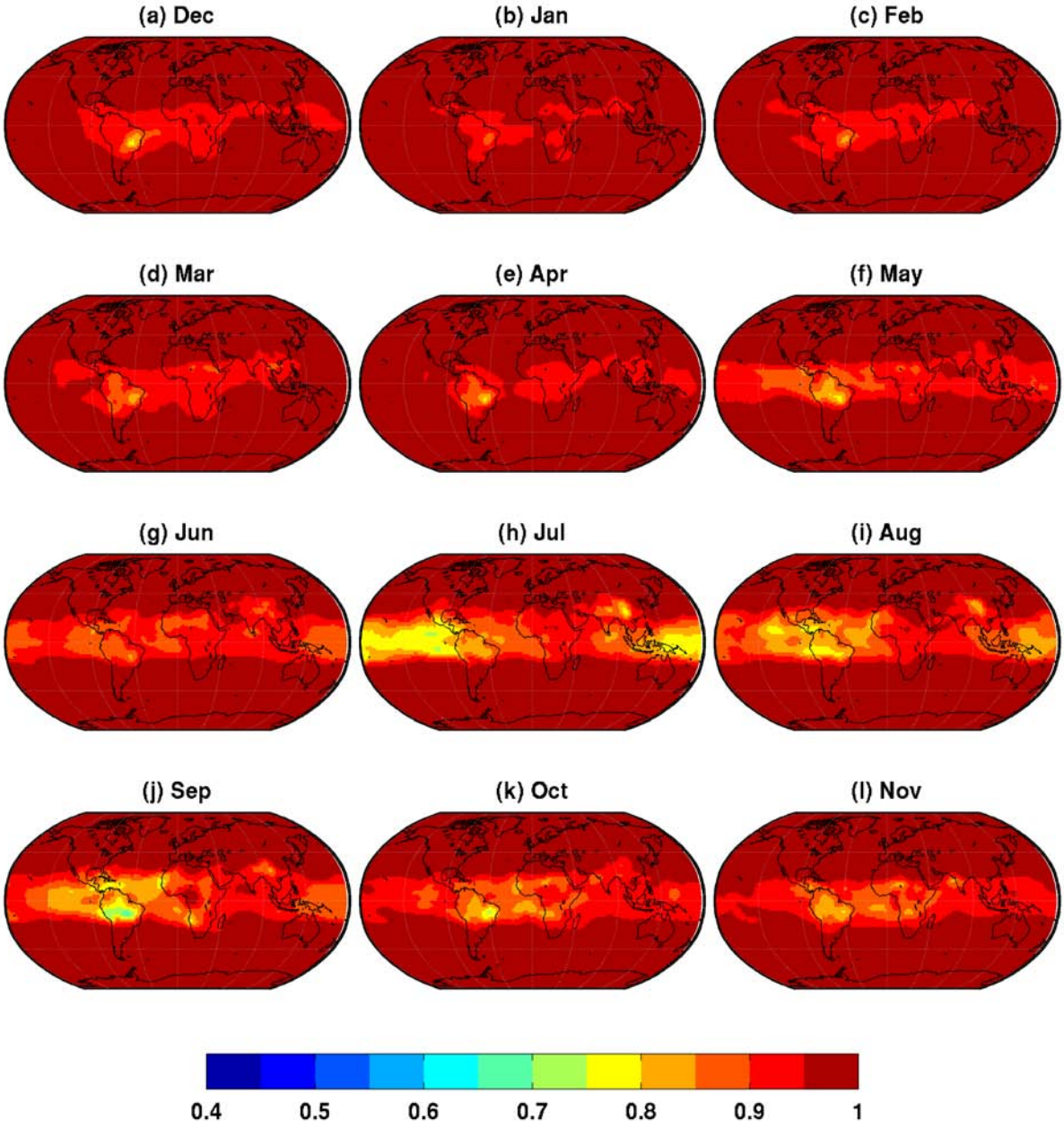

Figure 3.4: As Figure 3.3 but for at 500-hPa.

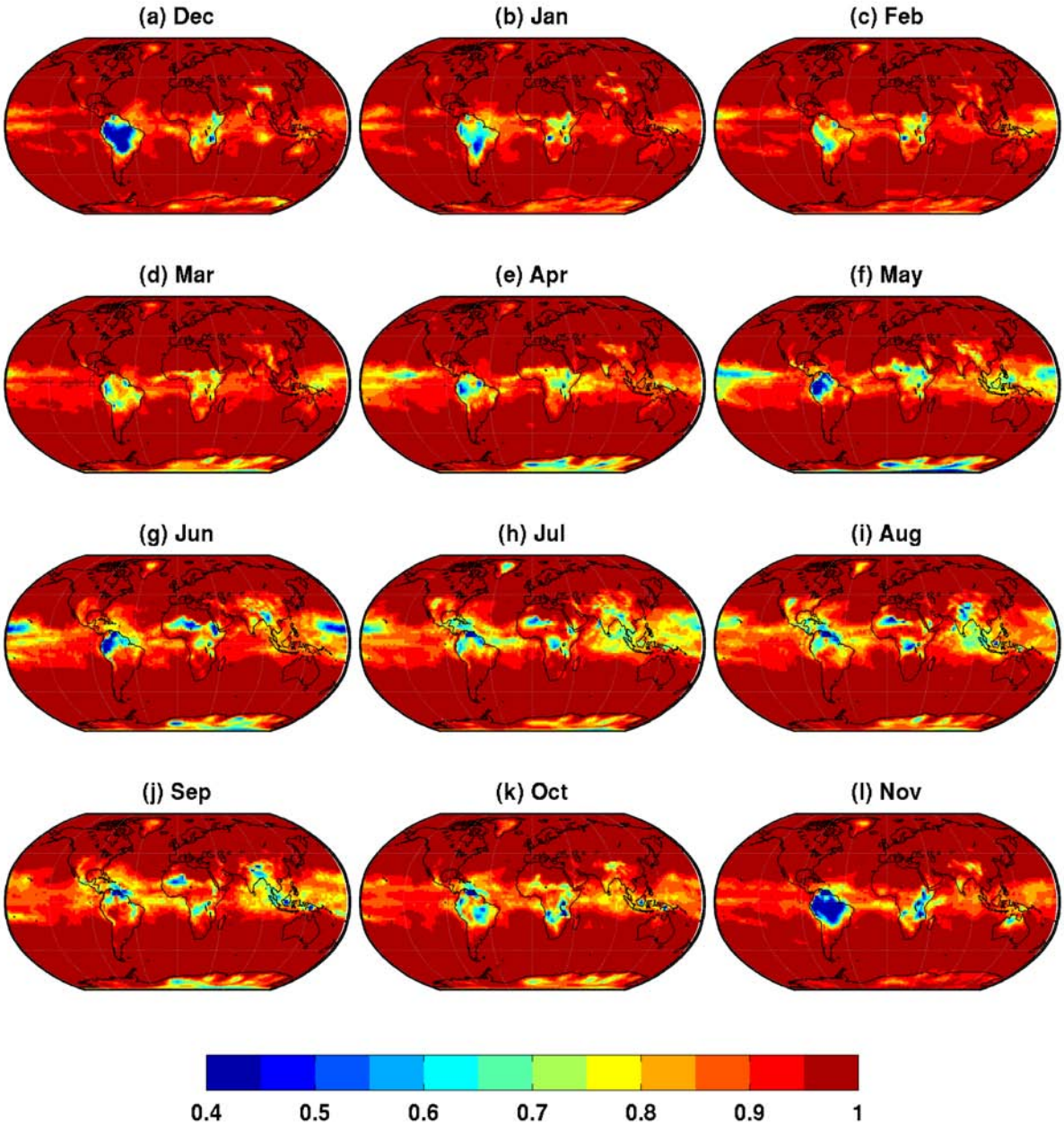

Figure 3.5: Correlations of monthly ERA-40 temperature at 850-hPa with corresponding fields from the nudged ECHAM5 simulation (1958-2001).

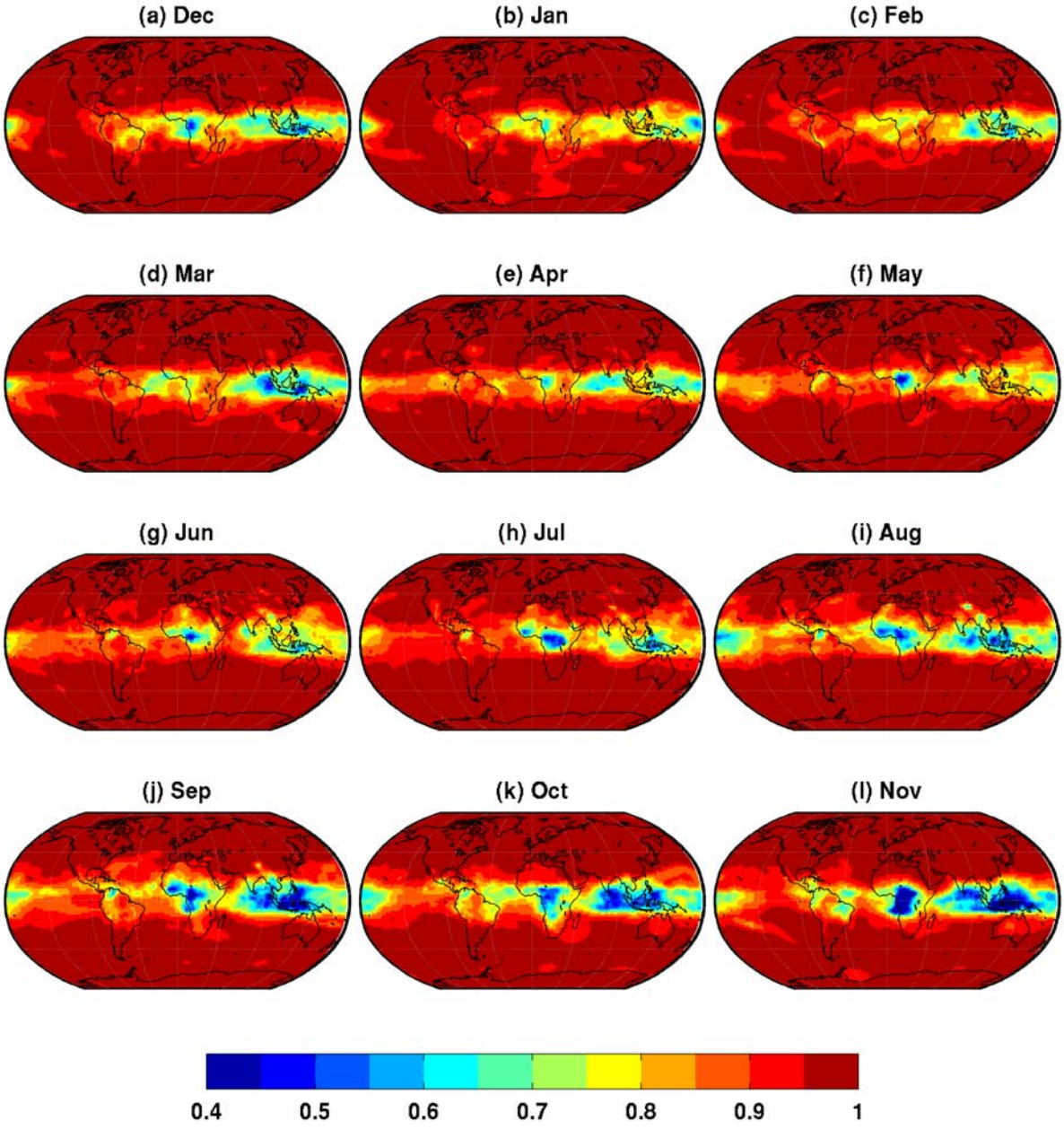

Figure 3.6: As Figure 3.5 but for at 500-hPa.

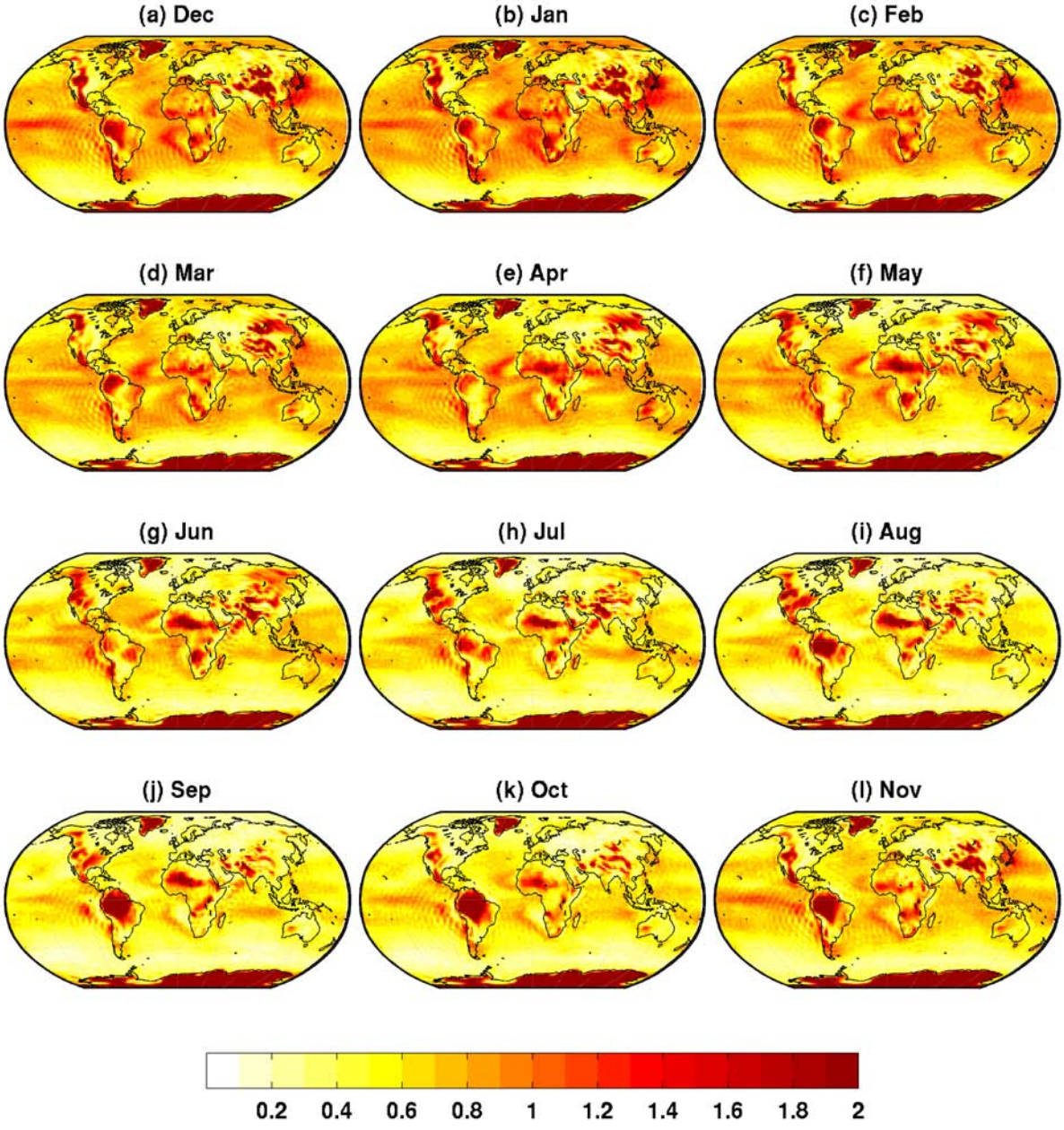

Figure 3.7: RMSE of monthly ERA-40 temperature (°C) at 850-hPa and corresponding fields from the nudged ECHAM5 simulation (1958-2001).

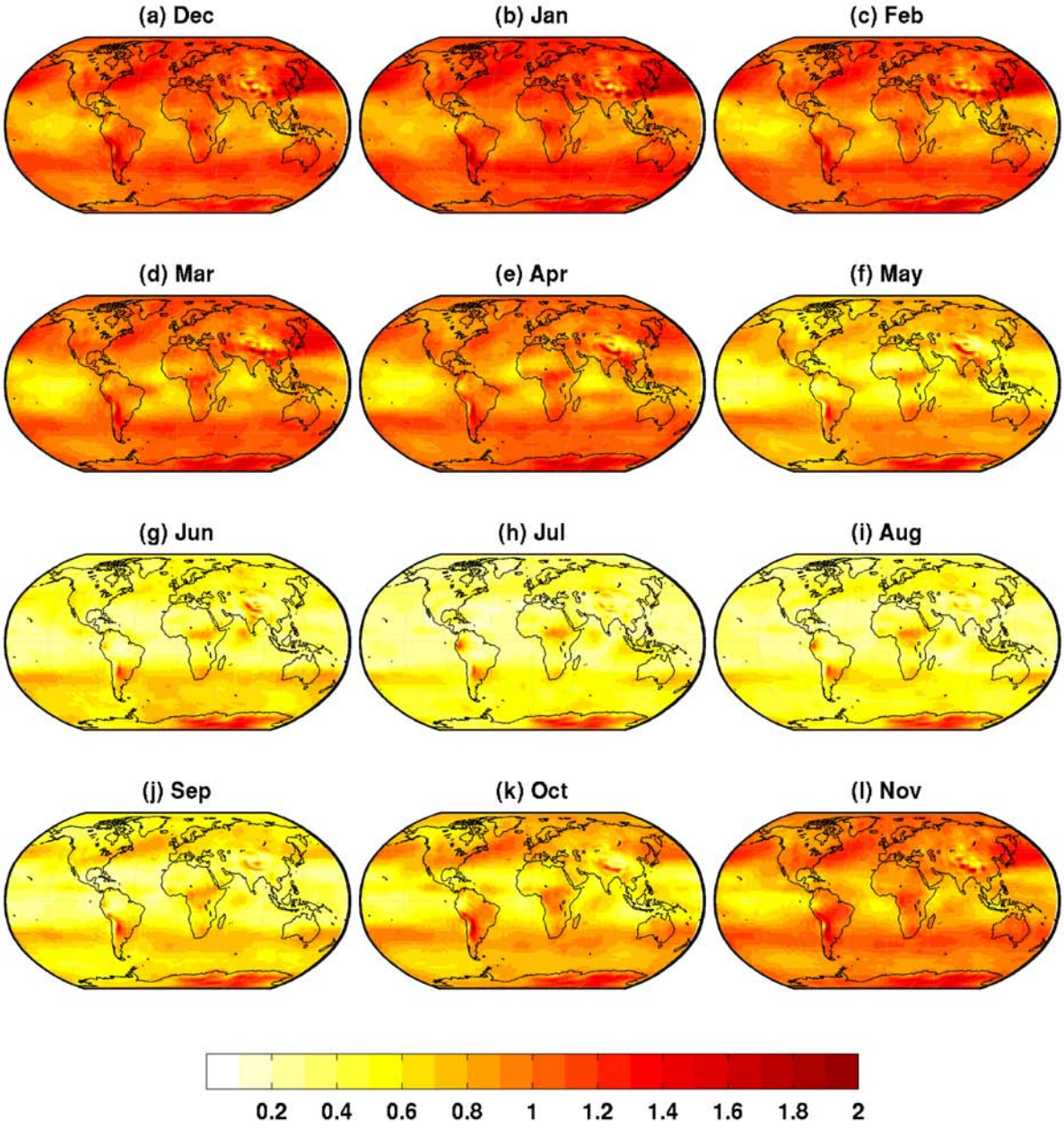

Figure 3.8: As Figure 3.7 but for at 500-hPa.

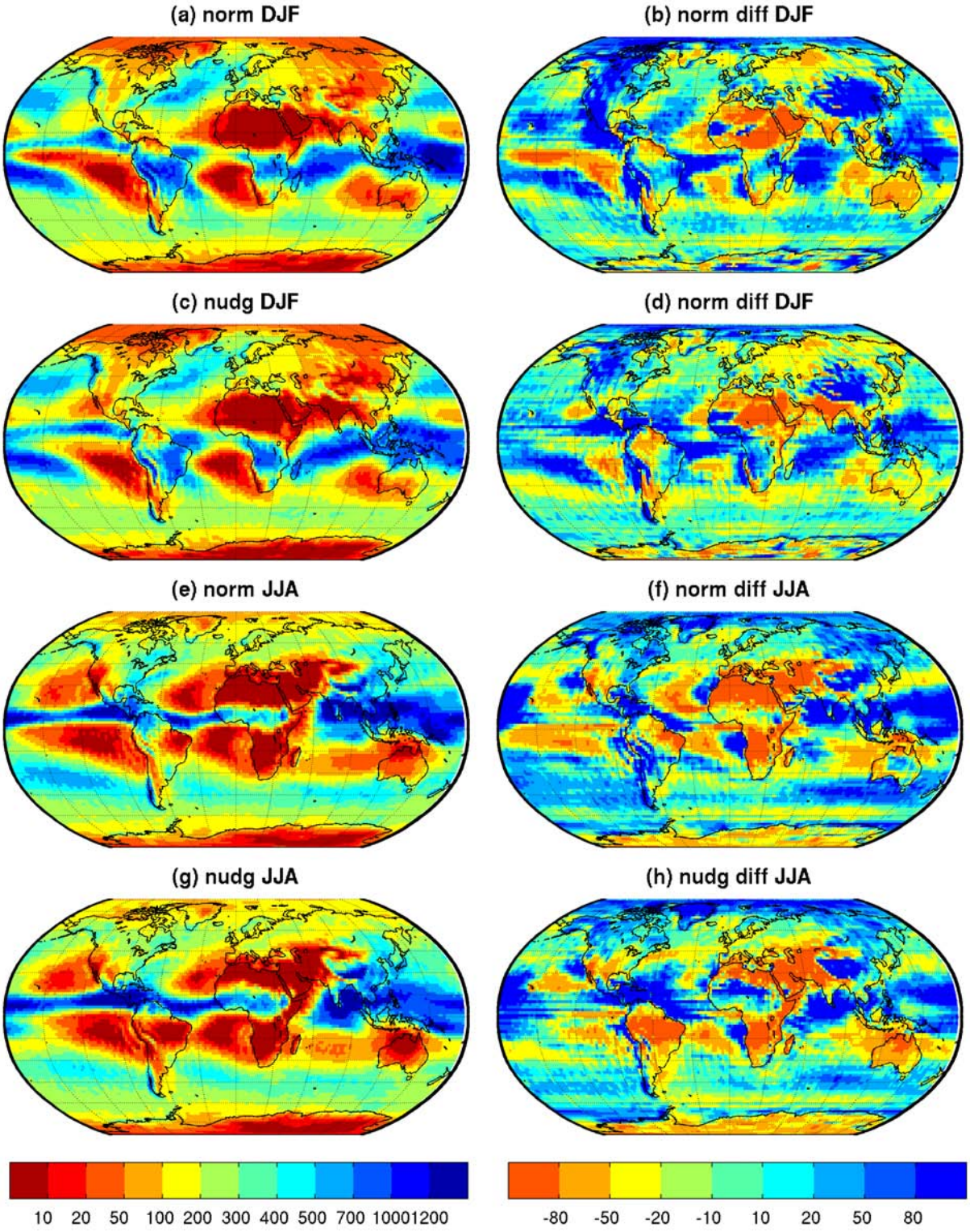

Figure 3.9: Seasonal precipitation (expressed as monthly means and in mm) from the standard ('norm') and nudged ('nudg') ECHAM5 simulations (a,c,e,g; left colour bar) and the respective percentage deviation from GPCP observations (1979-2001) (b,d,f,h; right colour bar).

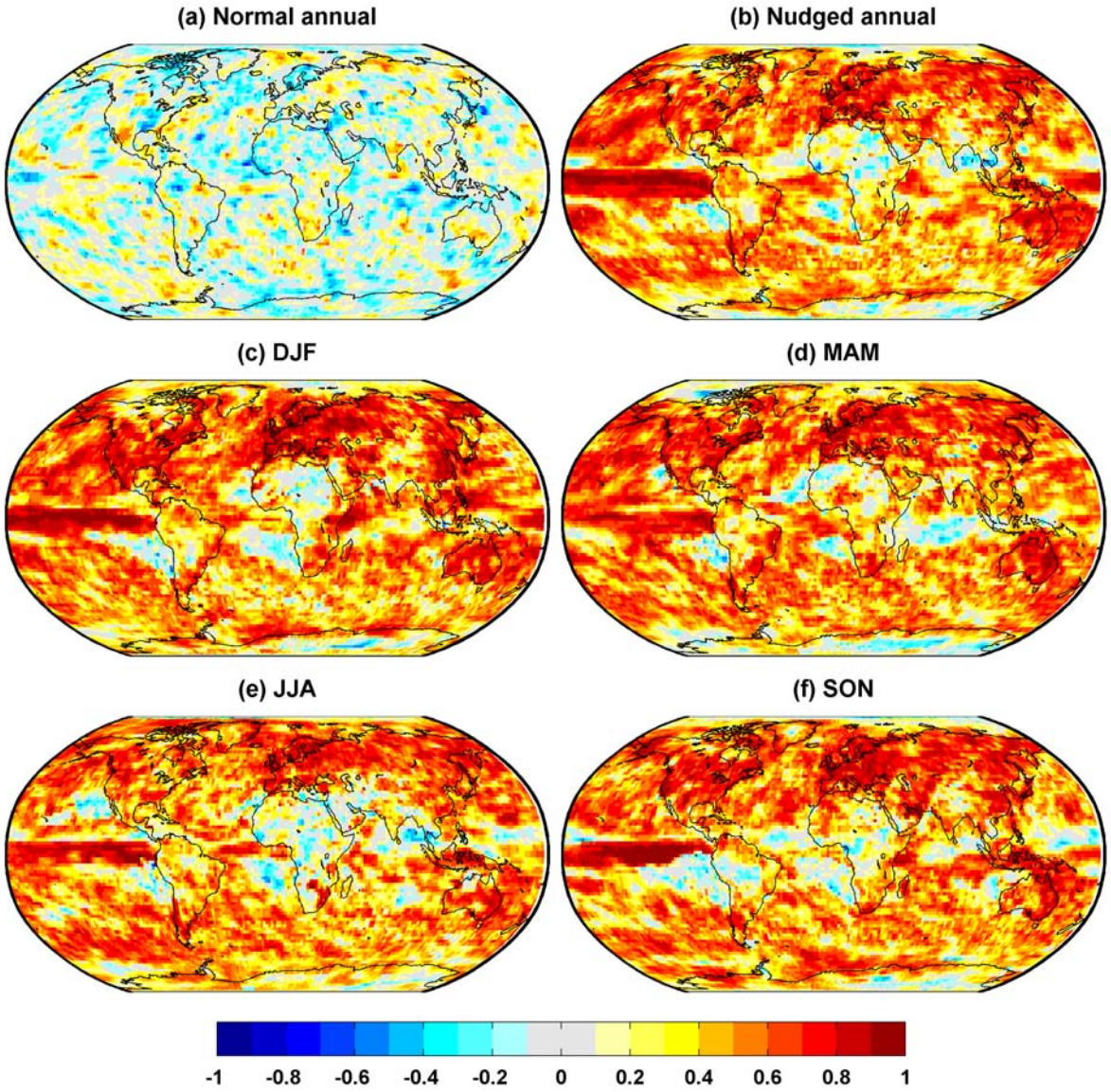

Figure 3.10: Correlation between observed and ECHAM5-simulated annual and seasonal precipitation (1979-2001). (a) Annual precipitation from the standard simulation. (b) Annual and (c-f) seasonal precipitation from the nudged ECHAM5 simulation.

northern hemispheric land mass for all seasons with the exception of summer (JJA). Over the ocean, correlations fluctuate more, with the exception of the eastern equatorial Pacific where they are consistently high, and over tropical land areas they are noticeably lower. Indeed, agreement is considerably poorer over the majority of the African continent than any other region, tropical or extratropical. In general, correlation is weak in regions of negligible precipitation, such as the maritime deserts of the sub-tropical Americas and south-west Africa, which can be attributed to a poorer model performance and less confidence in the quality of observational data.

This shows that given the correct large-scale atmospheric states ECHAM5 is able to successfully

reproduce interannual variability of seasonal precipitation means over large areas, but that there are also areas where this is not the case. Many of the low-skill regions are in the tropics where atmospheric circulation is not quasi-geostrophic and thus may be less accurately represented by the reanalysis and by the nudged ECHAM5 simulation. Similarly, the relatively low density of input data for the reanalysis over the tropics, over the oceans (particularly across the Southern Hemisphere), and over the Arctic and Antarctic may lead to significant errors in the large-scale atmospheric states, such that type 1 errors are not fully eliminated everywhere. In addition there are also errors in the observed precipitation which depend on the data density.

The correlation maps thus provide an upper estimate for the remaining error, which consists of both a parameterisation component (type 3) and a nudging-induced component (type 4). For the purposes of this work, the objective is not to isolate each component but to identify regions where the total remaining error is small and where the representation of temporal variability is skilful enough for the potential development of a statistical correction. Thus, correlation maps in this instance can be considered a lower estimate for the respective skill of ECHAM5 precipitation, but only in locations where correlation is sufficiently strong to realistically reject the influence of a nudging-induced error (type 4).

Nudging to the chosen set of variables (divergence, vorticity and temperature) improves the inter-annual representation of the observed large-scale circulation, which is the main driver of precipitation in extra-tropical regions. GCMs have previously shown skill in reproducing the links between precipitation and particular circulation types (Maheras et al., 2004; Tolika et al., 2006). The nudging procedure does not account for an explicit measure of atmospheric moisture (i.e. humidity), which may explain the weaker correlations in large parts of the tropics where precipitation formation is driven by convective processes to a greater extent than by the large-scale circulation (Reid et al., 2001). It should also be noted that for much of Africa (especially during the boreal summer) and South America, weaker correlations can be partly attributed to a sparse observational network compared with North America and Eurasia. Previous validation work focusing on reanalysed precipitation has also noted poor simulation agreement with observations in data-poor regions, such as the tropical Pacific (Janowiak et al., 1998; Wu and Xie, 2003). Simulated precipitation is thus shown to be dependent upon circulation conditions. Previous work has shown that the errors in the Atmospheric Model Intercomparison Project-type (AMIP) GCM runs are largely attributable to errors in large-scale circulation (e.g. Risbey and Stone, 1996; Leung and Ghan, 1999).

3.4.3 Bias correction

Even in areas where the correlations of simulated and observed precipitation are high, the simulated values may still systematically under- or over-estimate the real values. If these scaling errors in simulated

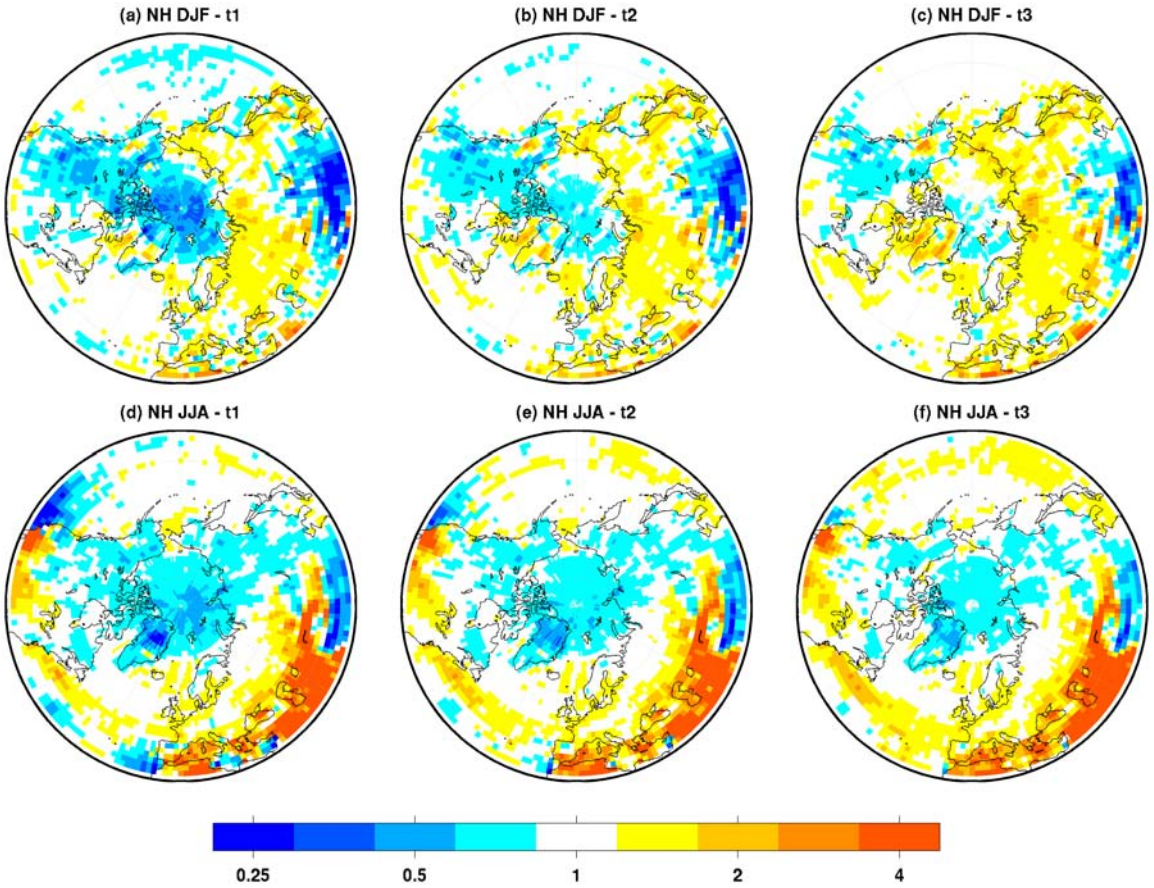

Figure 3.11: Factors required to correct the bias in long-term seasonal mean precipitation from the nudged ECHAM5 simulation across the northern hemisphere.

precipitation are stationary in time, they can be corrected (Widmann et al., 2003), and this simple correction has previously been used as a reference method for statistical downscaling (Schmidli et al., 2006). The correction is the scaling factor that would be required to transform simulated precipitation to accurately reflect observations. As discussed in Chapter 2 (section 2.5.2), bias corrections have previously been successfully developed for RCM simulations (e.g. Leander and Buishand, 2007; Graham et al., 2007; Engen-Skaugen, 2007) but there is also potential for application to GCMs (Widmann and Bretherton, 2000; Widmann et al., 2003). At this stage, scaling is conducted on the same grid cell resolution as ECHAM5 ($1.87° \times 1.87°$) and thus is not a downscaling correction.

In order to assess the sensitivity of a scaling approach to wet or dry situations, the scaling correction factors were derived separately between ECHAM5 output and the lower, middle and upper terciles (t1, t2 and t3) of seasonal precipitation means. That is, observed seasonal means from all years are split into thirds, the first third (t1; events below the 33rd percentile) consisting of the driest seasons. Here, seasonal scaling factors are shown for the extra-tropical regions only, north of 30°N (Figure 3.11) and south of 30°S (Figure 3.12). Northern hemisphere precipitation for the dry winter seasons (DJF

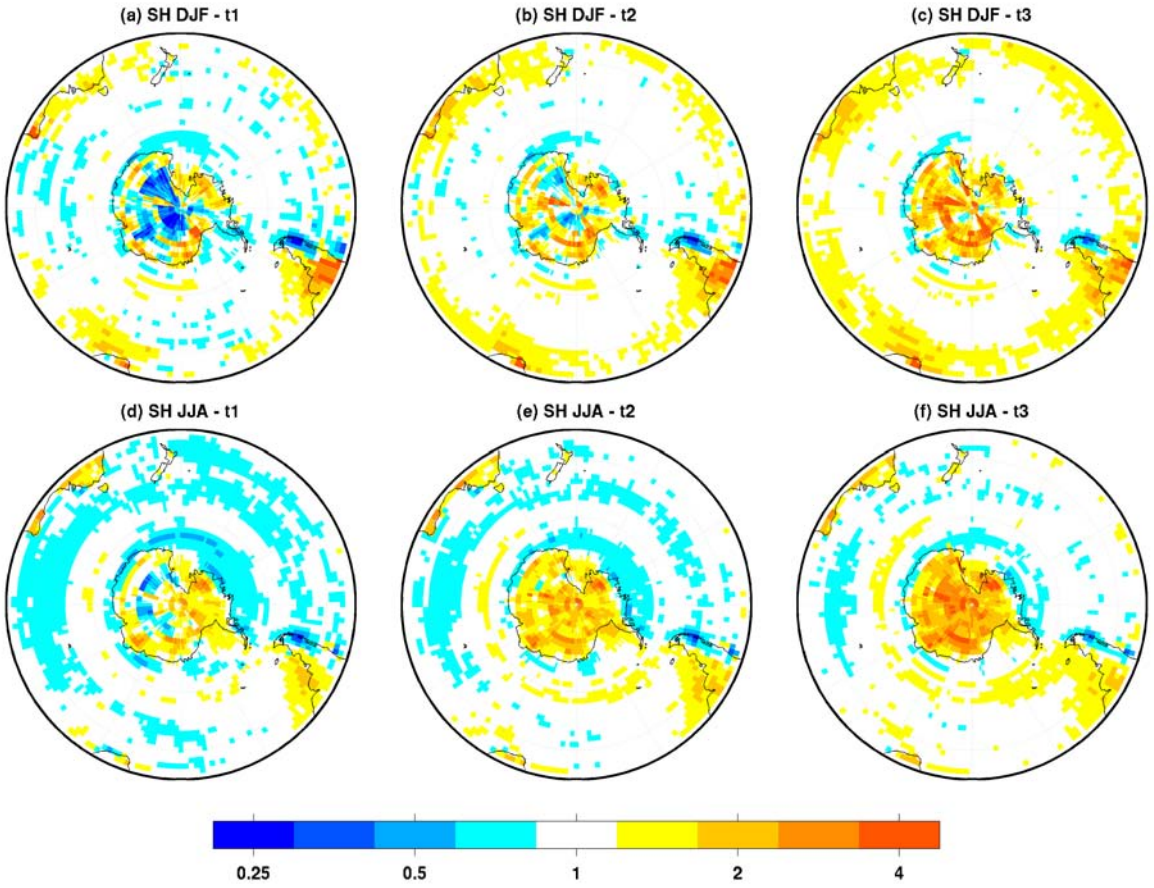

Figure 3.12: As Figure 3.11 but for the southern hemisphere.

- t1) is slightly overestimated by the model over the Arctic, western Canada, and parts of Asia, and slightly underestimated in most of the mid-latitudes. For the wet winter seasons (DJF - t3), under-/over-estimation is similar, with Arctic estimates very close to observations. Most of the areas with substantial scaling errors are very dry in winter, and even small absolute errors lead to considerable scaling factors. It is important to note that in areas of extreme aridity, such as northern Africa and the deserts of central Asia, large apparent errors may be ignored given that even a small absolute error may appear large relative to near-zero precipitation in such areas. In summer (JJA) the results are similar in the sense that scaling errors are mainly located over dry regions and that the t1 and t3 patterns are similar, while the magnitude of the scaling errors is larger than in winter. Thus the model performs similarly in both seasons in wet and dry years, whereas scaling factors are regionally and seasonally dependent. In both seasons the scaling factor is close to 1 over most areas of substantial precipitation, with East Asia in summer being the most important exception. A similar pattern emerges in the southern hemisphere of under- (over-) estimation of the wettest (driest) events. There is a more widespread underestimation of dry winter seasons (JJA - t1) compared to dry summer seasons (DJF - t1). Across the southern portion

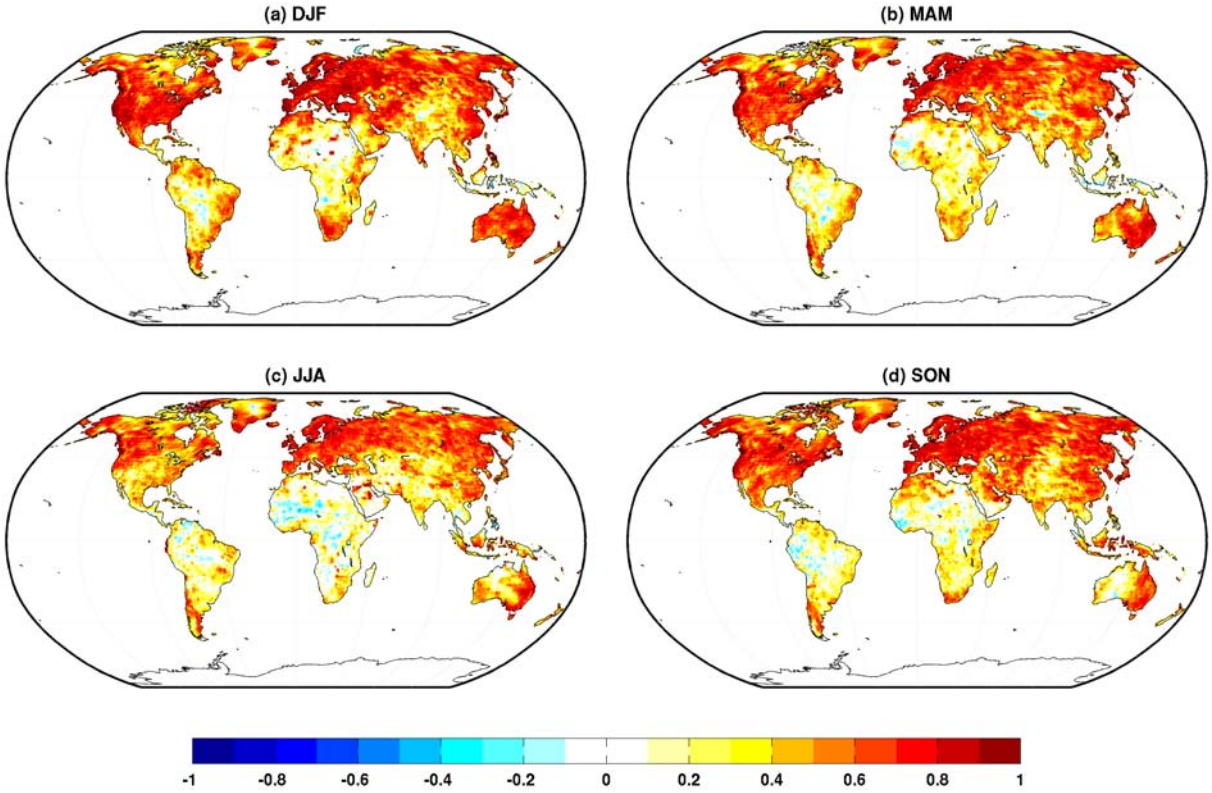

Figure 3.13: Correlation between ECHAM5 seasonal precipitation and GPCC observations (1958-2001). Simulated precipitation taken from the nudged ECHAM5 simulation and subject to linear interpolation to match the $0.5° \times 0.5°$ grid cell resolution of the GPCC observations.

of South America, events in all three terciles are consistently underestimated with the exception of the western coastline where precipitation is consistently overestimated.

3.5 Potential for a downscaling correction

The precipitation field from the nudged simulation is shown to be in good agreement with monthly and seasonal mean observations, both in terms of size of error and representation of temporal variability. The potential for deriving a statistical downscaling correction of the simulated precipitation is subsequently evaluated. ECHAM5 precipitation for the period 1958-2001 is compared with observations from the land-only GPCC dataset. Global features in correlation coefficients between ECHAM5 and GPCC monthly precipitation means are similar to those between ECHAM5 and GPCP data, and land-only correlations are shown to be stronger between ECHAM5 and GPCC (Figure 3.13). Correlations during the boreal winter (DJF) would appear stronger, particularly in the northern hemisphere.

Again, tropical regions exhibit weaker correlations, as discussed in section 3.4. In contrast to the satellite-gauge merged GPCP dataset, the GPCC data is based solely on gauge observations. No coverage exists over the oceans, and coverage over large parts of Africa and South America is very sparse. It is

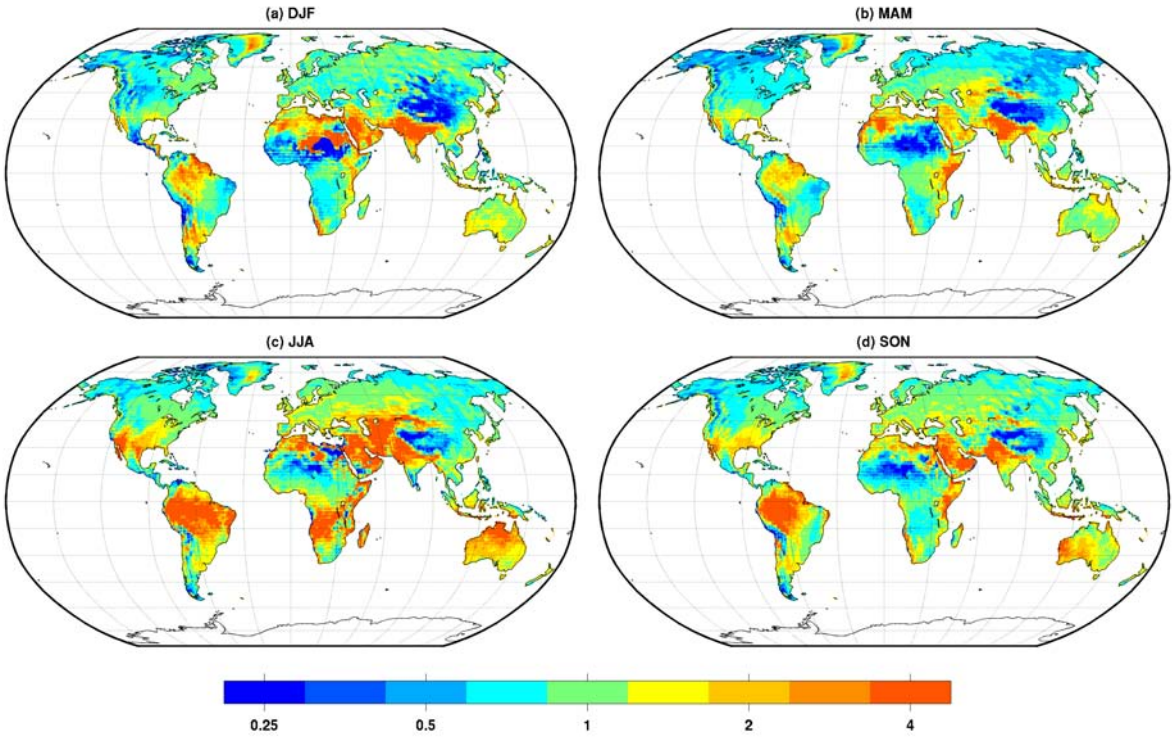

Figure 3.14: Local scaling factors required to correct seasonal precipitation from the nudged simulation (1958-2001).

therefore difficult to make meaningful comparisons between simulated and observed precipitation. Whilst the tropics are acknowledged as an extremely important region, the derivation of a robust downscaling function is limited to those areas with an excellent observational network. At present, these are the regions where informed inferences can be made about the model's performance and about the potential for a statistical correction.

The bias correction approach described in section 3.4.3 was extended to derive a downscaled correction for ECHAM5 precipitation based on the high-resolution GPCC gridded precipitation dataset. The correction in this case can be considered a basic 'local scaling' downscaling model, correcting large-scale simulated precipitation to derive an estimate for regional precipitation at a smaller spatial scale (e.g. Widmann et al., 2003). The distribution of scaling factors across the globe is shown in Figure 3.14. Again, large scaling factors in arid regions should be ignored and are only shown here so as to provide a holistic global view of the potential for a scaling correction. Scaling factors in the majority of extratropical regions are between 0.5 and 2, although this information is not sufficient to suggest the application of scaling in these regions will provide robust results. A further step is to identify areas where confidence in downscaling is high. In the analysis of these scaling factors, focus is limited to Europe where correlation is generally strong during most of the year (Figure 3.15).

As the scaling correction of simulated precipitation would only yield meaningful results in regions

where the correlation between simulated and observed values is high, the factors are shown only for areas in which the correlation for seasonal means (using the GPCC data) are higher than 0.7. The GPCP and GPCC-based correlation maps are very similar to one another on the coarser GPCP grid, and GPCC-based correlations over data-dense regions are as high as 0.9 (Figure 3.13). Considerable scaling errors include overestimation of spring and partly autumn precipitation over Scandinavia by a factor of 1.7, and underestimation in parts of southern Europe in all seasons up to a factor of 2. In autumn and winter almost all of Europe exhibits correlations above the 0.7 threshold (with the Alps being a noticeable exception), whereas in spring and particularly in summer the temporal variability is not simulated well enough to make a correction of simulated precipitation meaningful.

The downscaling correction was cross-validated using a variant of the leave-one-out approach. This approach allows for seasonal precipitation for each year to be estimated independently using a scaling factor derived from simulated and observed data from all other years between 1958-2001. A portion of data is held back during the fitting process and then used to independently validate the method. In practice, a leave-seven-out approach was used in which a scaling factor to estimate seasonal precipitation for given year was fitted using simulated and observed data from all other years aside from a seven-year period centred on the year to be estimated. A period of seven years was chosen to account for the influence of decadal variability and resulting auto-correlation in the calculation of independent scaling factors.

A cross-validated precipitation estimate was shown to be a substantial improvement over both the Mediterranean basin and Australia, where ECHAM5 consistently underestimates precipitation throughout the analysis period (Figures 3.16 and 3.17). Correction of Southern Africa precipitation, however, is less successful as the simulation does not offer a consistent representation of observed temporal variability. Additionally, the quality of ERA-40 (to which the simulation is nudged) is problematic over the Southern Hemisphere in the pre-satellite era, which appears to be reflected in the better skill of the nudged simulation in the second half of the time series. In all three cases good skill on interannual timescales is also associated with good reproduction of decadal variability. It is this property that makes the application of scaling correction factors to climate change simulations promising.

In the practical application of downscaling for future climate scenarios, there is an argument for scaling factors to be derived from standard (non-nudged) GCM simulations in order to consider the minor model-inherent errors in the simulated large-scale circulation (type 1 error) that continue to be present in the future simulation. However, only simulations in which the large-scale atmospheric states are brought close to reality by nudging (or potentially direct data assimilation) allow a derivation of scaling factors that are free from errors caused by differing internal variability in the simulation and in the real world. It is also only possible with these simulations to assess in which regions the simulated precipitation is

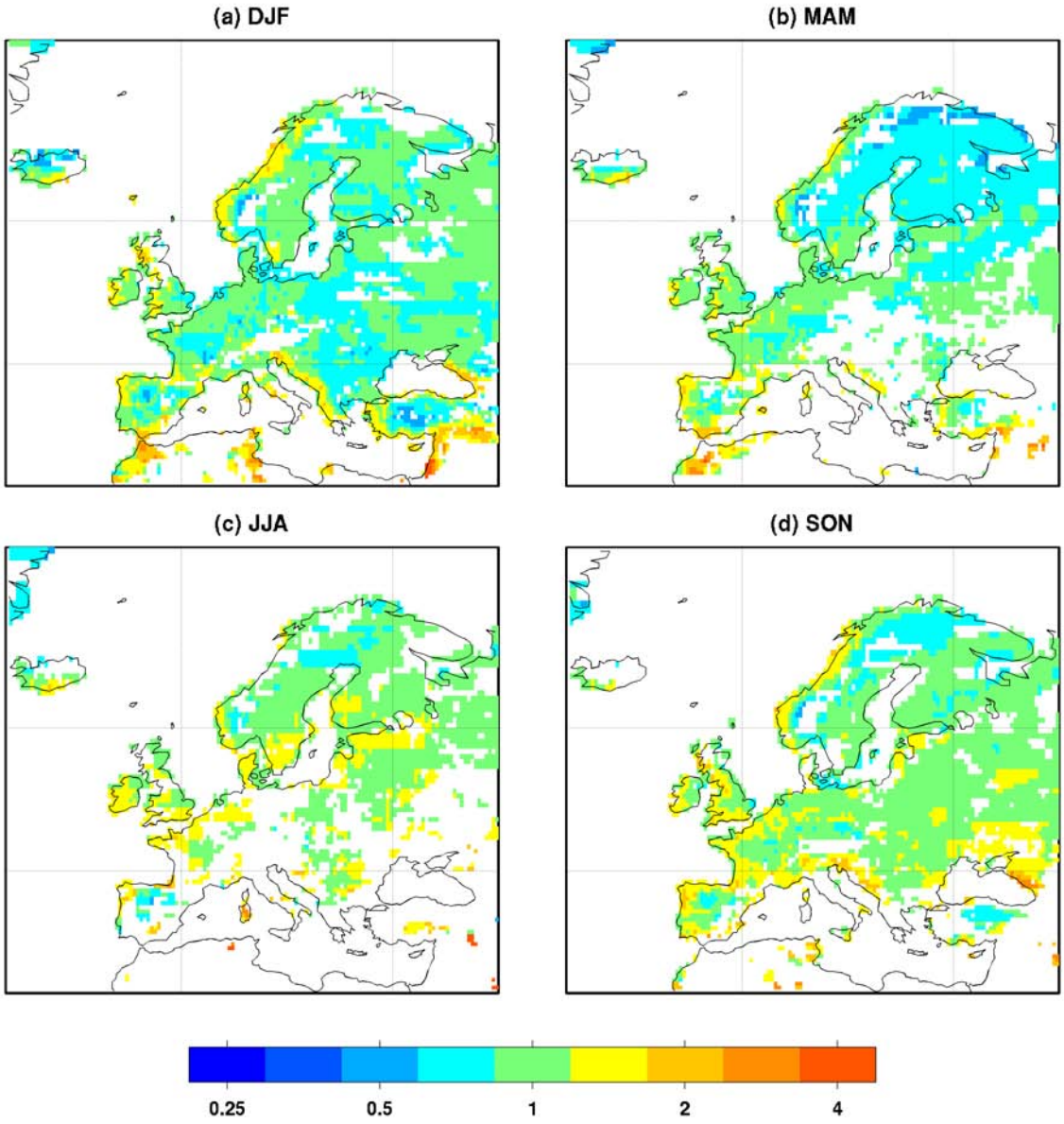

Figure 3.15: Local scaling factors required to correct seasonal precipitation from the nudged simulation over Europe (1958-2001). Scaling factors shown only in locations where correlation shown in Figure 3.13 is greater than 0.7.

highly correlated with real precipitation, which is a necessary condition for justifying the application of correction factors. The nudged simulation also offers the possibility to employ more sophisticated MOS methods that are based, for example, on regression.

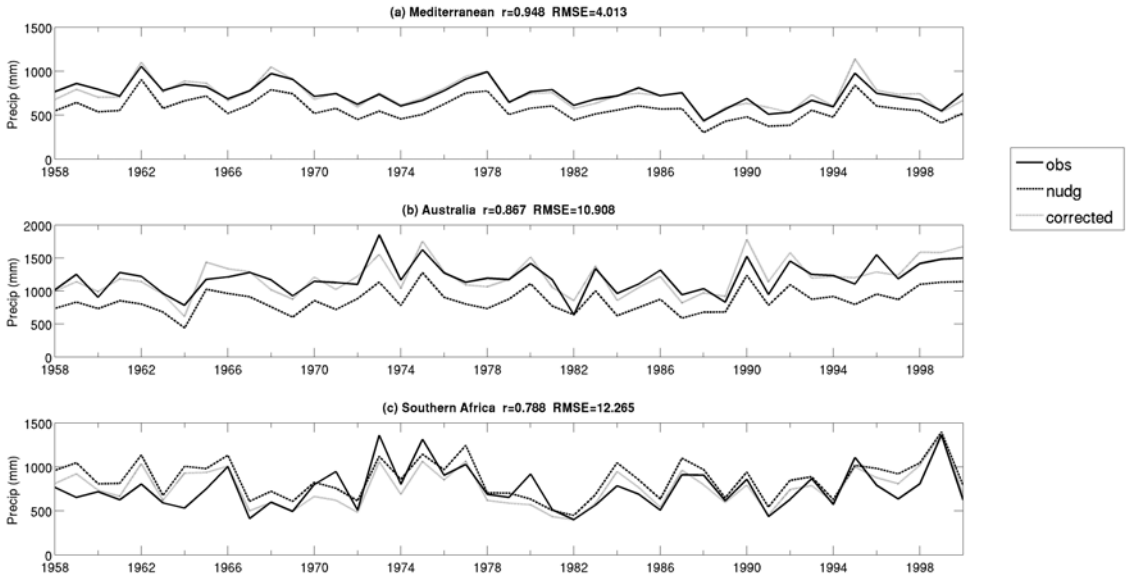

Figure 3.16: Observed, simulated and corrected DJF precipitation (1958-2001). (a) Mediterranean basin, (b) Australia and (c) Southern Africa. GPCC observations are represented by the solid line, nudged ECHAM5 precipitation by the dashed line and cross-validated correction of nudged ECHAM5 precipitation by the dotted line. Correlation coefficient (r) and root mean squared error (RMSE) of the observed and corrected time series are shown for each location.

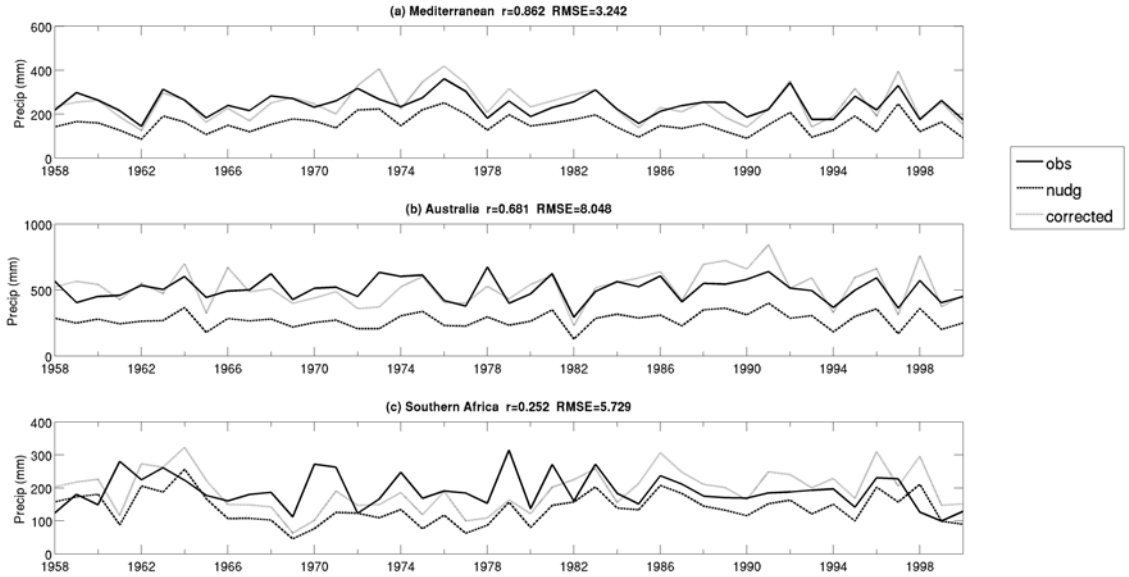

Figure 3.17: As Figure 3.16 but for JJA.

3.6 Summary and conclusions

The skill of GCM-simulated precipitation is difficult to assess through simulation-observation comparison as standard, free-running GCM simulations for historic periods cannot represent temporal variability on a day-to-day basis. Here, a distinction was made between the three sources of error in GCM-simulated precipitation and it was suggested that in removing the model's large-scale circulation errors (type 1) and random internally-generated variability (type 2) it is possible to isolate the error relating to parameterisation and resolution of real-world orography (type 3).

In this chapter, the skill of a reputable GCM to parameterise precipitation based on realistic large-scale conditions has been assessed. A nudged ECHAM5 simulation was conducted for the period 1958-2001 in which the simulated circulation and temperature variables were forced towards corresponding fields from ERA-40. The nudged simulation sufficiently reproduced the temporal variability in the extra-tropical observed large-scale atmospheric state, thus accounting for deficiencies in the large-scale circulation (type 1) and random internal variability (type 2), and permitting a direct assessment of the parameterisation of precipitation (type 3).

It was shown that GCM-simulated precipitation, despite being largely overlooked in downscaling research, can be highly informative. Seasonal precipitation from the nudged simulation was shown to be in excellent agreement with satellite-rain gauge observations in many regions of the world. Correlation maps were used to identify regions in which observed interannual variability was well-reproduced, and model skill was shown to be particularly high in the extra-tropics and over the northern hemispheric land surface. ECHAM5 precipitation was shown to be less useful over large parts of Africa and South America, although the success of nudging large-scale circulation and temperature fields in these regions may be limited by the quality of ERA-40. Subsequently, a statistical downscaling correction of the simulated precipitation field was proposed that is conceptually similar to the MOS approach used in weather forecasting. A simple scaling factor downscaling model was developed, demonstrating that simulated precipitation offers excellent value as a predictor for local precipitation in many regions.

It is noted that the nudging process may induce thermodynamic imbalance and possibly spurious precipitation (type 4 error), which is also identified in previous work (Jeuken et al., 1996). The proposed corrections are strictly applicable only to regions where temporal variability is well-represented and overall model bias is relatively low, and a detailed diagnosis of a type 4 error was not considered, as such an error is likely to be minimal in these areas. Application of a nudging approach to fully isolate the parameterisation deficiencies (type 3) and thus quantifying model skill in areas where nudging-induced spurious precipitation is more likely to cause problems (e.g. tropical regions) would require a more thorough analysis.

Until now, the skill of GCM-simulated precipitation has typically been assessed by means of a multi-model comparison (e.g. IPCC, 2007; Randall et al., 2007). Such comparisons inform as to where inter-model agreement is high, but cannot distinguish between the three types of error that may be responsible for the differences. With the correlation maps created here, it is possible to identify regions where agreement amongst models is backed up by good parameterisation skill in ECHAM5 (e.g. Europe and North America), and crucially regions where multi-model agreement may be due to an unknown error common to many models. An example of the latter is East Africa, a region where the most recent IPCC estimates are for an increase in winter (DJF) precipitation during the twenty-first century (Christensen et al., 2007), but where ECHAM5 skill was shown to be poor. Understanding the individual skill of each model using the approach developed here would appear to be an important step in interpreting a multi-model analysis.

Chapter 4

Development, validation and comparison of MOS and Perfect-Prog downscaling methodologies

4.1 Introduction

It was demonstrated in Chapter 3 that there exists potential for a direct statistical downscaling correction of the precipitation field simulated by ECHAM5. Development of the local scaling correction, based on a simple derivation of the difference between observed and simulated precipitation means, is a simple form of MOS used previously in correcting bias in RCM simulations, and is conceptually different to the traditional Perfect-Prog statistical downscaling approach. A more sophisticated MOS approach would be to calibrate a downscaling model on patterns of variability that exist in the time series of the predictor and predictand fields. The main barrier to the development of such corrections for climate change simulations is that, in order to make monthly or daily comparisons with observations, the required simulation of the past (on which the correction is conditioned) must reproduce the temporal variability in the observed record. Since GCMs are designed to evolve freely, there is typically no assimilation of historical data and model to real-world comparison of anything other than long-term means and distributions is not possible.

The benefits of the development of a nudged simulation described in Chapter 3, in which the prognostic variables of the ECHAM5 GCM are forced toward (though not replaced by) real world observations, can thus be considered two-fold. Firstly, as demonstrated in Chapter 3 (section 3.4), nudging ECHAM5 to ERA-40 permits a quantification of the skill of the precipitation field to simulate temporal variability given a realistic large-scale circulation. Previously, evaluation of GCM-simulated precipitation has been

region-specific and limited to a particular synoptic situation (e.g. Osborn et al., 1999) but a nudging approach provides, or at least the potential for, a global quantification of skill. Secondly, in the context of MOS downscaling, forcing the precipitation field into temporal phase with the observed record allows for statistical downscaling models to be calibrated on coupled modes of variability between simulated and observed precipitation.

In this chapter, the potential for MOS downscaling correction of GCM-simulated precipitation is explored further. In addition to the local scaling approach suggested in Chapter 3, more sophisticated regression-based downscaling techniques are developed and assessed in terms of their respective skill. These so-called 'non-local' techniques use a larger spatial degree of predictor information and are based on MCA and principal component multiple linear regression (PC-MLR) respectively. Following the MOS approach, ECHAM5-simulated precipitation is used as the sole predictor in both the development and implementation of each downscaling model. For comparison, equivalent downscaling models are developed using both non-local techniques as part of a Perfect-Prog approach, in which reanalysis fields are used as predictors.

The remainder of this chapter is structured as follows. Section 4.2 describes in detail the methods used here to downscale ECHAM5 precipitation. Each method is cross-validated in section 4.3 and a regional inter-comparison of method skill is discussed in section 4.4. A summary of key findings is given in section 4.5 .

4.2 Description of MOS and Perfect-Prog downscaling methodologies

In addition to the local scaling method developed in Chapter 3, two further methods are developed that seek to identify modes of variability within a field defined by a spatial grid surrounding the predictand that are used to define a regression model. These methods offer a direct comparison of skill of conceptually identical methods using differing predictor variables, and thus can operate under both MOS and Perfect-Prog frameworks. Under a MOS framework, this predictor field is the simulated precipitation, whereas under a Perfect-Prog framework, historical observations of large-scale circulation variables constitute the predictor field.

4.2.1 Scaling of simulated precipitation (MOS only)

The local scaling method is described in Chapter 3 (section 3.5) and is extended upon here. This method has previously been successfully applied to GCMs and also shown to have comparable skill to more

sophisticated methods (e.g. Widmann and Bretherton, 2000; Widmann et al., 2003). Scaling methods have also been applied to RCMs (e.g. Engen-Skaugen, 2007; Lenderink et al., 2007) and particularly in a hydrological context (e.g. Graham et al., 2007; Leander and Buishand, 2007).

If it is assumed that the precipitation field from the nudged ECHAM5 simulation is sound in terms of its representation of spatial distribution and temporal variability in the observed record, then it may only be necessary to 'scale' the simulated precipitation field to estimate precipitation on a spatial grid of higher resolution. In development, a scaling factor is derived at each grid point, defined as the ratio between mean observed precipitation and mean simulated precipitation for the period 1958-2001. In estimating future precipitation at a given location, simulated precipitation is multiplied by that particular location's scaling factor (Widmann et al., 2003).

At grid point i, j, scaling factor F is given by

$$F_{i,j} = \frac{\overline{Y}_{i,j}}{\overline{X}_{i,j}} \tag{4.1}$$

where $\overline{Y}_{i,j}$ and $\overline{X}_{i,j}$ are the means of observed Y and simulated X precipitation respectively at i, j. Future observed precipitation Y^* at i, j is thus given by

$$Y^*_{i,j} = X^*_{i,j} \cdot F_{i,j} \tag{4.2}$$

where $X^*_{i,j}$ is a future simulated precipitation value at i, j.

4.2.2 Regression-based methods (MOS and Perfect-Prog)

Scaling of the simulated precipitation is simple to perform, but considers only predictor information at the same geographical location as the predictand. Many downscaling methods seek to utilise neighbouring predictor information, patterns which may provide more realistic and stable precipitation estimates (Fowler et al., 2007). Non-local downscaling is more equipped to account for GCM errors due to spatial discrepancy rather than unresolved physical processes. For instance, a GCM may be able to resolve a particular precipitation characteristic in terms of magnitude and spatial extent, but in an incorrect location. In practical terms, this has particular benefits in accounting for a GCM's poor representation of complex topography, for example.

Two further downscaling methods were developed using linear regression models. Key in each method is the transformation of an initial two-dimensional predictor field so as to identify coupled modes of variability between the predictor and predictand fields. In practice, these methods are equivalent to one-dimensional CCA, itself equivalent to principal component multiple linear regression (PC-MLR), and one-dimensional MCA respectively. Linear regression models constructed using fields transformed

with these methods have been used frequently in the literature, and in some cases subjected to a direct comparison (e.g. Widmann et al., 2003; Tippett et al., 2008). It is not clear which method performs better on independent data and it is arguable that the two should be viewed as complementary rather than competing techniques (Cherry, 1996). In discussing the development of each downscaling model, it is first of all necessary to describe the procedure of data transformation in detail.

4.2.2.1 One-dimensional CCA and MCA

It is useful to discuss CCA and MCA in the context of the similarities and differences of each. CCA is a widely applied approach used to identify coupled patterns in two datasets and first implemented in a meteorological application by Glahn (1968). Projection of the original two data sets onto these patterns produces a new set of variables, linear combinations of the original fields, that have maximum possible correlation (Bretherton et al., 1992; Wilks, 2006; Tippett et al., 2008). All subsequent pairs of patterns are found such that projections of the data sets onto them are the most strongly correlated linear combinations of the original fields that are not correlated with the preceding set(s) of linear combinations (von Storch, 1999). In contrast to principal component analysis (PCA), which identifies variability patterns in a single dataset, the new variables produced in a CCA maximise the relationship between two datasets, and in this sense CCA can be considered a 'double-barrelled' PCA (Wilks, 2006).

The procedures involved in performing MCA are conceptually similar to CCA and subject to the constraint that the patterns within each dataset are orthogonal to each other. However, the goal of MCA is to extract modes of variability between two data sets that exhibit maximum covariance rather that maximum correlation (Tippett et al., 2008). MCA uses the singular value decomposition of the cross-covariance matrix between the two fields, and in earlier papers is also referred to as SVD analysis (Bretherton et al., 1992; Widmann et al., 2003; Widmann, 2005).

CCA and MCA transformation methods have often been compared (e.g. Bretherton et al., 1992; Cherry, 1996; Widmann, 2005; Tippett et al., 2008) although it remains unclear which approach yields the greater skill when used in, for example, a regression methodology to estimate the value of a predictand. Widmann (2005) notes that whilst estimating a vector from a scalar produces identical results regardless of the method used, estimating a scalar (one-dimensional predictand time series) from a vector (spatial predictor field) is method-dependent. In the example used by Widmann (2005), it is unclear whether the CCA or MCA is most skilful when estimating January Arctic Oscillation index (AOI) from northern hemispheric 850-hPa temperature and that each approach has advantages over the other.

Widmann (2005) and Tippett et al. (2008) have recently offered useful clarification on the relationship between CCA and MCA approaches for finding coupled patterns and multiple linear regression. In the case of Widmann (2005), focus was given to the relationship with regression maps which are often used

as a relatively simple way to capture the link between a particular time series and a meteorological field. The one-dimensional MCA method is also performed by Thompson and Wallace (1998); inter-dataset patterns were derived by regressing time-dependent geopotential height fields onto the AOI time series. Time series of 'expansion coefficients' were produced by orthogonally projecting the original fields onto the corresponding regression pattern. These so-called time expansion coefficients (TECs; Widmann, 2005) exhibit maximum covariance with the time series when the values in the initial spatial field are expressed as anomalies (i.e. with a mean of zero). This method is thus identical to performing one-dimensional MCA, although this connection was not discussed by Wallace et al. (1995) or Thompson and Wallace (1998) and was first identified by Widmann (2005).

Here, one-dimensional MCA was performed in the same way. For each location (i.e. each grid point) a map of coefficients, determined by regressing the spatial predictor field onto the predictand time-series, (in this case observed monthly mean precipitation) was produced to capture the signal of the time-series in the field (Widmann, 2005). TECs, defined by orthogonal projection of the original predictor field anomalies onto the regression map, formed the new variables entered in a linear regression model to estimate unknown values of monthly mean precipitation. For transparency in the approach taken in constructing the one-dimensional MCA, the method is referred to as SVD-regression map (hereafter SVD-RM) in the remainder of this and subsequent chapters. This is consistent with the terminology used in its recent application by Widmann (2005).

One-dimensional CCA was performed in a conceptually identical manner: linking point-scale observed precipitation at each grid point with a spatial meteorological field. In this case, the actual method used was principal component multiple linear regression (PC-MLR, or simply, PCR); that is, multiple linear regression between point-scale precipitation and principal components of the spatial field, which in the one-dimensional case is a procedure equivalent to CCA (Glahn, 1968; DelSole and Chang, 2003; Widmann, 2005; Tippett et al., 2008). Again, to maximise transparency in the construction and application of this downscaling model, the method is hereafter referred to as CCA-PCR. In contrast to SVD-RM, CCA-PCR required a decision on the number of principal components (PCs) to retain. Inclusion of fewer predictors in a regression model is likely to avoid overfitting issues and improve the representation of actual relationships (Tippett et al., 2008). Here, ten MLR models were produced with 1-10 retained PCs. Selecting the most skilful CCA-PCR model ultimately requires a subjective decision, and whilst such a decision is removed from the SVD-RM approach, it may be possible to isolate a CCA-PCR model with a number of PC predictors that consistently outperforms its SVD-RM equivalent.

4.2.2.2 MOS and Perfect-Prog application

The SVD-RM and CCA-PCR methods were implemented under both MOS and Perfect-Prog approaches to statistical downscaling (see Chapter 2, section 2.5). For MOS, this simply involved using the ECHAM5-simulated precipitation field as a predictor variable. In model development under a Perfect-Prog approach, a number of observed predictor variables were taken from ERA-40. These included geopotential height, temperature, and specific and relative humidity at 1000-hPa, 850-hPa and 500-hPa. Whilst there exist higher resolution observational data sets of some variables, ERA-40 provides a platform for the analysis to be performed globally.

An important consideration is the size of the spatial domain from which predictor information is taken (Fowler et al., 2007). Wilby and Wigley (2000) gave examples of precipitation exhibiting stronger correlation with mean sea level pressure (MSLP) in neighbouring grid cells than with the same grid cell. As the goal here is to downscale precipitation individually at each grid cell across the globe, the position and spatial extent of the predictor domain is constant, although altering these parameters may form an important part of future studies at more specific locations. In developing a MOS downscaling model, the domain is defined as 20° longitude× 10° latitude for each location. In a Perfect-Prog approach, for all observed predictors, the spatial extent of the domain was considerably larger than that used for the MOS approach (40° longitude× 20° latitude). This was to account for a greater spatial influence of these variables on precipitation and the coarser resolution of ERA-40 (2.5° grid cells) compared to ECHAM5 (approximately 1.875° grid cells). No attempt was made to derive an optimally sized spatial grid, which is likely to be season- and location-specific. The potentially global application of the downscaling models developed here means that such an optimisation step is impractical, although recommendations are made in section 4.5 as to where downscaling performance may be improved by a change in domain size. It was also important that a concurrent signal is sought between local precipitation and meteorological conditions within an area of relative proximity. As noted by Widmann (2005), the GCM grid cells included within the domain (spatial grid) do not represent equal areas and it is necessary for the data in each cell to be weighted. In the case of the domains used in both MOS and Perfect-Prog downscaling development, this is performed by weighting the simulated precipitation field with the square root of the cosine of the latitude (Widmann, 2005).

4.2.3 Observational data and downscaling model validation

All downscaling methods are developed using real world observational data for the predictand. For global analysis, monthly mean observational precipitation data is taken from the Global Precipitation Climatology Center (GPCC) gridded dataset. Land-only interpolated rain gauge data is available at a

resolution of 0.5 × 0.5 from 1900-2008 (see also Chapter 3, section 3.3.2). The quality of the data set is dependent on station density and completion of observational records, both of which are sufficient for the period 1958-2001 which is required in this analysis.

In order to evaluate the actual skill of each method, it is necessary to develop a downscaling model to estimate local precipitation for a period that is held back in the development of that model. In short, the fitting period is required to be independent of the validation period. Such cross validation strategies are common where simultaneous records of simulated and observed data are limited (e.g. Widmann et al., 2003). Often, the observed record is split into two (or more) datasets, say P1 and P2. A downscaling model developed by linking observed predictand data for period P1 with a simulation of the same period is then used to estimate values of the predictand for period P2. The process is then repeated with reversed fitting and validation periods. The result is a concatenated time series of the predictand which has been independently-estimated, and which can then be compared with the original observational data. An important assumption of this cross validation method is that the quality of the observed data is similar in P1 and P2. In the case of the GPCC data, it is known that a the amount of station data (and consequently, the quality of the data) increased chronologically throughout the record. Thus, data for P1 is unlikely to be as reliable as data for P2. Additionally, splitting the observed data in this way may separate important modes of variability or inter-decadal patterns. As an alternative, a variation of the 'leave-one-out' cross validation approach is used (e.g. Wilks, 2006).

The downscaling model for a given year, t, is derived from simultaneous values of simulated and observed seasonal means for the complete time series (n years), with the exception of a seven-year 'exclusion period' around t. Thus, for the complete time series, simulated and observed values are taken for all years up to t-3 and after t+3. Independent estimates for all n years are then concatenated to form a single time series which can then be directly compared to observations. Omitting a period of seven years from each model development was considered sufficient to account for autocorrelation of monthly precipitation between neighbouring years.

4.3 Validation of MOS and Perfect-Prog downscaling methods

4.3.1 MOS approach

The local scaling approach was introduced in Chapter 3 (section 3.5) and is now extended and fully assessed in terms of its skill for independently estimating local-scale precipitation. Cross-validated estimates are made for local-scale monthly mean precipitation for 1958-2001. Correlation between the 'corrected' simulation and corresponding observations are shown in Figure 4.1. Across much of the northern hemisphere, correlation coefficients are strongly positive for most of the year. In Europe and North America,

correlation is particularly high during the winter months (DJF), and greater than 0.8 in some areas. Correlation is noticeably weaker during summer (JJA) across the northern hemisphere, and the difference is particularly apparent in central Asia and, to a lesser extent, the central United States. Throughout most of Africa and South America, correlation is weak for most of the year, with the exception of parts of South Africa and eastern South America during the austral summer months (DJFM). Correlation coefficients in eastern Australia are strong (> 0.7) for all months of the year, and in much of the rest of Australia during the austral summer and autumn months (DJFMAM). The spatial distribution of long-term bias in corrected precipitation is similar to that of correlation (Figure 4.2). Realistic monthly mean precipitation is well-reproduced throughout the extra-tropics and performance is poorest across the driest parts of the sub-tropics. As mentioned in Chapter 3, the apparent poor performance of any downscaling model in extremely arid regions may be attributable to small absolute errors appearing large relative to low precipitation totals. Correlation and bias results in such areas should be largely ignored in the remainder of this chapter.

ECHAM5-simulated precipitation downscaled using non-local regression methods is also directly compared to corresponding observations. Using a SVD-RM approach, the geographical distribution of correlation coefficients is relatively similar to that for the local scaling approach (Figure 4.3), with the strongest correlation occurring in the northern hemisphere during the boreal winter (DJF). The regions of smallest long-term bias also appear similar, with the majority of Eurasia and North America well-represented during all months of the year (Figure 4.4).

With the CCA-PCR method, correlation maps between reconstructed and observed local-scale precipitation precipitation once again appear similar to the previous two methods. CCA-PCR skill is dependent on the number of retained predictor PCs. Here, correlation coefficients for methods retaining 2 (Figure 4.5), 5 (Figure 4.6) and 10 (Figure 4.7) PCs are shown. In general, the spatial distribution is similar to the local scaling and SVD-RM methods and there is little discernible difference between different numbers of retained PCs. In terms of bias, CCA-PCR with 2 PCs (Figure 4.8) tends to estimate precipitation with good spatial consistency in regions where bias is generally small. Bias in the precipitation estimates based on 5 or 10 PCs (Figures 4.9 and 4.10) show greater spatial variability in the same regions. It is noted, however, that the optimal number of PCs to retain in a CCA-PCR model is location- and season-specific.

A comparison of the skill of each method on a regional basis provides a more in-depth analysis. Figure 4.11 gives a detailed overview of average correlation, root mean squared error (RMSE) and bias statistics during January and July across three regions where MOS application shows the greatest potential: Europe, North America and Australia. As is also evident in the previous plots in this section (Figures 4.1-4.10), model performance is greater during winter than during summer. Local scaling of-

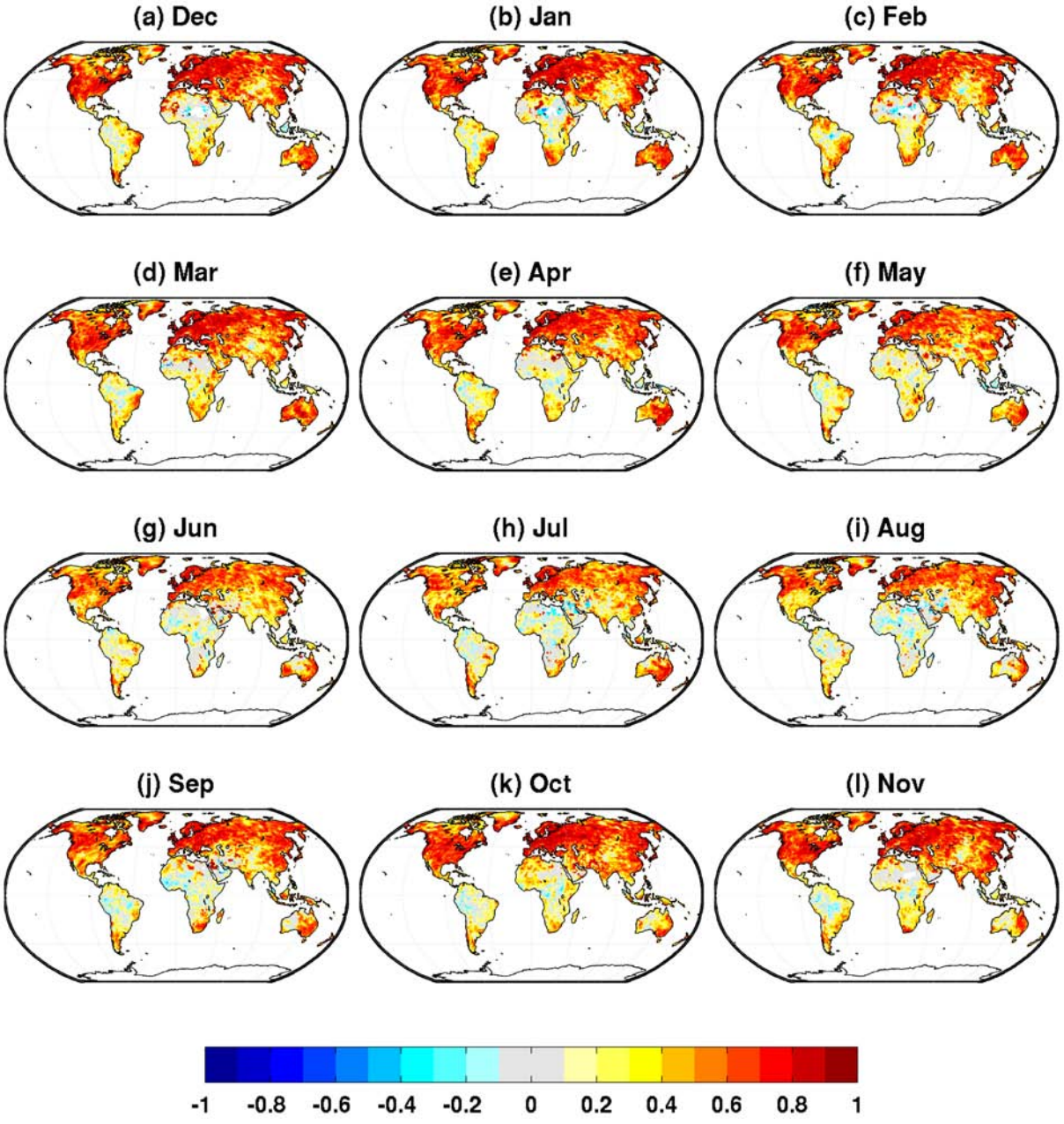

Figure 4.1: Correlation between monthly mean precipitation GPCC observations and cross-validated estimate based on local scaling.

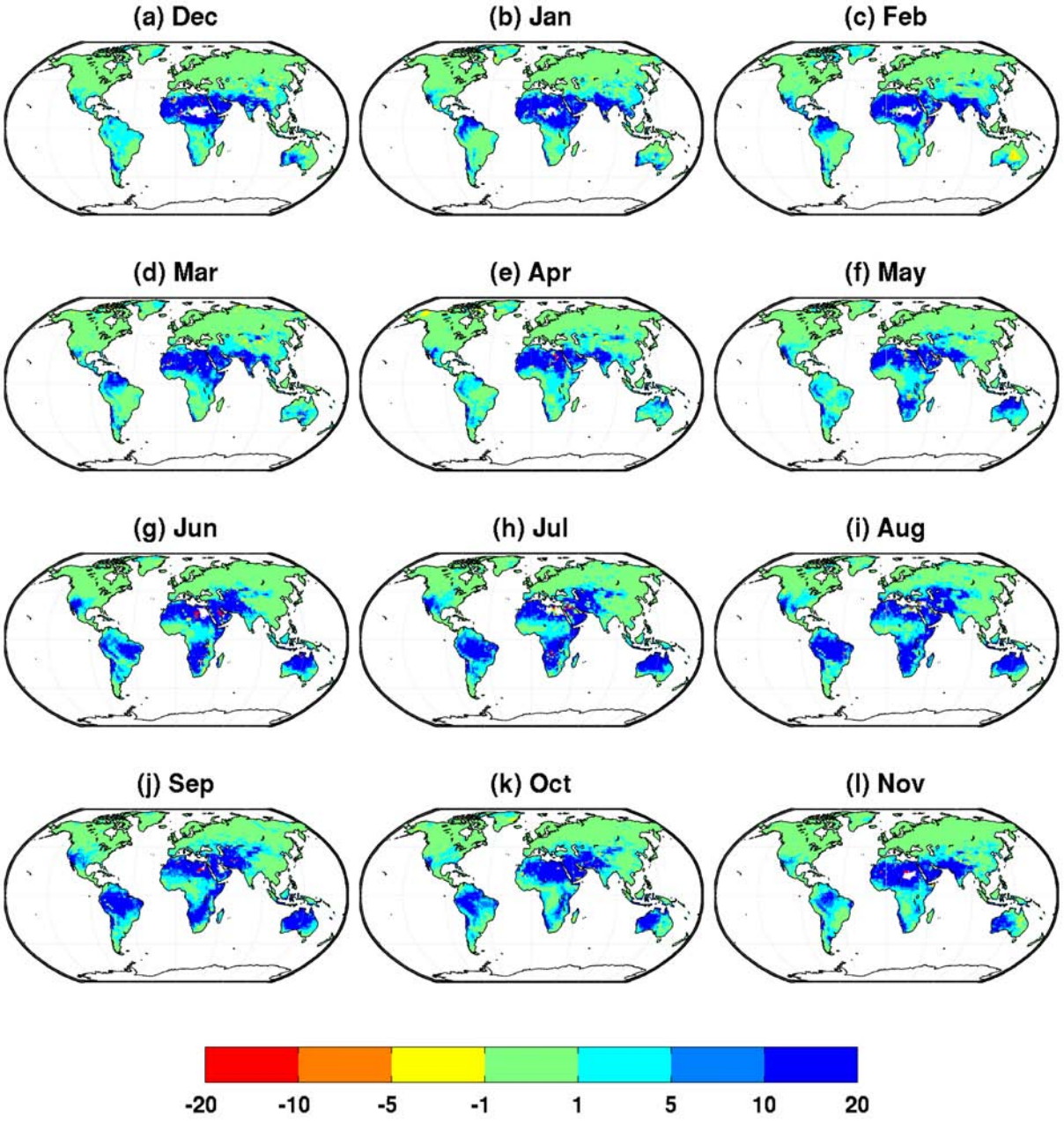

Figure 4.2: Bias (percentage) between monthly mean precipitation GPCC observations and cross-validated estimate based on local scaling.

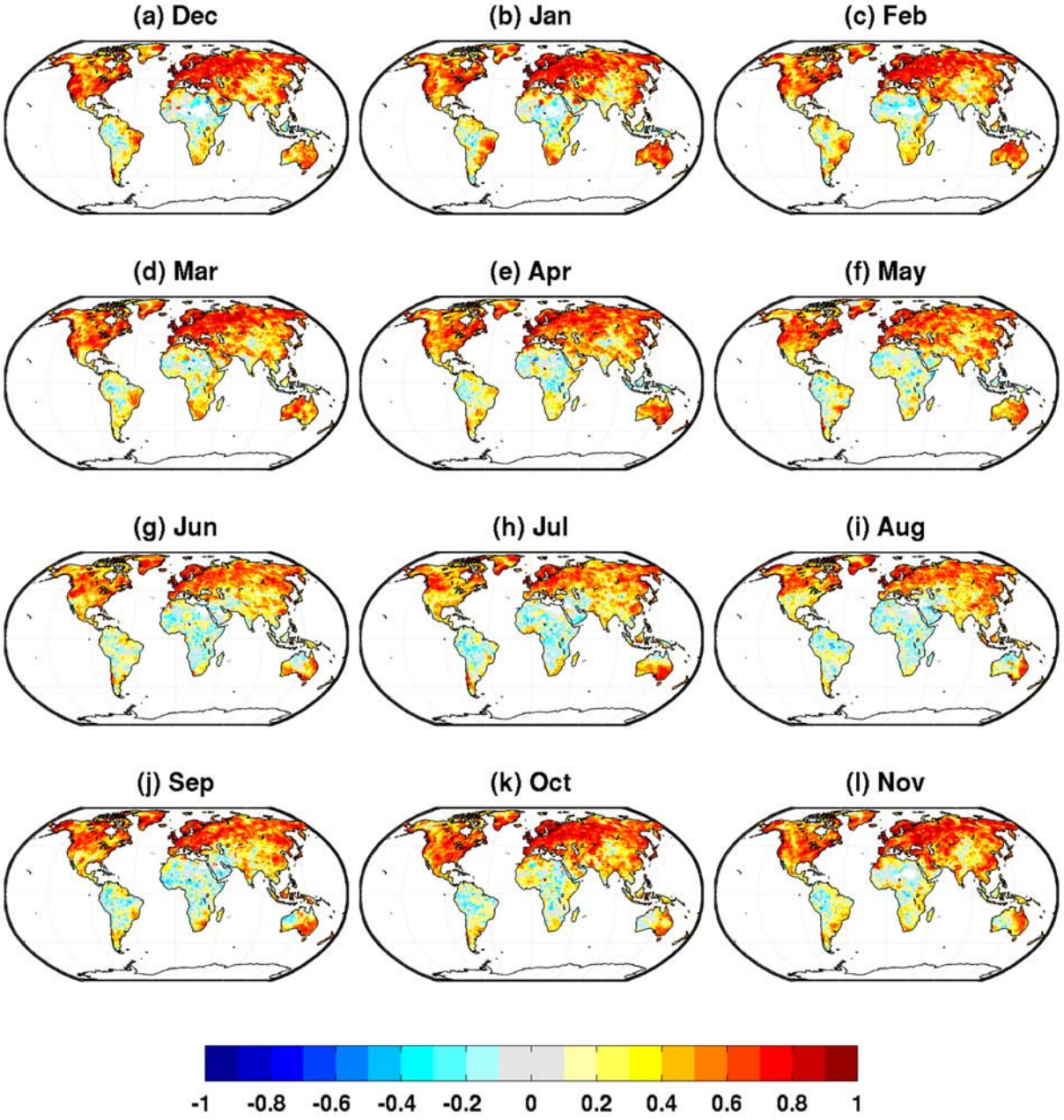

Figure 4.3: Correlation between monthly mean precipitation GPCC observations and cross-validated estimate based on SVD-RM.

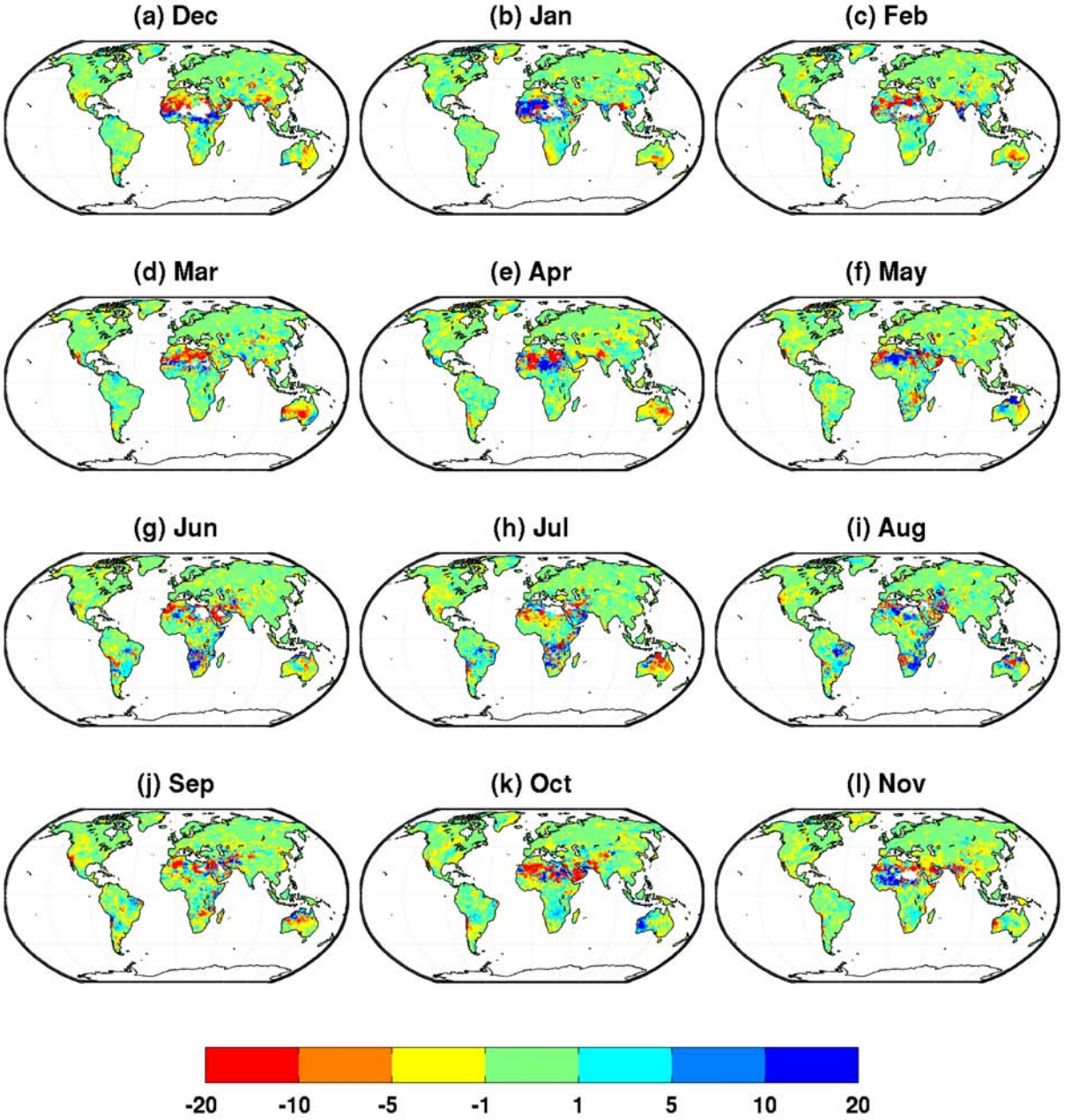

Figure 4.4: Bias (percentage) between monthly mean precipitation GPCC observations and cross-validated estimate based on SVD-RM.

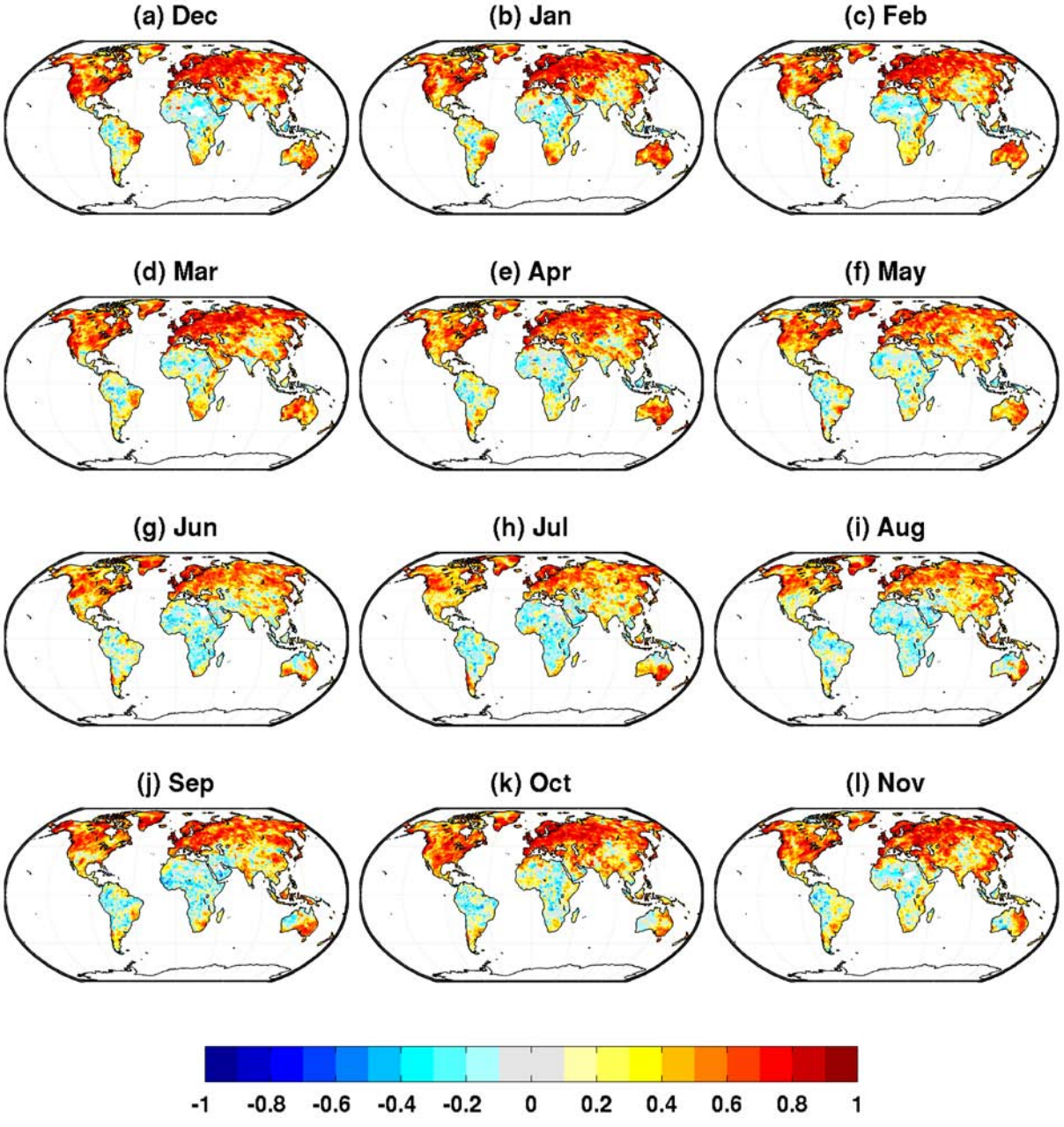

Figure 4.5: Correlation between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (2PCs).

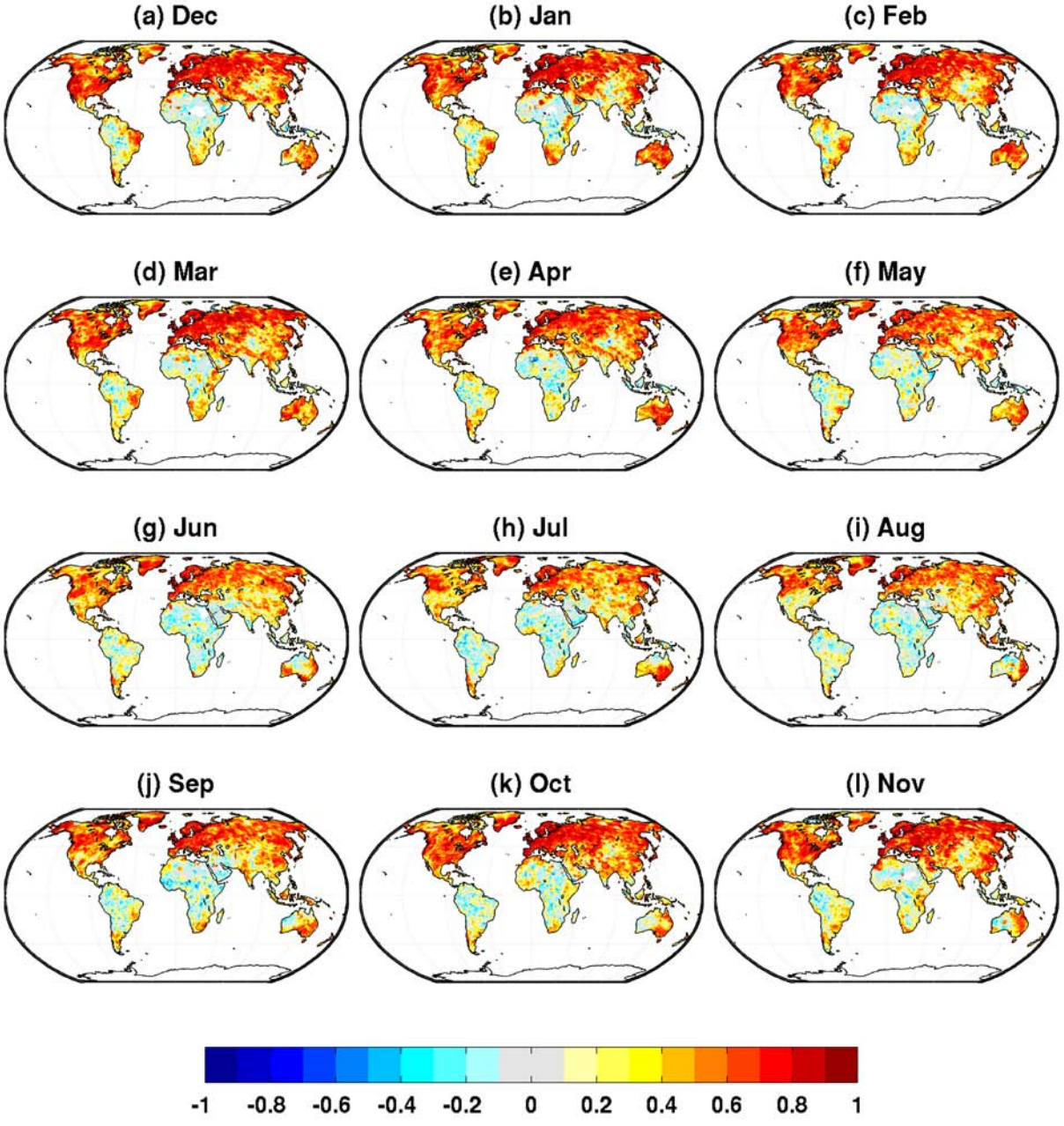

Figure 4.6: Correlation between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (5PCs).

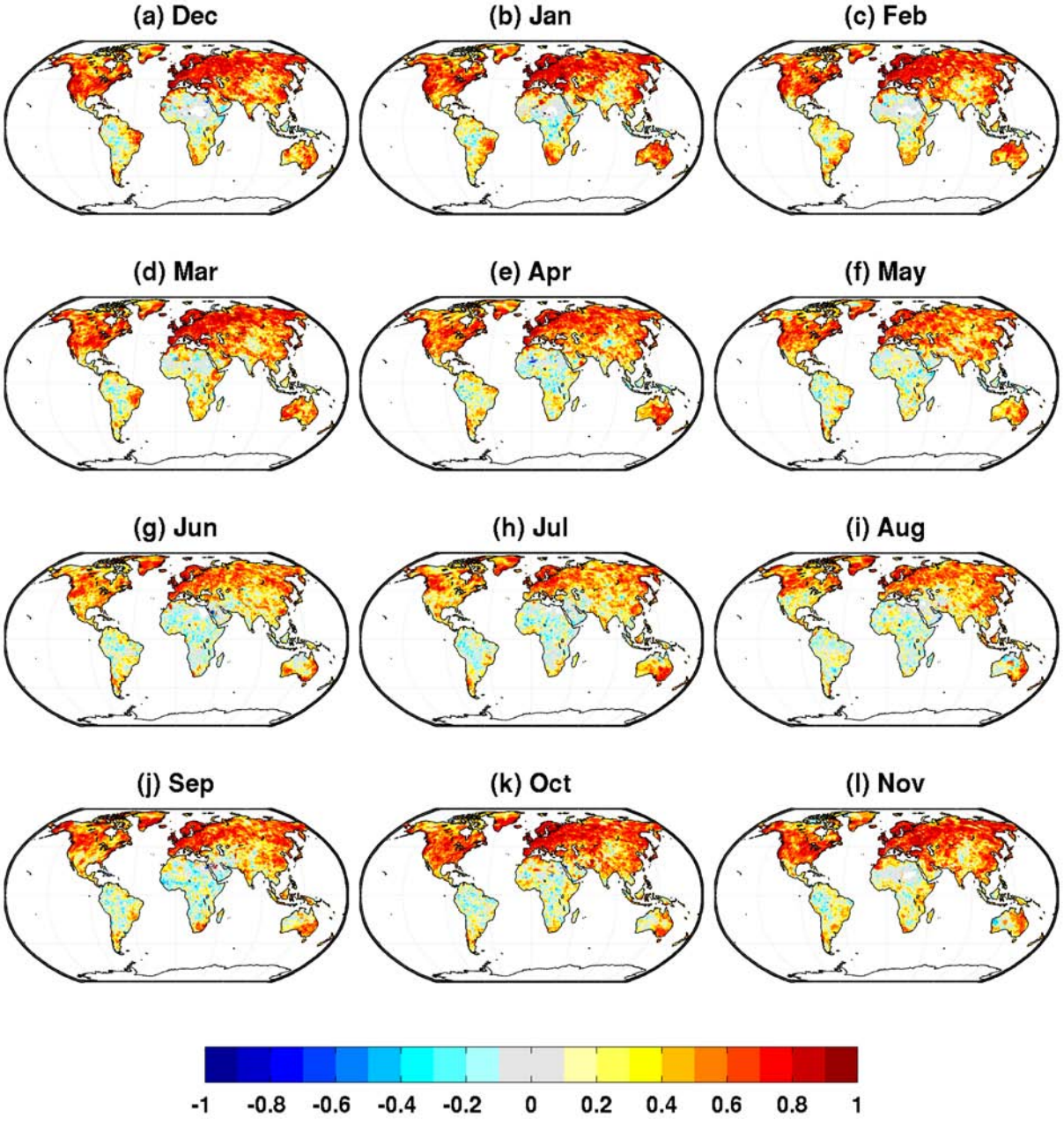

Figure 4.7: Correlation between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (10PCs).

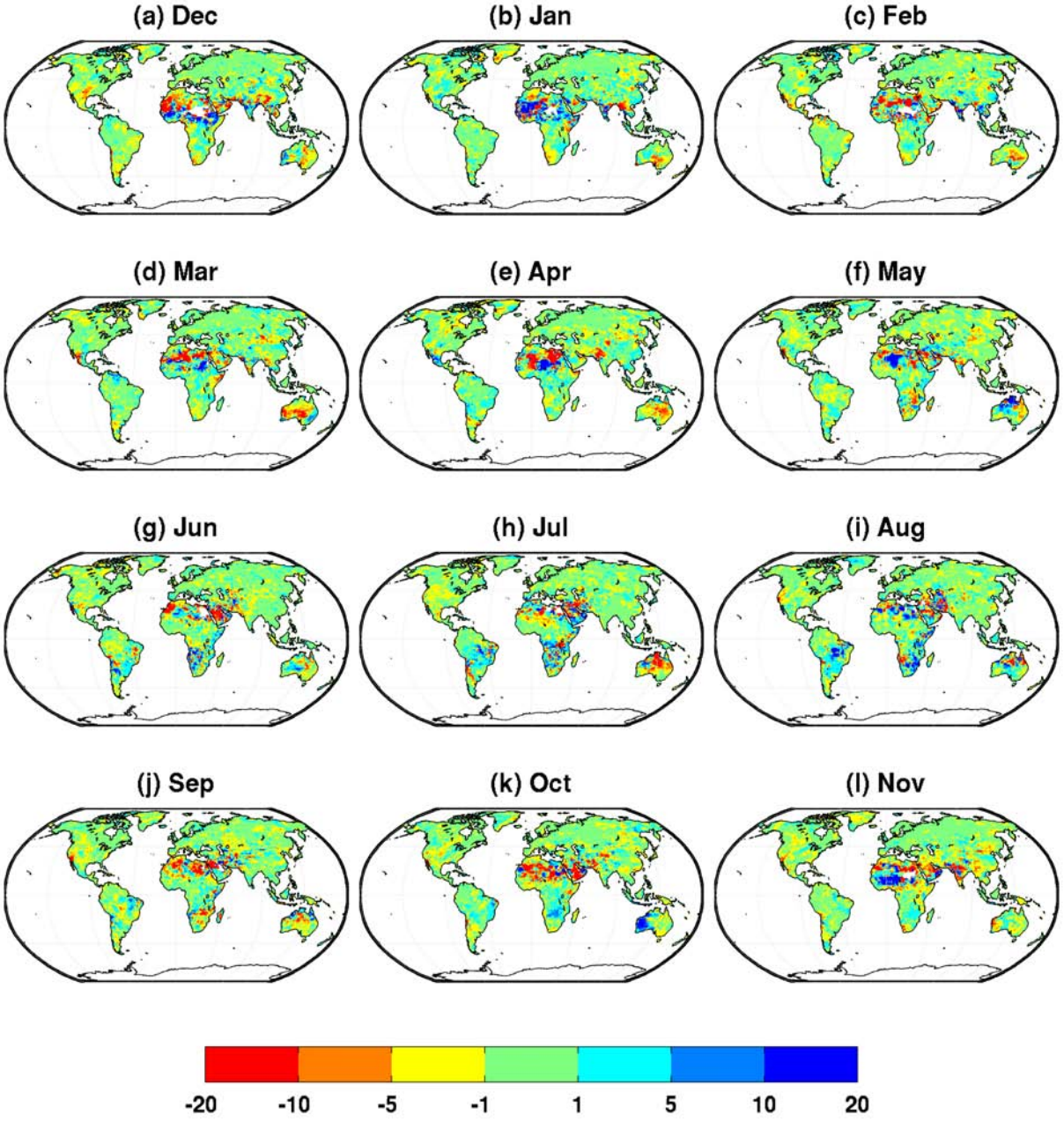

Figure 4.8: Bias (percentage) between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (2PCs).

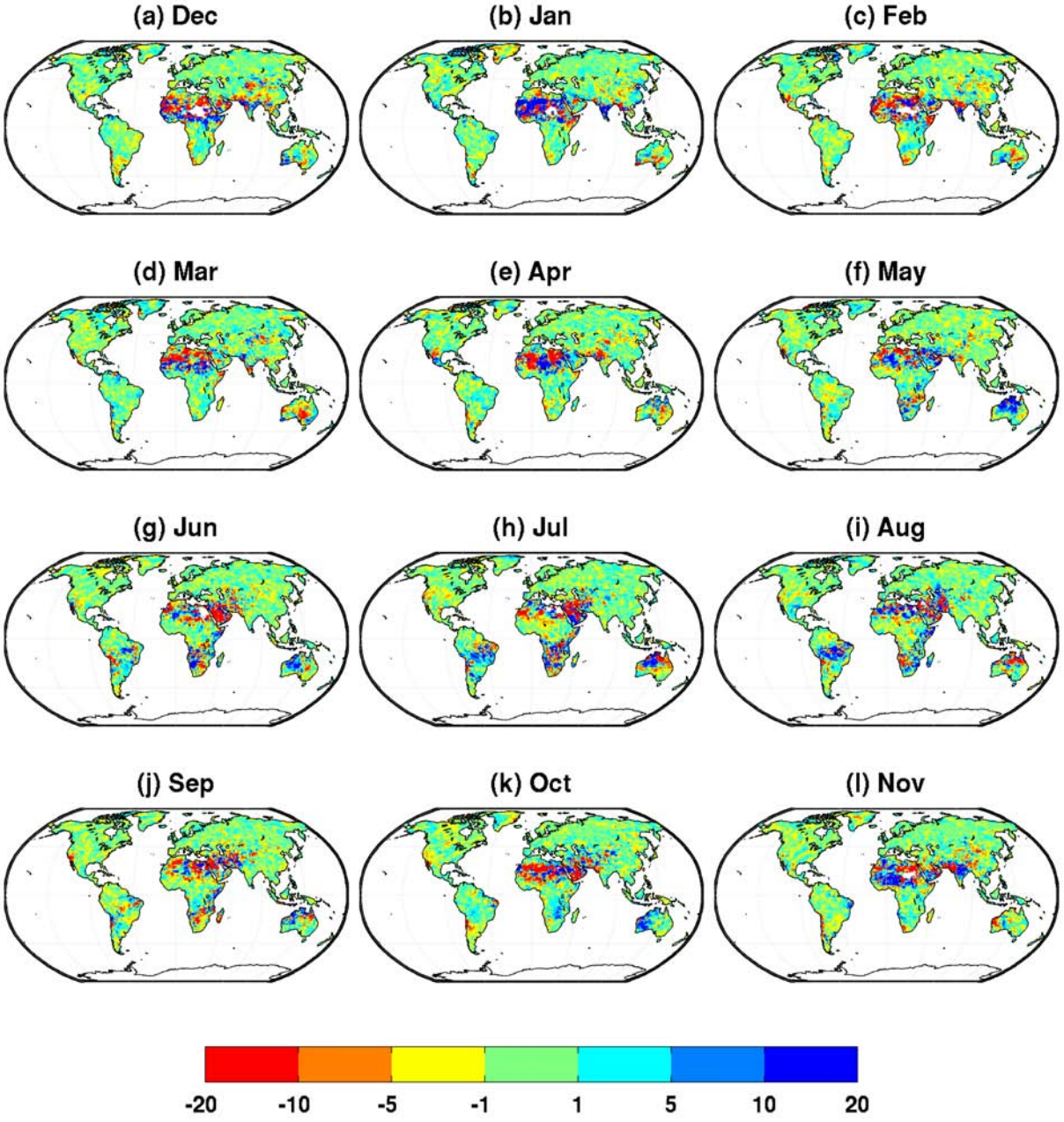

Figure 4.9: Bias (percentage) between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (5PCs).

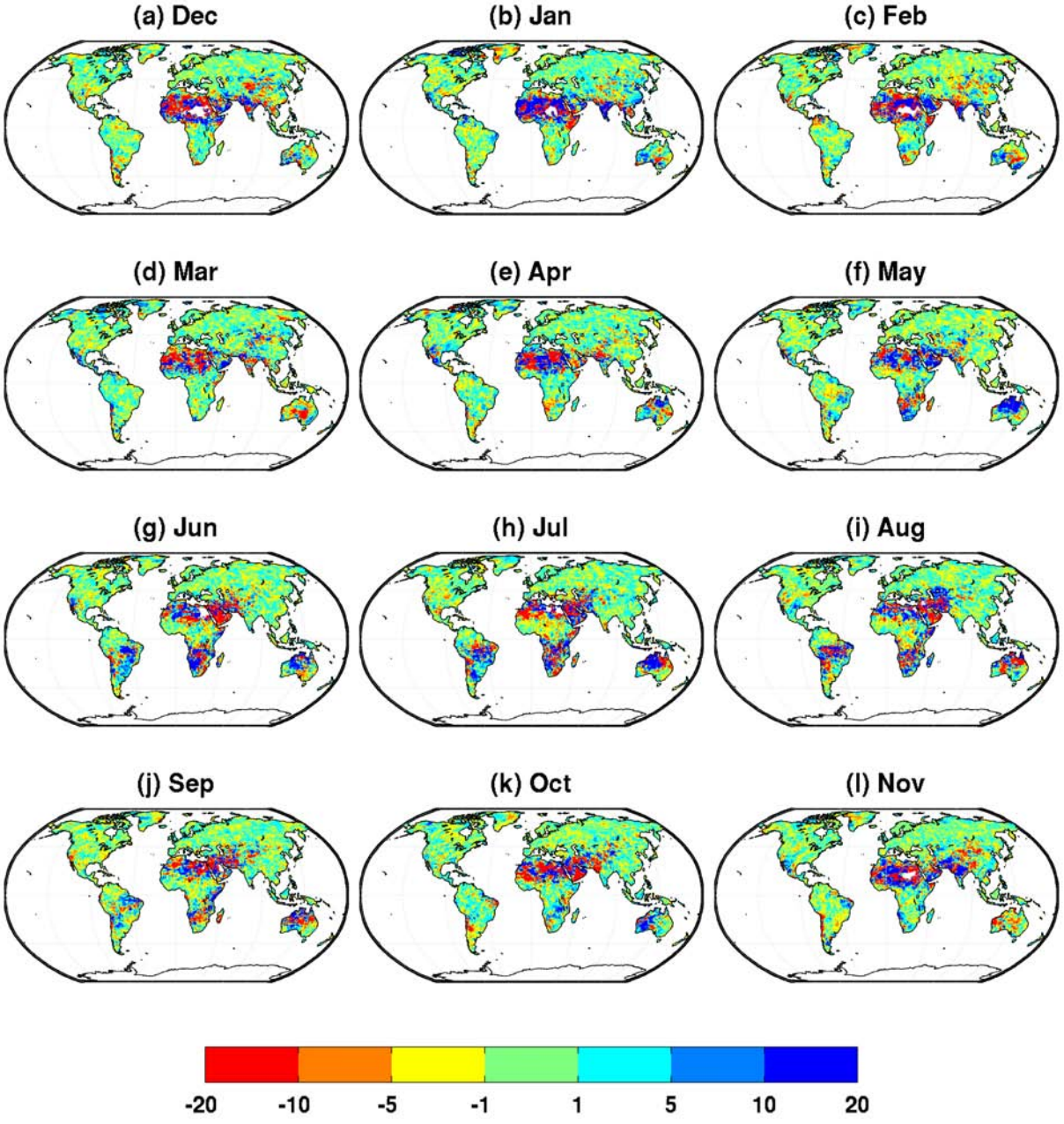

Figure 4.10: Bias (percentage) between monthly mean precipitation GPCC observations and cross-validated estimate based on CCA-PCR (10PCs).

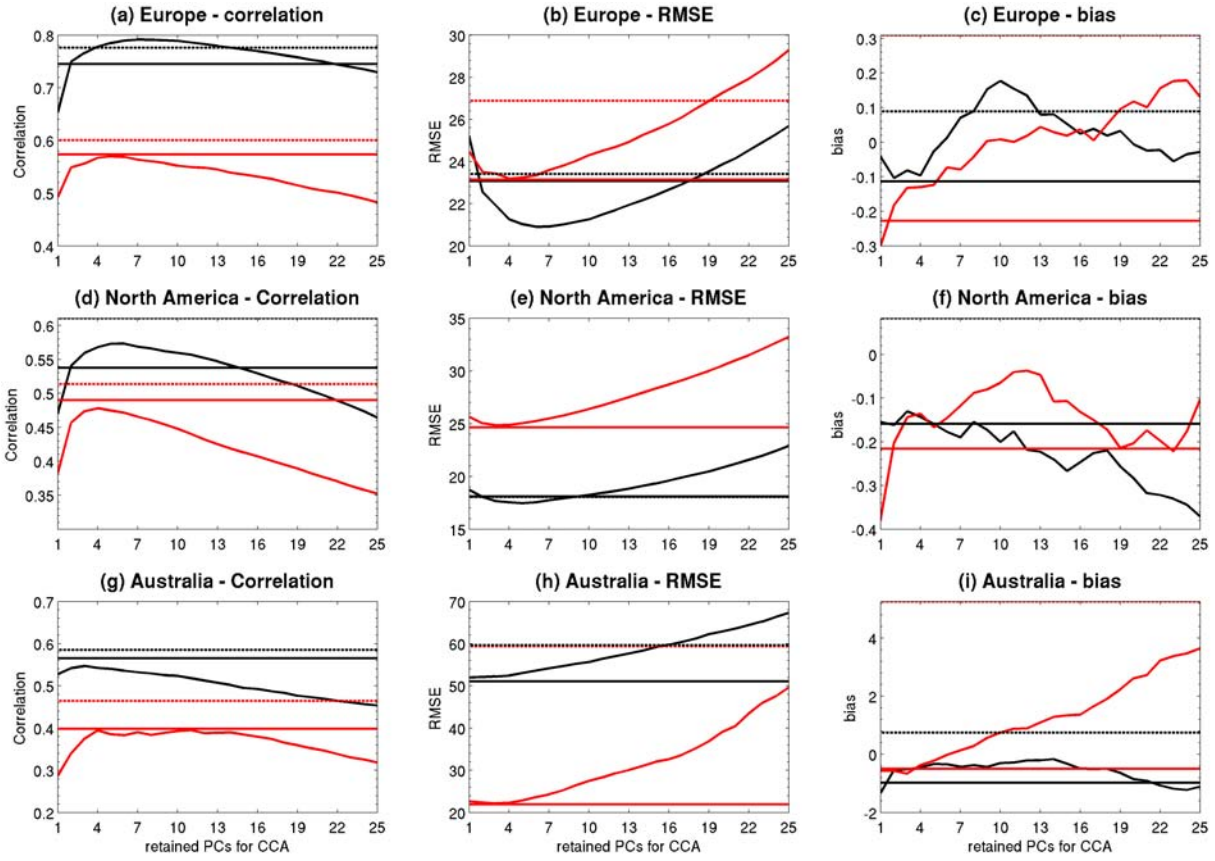

Figure 4.11: January (black lines) and July (red lines) correlation, RMSE (mm) and bias statistics (mm) for different downscaled corrections in three different regions. Values shown are averages of grid-cell statistics. Horizontal lines represent local scaling (dashed lines) and SVD-RM (solid lines) corrections; varying lines represent CCA-PCR corrections with different numbers of retained PCs. Note: local scaling RMSE (July only) and bias statistics are larger than the scale used.

fers a stronger mean correlation than SVD-RM in all case, and also outperforms CCA-PCR in North America and Australia during both January and July (Figure 4.11a,d,g). In Europe, CCA-PCR is able to outperform local scaling and SVD-RM during January when an optimal number of PCs is retained. With CCA-PCR, correlation becomes stronger with an increasing number of retained PCs up to a certain point, which is between 4 and 7 PCs across all regions. Retaining more than 7 PCs does not improve overall correlation and the model is at risk from overfitting. The optimal number of PCs appears smaller during the summer months when predictability in each region is generally poorer.

In terms of RMSE statistics, SVD-RM outperforms local scaling in all cases, which is a reversal of the apparent skill in representing temporal variability (Figure 4.11b,e,h). An optimal range of PCs to retain for CCA-PCR is again evident, with fewer PCs required during July. General conclusions are more difficult to draw from bias statistics (Figure 4.11c,f,i), although bias in the SVD-RM correction is generally smaller than for local scaling. This is especially the case in North America and in Australia during July when large errors in local scaling estimates appear to distort the mean statistics. For CCA-

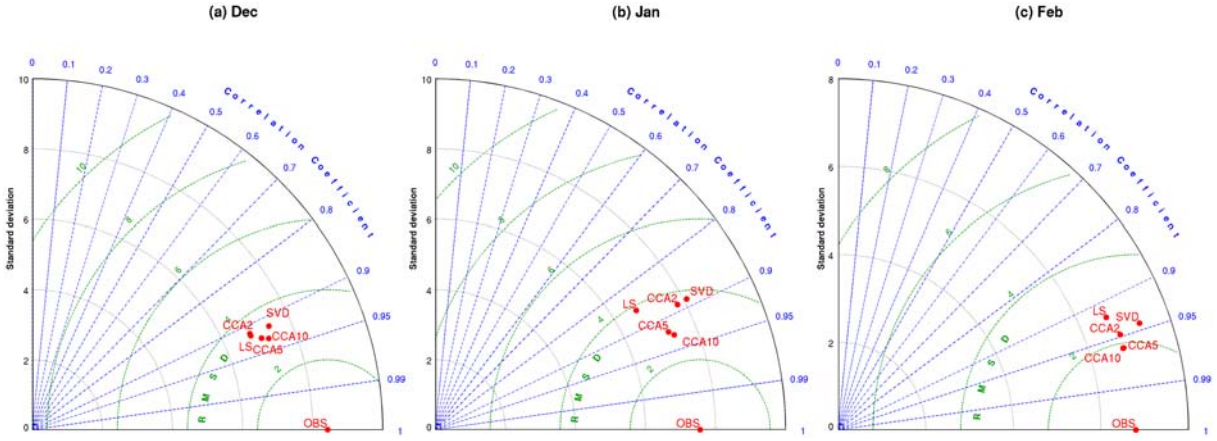

Figure 4.12: Taylor diagrams showing performance of local scaling (LS), SVD-RM and CCA-PCR (with 2, 5 and 10 retained PCs) methods to estimate European local-scale winter (DJF) precipitation. Points plotted correspond to correlation of estimated and observed time series and the standard deviation. Observed (OBS) precipitation is plotted as a reference.

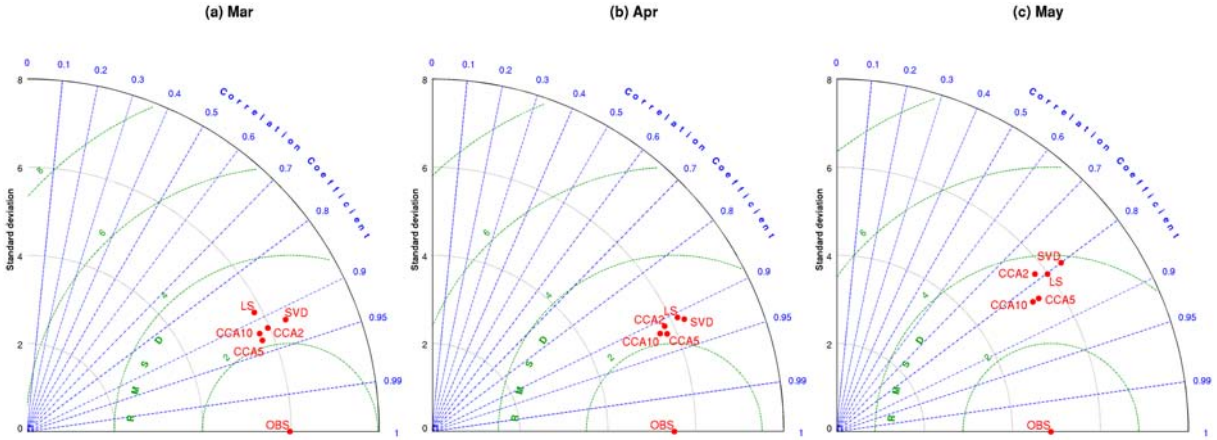

Figure 4.13: As Figure 4.12 but for spring (MAM).

PCR, optimisation in the number of PCs to retain is more difficult, although a similar range to that noted for correlation and RMSE is apparent (between approximately 4 and 7 PCs).

It is useful to further compare downscaled precipitation from each method and how closely they match observations. Taylor diagrams (Taylor, 2001) provide a visualisation of the similarity between two patterns, expressed in terms of their correlation, root-mean squared error and standard deviation. Here, Taylor diagrams are used to quantify the relative skill of the different downscaling methods in estimating seasonal European precipitation (Figures 4.12-4.15). It should also be noted that the statistics in these diagrams are calculated directly from European monthly mean precipitation, rather than as an average of grid-point correlation and bias statistics which are used in Figure 4.11.

In the majority of months, points representing precipitation estimated using the CCA-PCR method (with 2, 5 and 10 retained PCs respectively) lie nearest to the 'observed' point. These estimates show

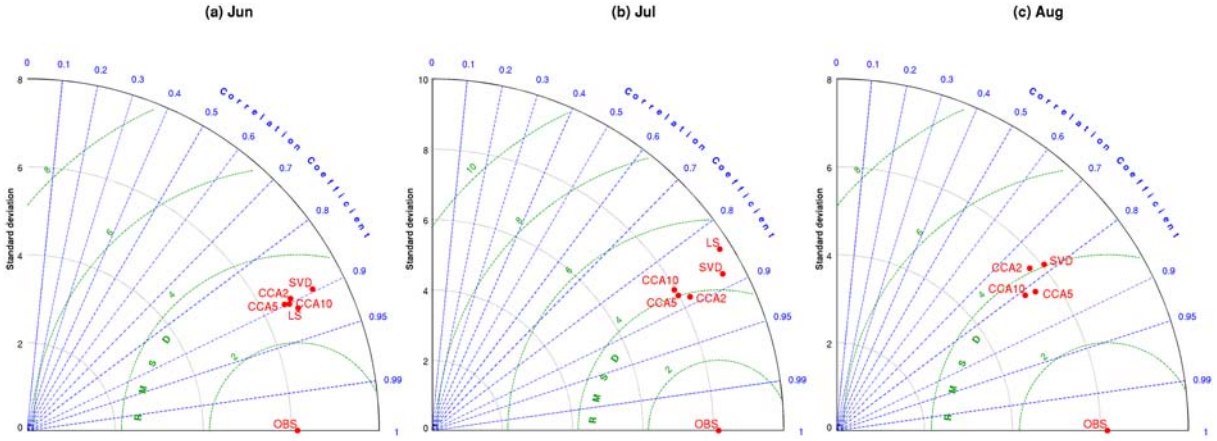

Figure 4.14: As Figure 4.12 but for summer (JJA). Note: local scaling estimate for August is very large and appears off the shown diagram.

strongest correlation with observations and lowest RMSEs, and this method would thus appear to exhibit the greatest skill. Additionally, CCA-PCR estimates show largely similar standard deviations to the observed time series, suggesting that the patterns in temporal variation are of the correct amplitude (Taylor, 2001). During winter, both the local scaling and SVD-RM methods produce estimates that correlate strongly with observations, but with smaller and larger standard deviations respectively (Figure 4.12). During spring and summer, the SVD-RM correction consistently over-estimates standard deviation, and thus has a greater RMSE than the CCA-PCR corrections despite exhibiting a similar correlation with observations (Figures 4.13 and 4.14). This over-estimation is curious; any regression method would be expected to represent less variance than is apparent in the real world. It is important to note that the estimates presented here have been constructed using a rigorous cross-validation that is likely to account for the over-estimation. Local scaling is also associated with an over-estimation of variance between April and September, but standard deviation is better represented from October onwards (Figure 4.15). During spring and summer, SVD-RM precipitation estimates have consistently stronger correlation with observations and smaller RMSEs compared to local scaling estimates. However, the SVD-RM estimate shows far greater temporal variability than observations during summer (greater than 9 mm for July; Figure 4.14). Spatial variability is more accurately estimated by SVD-RM method during Autumn but is still out-performed by CCA-PCR (Figure 4.15).

Analysis of average statistics provides an indicator of downscaling model skill but the global correlation plots in Figures 4.1-4.10 clearly demonstrate the need to consider each region individually. This is obviously of particular importance when assessing the ability of a downscaled model to resolve precipitation processes in regions of complex topography or land-sea contrasts, or in areas where other downscaling methodologies (both dynamical and statistical) have previously been shown to be lacking in skill. In-depth analysis of MOS downscaling potential in Europe, North America and Australia will thus follow

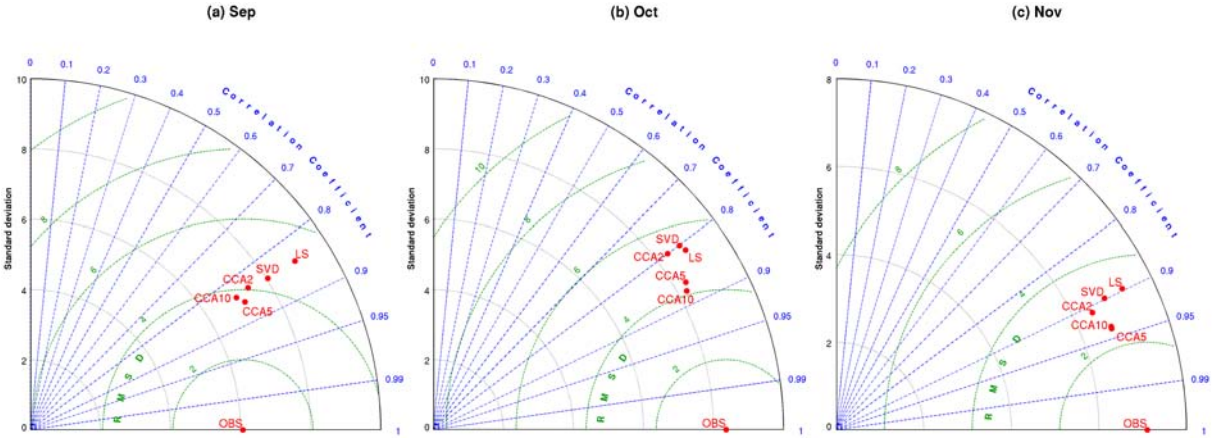

Figure 4.15: As Figure 4.12 but for autumn (SON).

in section 4.4.

4.3.2 Perfect-Prog approach

A number of observed atmospheric fields were used as predictor variables for local-scale precipitation as part of a Perfect-Prog framework. Geopotential height, temperature and specific and relative humidity at the 1000-hPa, 850-hPa and 500-hPa levels were taken from the ERA-40. Variables are used as single predictors and as combinations of two predictors. As with the formulation of MOS downscaling models, a leave-seven-out cross-validation technique is used to estimate precipitation between 1958-2001, with correlation and bias statistics calculated between the estimated series and the observed record. For brevity, tables and figures in this section illustrate results for January and July, which can be considered representative of the amplitude of intra-annual variability.

Table 4.1 provides a summary of mean correlation statistics for all variables. It is immediately apparent that the skill of each predictor is season- and method-dependent. CCA-PCR would appear to offer slightly more predictive skill overall than SVD-RM for most variables, and particularly for geopotential height and temperature. It is notable that, whilst for the MOS downscaling models the optimal number of principal components (PCs) to retain for inclusion in a multiple linear regression model was between 4 and 7, for the majority of Perfect-Prog models the optimal number is around 10. The larger spatial domain used for non-local Perfect-Prog compared with MOS downscaling models means a covariance matrix is constructed with information from a larger number of grid points (i.e. individual time series; considered 'variables' in this case). It is unsurprising, therefore, that a larger number of PCs is required to produce the optimal level of model skill. The SVD-RM method appears better suited for those models based on atmospheric moisture predictors. Whilst crude, the summarised results in Table 4.1 help justify the choice for detailed graphical representation of the best performing variables.

Table 4.1: Summary correlation statistics for cross-validated correction of January and July precipitation using varying predictors and downscaling methods. Statistics shown are global mean of local correlations (standard font) and percentage of the globe with correlation greater than 0.5 and 0.7 (in parentheses) respectively (both shown in italics). The predictors are geopotential height (Z), temperature (T), specific (q) and relative humidity (rh) at the 1000-hPa, 850-hPa and 500-hPa levels. Methods are SVD-RM and CCA-PCR with 2 and 10 retained PCs.

	January			**July**		
	SVD-RM	CCA-PCR(2)	CCA-PCR(10)	SVD-RM	CCA-PCR(2)	CCA-PCR(10)
Z_{1000}	0.157 *9.56 (1.24)*	0.123 *11.37 (2.43)*	0.347 *27.43 (7.54)*	0.151 *4.92 (0.18)*	0.089 *5.78 (0.235)*	0.274 *16.15 (2.28)*
Z_{850}	0.157 *10.05 (1.58)*	0.126 *12.45 (2.61)*	0.338 *27.44 (6.86)*	0.162 *6.20 (0.22)*	0.106 *6.07 (0.29)*	0.280 *17.94 (2.63)*
Z_{500}	0.153 *9.87 (1.74)*	0.091 *10.55 (2.05)*	0.276 *19.20 (4.00)*	0.165 *6.07 (0.24)*	0.106 *5.31 (0.29)*	0.242 *12.42 (0.82)*
T_{1000}	0.205 *9.97 (0.72*	0.151 *13.65 (1.25)*	0.301 *24.80 (3.78)*	0.199 *5.57 (0.19)*	0.112 *5.34 (0.24)*	0.242 *12.88 (0.52)*
T_{850}	0.171 *7.61 (0.39)*	0.130 *8.45 (18.34)*	0.273 *18.34 (2.19)*	0.187 *5.96 (0.18)*	0.115 *5.20 (0.23)*	0.224 *11.20 (0.46)*
T_{500}	0.124 *6.41 (0.62)*	0.071 *6.65 (0.74)*	0.210 *12.81 (1.40)*	0.121 *3.28 (0.01)*	0.067 *3.89 (0.11)*	0.164 *6.06 (0.15)*
q_{1000}	0.240 *14.58 (1.12)*	0.179 *15.05 (1.68)*	0.314 *24.94 (4.81)*	0.155 *4.30 (0.05)*	0.067 *3.43 (0.07)*	0.200 *9.49 (0.27)*
q_{850}	0.227 *14.25 (0.93)*	0.170 *13.81 (1.35)*	0.303 *23.40 (3.73)*	0.166 *4.30 (0.06)*	0.074 *4.09 (0.10)*	0.192 *7.87 (0.24)*
q_{500}	0.190 *10.97 (1.00)*	0.130 *10.58 (1.14)*	0.242 *17.94 (2.09)*	0.147 *4.08 (0.03)*	0.063 *3.490 (0.07)*	0.177 *7.49 (0.19)*
rh_{1000}	0.210 *13.14 (0.86)*	0.105 *10.20 (0.82)*	0.246 *18.56 (2.64)*	0.200 *10.65 (0.40)*	0.096 *7.12 (0.27)*	0.238 *15.47 (1.10)*
rh_{850}	0.216 *14.55 (1.07)*	0.119 *12.41 (0.94)*	0.255 *19.46 (2.62)*	0.222 *9.08 (0.16)*	0.134 *6.81 (0.16)*	0.239 *12.58 (0.68)*
rh_{500}	0.131 *8.93 (0.68)*	0.044 *7.35 (0.68)*	0.175 *11.79 (1.05)*	0.149 *4.38 (0.02)*	0.048 *2.51 (0.02)*	0.175 *8.85 (0.20)*
Z_{1000} T_{1000}	0.248 *16.95 (1.74)*	0.106 *11.52 (1.36)*	0.359 *31.25 (7.37)*	0.234 *9.57 (0.32)*	0.099 *5.44 (0.19)*	0.271 *16.40 (1.32)*
q_{1000} T_{1000}	0.248 *14.49 (1.22)*	0.167 *14.49 (1.22)*	0.331 *27.69 (5.47)*	0.219 *8.42 (0.23)*	0.087 *5.76 (0.30)*	0.254 *15.38 (0.82)*
Z_{1000} q_{1000}	0.264 *18.18 (2.16)*	0.139 *12.46 (1.53)*	0.358 *30.78 (7.85)*	0.217 *8.44 (0.28)*	0.083 *5.16 (0.21)*	0.249 *14.60 (1.26)*
Z_{1000} q_{850}	0.260 *18.95 (2.04)*	0.139 *13.60 (1.27)*	0.363 *31.14 (7.61)*	0.221 *8.34 (0.22)*	0.086 *5.56 (0.19)*	0.261 *14.43 (1.21)*

The presentation of results in this section is designed to describe the seasonal and spatial distribution of downscaling model skill and thus analysis at this stage is confined to an interpretation of correlation and bias statistics. Focus is initially given to downscaling models constructed using single predictor variables, with these results contributing to the formulation of models with paired predictors. CCA-PCR models constructed using 2 and 10 PCs only are shown in the following figures.

4.3.2.1 Single predictors

In terms of correlation, and thus representation of temporal variability, skill is generally highest in Europe and North America during the winter months (inferred from January) that are characterised by greater rainfall. For European precipitation, the most powerful predictor is geopotential height at 1000-hPa (Z_{1000}) and to a lesser extent at 850-hPa (Z_{850}) (in Figure 4.16). Specific humidity at 850-hPa (q_{850}; Figure 4.17) and temperature at 1000-hPa (T_{1000}; Figure 4.18) shows good predictive power across much of southern and Mediterranean Europe. Relative humidity at higher levels (850-hPa and 500-hPa) shows reasonable skill across parts of western and central Europe and also across parts of North America, particularly west of the Rocky Mountains (Figure 4.19). Temperature and geopotential height (at 1000-hPa and 850-hPa) also offer sound skill in these areas during January, although skill is noticeably poor for all predictors across much of the eastern USA and Canada. This is surprising given a dense rain gauge network and completeness of station records. Additionally, the good performance of MOS downscaling models in this region is similar to that in Europe, a pattern not evident amongst Perfect-Prog predictors.

For a number of predictors, correlation is also strong across much of Australia, southern Africa and parts of eastern South America during the Austral summer. Parts of Asia and eastern Africa also show reasonable skill although with limited consistency in terms of seasonality. Specific and relative humidity exhibit good skill across much of northern Australia during the wetter summer months (Figures 4.17 and 4.19), and temperature at 1000-hPa and 850-hPa (T_{1000} and T_{850}) also shows reasonable, if not outstanding, skill (Figure 4.18). Rainfall during these months is largely convective and related to variations in atmospheric moisture and temperature as a result of the southward migration of the Intertropical Convergence Zone (ITCZ). One might expect some relationship between precipitation and geopotential height at this time of year to reflect the movement of the trough of low pressure associated with the ITCZ. In this case, it is unlikely that size and orientation of the domain size used to conduct non-local estimates for local precipitation is optimal. The temperature and humidity predictors are also associated with strong correlation in southern Africa during the summer months (DJF), and q_{500} shows decent predictive power for parts of north-eastern Brazil.

In the Northern Hemisphere, precipitation estimates during July (i.e. indicative of summer) are generally associated with weaker correlation, a finding consistent with the validation of MOS models

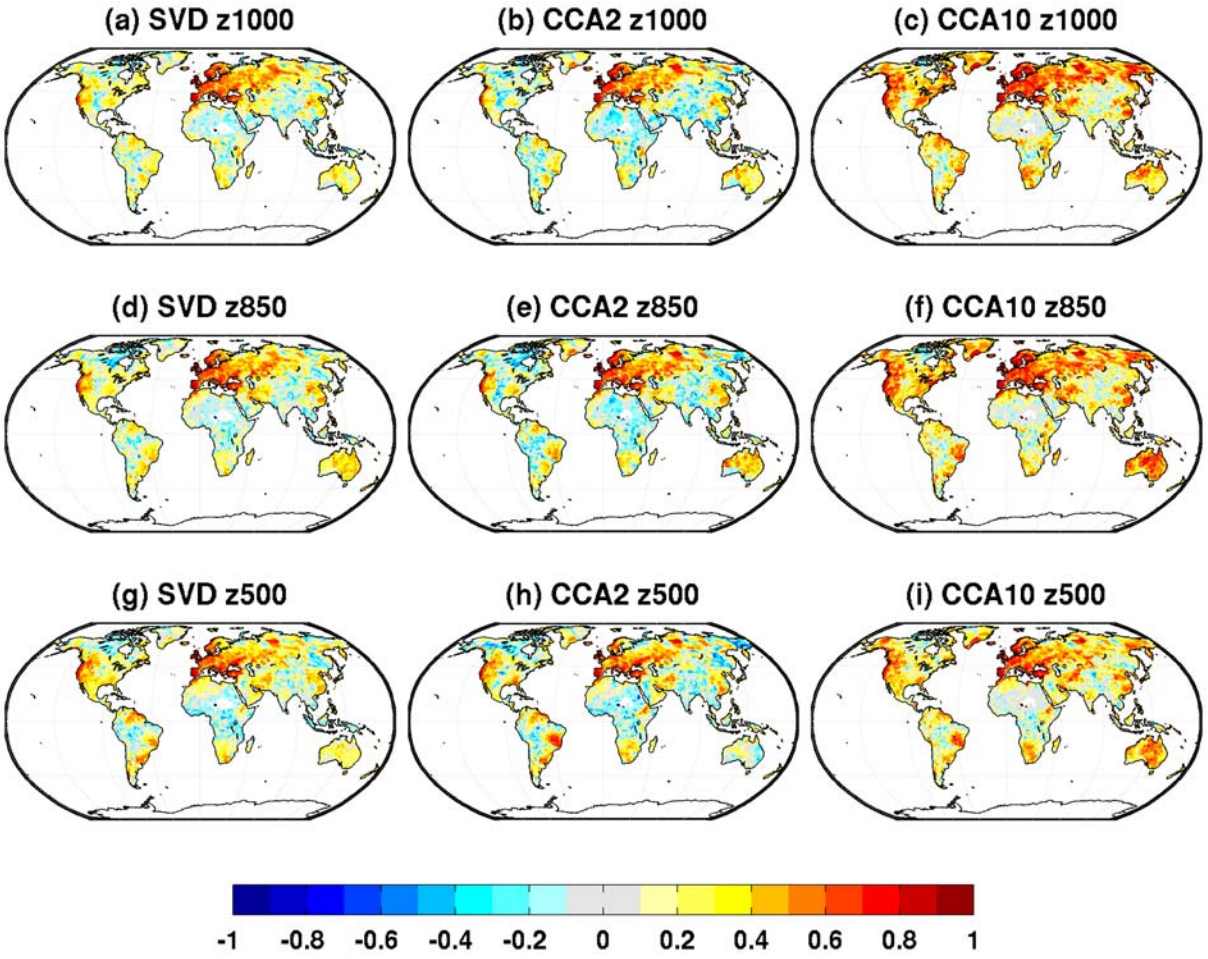

Figure 4.16: Correlation between January mean precipitation GPCC observations and cross-validated downscaled estimates using geopotential height as a predictor.

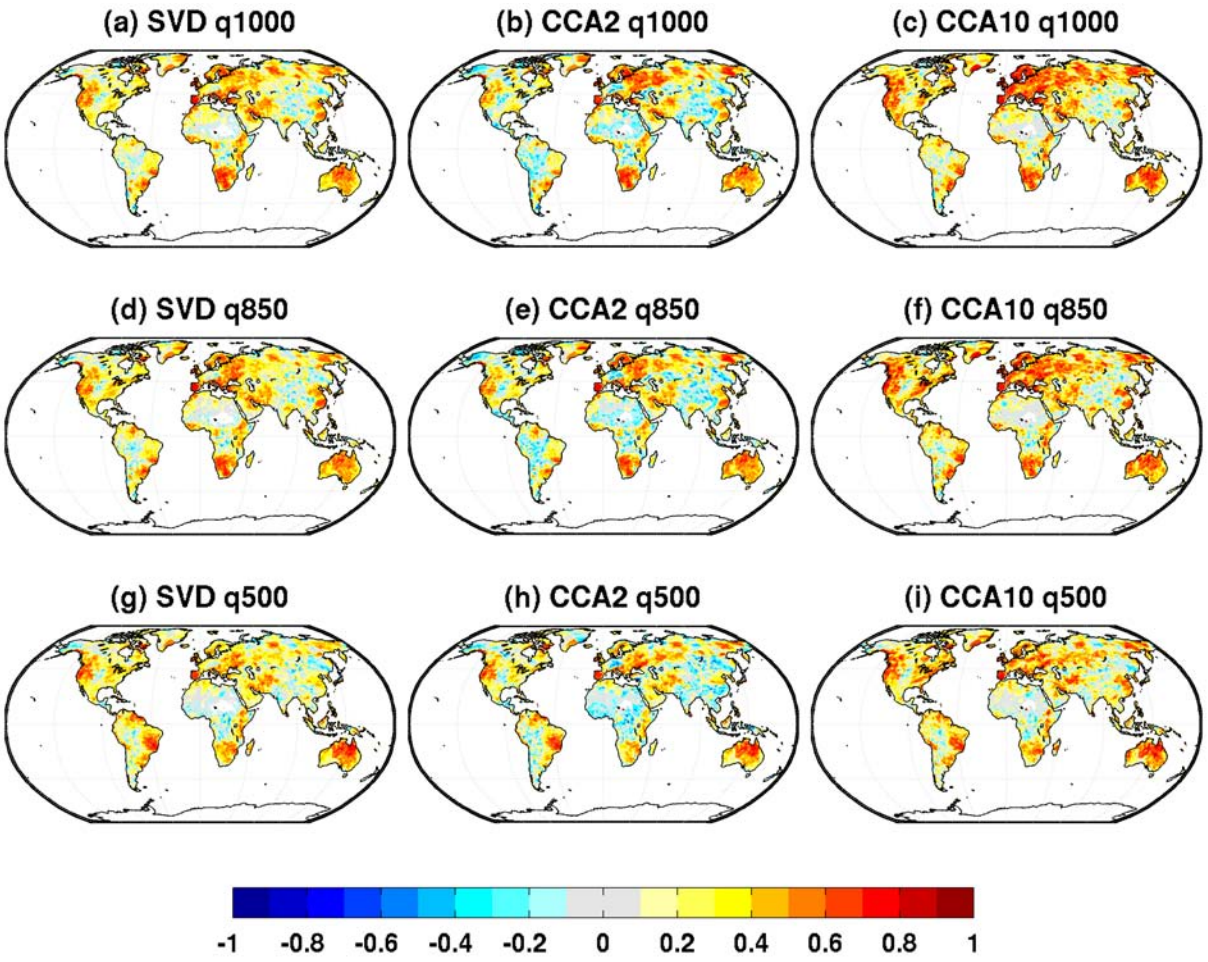

Figure 4.17: As Figure 4.16 but using specific humidity as a predictor.

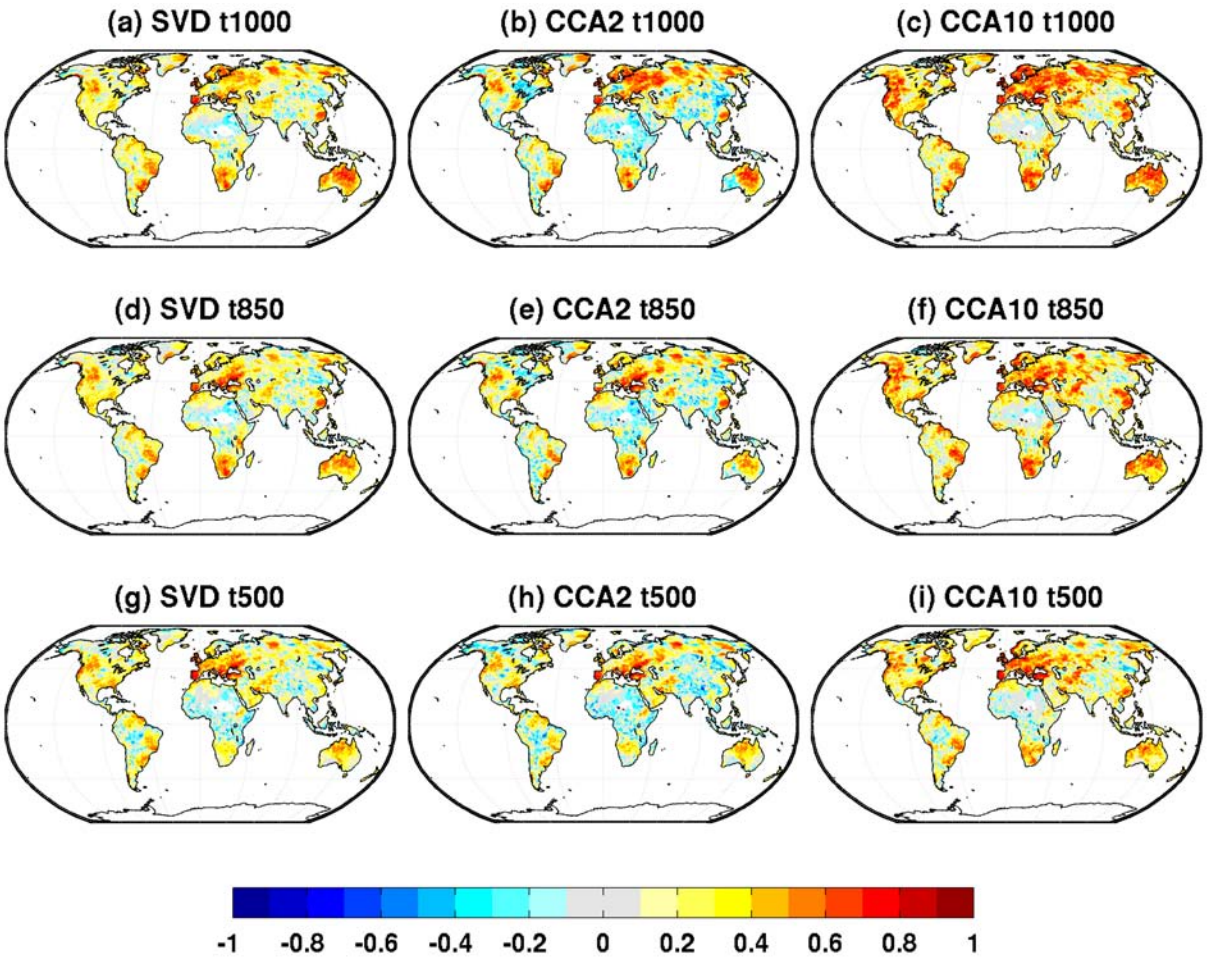

Figure 4.18: As Figure 4.16 but using temperature as a predictor.

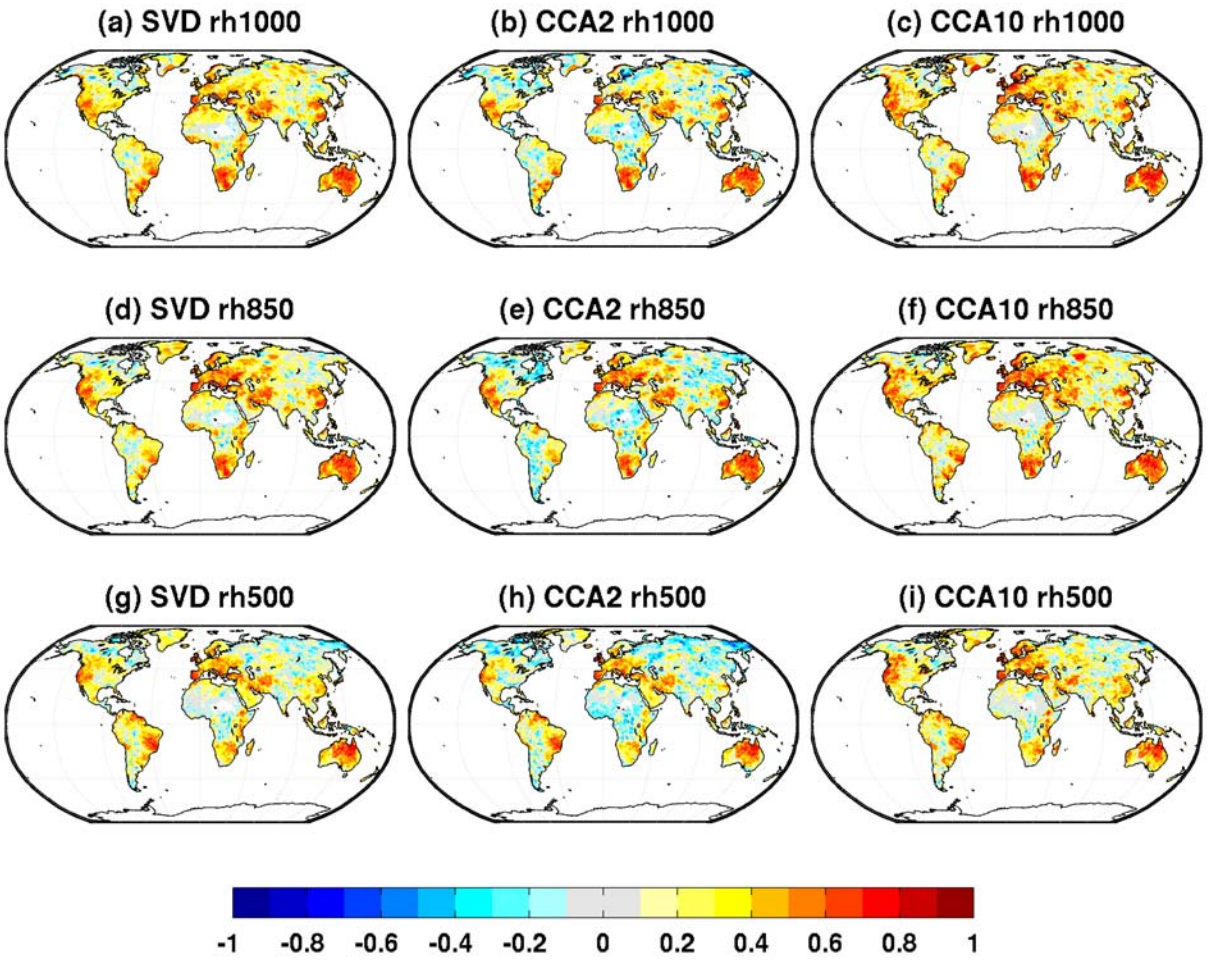

Figure 4.19: As Figure 4.16 but using relative humidity as a predictor.

in section 4.3.1 and in previous work in downscaling European and North American precipitation. In northern Europe, geopotential height and, to a lesser extent, temperature at 1000-hPa and 850-hPa show decent predictor skill (Figures 4.20 and 4.22), but southern Europe precipitation appears difficult to predict. Similarly, correlation amongst all variables is sporadic in North America, although relative humidity does show some skill in the Pacific north-west. In the Southern Hemisphere, there is evidence of greater predictability, geopotential height at 850-hPa showing reasonably strong correlation across the southern fringe of Australia (Figure 4.20). Specific and relative humidity also offer decent skill in other parts of the Australian continent, although there is little consistency as to which particular areas are best represented (Figures 4.21 and 4.23).

In general, long-term means are well resolved by the majority of downscaling models throughout the extra-tropics and in areas where correlation is high. In the world's tropical regions, long-term bias can be significant, particularly across Africa and South America. Global biases for each individual variable at 1000-hPa during January and July are shown in Figures 4.24 and 4.25 respectively. Bias in precipitation estimated from CCA-PCR with 2 retained PCs is similar to that estimated from SVD-RM. CCA-PCR developed with 10 PCs appears to induce over-estimation of January precipitation over large parts of Eurasia. Moreover, inconsistencies in spatial distribution of long-term bias are common in all CCA-PCR downscaling models which include large numbers of retained PCs. More attention is given to bias in downscaled estimates in section 4.4 for particular regions, where the effect of real-world orography is discussed in more detail.

There is little noticeable difference in correlation between methods on a global scale, although CCA-PCR is generally shown to outperform SVD-RM. Furthermore, CCA-PCR would appear to perform better for low-level predictor variables (e.g. T_{1000} in January; Figure 4.18). In most cases, there is little extra predictive skill that the SVD-RM method can offer over CCA-PCR with 2 retained PCs. Given that 10 is clearly a more appropriate number of PCs to retain, it stands to reason that pre-filtering is a necessary and worthwhile additional step in the CCA-PCR method. The relative benefits of each method is discussed in greater detail in section 4.4

For regions where downscaling models perform consistently well there is little difference in the overall skill between SVD-RM and CCA-PCR methods. Taylor Diagrams (Taylor, 2001) can be used again to reveal more information in regions where downscaling models perform consistently well (Figure 4.26). Estimates of precipitation made using CCA-PCR with 10 retained PCs are consistently more strongly correlated with observations than equivalent SVD-RM estimates and exhibit smaller RMSEs. Additionally, the precipitation time series reconstructed with SVD-RM are generally associated with an unrealistic standard deviation; CCA-PCR estimates are a more consistent match for observations in this sense. Six (four) CCA-PCR estimates for January (July) have a correlation greater than 0.8 and a RMSE less than

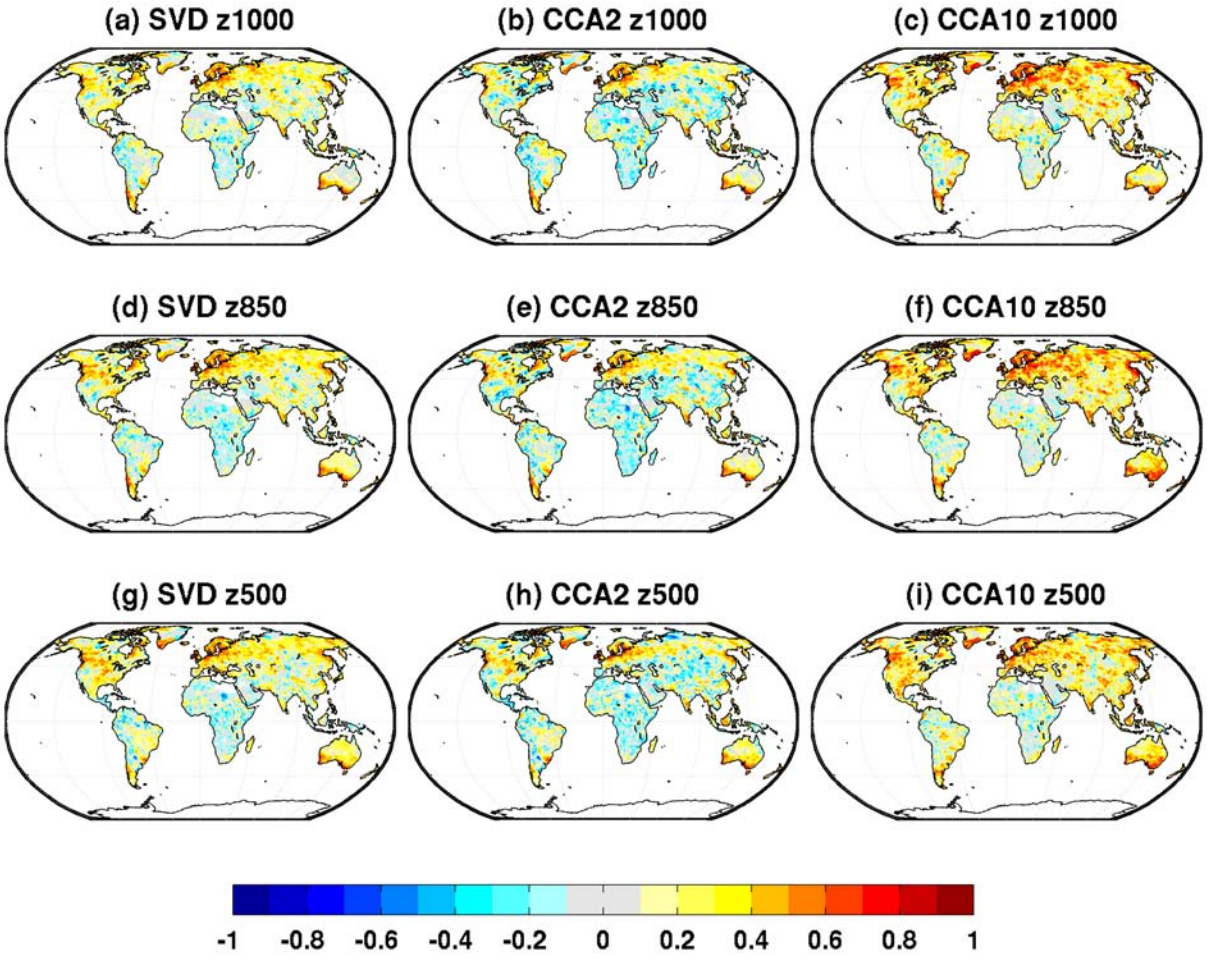

Figure 4.20: Correlation between July mean precipitation GPCC observations and cross-validated downscaled estimates using geopotential height at 1000-hPa, 850-hPa and 500-hPa as a predictor.

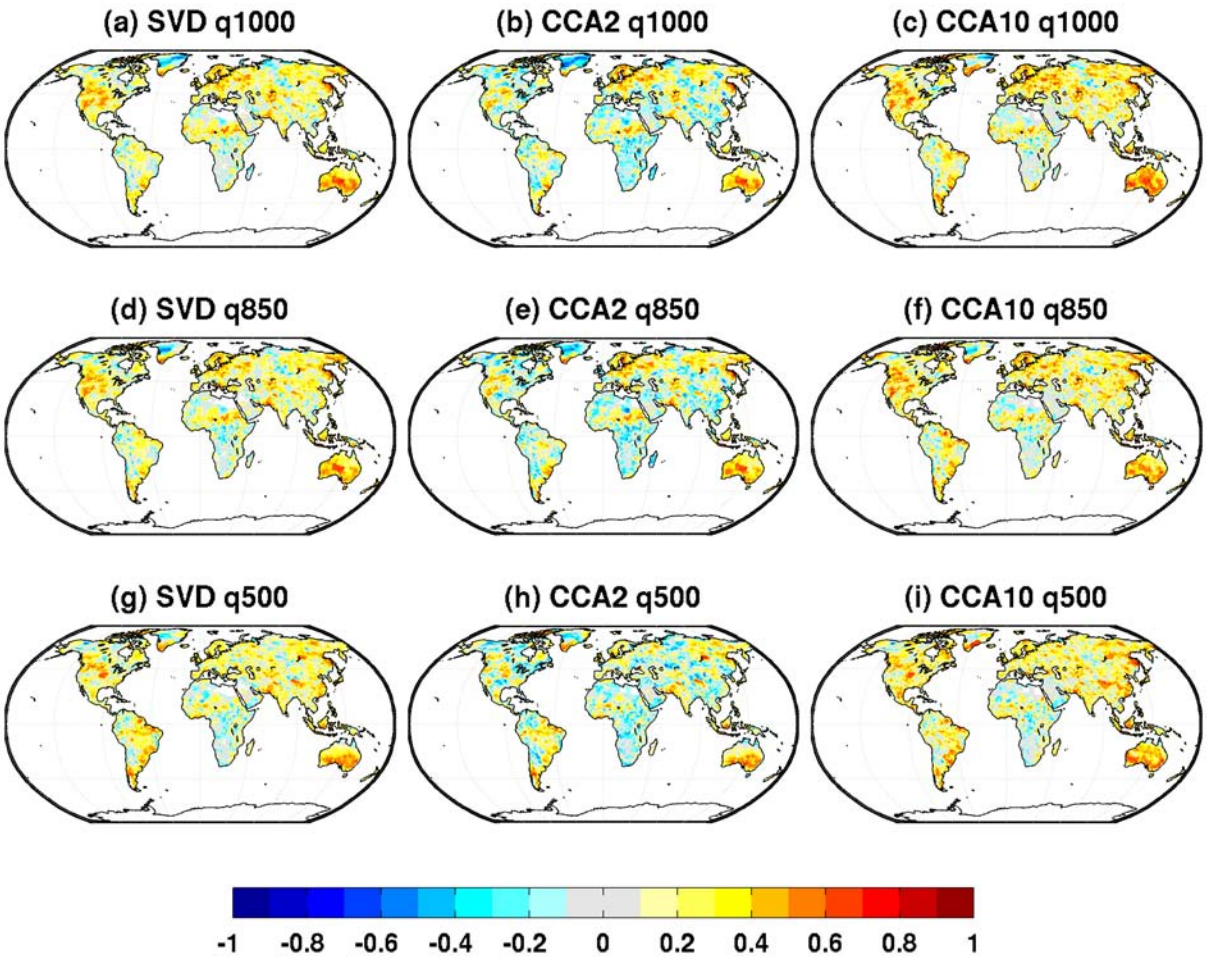

Figure 4.21: As Figure 4.20 but with specific humidity as a predictor.

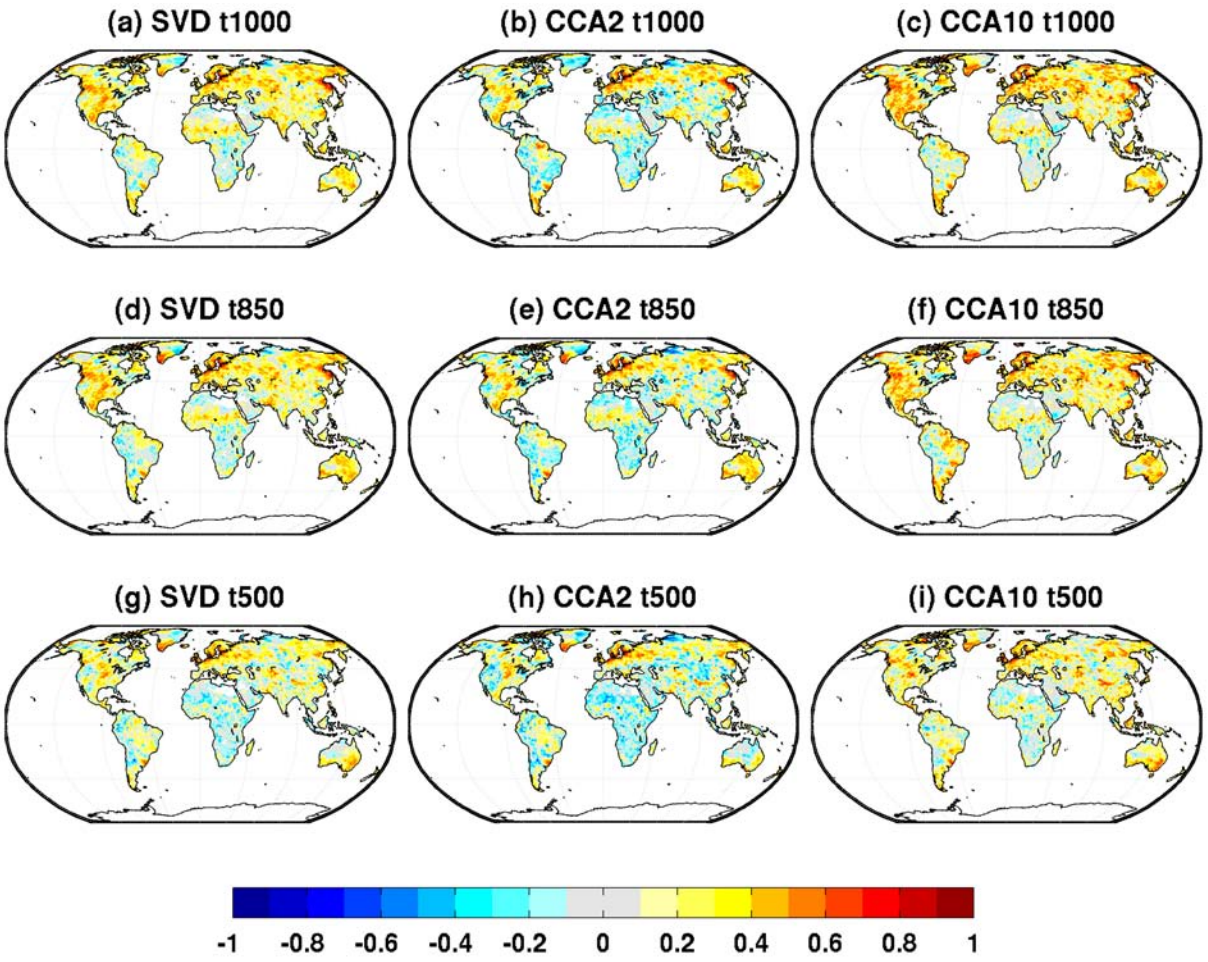

Figure 4.22: As Figure 4.20 but with temperature as a predictor.

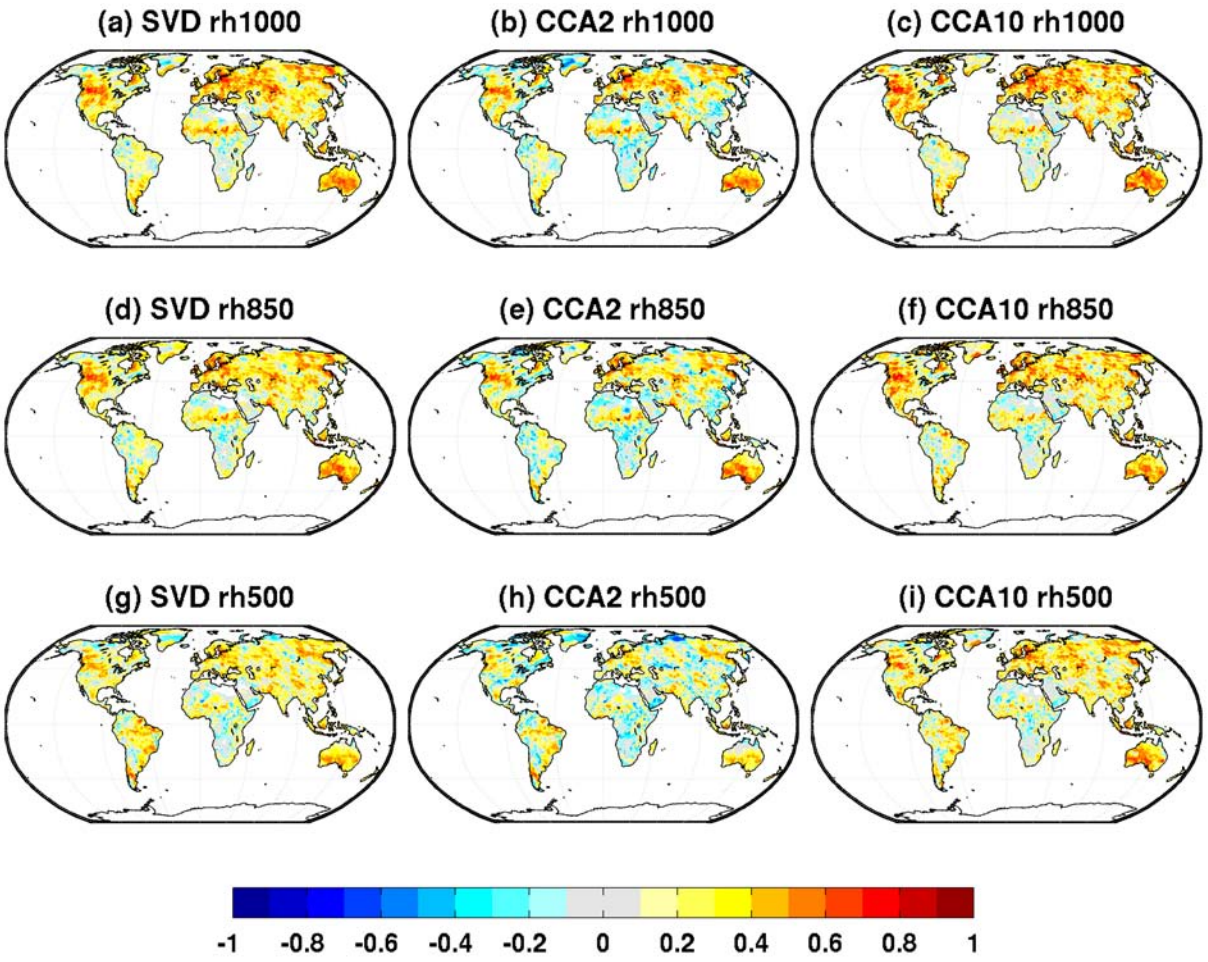

Figure 4.23: As Figure 4.20 but with relative humidity as a predictor.

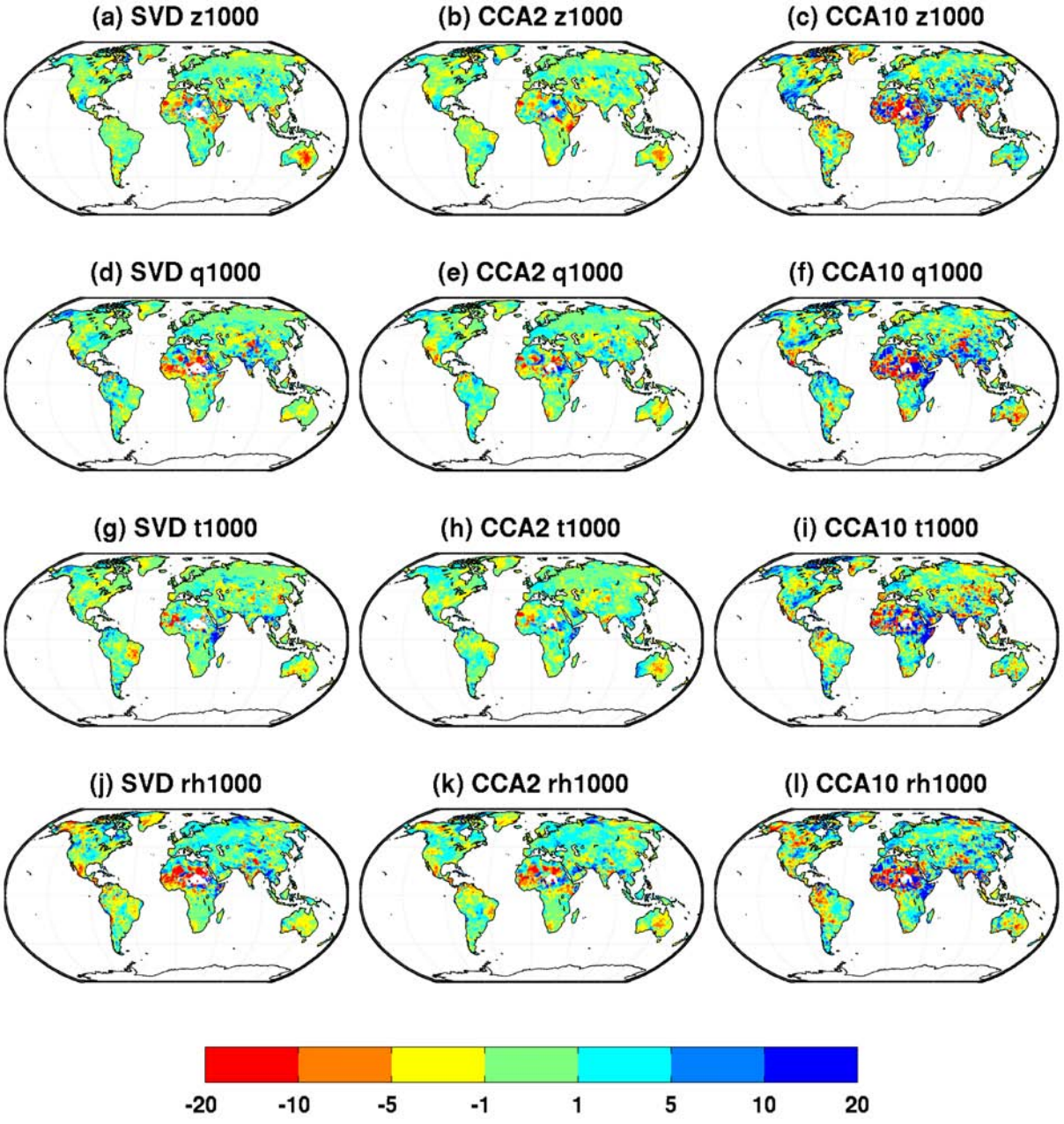

Figure 4.24: Bias (percentage) between January mean precipitation GPCC observations and cross-validated downscaled estimates using various predictors at 1000-hPa.

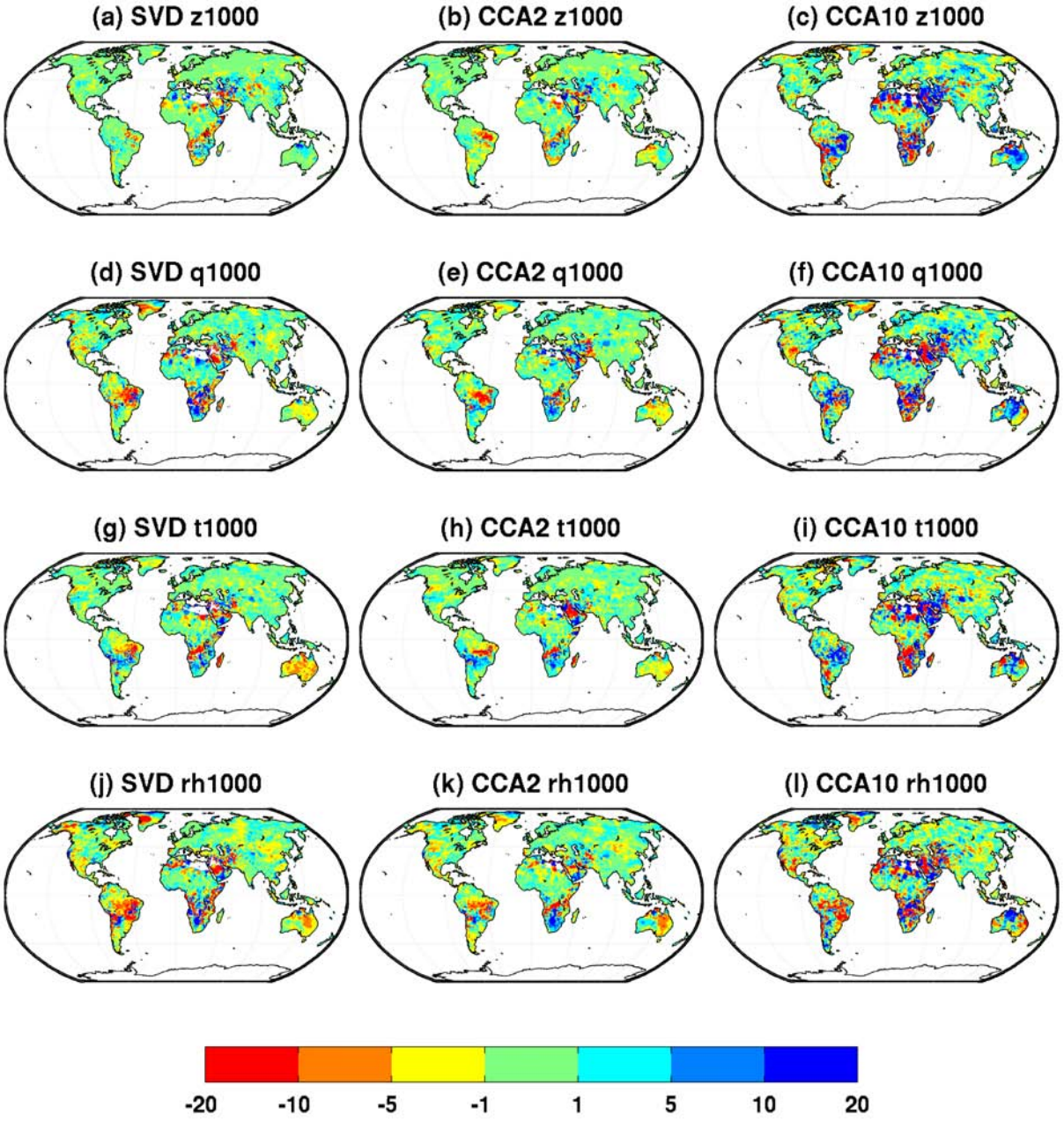

Figure 4.25: Bias (percentage) between July mean precipitation GPCC observations and cross-validated downscaled estimates using various predictors at 1000-hPa.

5mm per month compared, whereas no SVD-RM estimates meet these thresholds. More detailed comparisons are made in section 4.4, but inconsistencies in the optimal number of PCs to retain may count against the CCA-PCR method, in terms of its potential for global application. It is also worth reiterating that much work has focused on developing statistical downscaling models using predictors such as these on much smaller, and ultimately more specific, scales. The purpose here is for the Perfect-Prog downscaling models to act as a reference and to gauge the relative performance of MOS downscaling corrections across the globe.

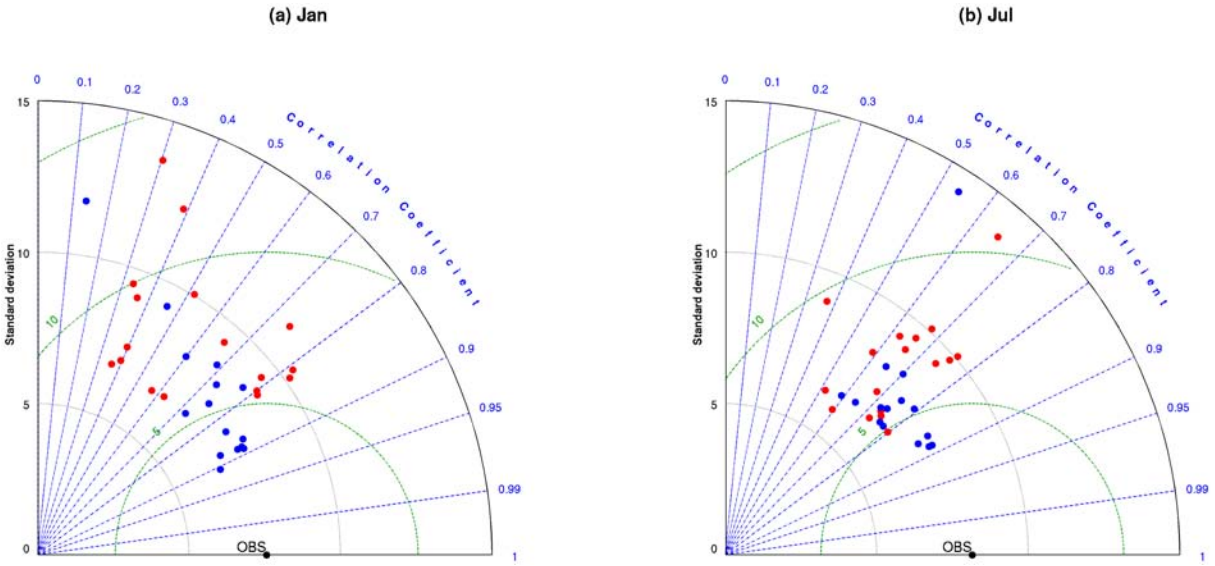

Figure 4.26: Taylor diagrams showing all SVD-RM and CCA-PCR (10 PCs) estimates from Perfect-Prog predictors, Europe only. Red = SVD-RM; Blue = CCA-PCR.

4.3.2.2 Paired predictors

SVD-RM and CCA-PCR downscaling models were also constructed with four paired combinations of predictors: Z_{1000}/T_{1000}, q_{1000}/T_{1000}, Z_{1000}/q_{1000} and Z_{1000}/q_{850}. Pairings were based on the strongest performing predictors (Table 4.1) when used in isolation. Relative humidity was not used in a pairing due to its dependence on temperature; specific humidity is instead used as a measure of atmospheric moisture. Correlation between downscaled estimates and observations is strong in a number of regions. During January, this includes the majority of Europe, western North America, southern Africa and parts of Australia (Figure 4.27). The spatial distribution in correlation strength is approximately similar amongst all methods, although CCA-PCR with 10 PCs outperforms SVD-RM on a global scale. The distribution of bias is similar to that amongst the most skilful solitary predictors; that is, a greater degree of spatial consistency in the SVD-RM and CCA-PCR(2) estimates than in the CCA-PCR(10) estimates (Figure 4.28). With all methods, there appears to be a greater over-estimation of precipitation in western

and central Europe, a region where bias is generally low across all other models. During July, correlation is again generally weaker with the only areas in the Northern Hemisphere showing skill limited to northern Europe, isolated patches of North America and southern Australia (Figure 4.29). Bias is again smallest in the SVD-RM and CCA-PCR(2) estimates.

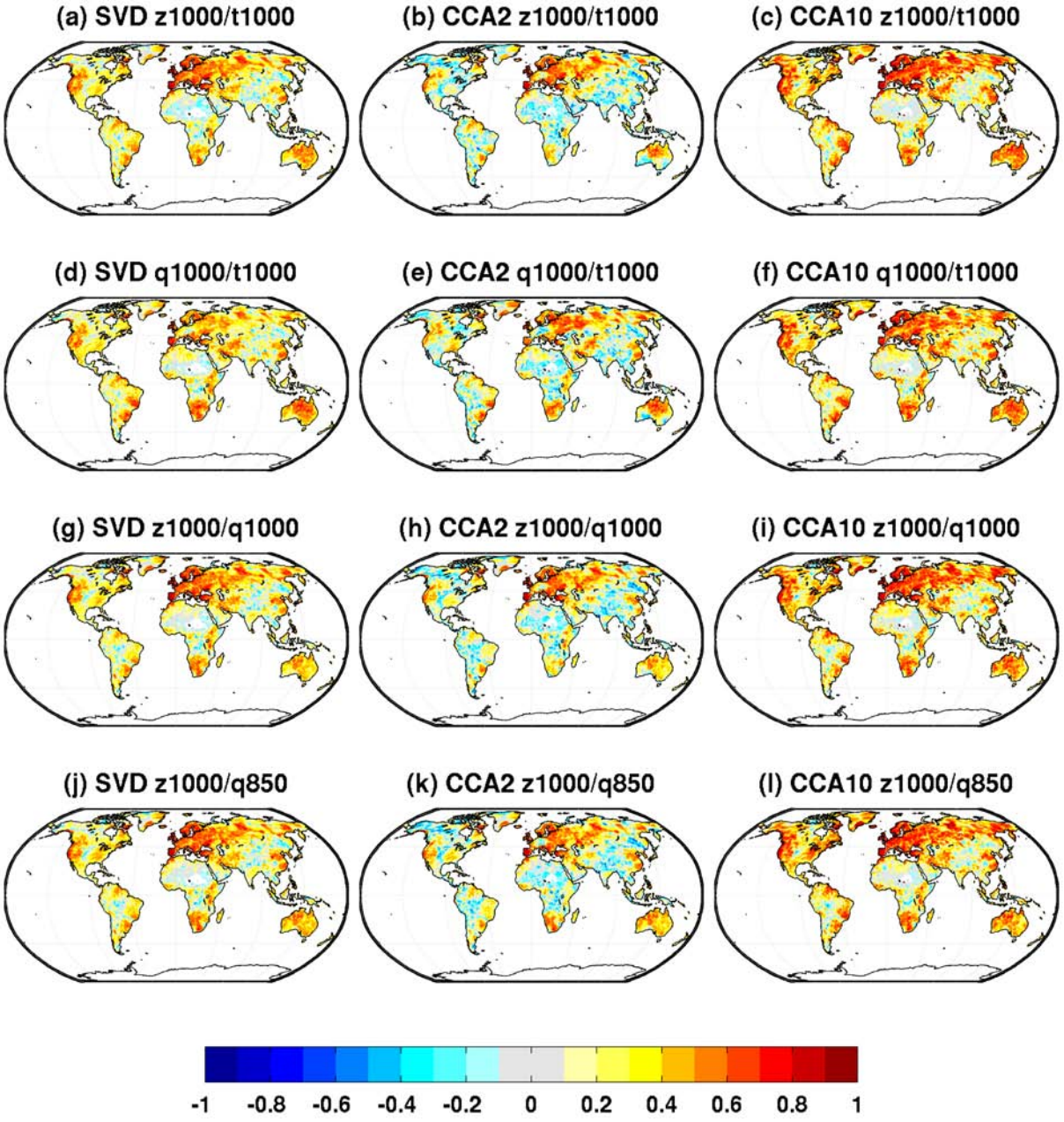

Figure 4.27: Correlation between January mean precipitation GPCC observations and cross-validated estimates with various paired combinations used as predictors.

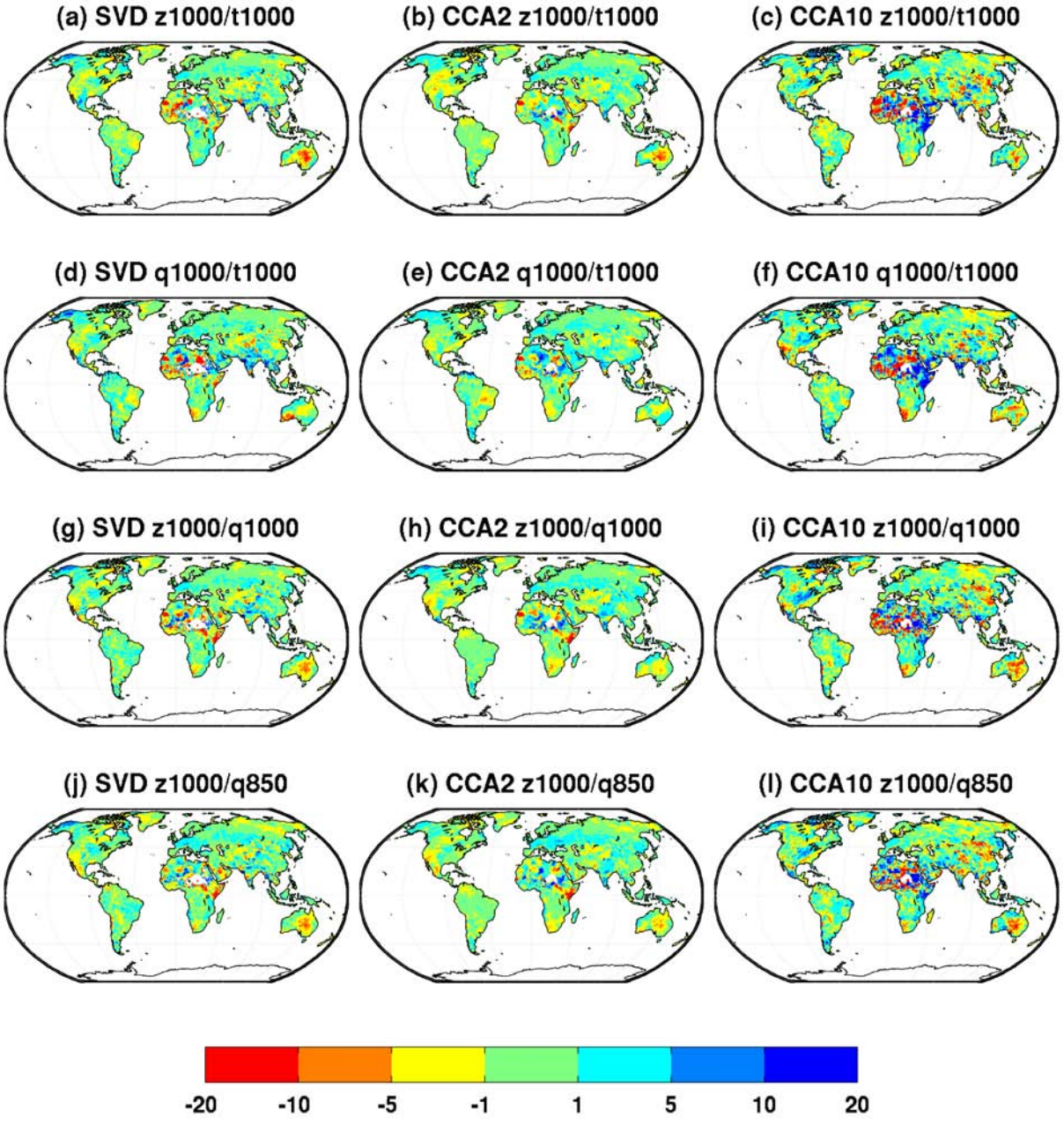

Figure 4.28: Bias (percentage) between January mean precipitation GPCC observations and cross-validated estimates with various paired combinations used as predictors.

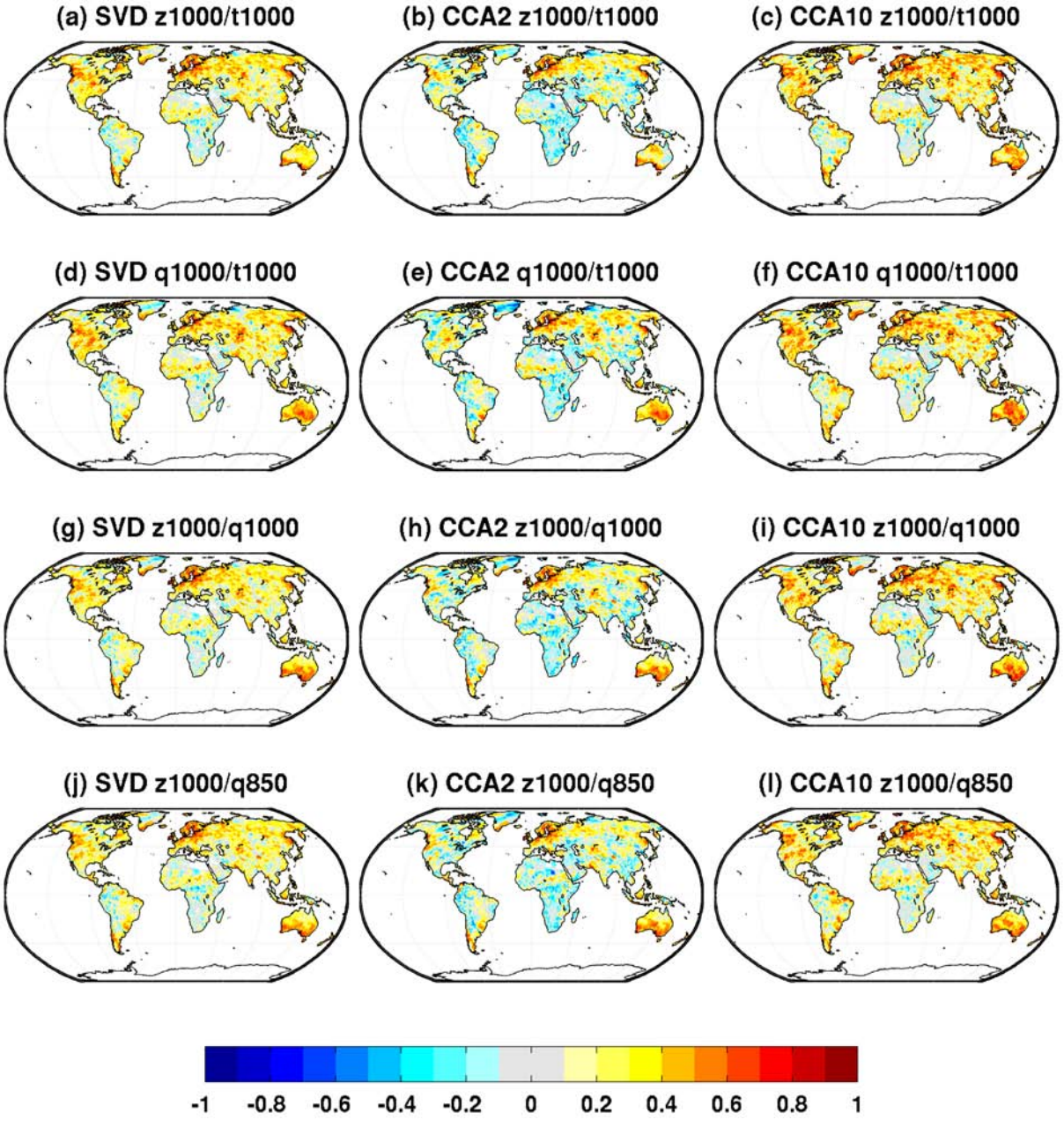

Figure 4.29: Correlation between July mean precipitation GPCC observations and cross-validated estimates with various paired combinations used as predictors.

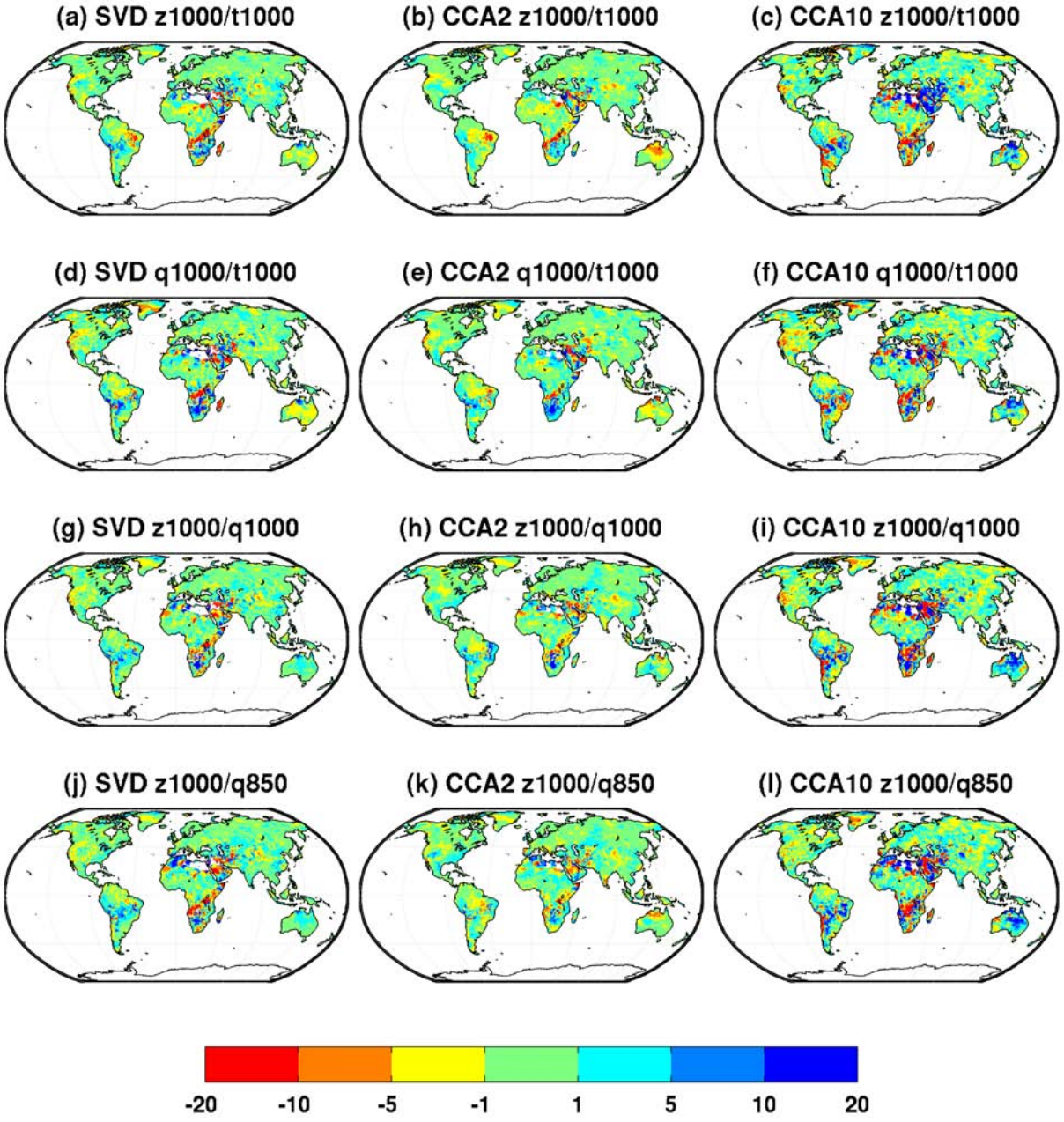

Figure 4.30: Bias (percentage) between July mean precipitation GPCC observations and cross-validated estimates with various paired combinations used as predictors.

4.4 Regional comparison of MOS and Perfect-Prog downscaling methods

Methods to estimate local-scale precipitation following both MOS and Perfect-Prog approaches have been presented in the previous section. The relative merits of each approach's methods are now discussed in detail. Amongst the main points to conclude from section 4.3 are that the skill of each downscaling method varies greatly across different regions. It is necessary, therefore, that interpretation of model skill considers individual regions separately. Clearly, many methods (in some cases all) lack any discernible skill over large parts of the globe. In most cases, this is limited to the tropics and otherwise areas where observation network quality is poor. However, it is possible to identify a number of regions where the predictability of precipitation is high enough that a number of downscaling models show good skill. Here, the relative skill of the downscaling models across Europe, North America and Australia is annotated and discussed in detail. It is important not only to understand where MOS downscaling is appropriate and skilful, but also to ask, "Where does MOS downscaling offer greater potential than Perfect-Prog approaches?" To date, the majority of downscaling research has focused on Europe and North America (Maraun et al., 2010), but given the potential global applicability of MOS downscaling methods it is possible that MOS can add value (even if only minor) to precipitation estimations in regions where downscaling has proved difficult.

In the remainder of this and subsequent chapters, the CCA-PCR MOS correction based on 5 retained PCs is used to make comparisons with local scaling and SVD-RM MOS corrections. The number of PCs is independent of location or season, although is intended to be have global applicability.

4.4.1 Europe

Taylor diagrams are used again to illustrate how closely the cross-validated mean European precipitation estimate from a number of methods matches observations. In this case, all downscaling models are labelled to enable the identification of the most skilful method/predictor combinations. In January, both MOS regression-based methods (henceforth P-SVD and P-CCA5), in addition to the local scaling estimate (P-LS), outperform Perfect-Prog estimates (Figure 4.31). CCA-PCR using Z_{1000} and q_{1000} and a combination of both are the strongest performing Perfect-Prog models. During July, P-LS and P-SVD again outperform all Perfect-Prog estimates, both in terms of correlation and RMSE, but overestimate the amplitude of variation in the observed record (Figure 4.32). This is further highlighted in the time series (right panel; top) where both P-SVD and P-LS greatly over- and underestimate on a number of occasions between 1958 and 2001. P-CCA5, on the other hand, offers both a strong positive correlation and a realistic representation of observed variance. The Z_{1000}, T_{1000}, q_{1000} and q_{1000}/T_{1000} CCA-PCR

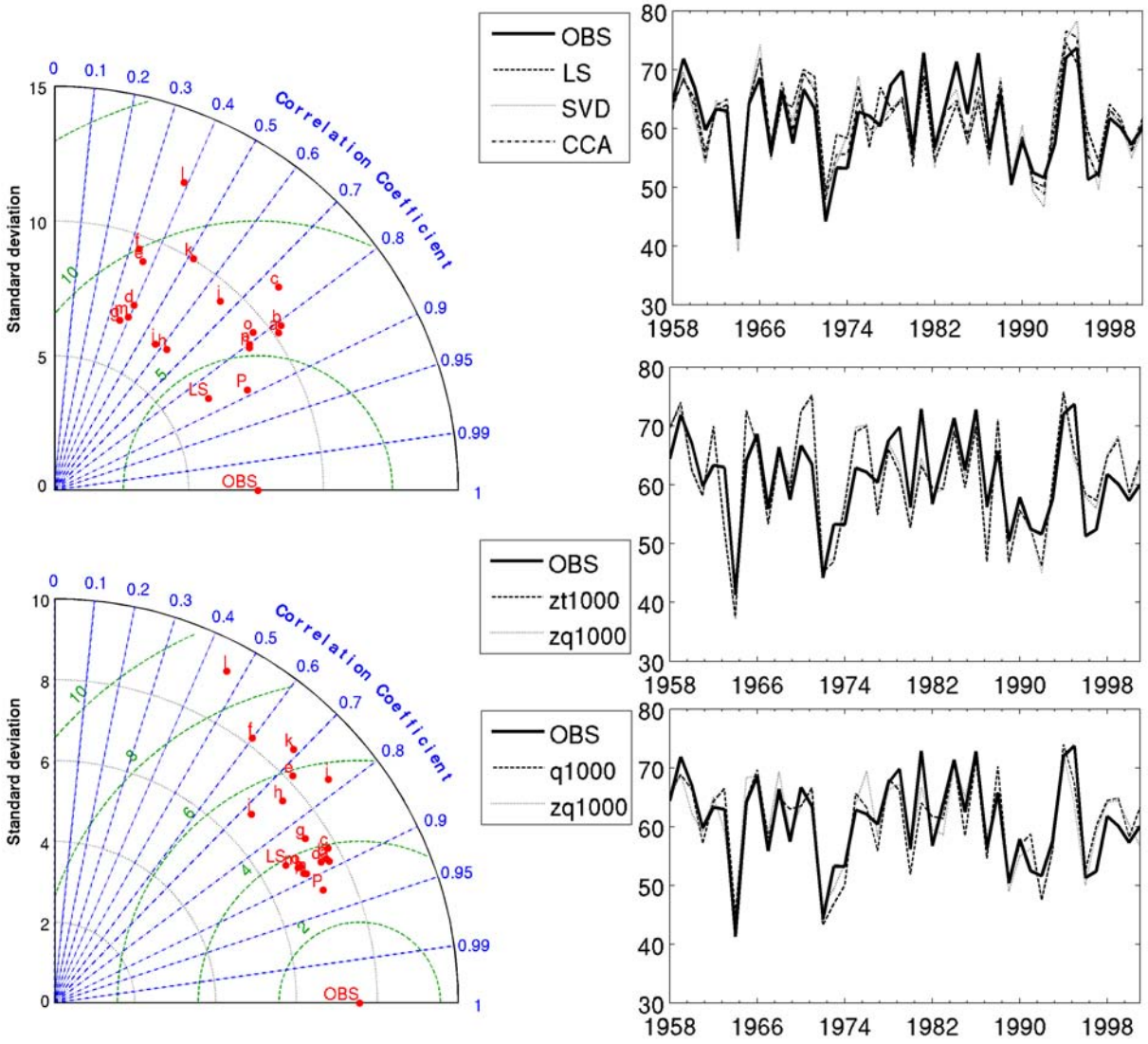

Figure 4.31: Summary of relative downscaling model skill in estimating mean January European precipitation (1958-2001). Left panels show Taylor diagrams detailing correlation, standard deviation and centred RMSE for all SVD-RM (top) and CCA-PCR (bottom) methods. Green arcs centred on the observation (OBS) indicate RMSE. Equivalent MOS downscaling models denoted by P (precipitation) in each case; local scaling estimates denoted by LS. Right panels show observed precipitation against cross-validated reconstructions based on MOS methods (top panel), and the most skilful SVD-RM (middle panel) and CCA-PCR Perfect-Prog methods (bottom panel). All precipitation values expressed in mm. CCA-PCR Perfect-Prog estimates based on 10 retained PCs; MOS estimates on 5 retained PCs. Variable codes: a-c = Z_{1000}, Z_{850}, Z_{500}; d-f = q; g-i = T; j-l = rh; m = q_{1000}/T_{1000}; n = Z_{1000}/q_{1000}; o = Z_{1000}/q_{850}; p = Z_{1000}/T_{1000}.

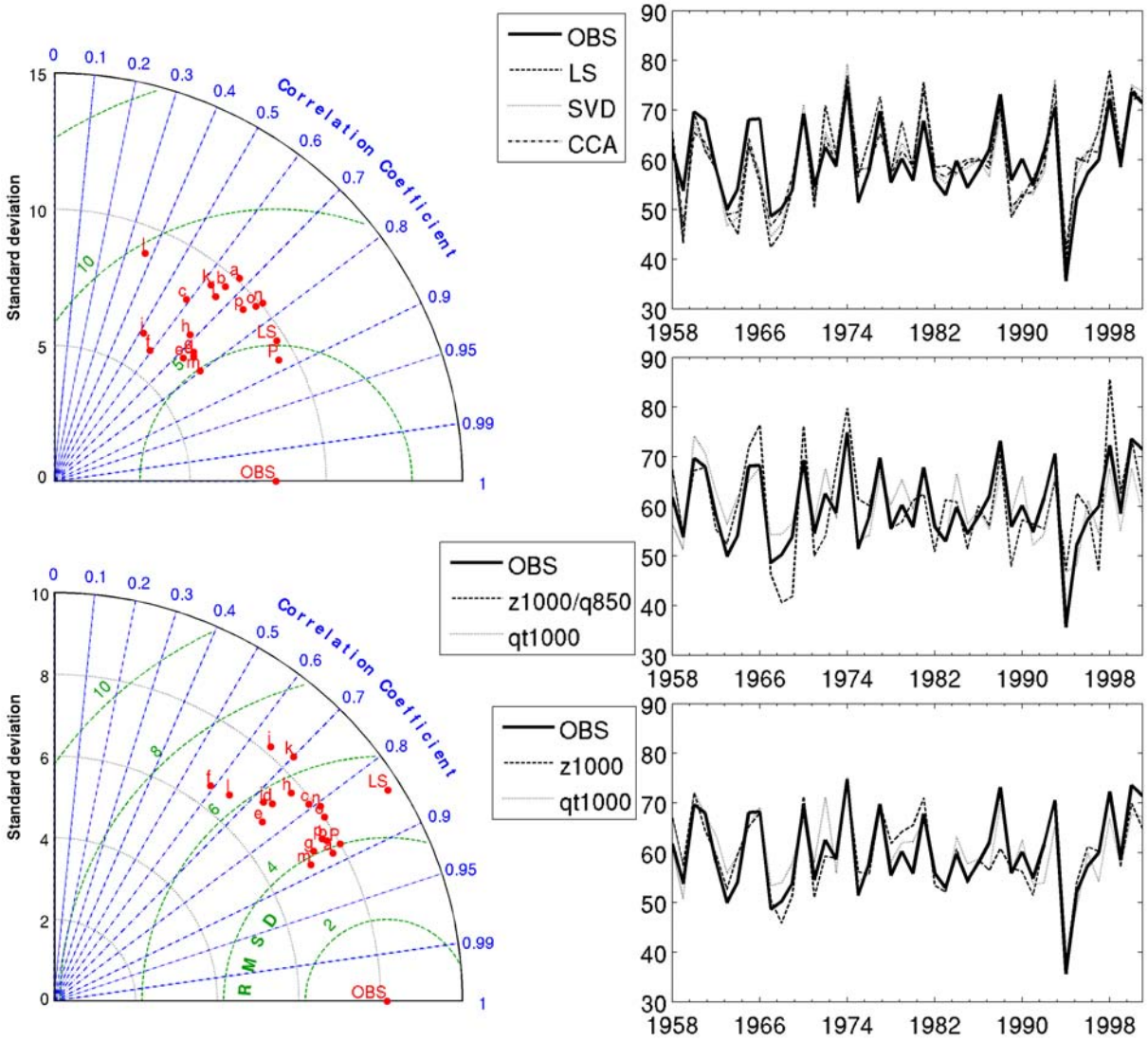

Figure 4.32: As Figure 4.31 but for July.

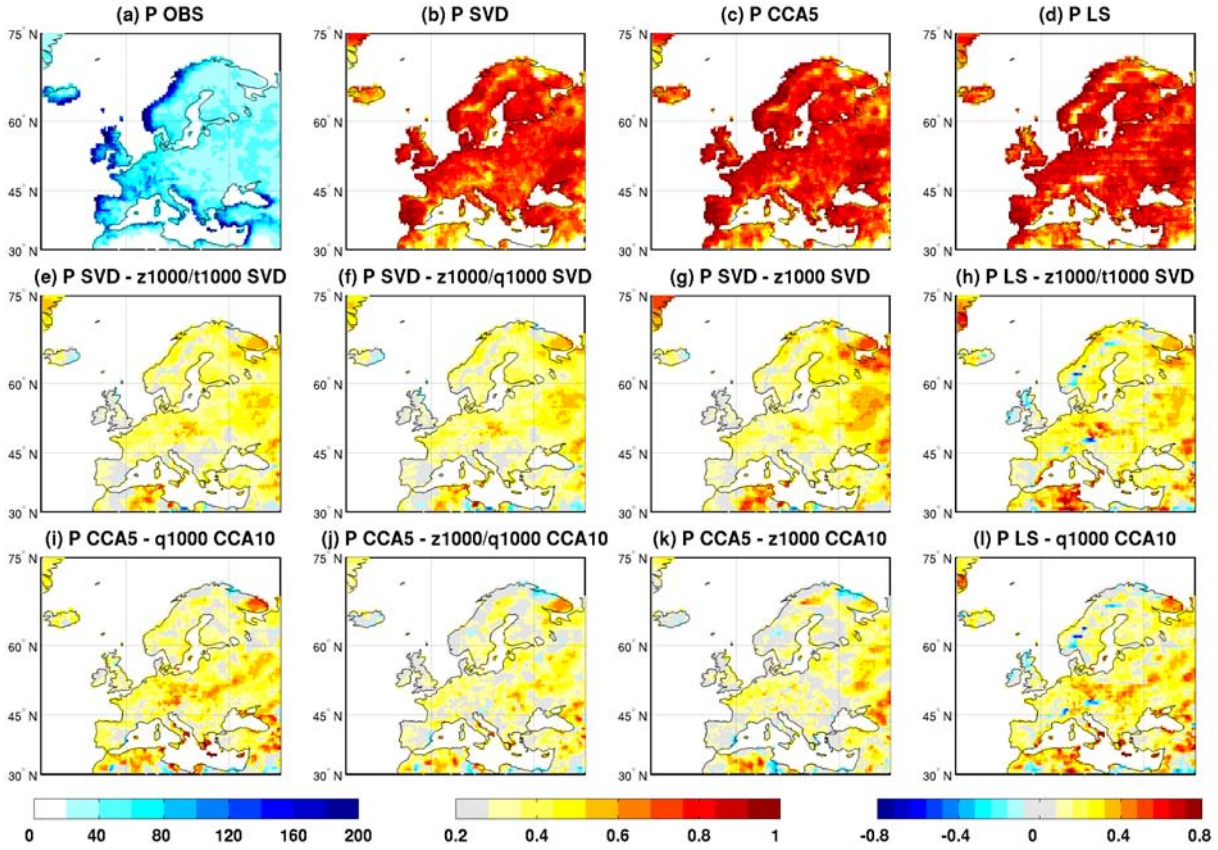

Figure 4.33: Correlation statistics for observed and estimated January precipitation in Europe (1958-2001). (a) January mean precipitation (mm; left colour bar). (b-d) Correlation between observed and MOS estimated (SVD, CCA5 and LS) precipitation (centre colour bar); correlations shown where significant at 5% level. (e-l) Differences in correlation of MOS and Perfect-Prog precipitation estimates (right colour bar).

estimates all exhibit strong correlations which are comparable to P-CCA5. At first, this would seem to contradict findings in section 4.3.2 where correlation coefficients for these variables only appear strong over northern Europe, but precipitation in this region forms the main contribution to a European mean.

European precipitation can generally be described by two dominant regimes, with heavier precipitation in the north than the south, with this north-south difference expected to accentuate during the next century (Christensen et al., 2007; van der Linden and Mitchell, 2009). A number of studies have demonstrated the potential of Perfect-Prog statistical methods to estimate local-scale precipitation, with skill usually higher in northern Europe than in southern Europe (Murphy, 1999; Goodess et al., 2010). MOS downscaling models have also shown good potential when applied to output from RCM simulations, an approach widely taken in the PRUDENCE intercomparison project (e.g. Boe et al., 2007; Deque et al., 2007). Differences between the correlations of the MOS and the most skilful Perfect-Prog downscaling models are shown in Figures 4.33 and 4.34. During January, P-SVD offers consistently stronger correlations than the best Perfect-Prog SVD-RM downscaling models (based on Z_{1000}/T_{1000}

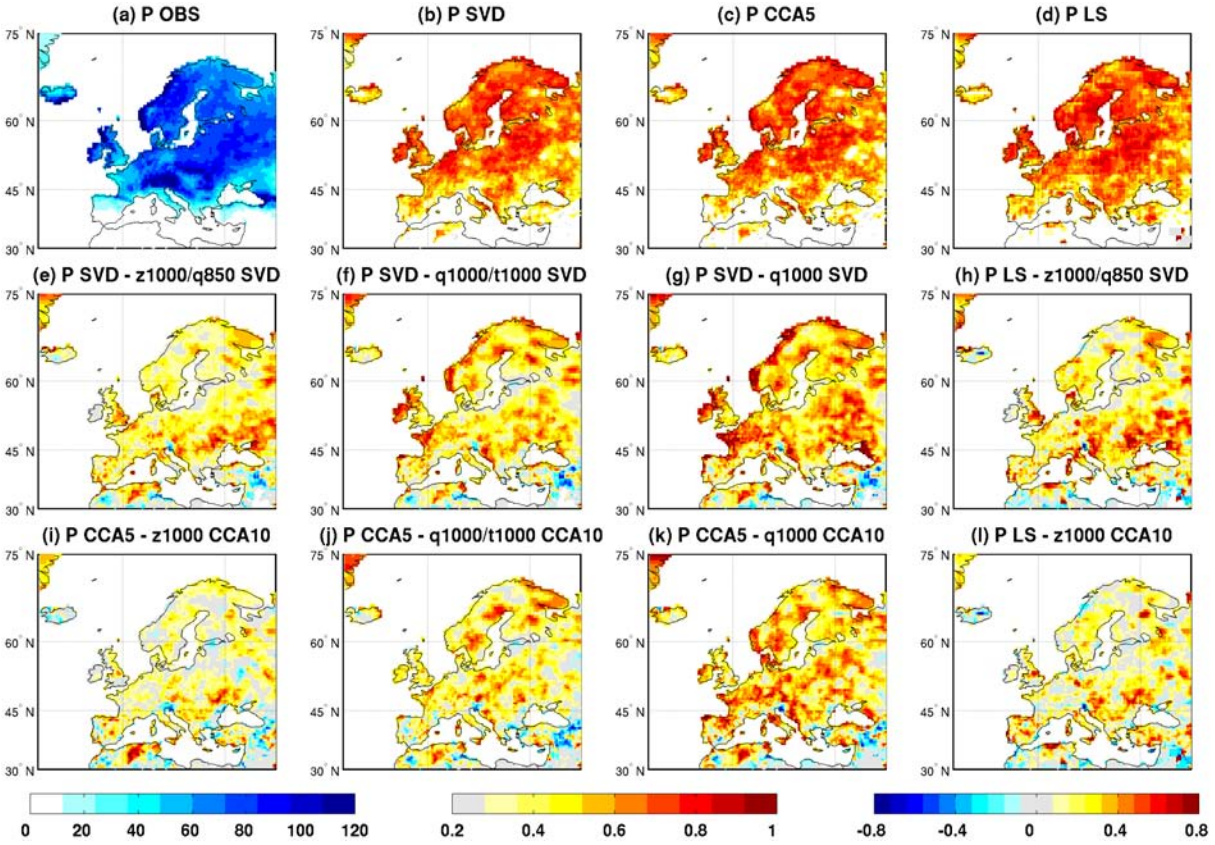

Figure 4.34: As Figure 4.33 but for July

and Z_{1000}/q_{1000}) across much of continental Europe. In eastern Europe, particularly of parts of Ukraine and western Russia, P-SVD correlations are around 0.4 greater than geopotential height (Z_{1000}) SVD estimates. Geopotential height performs well as a predictor in much of southern Europe, including the Iberian peninsula and western Balkans. The P-LS method, whilst offering higher correlations than the most skilful SVD and CCA Perfect-Prog estimates across much of Europe, is actually outperformed by both methods in particularly wet, windward regions of northern Europe and the northern Alps. P-CCA5, like P-SVD, offers more skill than its Perfect-Prog equivalents, and particularly so in some of Europe's mountainous regions including the Alps, Carpathians and Pyrenees. Schmidli et al. (2007) found a number of Perfect-Prog downscaling models underestimated the magnitude of interannual variations in daily precipitation across the Alps and so the ability to downscale precipitation successfully in regions of complex topography is a promising trait of the P-CCA5 model. It has already been shown that whilst the P-LS model is easily applicable and relatively successful across continental Europe, its skill is lacking where precipitation is variable in intensity and frequency. Crucially, this includes many mountainous areas, the climate of which is widely considered to be poorly represented in GCMs in the first instance.

In the summer months, the consistent north-south divide in European precipitation distribution would appear to impinge on the overall skill of all downscaling methods. This seasonal skill dependence is known

to extend to dynamical downscaling, with the suite of PRUDENCE RCMs showing greater skill in winter than in summer (Christensen and Christensen, 2007; Jacob et al., 2007). However, MOS methods for July continue to show greater skill than Perfect-Prog techniques, and the difference is generally larger than during January. In Figure 4.34, P-SVD models produce stronger correlation coefficients than those based on q_{1000}/T_{1000} and q_{1000}. This greater skill is particularly evident in northern and western Europe and greatest along the western coastlines of France, Norway and the United Kingdom. It is possible that the observed predictors are not able to sufficiently explain the variability of precipitation associated with the prevalence of maritime air mass types originating in the tropics. Cavazos and Hewitson (2005) note that precipitation processes during summer may be linked to the lower- and upper- tropospheric circulation more typical of convective, tropical regimes.

4.4.2 North America

Precipitation across the North American continent is controlled by climate regimes on different scales, ranging from the interannual variability linked to large-scale ENSO and NAO patterns to small-scale dynamical features which exact influence on particular regions (Trenberth et al., 2007). Understanding variability in storm tracks and frequency of mid-latitude cyclones that influence central and northern parts of North America is important, but not sufficient, in making reliable estimates of future changes in precipitation (Christensen et al., 2007). Dynamical downscaling models have been shown to offer little additional information on local-scales, and in some cases RCM simulations lack agreement on the sign of precipitation changes (e.g. Chen et al., 2003). The success of various statistical downscaling methods in North America has previously been shown to be highly variable and dependent on the climatology of the study region, and there is little consensus on the most appropriate method/predictor combination (Wilby and Wigley, 1997; Wilby et al., 1998).

Any improvement that MOS methods can provide may thus prove crucial in areas where projections of precipitation changes are uncertain. On first inspection, during the winter months, and to a lesser extent spring and autumn, MOS methods appear reasonably skilful across the majority of North America, with the exception of the Rocky Mountains and northern Canada (Figures 4.1-4.10). Interestingly, the local scaling method would appear to show greater skill than SVD-RM and CCA-PCR methods in representing temporal variability in the areas west of the Rocky Mountains and across the great plains of the United States (Figure 4.1). During summer, MOS methods exhibit far less skill, particularly across the dry western United States. Rainfall in western Canada is well-represented, however, both in terms of correlation and bias.

Taylor diagrams provide a generalised diagnostic of inter-model skill and again the results presented focus on January and July only (Figures 4.35-4.36). During January, mean precipitation across North

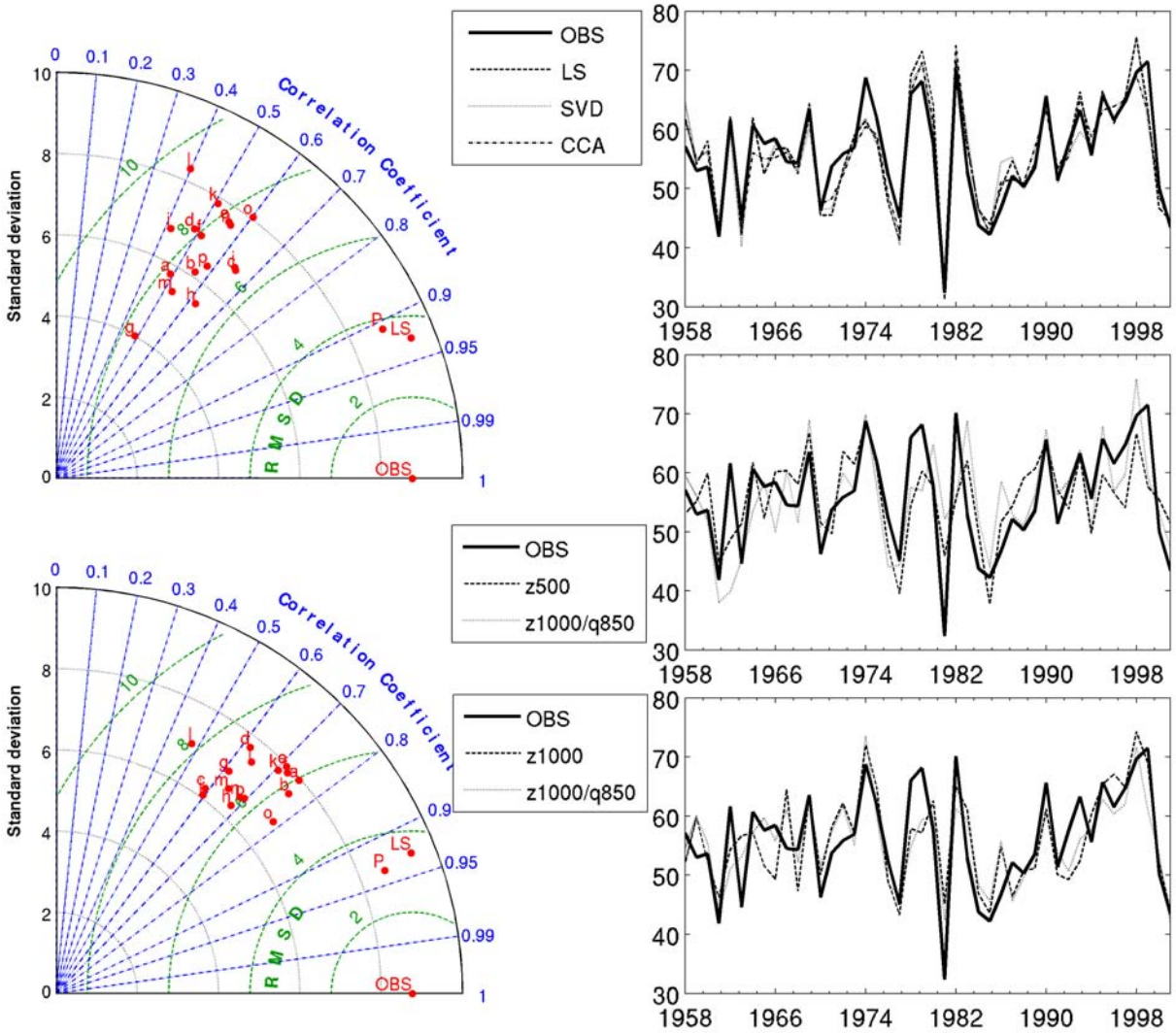

Figure 4.35: Summary of relative downscaling model skill in estimating mean January North America precipitation (1958-2001). Left panels show Taylor diagrams detailing correlation, standard deviation and centred RMSE for all SVD-RM (top) and CCA-PCR (bottom) methods. Green arcs centred on the observation (OBS) indicate RMSE. Equivalent MOS downscaling models denoted by P (precipitation) in each case; local scaling estimates denoted by LS. Right panels show observed precipitation against cross-validated reconstructions based on MOS methods (top panel), and the most skilful SVD-RM (middle panel) and CCA-PCR Perfect-Prog methods (bottom panel). All precipitation values expressed in mm. CCA-PCR Perfect-Prog estimates based on 10 retained PCs; MOS estimates on 5 retained PCs. Variable codes: a-c = Z_{1000}, Z_{850}, Z_{500}; d-f = q; g-i = T; j-l = rh; m = q_{1000}/T_{1000}; n = Z_{1000}/q_{1000}; o = Z_{1000}/q_{850}; p = Z_{1000}/T_{1000}.

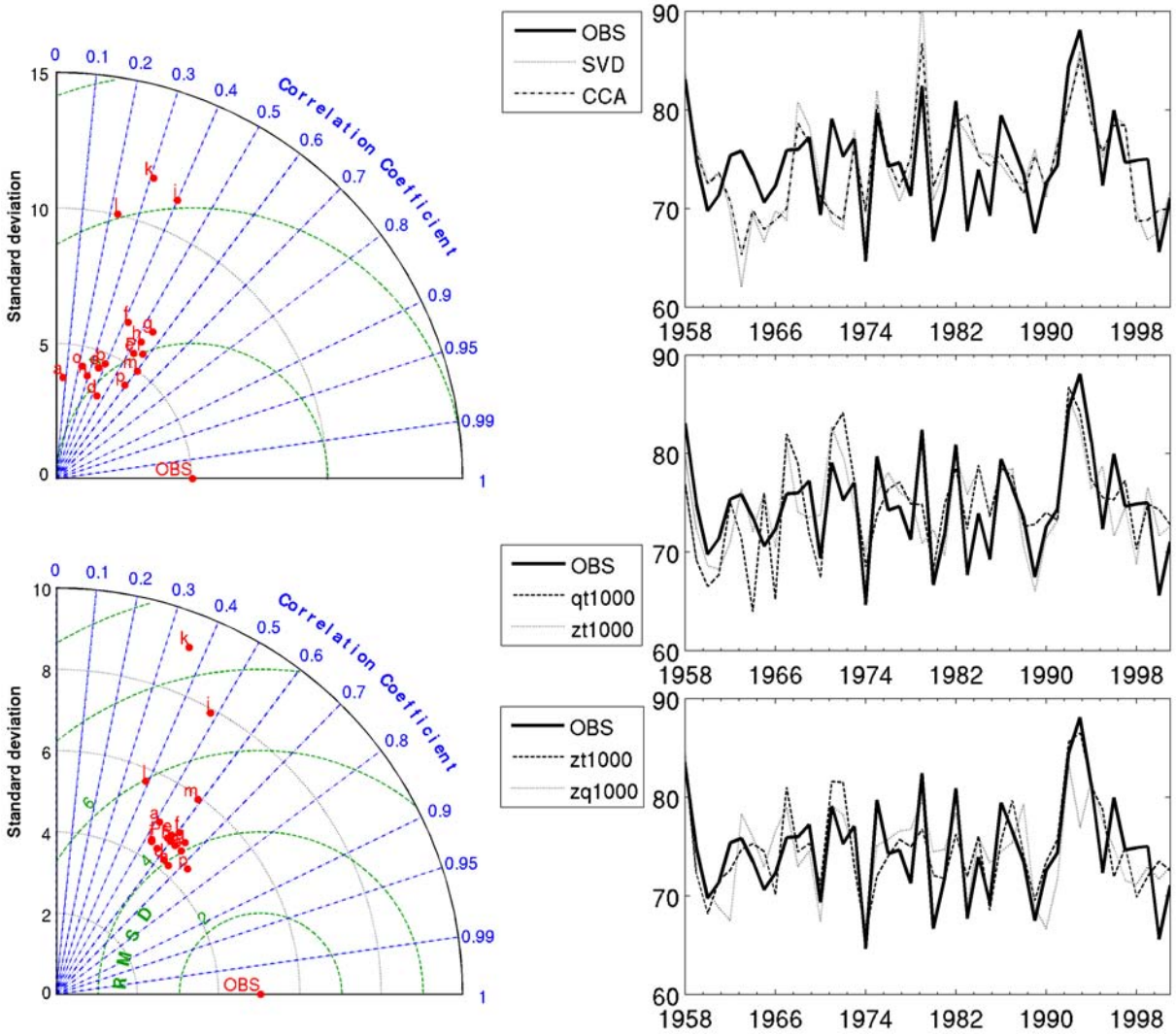

Figure 4.36: As Figure 4.35 but for July precipitation. Note: local scaling estimated time series associated with a number of outliers and is omitted from the plots shown here.

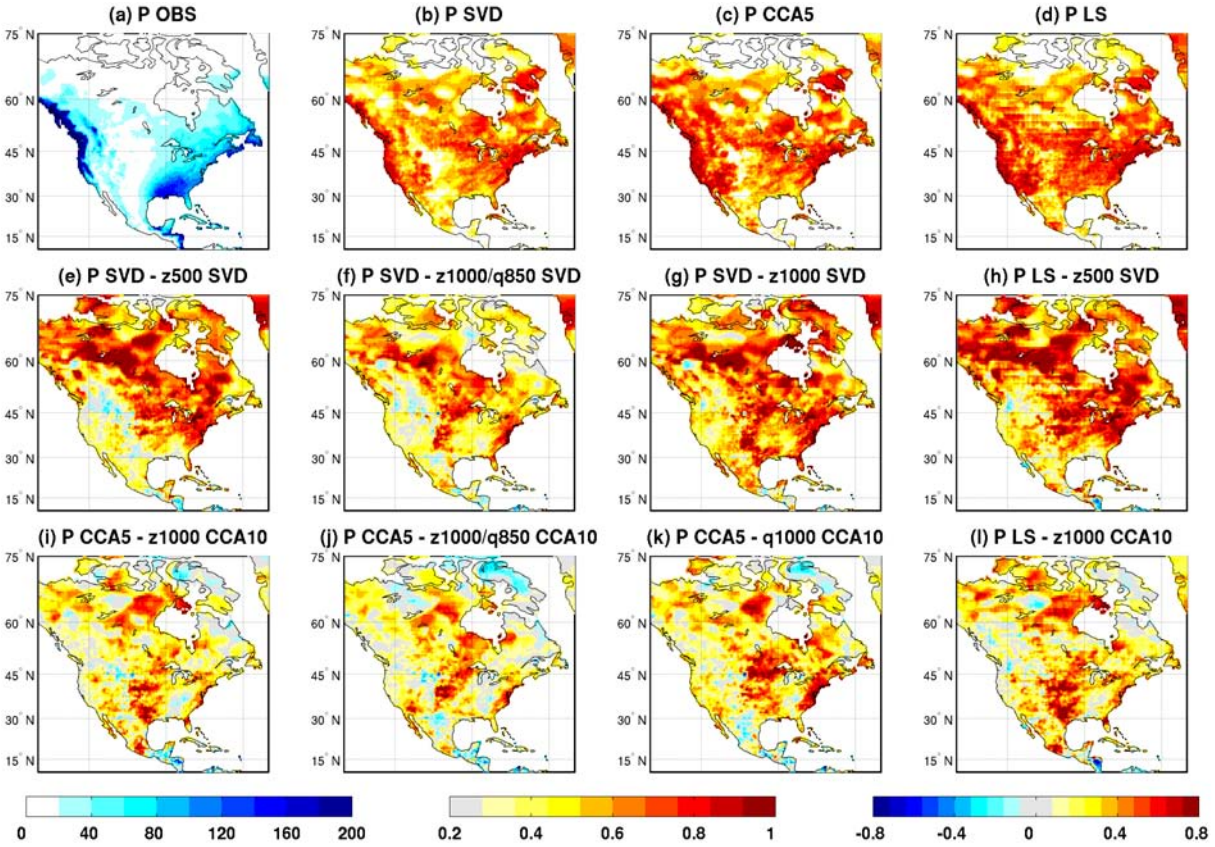

Figure 4.37: Correlation statistics for observed and estimated January precipitation in North America (1958-2001). (a) January mean precipitation (mm; left colour bar). (b-d) Correlation between observed and MOS estimated (SVD, CCA5 and LS) precipitation (centre colour bar); correlations shown where significant at 5% level. (e-l) Differences in correlation of MOS and Perfect-Prog precipitation estimates (right colour bar).

America is well represented by the MOS models, with P-SVD showing greater greater correlation than any Perfect-Prog SVD-RM methods. Performance of CCA-PCR Perfect-Prog models is stronger than SVD-RM, with Z_{1000} appearing to be the most useful predictor when used independently or with q_{850}. During July, mean precipitation is much more challenging to realistically estimate, particularly the extremes in the observed record indicated by a large standard deviation. Performance amongst all models (both MOS and Perfect-Prog) is similar, with standard deviation well-represented by many CCA-PCR models. The local scaling correction is associated with a number of outliers throughout the time series and is not shown in Figure 4.36. Low summer precipitation totals in large parts of North America limit the usefulness of the local scaling correction, which appears more useful areas associated with more consistent, non-convective precipitation features.

The strongest performing models are assessed in depth in Figures 4.37-4.38. In January, the MOS models all produce strong correlations across most of the wettest parts of North America. Along much of the east coast of the United States, MOS methods are able to offer substantial improvement over

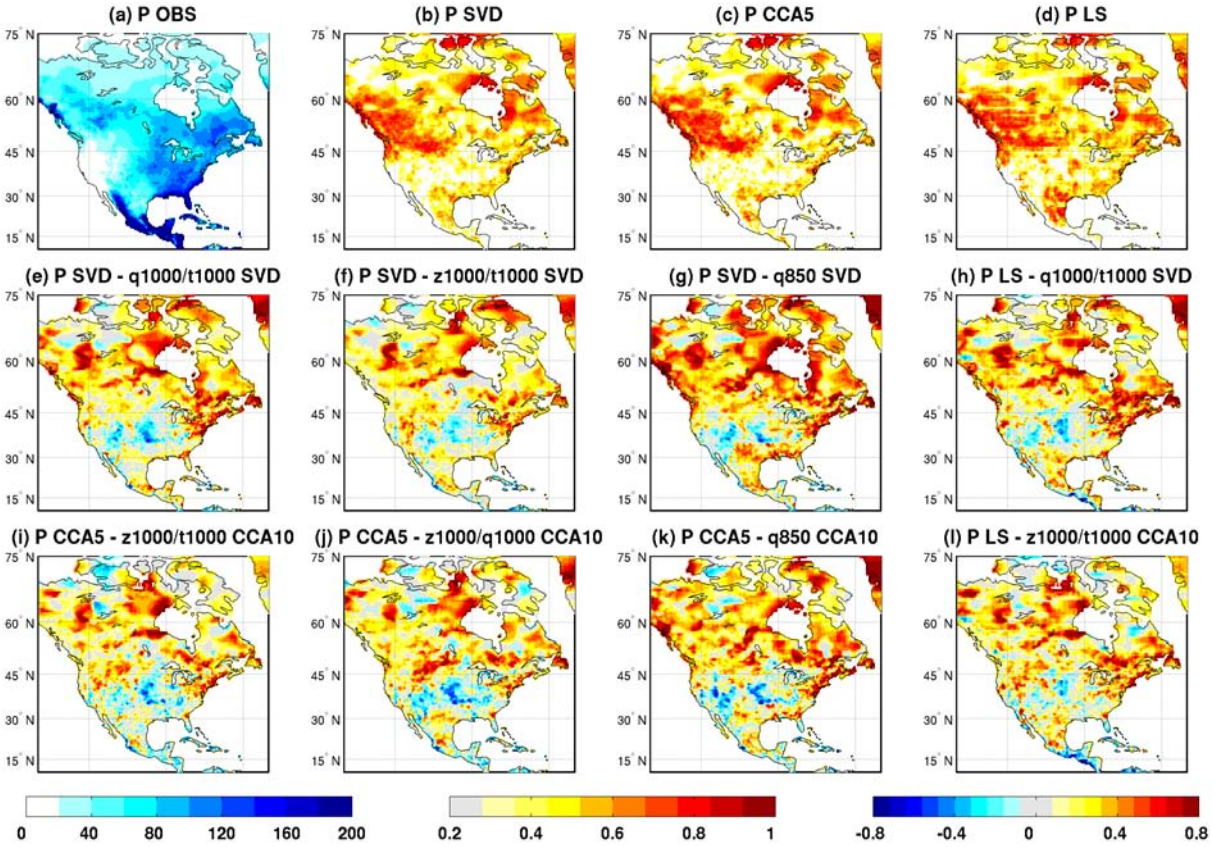

Figure 4.38: As Figure 4.37 but for North America July precipitation.

equivalent Perfect-Prog models. This is also true in areas of the mid-west United States, and in parts of central and western Canada. During July, MOS skill is limited to the Pacific north-west, parts of western Canada and the eastern and northern shores of Hudson Bay, and there is little consensus on where MOS offers an improvement over Perfect-Prog. Parts of central United States appear to be better represented by some Perfect-Prog methods but the estimation of realistic precipitation amounts in these areas has already been shown to be inconsistent (Figure 4.25). There is some improvement over Perfect-Prog models along parts of the north-east coast of the United States but this skill diminishes with distance from the coast.

4.4.3 Australia

The MOS downscaling models show some encouraging skill over Australia which merits closer inspection. Previous work has shown that spatial patterns of annual rainfall changes in Australia over the last few decades can generally be split into two regions, with rainfall decreasing in the east and increasing in the west (Taschetto and England, 2009). The south-east is particularly important as the factors responsible for long-term drought in this region are not fully understood (Sohn, 2007; LeBlanc et al., 2009; Ummenhofer et al., 2011). Statistical downscaling based on reanalysis-derived predictors have shown some good promise in reproducing interannual variability in winter (JJA) precipitation but there is scope for improvement (Charles et al., 2004).

The contribution of tropical rainfall to the Australia total during winter (DJF) is a result of the southerly migration of the inter-tropical convergence zone (ITCZ) over the northerly extent of the continent. Australian climate is strongly affected by a number of tropical processes, including ENSO, the Australian monsoon and the south-east trade wind circulation (Christensen et al., 2007). Charles et al. (2004) suggested that using a predictor that accounts for variability in ENSO, and other processes that evolve on intra-decadal time-scales, may improve downscaling performance. Extreme rainfall, particularly in the north of the continent, is usually associated with tropical cyclones. Already, the skill of MOS models in the tropics is limited both by the quality of ECHAM5 precipitation in these regions and the nudging procedure's representation of processes responsible for the formation of convective precipitation (see Chapter 4). As a result, MOS skill in Australia is highly variable by location and season, but in general large parts of eastern Australia, including the Murray-Darling basin, are associated with realistic estimates for much of the year.

Interpreting Taylor diagrams (Figures 4.39 and 4.40), which describe the spatially-averaged skill of a number of downscaling models, is challenging in a region of large spatial variability. However, it is still possible to identify those models with good skill which can then be assessed in more detail. During January, SVD-RM models based relative and specific humidity show reasonable skill, but are outperfomed by both P-SVD and P-LS. The Z_{850} CCA-PCR models exhibits comparable correlation to that of P-CCA5 and also appears reproduce realistic amounts for the majority of the time series (Figure 4.39, bottom-right panel). Additionally, the potential shown by moisture content variables at higher atmospheric levels (i.e. rh_{500} and q_{500}) to be skilful predictors is attributable to the processes involved in the formation of convective rainfall over the north of the continent.

During July, the heaviest precipitation is concentrated along the southern coastlines of Australia. The SVD-RM and CCA-PCR MOS models continue to offer good skill (Figure 4.40, top right panel); P-LS is shown to over-estimate rainfall in a number of years (not shown). Extra-tropical influences

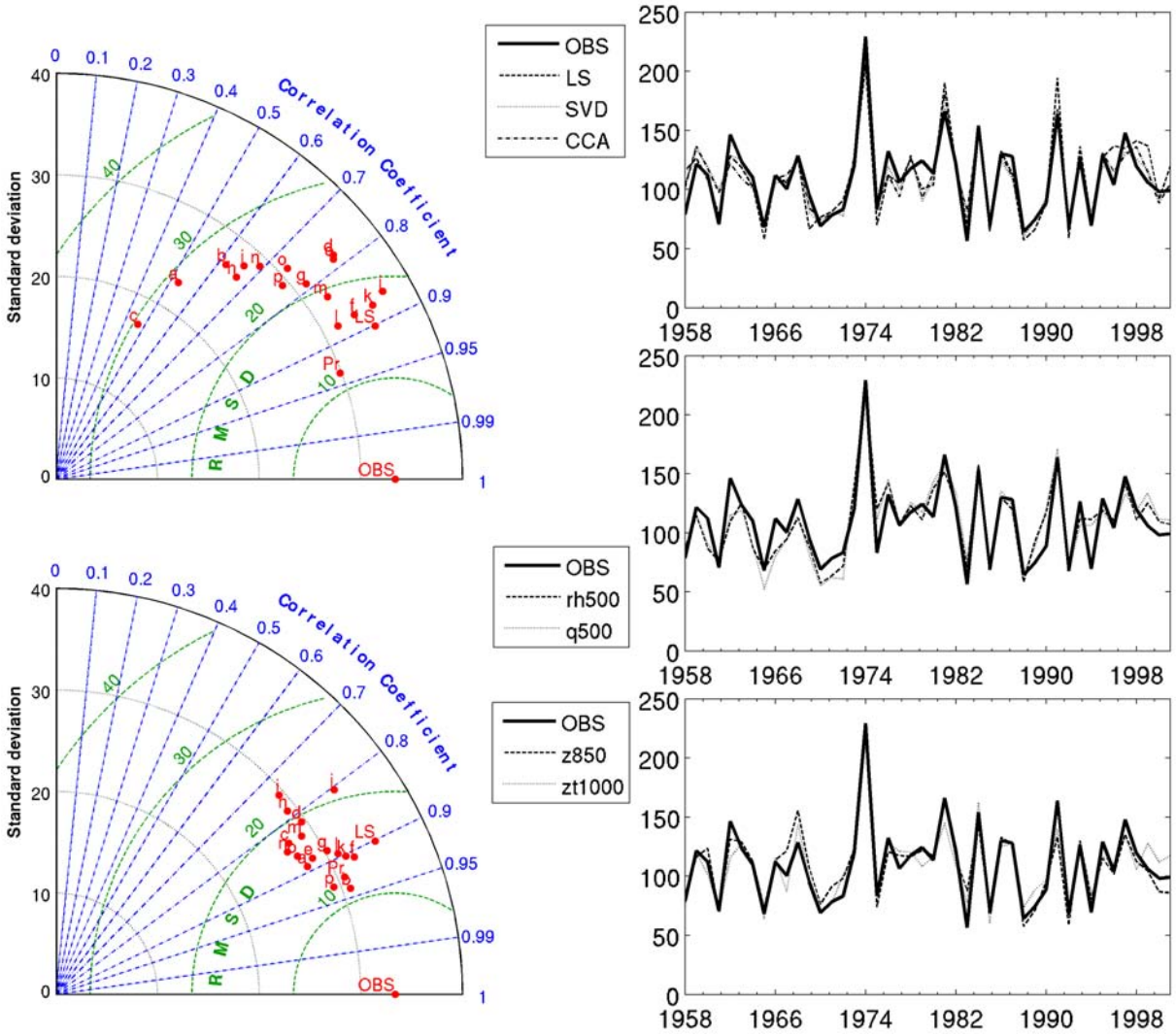

Figure 4.39: Summary of relative downscaling model skill in estimating mean January Australia precipitation (1958-2001). Left panels show Taylor diagrams detailing correlation, standard deviation and centred RMSE for all SVD-RM (top) and CCA-PCR (bottom) methods. Green arcs centred on the observation (OBS) indicate RMSE. Equivalent MOS downscaling models denoted by P (precipitation) in each case; local scaling estimates denoted by LS. Right panels show observed precipitation against cross-validated reconstructions based on MOS methods (top panel), and the most skilful SVD-RM (middle panel) and CCA-PCR Perfect-Prog methods (bottom panel). All precipitation values expressed in mm. CCA-PCR Perfect-Prog estimates based on 10 retained PCs; MOS estimates on 5 retained PCs. Variable codes: a-c = Z_{1000}, Z_{850}, Z_{500}; d-f = q; g-i = T; j-l = rh; m = q_{1000}/T_{1000}; n = Z_{1000}/q_{1000}; o = Z_{1000}/q_{850}; p = Z_{1000}/T_{1000}.

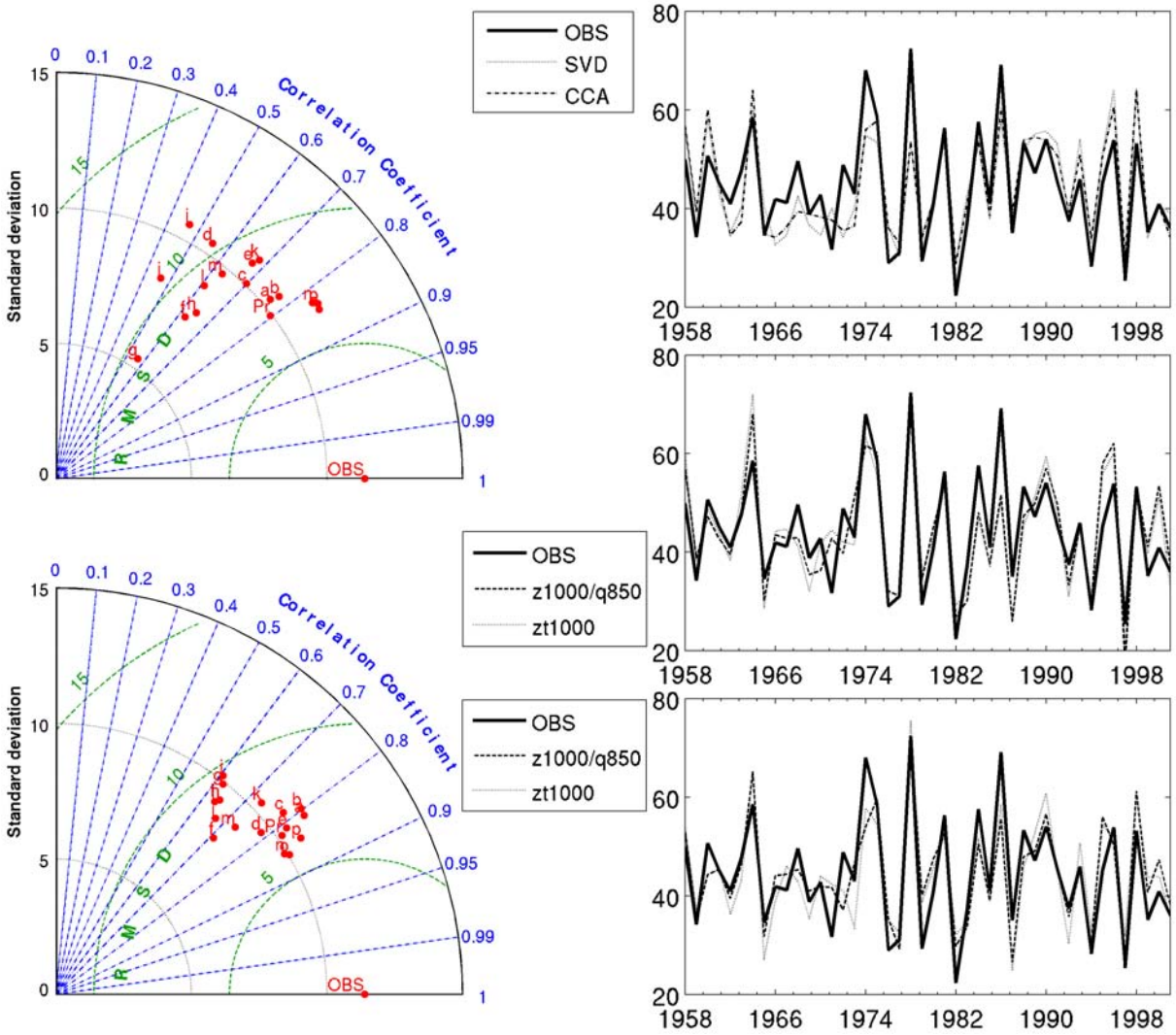

Figure 4.40: As Figure 4.39 but for July precipitation. Note: local scaling estimated time series associated with a number of outliers and is omitted from the plots shown here.

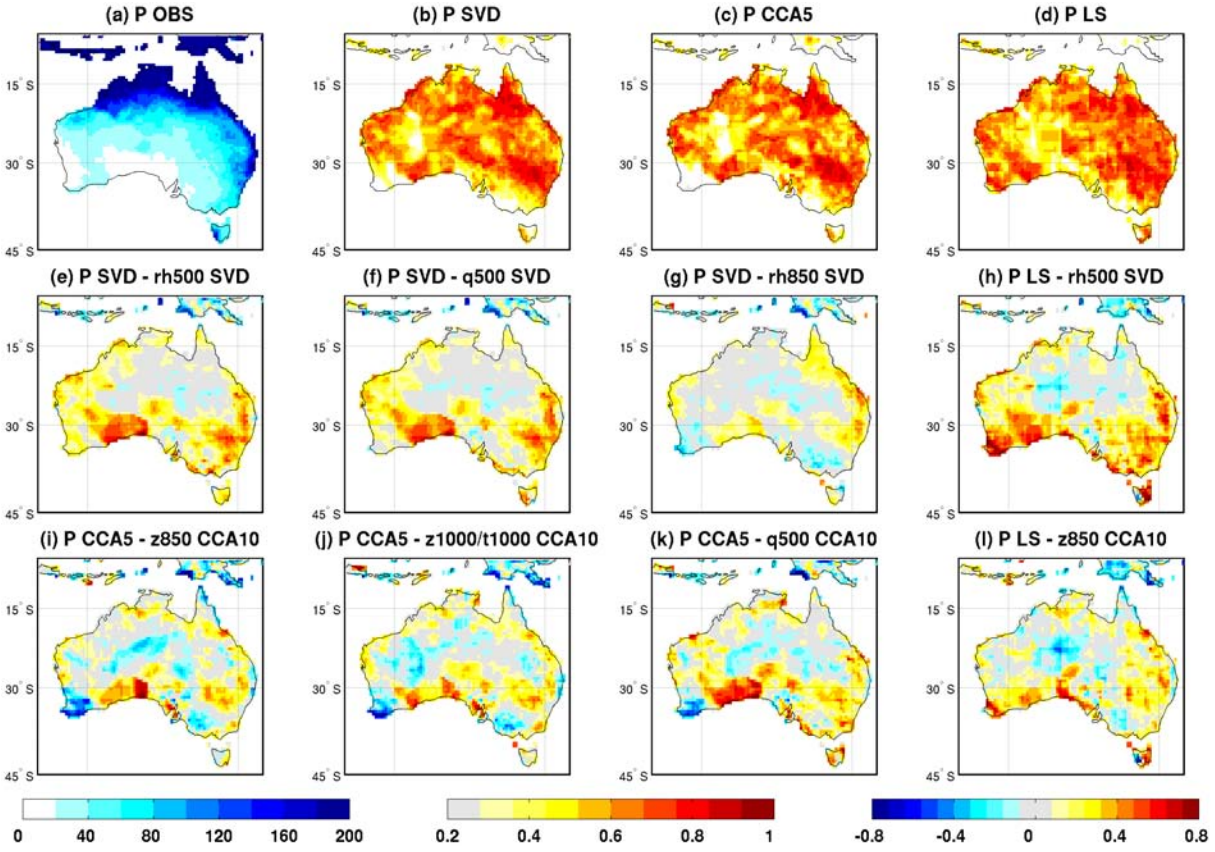

Figure 4.41: Correlation statistics for observed and estimated January precipitation in Australia (1958-2001). (a) January mean precipitation (mm; left colour bar). (b-d) Correlation between observed and MOS estimated (SVD, CCA5 and LS) precipitation (centre colour bar); correlations shown where significant at 5% level. (e-l) Differences in correlation of MOS and Perfect-Prog precipitation estimates (right colour bar).

on Australian rainfall result in greater explanation of variability from geopotential height and specific humidity at lower atmospheric levels, and models based on Z_{1000}/q_{1000} Z_{1000}/q_{850} and Z_{1000}/T_{1000} all marginally outperform the non-local MOS methods.

In January, whilst correlation amongst MOS models is reasonably strong across much of the continent, it is only parts of the eastern and southern coasts where MOS offers an improvement over the most successful Perfect-Prog methods (Figure 4.41). Indeed in the dry south-west corner of Western Australia the Z_{850}, q_{500} and rh_{850} CCA-PCR methods yield stronger correlation coefficients than the P-CCA equivalent. This is also the case across the continental interior, although it is known that the station network is poor here and these results may not be significant. In July, MOS skill is highly concentrated in the south-east of Australia and does offer a significant improvement on the skill of Perfect-Prog models across much of this region (Figure 4.42). In particular, the P-SVD model performs substantially better than any Perfect-Prog SVD-RM models across the Murray-Darling basin away from the coastlines (where performance is still greater but not as much so). Conversely, the Perfect-Prog models using geopotential

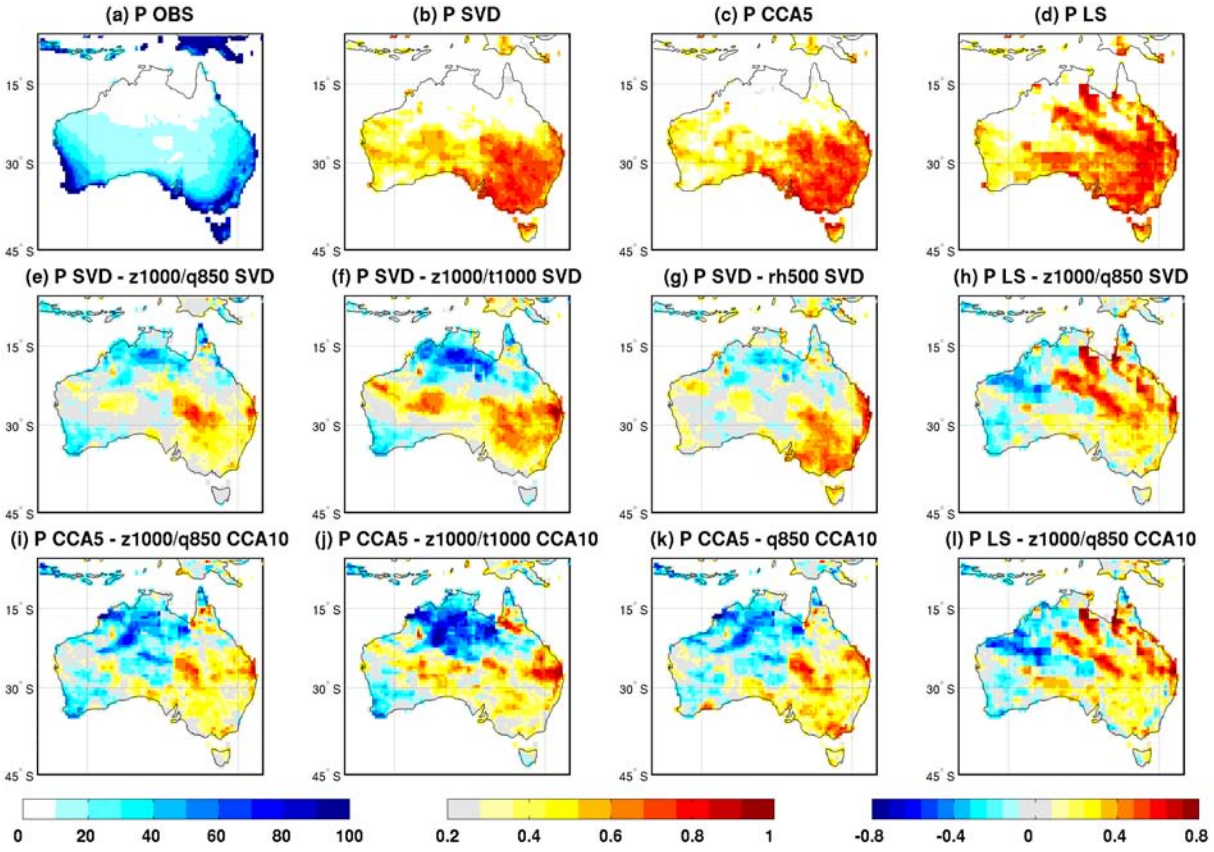

Figure 4.42: As Figure 4.41 but for July precipitation.

height and humidity predictors (particularly the CCA-PCR models) appear to better reproduce temporal variability across the drier parts of northern Australia, although representation of bias remains an issue.

4.5 Summary and conclusions

A set of statistical downscaling models have been developed under MOS and Perfect-Prog frameworks and compared in terms of their skill to estimate local-scale precipitation across the globe. MOS downscaling models use GCM-simulated precipitation as a predictor variable, and use both local-scaling and non-local regression-based techniques, namely one-dimensional MCA and one-dimensional CCA. The MCA method, sometimes known as SVD analysis and referred to here as SVD-RM, used TECs of a regression map as the predictor in a linear regression model for estimating monthly mean precipitation. The CCA method was implemented as principal component multiple linear regression and referred to here as CCA-PCR. Both the SVD-RM and CCA-PCR techniques seek to derive a link between simulated precipitation across a fixed spatial field and observed precipitation at a given location. Perfect-Prog models were developed using the same SVD-RM and CCA-PCR techniques with observed (reanalysis-derived) circulation, temperature and humidity fields as predictors. The skill of each method to estimate

local-scale precipitation was assessed using a leave-seven-out cross-validation technique.

MOS downscaling models are shown to outperform Perfect-Prog counterparts in many regions of the world, in terms of reproducing temporal variability, actual precipitation amounts and a realistic amplitude of variance. The relative skill of each MOS model, and likewise the improvement over those formulated using Perfect-Prog methods, varies greatly by location and season. Three key continental regions (Europe, North America and Australia) have been identified where simulated precipitation is well-represented and estimates of local precipitation are, in general, skilful. Detailed analysis of each region would be preferable and certainly a focus for future work.

All MOS models tend to show similar skill in reproducing temporal variability in the observed record, although CCA-PCR generally produces less bias and more realistic variance. The local-scaling approach is very simple to apply and offers a good alternative to most Perfect-Prog methods. Additionally, in many data-rich regions of the mid-latitudes (i.e. large parts of Europe) there is often little, if any, improvement to be offered by the far more computationally intensive non-local techniques. An exception is across mountainous areas where SVD-RM and (particularly) CCA-PCR are able to better capture small-scale variability not represented in local scaling estimates. The success of non-local techniques in such regions is a key finding in this work. In principle, the spatial patterns in simulated precipitation identified by SVD-RM and CCA-PCR allow for a correction of unrealistic spatial structure of the simulation (Widmann et al., 2003; Maraun et al., 2010). The poor representation of orography in GCMs, and subsequent difficulty in parameterising orographic processes responsible for precipitation formation, is a major cause of spatial incoherence of the simulated field precipitation.

Whilst useful in large parts of the mid-latitudes, local-scaling is not appropriate in tropical and sub-tropical regions where long-term precipitation means are greatly over-estimated. The dominance of convective processes in rainfall formation in these regions makes a fixed factorial scaling of precipitation unsuitable. The over-estimation is particularly prominent within the dry sub-tropical bands of high pressure and at all times of year. It is also uncertain how suitable the application of a scaling factor to simulated precipitation in future climates. This is investigated in more detail in Chapter 5.

The most skilful CCA-PCR models are based on around five principal components (PCs) retained in the pre-filtering process. The inclusion of an increasing number of PCs results in a weaker correlation between estimated precipitation and observations. This is especially so when more than about ten PCs are retained. On a global scale there is no optimal number of PCs, although such an optimisation could be possible on a regional basis and would depend on the prevalence of modes of variability. As discussed in section 4.2.2, the obvious advantage of SVD-RM is that there is no pre-filtering procedure, as there is in CCA-PCR. However, regression models based on SVD-RM appear susceptible to overfitting and there are few locations were model performance is strong enough not to necessitate the use a CCA-PCR model,

the development of which requires an additional pre-filtering step.

This work represents the first attempt to develop local and non-local MOS downscaling models using simulated precipitation from a nudged GCM simulation and to make a detailed comparison with Perfect-Prog downscaling models. A rigorous cross-validation procedure has permitted a quantification of the skill of each downscaling model, and an understanding of where simulated precipitation is likely to be reliable is hugely important when downscaling projections of future scenarios. In applying the MOS models to the future, the reliability of downscaled information is a product of both the skill of the downscaling model itself and the overall skill of ECHAM5 precipitation. More attention will be given these concepts in Chapters 5 and 6.

Chapter 5

Applicability of MOS downscaling models to future simulations

5.1 Introduction

The skill of local and non-local MOS downscaling models has been fully assessed in Chapter 4. This assessment included a comparison with equivalent Perfect-Prog downscaling techniques using different (or combinations of) quasi-observed variables as statistical predictors. MOS downscaling corrections were shown to exhibit excellent potential, and in many regions show greater skill than downscaling models constructed using a Perfect-Prog approach. The remaining empirical component of this thesis is concerned with the application of the MOS corrections developed in Chapter 4 to climate simulations for the twenty-first century.

Downscaling of climate change simulations requires additional considerations pertaining to the applicability of a given downscaling function in a perturbed climate. Cross-validation in Chapter 4 permitted an assessment downscaling skill in the current climate, but this may not be a sound indicator of skill in a future climate (Charles et al., 1999; Christensen et al., 2007; Maraun et al., 2010). In a recent review paper, Maraun et al. (2010) cite two major requirements, or assumptions, for a downscaling model in the context of application to future climates. The first assumption is that the predictor variable(s) must be well-simulated in the driving GCM. The second assumption is that the statistical relationship underpinning the downscaling model must be transferable to alternative climates, a concept usually referred to as the model's 'stationarity'

In the context of Perfect-Prog downscaling, it is possible to verify the first assumption (in part) by comparing long-term simulated and observed climatologies of large-scale variables that will act as

predictors. Under a MOS approach, this assumption applies to GCM-simulated precipitation itself. It was demonstrated in Chapter 3 that by conducting a GCM simulation nudged to a reanalysis, and forcing the simulation into temporal phase with the real world, it is possible to compare simulated and observed precipitation time series and to quantify the skill of the GCM to reproduce long-term temporal variability. The second assumption is rather more troublesome to assess and the issue of stationarity is not always given credence in downscaling validation studies. In the Perfect-Prog approach (section 2.5.1), a statistical downscaling link is established between observations of precipitation and large-scale predictors, typically pertaining to atmospheric circulation, temperature or humidity. In application to a future climate, possible changes in the physical predictor-predictand relationship may render a stable relationship invalid. That is, in regions where the relationship between large-scale circulation and temperature variables and precipitation is likely to alter in a future climatic state, a statistical downscaling model may be compromised. This issue may be exacerbated when multiple predictors are included in a downscaling model's development. Alternatively, when using the MOS approach, GCM simulated precipitation is used as the sole predictor variable. Any shift in the response of precipitation to changes in other atmospheric variables is, theoretically, resolved by the GCM. This concept pertains to a third crucial assumption that must be satisfied in downscaling climate change simulations: that the chosen predictor variables are able to sufficiently capture the climate change signal. The role of post-simulation statistics would thus be to merely correct the simulated precipitation field. However, it should be noted that there may be a 'state-dependency' of any systematic biases in the simulated precipitation field. That is, bias may be dependent on the synoptic situation or, equally, or the contribution of large-scale (frontal) and convective precipitation to simulated totals. The changing nature of these controlling factors in a altered climate may render any seemingly robust statistical bias correction non-stationary. Therefore, the applicability of statistical corrections to a future climate should ideally be thoroughly tested on (quasi-) perturbed conditions.

It is the purpose of this chapter to provide an assessment of the potential of the applicability of MOS downscaling for correction of GCM simulation of future precipitation. The content of this chapter can be considered an extension of the general evaluation of model skill detailed in Chapter 4 and does not include actual MOS application to future GCM simulations, which is covered comprehensively in Chapters 6 and 7. Analysis and discussion in this chapter is undertaken in two parts. Firstly, focus is given to the projected response of precipitation to future changes in large-scale circulation and temperature as simulated by ECHAM5. With this in mind, the limitations of Perfect-Prog downscaling for future simulations are discussed, with recommendations made as to the regions where MOS application may be particularly beneficial. Secondly, analysis is undertaken in the assessment of stationarity in the MOS downscaling models developed in Chapter 4.

The remainder of this chapter is structured as follows. Section 5.2 describes the methods used in the two-part analysis, and includes a brief summary of previous attempts to assess stationarity in Perfect-Prog downscaling models. In section 5.3, the precipitation response to a changing climate is discussed. This includes analysis of relationships of precipitation with large-scale circulation and temperature variables as simulated by ECHAM5 for a future period and a discussion of why stationarity of Perfect-Prog downscaling is limited. Further analysis in section 5.4 aims to establish whether MOS downscaling corrections are potentially transferable to future climates. A summary is provided and conclusions drawn in section 5.5.

5.2 Methods of analysis

5.2.1 Assessing the changing response of precipitation to large-scale circulation and temperature

Focus is given to how the relationship of precipitation with large-scale circulation and temperature variables is expected to change in future climates, and how well these changes are simulated by ECHAM5. Such relationships underpin Perfect-Prog downscaling models, and understanding future changes gives an idea of the possible limitations of a given downscaling model (Lenderink and van Meijgaard, 2008). Seasonal means of ECHAM5-simulated sea level pressure and surface temperature are compared with simulated precipitation for the periods 1980-1999 and 2080-2099. The properties of good predictor variables were discussed in Chapter 2 (section 2.6.1) and in the construction of Perfect-Prog downscaling models in Chapter 4. It is arguable that atmospheric variables at higher levels have a greater physical connection to precipitation formation, but these may be exacerbated in a perturbed climate to an extent that is not consistent with overall change in circulation (Zorita et al., 1995; Frias et al., 2006). Therefore, high-level climate variables may not reflect the same climate change signals that are responsible for precipitation changes. All future projections are taken from the ECHAM5 SRES A1B simulation (Nakicenovic and Swart, 2000; Roeckner et al., 2006), which is described in more detail, and in the context of its sister simulations, in Chapter 6.

5.2.2 Assessing stationarity of MOS downscaling corrections

As discussed in Chapters 2-4, development of any downscaling model requires a thorough validation of its skill (e.g. Murphy, 1999; Haylock et al., 2006; Frias et al., 2006). Ideally, this should also consider the model's stationarity, or its transferability to a future, perturbed climate. In the formulation of a Perfect-Prog downscaling model, which is invariably fitted on reanalysis fields, this usually does not

extend beyond an assessment of the model's ability to reproduce the long-term climatology when driven by the same reanalysis (e.g. Huth, 1999; Robertson et al., 2004) or by a hindcast GCM simulation (e.g. Wilby and Wigley, 2000; Charles et al., 2004).

Any true model validation requires a comparison with observations, and understanding the behaviour of downscaling models when applied to climate change simulations remains challenging. Previously, comparisons have been made between past and future simulations from coarse-scale GCM fields (Frias et al., 2006) or higher resolution RCM output (Wood et al., 2004; Busuioc et al., 2006; Haylock et al., 2006). To some extent, such comparisons are able to identify where a given downscaling model is likely to be robust in a future climate. However, the extent to which model applicability is dependent on changes in large-scale climate forcing on precipitation is not fully diagnosed.

It is possible to draw conclusions about model transferability by identifying consensus between statistical and dynamical downscaled estimates. Vrac et al. (2007) compared downscaled estimates of future quantities of local-scale precipitation with, what can be considered, 'pseudo observations' from RCM simulations driven by the same GCM. Whilst this does permit a comparison of two independently-derived estimates, the process introduces additional errors inherent to the RCM itself. It is also noted by Vrac et al. (2007) that agreement between statistical and dynamically (i.e. RCM-simulated) downscaled estimates is dependent on the strength of the perturbation of the future climate scenario. Agreement is good under the B1 SRES scenario (associated with low emissions) and poorest in the highly-forced A1F1 scenario.

So, as any true model validation requires a comparison with observations, the length of observed record (in this case, 1958-2001) is the obvious limitation. In seeking to make maximum use of the observations available, one possibility is an advanced cross-validation in which the downscaling model is fitted on a time period with particular precipitation characteristics (such as years with lightest precipitation) and then validated on an alternative period. It is widely acknowledged that, on a global scale, an enhanced hydrological cycle resulting from an increase in global temperature is likely to lead to increased precipitation (see section Chapter 2, section 2.2) (Allen and Ingram, 2002; Meehl et al., 2007). However, the extent to which precipitation will change in a future climate is highly variable and dependent on both location and season. A more appropriate approach is to use global mean temperature to define fitting and validation period.

Here, an approach suggested by Maraun et al. (2010) is first of all implemented in which the downscaling model is fitted on simulated and observed precipitation during the coolest (and thus, in general, driest) years, according to ERA-40 surface temperature, and validated on the warmest (wettest) years. The fitting and validation process is undertaken on the three downscaling methods developed in Chapter 4, namely local scaling, SVD-RM and CCA-PCR. Model skill is assessed in terms of its ability to simulate

the mean and variance of the observed record.

Additionally, given that an increase in the magnitude and frequency of extreme precipitation events is widely anticipated in a future climate (e.g. Hegerl et al., 2004; Christensen et al., 2007), it is important to understand how well heavy precipitation is simulated in the present climate. Focus was given to this to some degree in Chapter 3 (section 3.4) where ECHAM5-simulated seasonal mean precipitation was split into terciles and assessed on its ability to accurately reflect different magnitudes of monthly mean precipitation. Here, 'heavy' precipitation for a given month is defined as the mean of the five months with heaviest precipitation between 1958 and 2001, and the representation of these quantities in each downscaling model is assessed. It should be noted that the three downscaled estimates for each event are taken from the cross-validated ('leave-seven-out') estimate detailed in Chapter 4. All observations are taken from the GPCC dataset (described in detail in Chapter 3, section 3.3.2).

5.3 Response of precipitation to future changes in large-scale circulation and temperature

Traditional statistical downscaling methods, following a Perfect-Prog approach, rely on the future simulation of large-scale circulation, temperature and often some measure of atmospheric moisture. Additionally, the statistical models that describe the link between these variables and local-scale precipitation are expected to be robust in a perturbed climate. It is therefore important to consider how the relationship of precipitation with large-scale circulation and temperature is projected to change over the course of the next century. Figure 5.1 shows correlation between seasonal precipitation and concurrent mean sea-level pressure (MSLP) in the last two decades of the 20th and 21st centuries. Some key differences in the tropics are immediately identifiable. Across the northern part of South America, the strong positive relationship between MSLP and precipitation is substantially weaker in 2080-2099 during both DJF and JJA. The inverse relationship between surface temperature and precipitation across tropical inner-continental regions also appears far weaker (Figure 5.2).

Whilst the differences in the tropics are notable, it is already been shown in Chapter 3 that ECHAM5 lacks the skill to simulate temporal variability in precipitation across these regions. Of greater interest are relationship changes in regions where confidence in model-simulated precipitation is usually high, including Europe, North America and parts of Australia, as discussed in Chapter 4. During winter to the west of the European mainland, the region of strongest (negative) correlation between MSLP and precipitation would appear to extend in a south-westerly direction out into the Atlantic, a possible consequence of increased westerly flow and a greater contribution to total precipitation from depressions originating in the North Atlantic (Rummukainen et al., 2004; Christensen et al., 2007). Additionally,

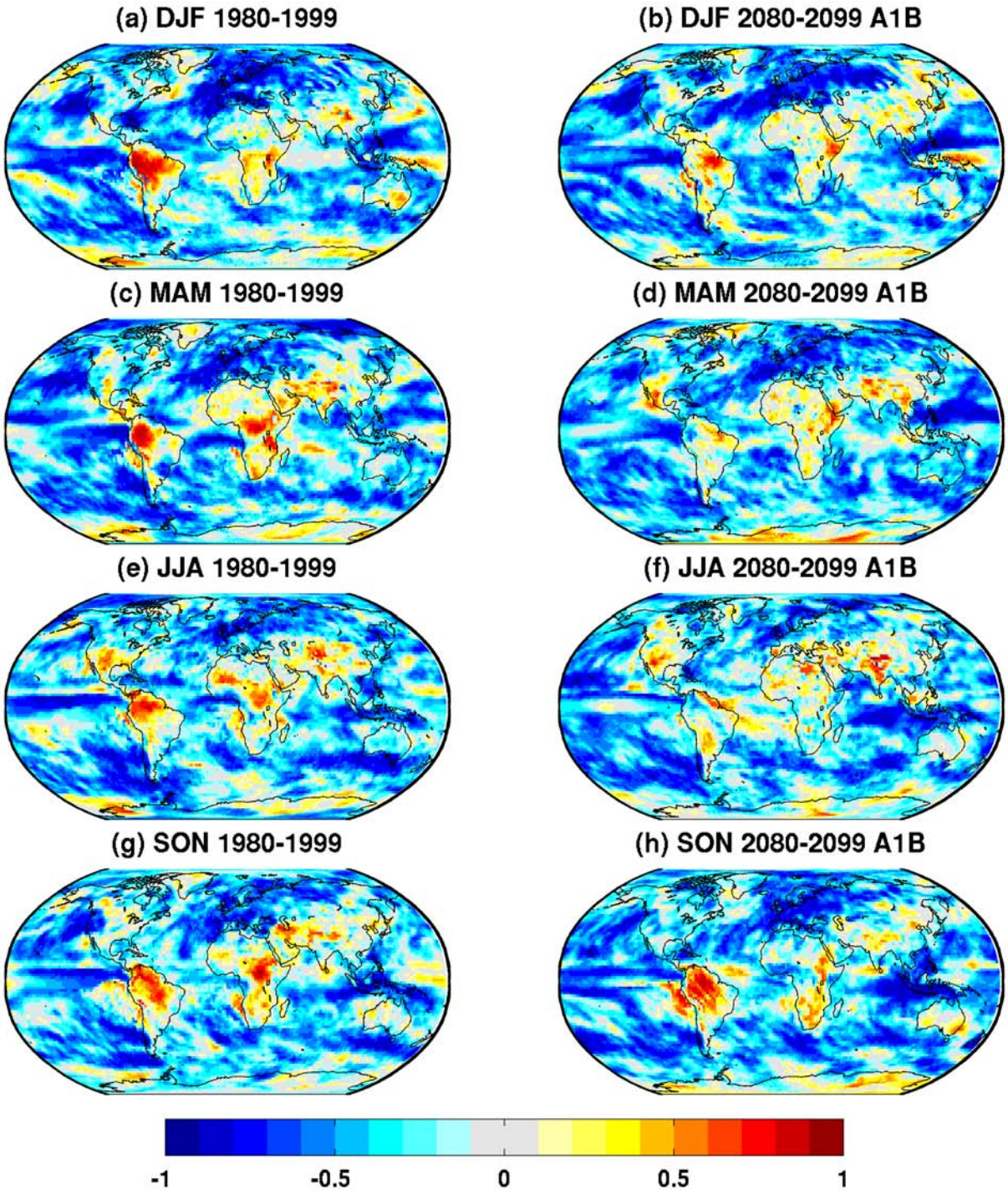

Figure 5.1: Historical (1980-1999) and projected (2080-2099) correlation between ECHAM5-simulated seasonal precipitation and mean sea level pressure. Historical fields taken from the nudged simulation; future simulation conducted according to IPCC SRES scenario A1B.

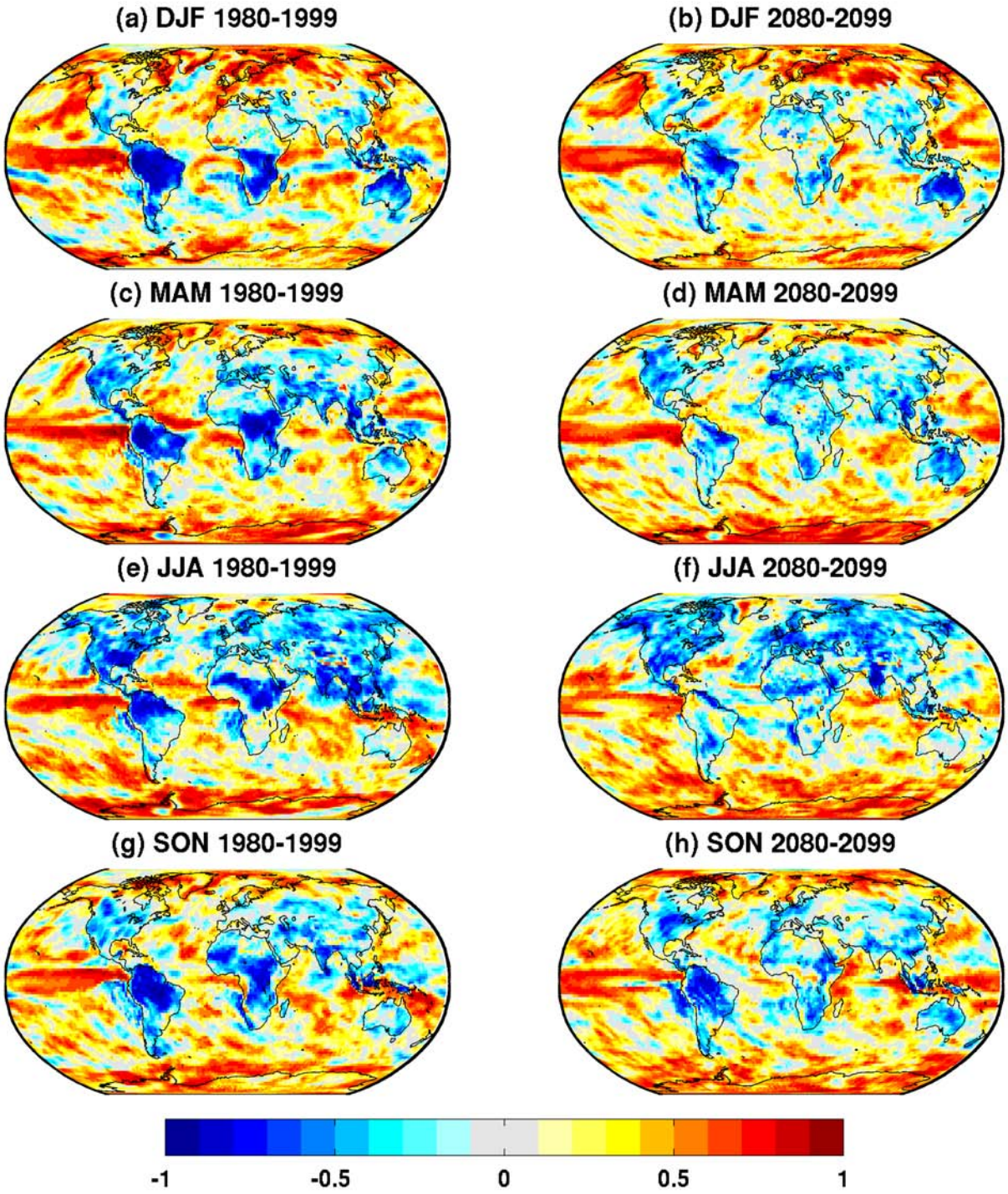

Figure 5.2: Historical (1980-1999) and projected (2080-2099) correlation between ECHAM5-simulated seasonal precipitation and surface temperature. Historical fields taken from the nudged simulation; future simulation conducted according to IPCC SRES scenario A1B.

the dominance of a similar Europe-wide inverse relationship during summer is not as evident in a future climate, with such a relationship substantially weakened in southern parts of Europe. When considering regional precipitation response to climate change forcing, one must also recognise the influence of a change in atmospheric circulation (Hurrell, 1995; Carnell and Senior, 1998). Particularly relevant for European precipitation is an intensification of the North Atlantic Oscillation (more positive on average), which is a major driving factor in the projected increase in precipitation across Eurasia and other mid-latitude land masses (Jones and Conway, 1997; Osborn et al., 1999; Allen and Ingram, 2002).

In North America, the spatial distribution of MSLP- and temperature-precipitation relationships during winter is relatively unchanged between the two periods. During summer in Canada, the negative relationship evident in 1980-1999 is projected to be far weaker at the end of the 21st century. Relationships in Australia do not appear to be of substantial strength for significant differences to be noted.

The changing nature of relationships between precipitation and other atmospheric variables in a future climate obviously has significant implications for Perfect-Prog downscaling. A downscaling model constructed using simultaneous observed quantities of large-scale circulation or temperature and local-scale precipitation describes a relationship which may not be stationary in time. In regions where such relationships are likely to change, an estimate from the same downscaling model may be flawed.

5.4 Stationarity of MOS relationships in future climates

5.4.1 Downscaling precipitation under a warmer global climate

For brevity, results are shown for January and July only (Figures 5.3 and 5.4). Bias in downscaled precipitation during the years of warmest global mean temperature is smallest across the mid- to high-latitudes in the northern hemisphere (Figure 5.3). This is consistent with findings in Chapters 3 and 4 that large parts of Eurasia and North America tend to be associated with good skill of ECHAM5 precipitation in general and of downscaling corrections of monthly mean precipitation. There is little decipherable difference between bias levels in December to February and June to August nor between the downscaled corrections across these areas. In other parts of the world, seasonal differences and intra-model differences are more apparent.

The non-local downscaling models continue to perform slightly differently to the local scaling models, particularly in parts of South America and Australia. In particular, during the winter months (June to August), bias of south-east Australian precipitation in the non-local corrections is far smaller than in the local scaling correction. During summer (December to February), the same pattern is evident in the majority of eastern Australia. Global patterns of the ratio of corrected and observed standard deviations are similar to the distribution of bias, with the amplitude of observed variance most realistically repro-

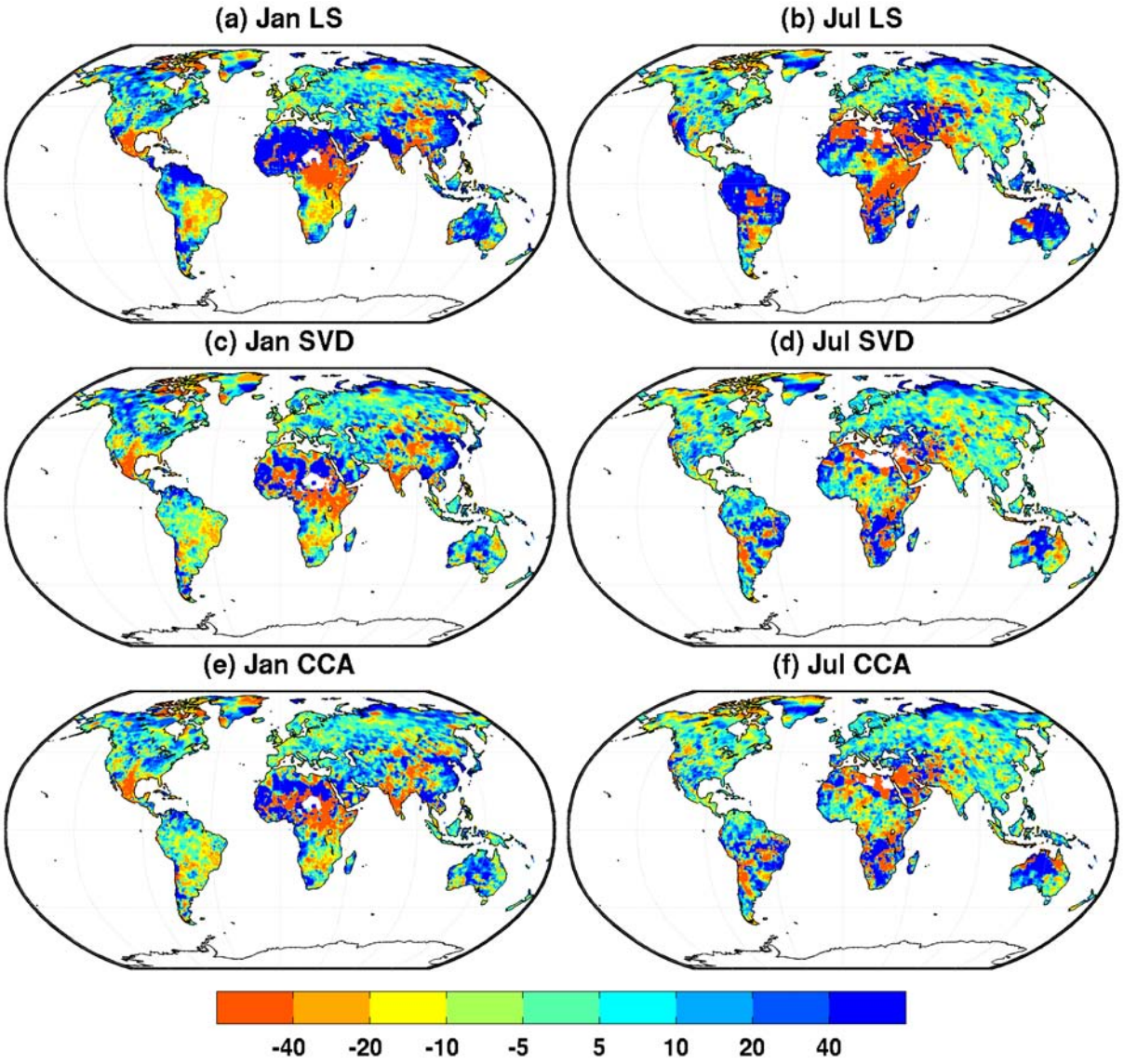

Figure 5.3: Bias (expressed as a percentage) in local scaling, SVD-RM and CCA-PCR downscaling corrections of monthly mean precipitation in the twenty-two warmest years between 1958 and 2001. Downscaling relationships are derived from simulated and observed precipitation from the remaining (coolest) twenty-two years.

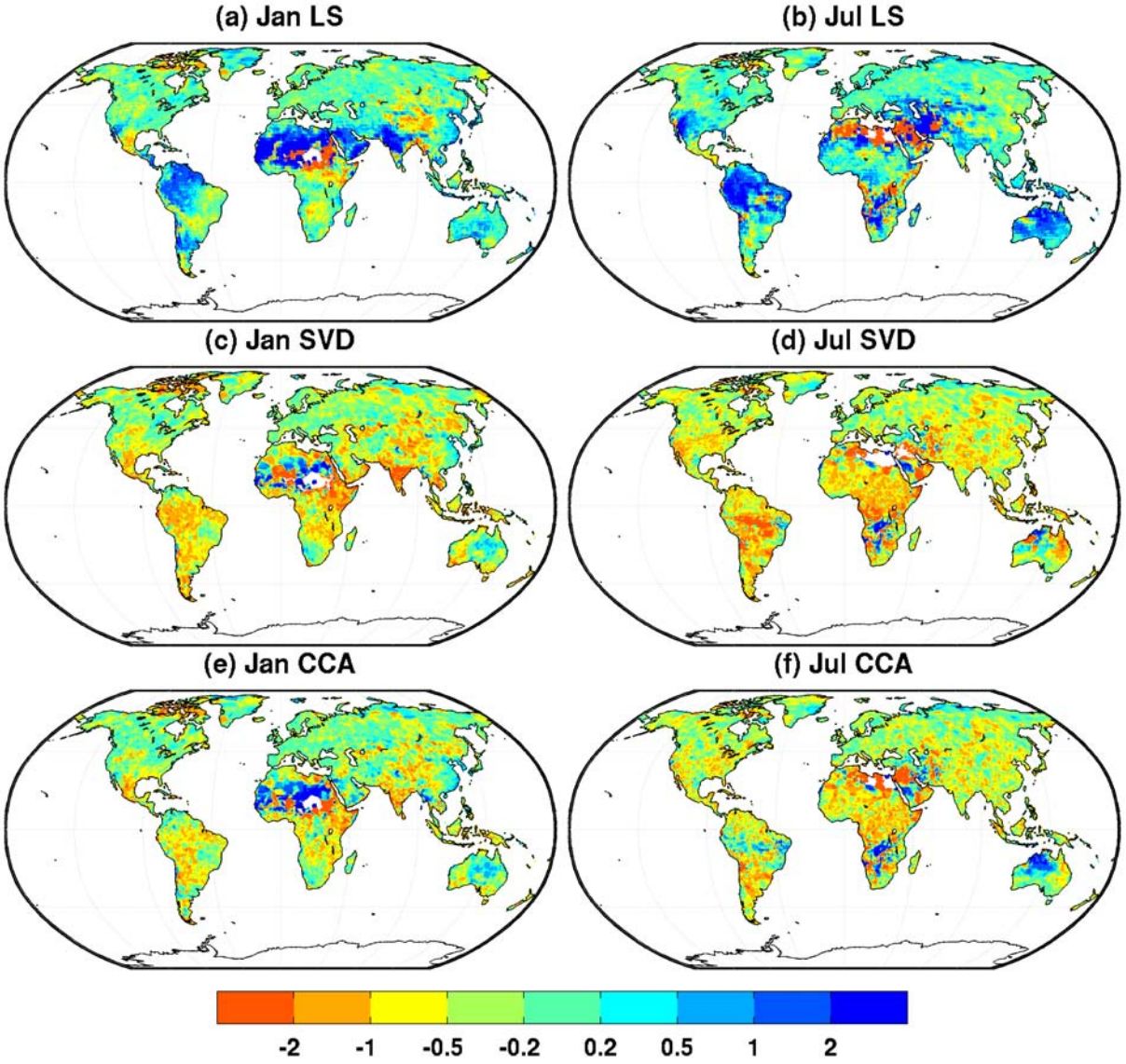

Figure 5.4: Log of ratio of the standard deviations of corrected and observed monthly mean precipitation in the twenty-two warmest years between 1958 and 2001. Downscaling relationships are derived from simulated and observed precipitation from the remaining (coolest) twenty-two years.

duced in the mid- to high-latitudes of the northern hemisphere (Figure 5.4). Realistic standard deviation is more widespread in the local scaling correction. As shown in Chapter 3, the local scaling correction is particularly useful in regions where precipitation is considered skilful (i.e. where a type 3 parameterisation error is small). Precipitation across the mid- to high-latitudes of the northern hemisphere is widely anticipated to increase throughout the next century. Confidence in these projections is largely due to the well-established and consistent processes that link temperature and precipitation at these latitudes. Long-term variance continues to be well-explained by a local scaling correction in warmer climates.

The results shown in Figures 5.3 and 5.4 ideally require an additional interpretation based on statistical significance. Further analysis may include a t-test to determine the degree of similarity between long-term means, which in this case defines the bias statistics presented in Figure 5.3. Similarly, a variance ratio (F) test would provide a suitable metric through which to ascertain whether a significant difference exists between the observed and downscaled standard deviation statistics presented in Figure 5.4. However, at this stage the purpose of the analysis is to gauge an understanding of the potential performance of each downscaling model in a perturbed climate, and to identify areas where performance could be strong. Further analysis, particularly when focused on a given region, should seek to interpret results based on statistical significance.

5.4.2 Downscaling heavy precipitation

The skill of each correction to simulate the heaviest precipitation events is assessed in Figure 5.5. Bias statistics are shown only in areas where correlation between simulated and observed precipitation is greater than 0.6, which discards the majority of the tropics and southern hemisphere (with the exception of parts of Australia and southern Africa) and so for brevity only the northern hemisphere is analysed in detail. Bias in the local scaling correction is generally smaller than that in the non-local corrections. During the winter months (only January is shown in Figure 5.5a-c), much of northern Eurasia and eastern North America is well-represented by local scaling. There is evidence of over-estimation of extreme monthly totals in some regions bordering on the sub-tropics, including southern Iberia, eastern Mediterranean and south-western United States. A possible explanation is that heavier precipitation events at this time of year associated with tropical maritime systems are relatively infrequent. Precipitation simulated under such conditions may be more realistic than that simulated when other systems are prevalent, and so the long-term scaling factor may be too strong a correction. The improved performance of the non-local models, particularly CCA-PCR in some of these regions supports this explanation. In southern and western Iberia and along the Pacific coast of the United States, bias in the CCA-PCR corrections is much smaller than the local scaling equivalent. It is possible that in these and other regions where heavy precipitation events are associated with infrequent and distinct synoptic situations, a non-local correction

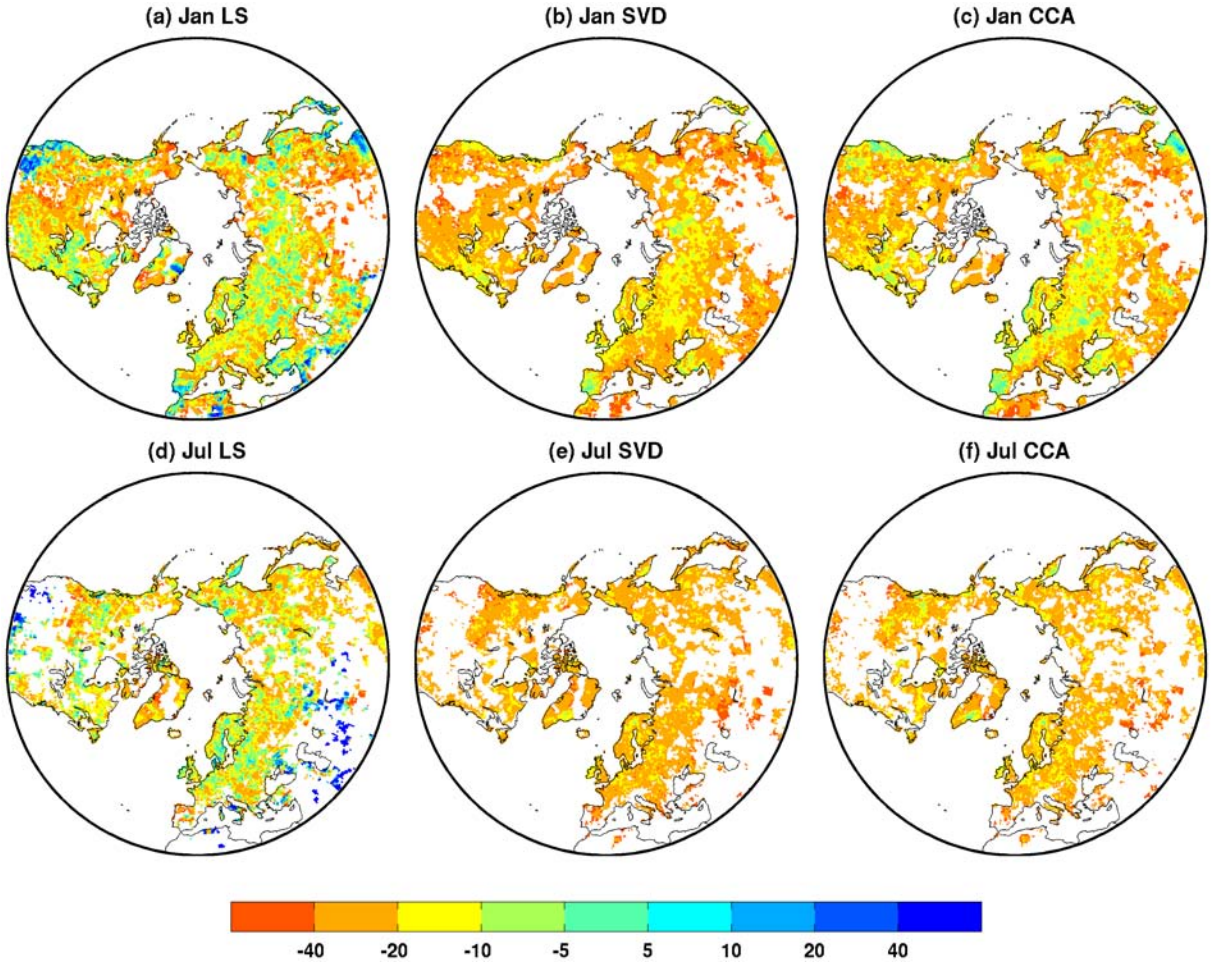

Figure 5.5: Bias (expressed as a percentage) in the correction of mean precipitation during the five wettest monthly events (1958-2001).

approach is more suitable to include neighbouring precipitation information.

Again, CCA-PCR consistently outperforms SVD-RM. CCA-PCR performance is particularly strong across northern Europe and western North America, and in some areas precipitation estimates are more realistic than those made in the local scaling correction. An obvious example is along the western coastline of Scandinavia, where precipitation during winter is consistently heavier than over the rest of Europe. As a result, precipitation variance is far less and what is defined here as 'heavy' precipitation may be nothing out of the ordinary. Given a smaller variance is these areas, it is surprising that the local scaling model does not estimate heavy precipitation with smaller bias. Additionally, a possible explanation for the comparative success of the CCA-PCR model is the distinct precipitation gradient in western Scandinavia, a function of land-sea contrast and the orography of the coastal mountains. Spatial patterns in heavy winter precipitation in this region appear consistent and thus easier to identify.

During the summer months, the performance of non-local models is considerably poorer than local scaling (Figure 5.5d-f). In areas where temporal variability is well-represented (i.e. correlation greater

than 0.6), including most of northern and central Europe, western Russia and western Canada, heavy precipitation events are generally underestimated by 20-40%. A possible explanation is the locality of convective precipitation during summer, and long-term temporal variability in a given location not necessarily matching a neighbouring location. Consequently, the inclusion of precipitation information in a surrounding area in this case does not complement model-simulated precipitation in the location being estimated. Therefore, local scaling appears a more appropriate correction for heavy summer precipitation means.

Although the results inferred from Figure 5.5 suggest local scaling is the most appropriate method in downscaling instances of heavy precipitation, the spatial distribution of bias quantities in some areas is difficult to explain in physical terms. That is not to say the bias statistics are random; in fact, on a continental scale for example, it is possible to identify a gradient in Europe during summer when estimates for heavy precipitation are relatively accurate in the north and generally underestimated in the south. But there is a lack of coherence on smaller-scales, suggesting that components of spatial variability may be misrepresented. Small-scale variability in bias is also consistent with random variability, which cannot be estimated from large-scale fields in a deterministic model. This may result, for example, in two vastly different precipitation estimates in neighbouring locations with similar climate and physical characteristics. Such a difference may even be intensified in a future climate.

5.5 Summary and conclusions

Statistical precipitation downscaling of future scenarios as simulated by GCMs requires a number of considerations on the applicability of a given downscaling methodology to future climates. Such considerations are largely concerned with the robustness or stationarity of predictor-predictand relationships in a changing climate. By definition, the application of Perfect-Prog downscaling models assumes a stationarity in the response of precipitation to future changes in large-scale circulation and temperature. As shown in section 5.3, the relationship of model-simulated precipitation with surface temperature and sea-level pressure (variables that can be considered representative of core predictors used in Perfect-Prog downscaling) is expected to change with a strong degree of regional dependency. This is corroborated by the IPCC AR4 (Meehl et al., 2007) and additional work over the past decade (e.g. Allen and Ingram, 2002; Held and Soden, 2006; Allan and Soden, 2008). It is suggested here, therefore, that a reliance on just one or two observed variables may be insufficient to capture the complete signal of climate change that is necessary to describe future variability in precipitation. As discussed previously in Chapter 4, inclusion of too many fields as predictors may introduce over-fitting issues and erroneous relationships.

The conceptually different MOS approach relies on the relationship between regional-scale observed

precipitation and GCM-simulated precipitation. This variable is a parameterisation with input from the GCM's prognostic variables. That is, the numerically resolved large-scale circulation and temperature fields in a future climate state, which is forced by an emissions scenario. Simulated precipitation is constructed from prognostic fields in physical consistency, and thus, conceptually at least, includes the combined signal of climate change captured in the collation of these fields.

Moreover, a MOS downscaling model is not reliant on the simulation of one or two variables, but ultimately on the simulation of the entire climate state. In this way, the approach requires fewer statistical inferences and places a greater reliance on climate model dynamics. Whilst there are many limitations amongst the current suite of GCMs, it is possible to place greater confidence in future estimations were model skill is shown to be good. Furthermore, downscaling with MOS can be considered an 'evolutionary' approach, being able to run in parallel with the continued development of GCMs and the acquisition of improved observational datasets.

Each downscaling method developed in Chapter 4 has been tested in terms of its applicability in a perturbed climate. Models are fitted with simulated and observed precipitation during the coolest years between 1958-2001 and used to estimate downscaled corrections of precipitation during the warmest years. Corrected precipitation from all three models is similar in terms of large-scale patterns. The local-scaling model appears to show the strongest skill and particularly in the estimatation of heavy precipitation events. Local scaling also appears to add the most value in regions where large-scale frontal precipitation is dominant, often coinciding with a low amplitude of variance and good skill in raw ECHAM5 output as discussed in Chapter 3.

Non-local models, of which CCA-PCR is consistently the best performing, certainly offer value in mountainous regions, and also in regions where precipitation features are forced by a dominant feature of the environment such as land-sea contrast and orographic blocking (i.e. western Scandinavia). A crucial parameter in defining the non-local models, which is not varied in this work, is the size of the spatial domain on which SVD-RM or CCA-PCR is performed. As is discussed in Chapter 4 (section 4.2.2), the domain size is fixed at each grid point for all seasons. Whilst the size used (20° longitude $\times$ 10° latitude) is sufficient in capturing variability in large-scale precipitation characteristics, such a domain size does not appear as useful when small-scale convective events exert a great influence on precipitation totals. Such events are dominant during summer and in regions of the mid-latitudes impacted by tropical air masses. This is especially evident in parts of southern Europe were correction skill is sporadic. Given that an increase in future occurrences of heavy precipitation is likely in a future climate, and that zones of transition between the extra- and sub-tropics are expected to shift, it is important to make allowances for the changing nature of precipitation events themselves that contribute to monthly or seasonal totals. In the context of the MOS models constructed in Chapter 4 and tested further here, a smaller spatial

domain may yield more skilful estimates during the summer months.

Chapter 6

MOS downscaling for climate change scenarios

6.1 Introduction

The IPCC AR4 published the results of multiple simulations from a suite of GCMs for different emission scenarios, with the best estimates for future conditions given by multi-model mean values. Generally, model skill and thus the reliability of these estimates is defined by model consensus. That is, the level of agreement amongst models as to the sign of the change. This is demonstrated in the commonly-used global plots which project an indicator of model agreement, often in the form of 'stippling', onto multi-model mean change maps (Meehl et al., 2007). However, as mentioned in Chapter 3 (section 3.2), model consensus is not a sound measure of reliability since there may exist common errors or deficiencies that occur in all models (Maraun et al., 2010).

Chapter 3 presented a new approach in assessing the skill of GCM-simulated precipitation. Here, a hindcast ECHAM5 simulation was performed in which the circulation and temperature variables were nudged to (and thus forced to, but not replaced by) fields from ERA-40. In comparing the monthly mean precipitation from the nudged simulation with temporally corresponding real-world observations, it is possible to understand and quantify the spatial variability in the GCM's ability to simulate precipitation given an realistically simulated large-scale circulation. Areas of good skill can be identified and, as in shown later in Chapter 3 and in more detail in Chapter 4, it is possible to implement a MOS correction of the simulated precipitation field as part of a statistical downscaling strategy.

In this chapter, MOS downscaling models are used to estimate changes in local-scale precipitation over the course of the next century. Section 6.2 details the ECHAM5 climate change simulations in accordance

with the SRES scenarios and specifies the MOS downscaled models used here. In section 6.3, projections of future precipitation are assessed with reference to the spatial and seasonal quantification of skill diagnosed in Chapter 3. Comparisons are made between the skill of these projections and the collective 'skill' of the GCM ensemble used in the IPCC AR4, defined by the degree of model consensus on century-long precipitation trends. In section 6.4 the MOS downscaling models developed and cross-validated in Chapter 4 are then applied to output from the ECHAM5 projections. Downscaled projections are made for regional precipitation and discussed in terms of their additional value over raw GCM simulations. A summary is given in section 6.5 along with some brief conclusions.

6.2 Future simulations and downscaling methods

6.2.1 ECHAM5 21st century simulations

Climate projections for the 21st century (2001-2100) from a three-member ensemble of ECHAM5/MPI-OM are used in this analysis. Simulations of the future are conducted according to the suite of IPCC Special Report of Emissions Scenarios (SRES) scenario which describe and distinguish between different eventualities of demographic, economic and technological driving forces of greenhouse gas and sulphur emissions (Nakicenovic and Swart, 2000). The B1, A1B and A2 scenarios represent 'low', 'medium' and 'high' scenarios, with respect to predefined greenhouse gas concentration and subsequent radiative forcing, within the full SRES suite (Meehl et al., 2007). This is the subset typically used in the IPCC AR4 and the majority subsequent climate change studies, and these scenarios are included in this analysis also.

6.2.2 Methods

Output from the future simulations is downscaled using each of the three MOS models developed in Chapter 4: local scaling, SVD-RM and CCA-PCR. As in Chapter 4, analysis is again global to begin with and then focused on the regions of Europe, North America and Australia. The rigorous cross-validation performed in Chapter 4 permitted a full assessment of each downscaling model's bias tendencies and ability in reproducing the temporal variability and long-term variance in the observed record. An understanding of the spatially and temporally varying skill of each model is crucial to interpreting the downscaled output of future climate projections, for which, of course, there are no observations with which to make comparison. Application of a downscaling model to some future climatic state assumes stationarity of the statistical relationships underpinning each model. That is, the predictor-predictand relationship, on which the downscaling model is calibrated, is assumed to be valid in a perturbed climate. The stationarity of each MOS downscaling model was assessed in Chapter 5, and was shown to have a

regional and seasonal dependence similar to other measurements of skill.

6.3 Interpreting the skill of ECHAM5 projections

ECHAM5/MPI-OM estimates a global warming over the next century, with mean global surface temperature expected to increase by between 2.5°C (B1) and 4.1°C (A2) relative to 1961-1990 (Roeckner et al., 2006). An increase of 3.7°C is expected in the A1B scenario, which is notably similar to the A2 scenario despite differing CO_2 concentrations. An increased water-holding capacity of a warmer atmosphere will lead to an intensification of the hydrological cycle and a consequent increase in global mean precipitation during all seasons (Douville et al., 2002; Meehl et al., 2007). There are substantial spatial and seasonal variations in projections of precipitation from various GCMs, and model consensus is far less refined than that for projections of global mean temperature (IPCC, 2007; Meehl et al., 2007). Amongst the features most consistently projected are increased precipitation at high latitudes and over the tropical oceans, and decreased summer precipitation in the mid-latitudes (Wang, 2005; Neelin et al., 2006). The relationship between global mean precipitation and increasing global temperatures is similar in each of the SRES scenarios; global mean precipitation is expected to increase by 7% in both A1B and A2, and by 5% in B1.

Assessing the reliability of climate change simulations usually involves an assessment of the level of consensus amongst a number of models. This may be limited to identifying those models that agree on the sign (positive or negative) of a change over time, or the magnitude of that change. With an understanding of the skill of ECHAM5-simulated precipitation to reproduce temporal variability, it is possible to objectively assess the reliability of climate change simulations. Figures 6.1-6.3 show projected changes in global precipitation, according to three SRES scenarios, between the last two decades of the 20th and 21st Centuries. The stippling overlay in these figures indicates areas where ECHAM5 shows skill in reproducing temporal variability is high. Although the interpretation of model skill is different, it is possible to make comparisons with model consensus plots published in the IPCC AR4 (IPCC, 2007; Christensen et al., 2007), which highlight areas where 90% of the models used in the AR4 agree on the sign of the change.

It is useful to identify, first of all, areas where strong multi-model consensus is matched with good skill in ECHAM5 precipitation. This is the case across the majority of the northern hemispheric land mass during the Boreal winter months (DJF) where each projection estimates a large-scale increase in precipitation, with the strongest increases occurring in northern and eastern Canada and Siberia. High confidence in multi-model projections of decreasing precipitation in the eastern Pacific either side of the equator is also matched by good skill of ECHAM5 in these regions. A likely decrease in precipitation in

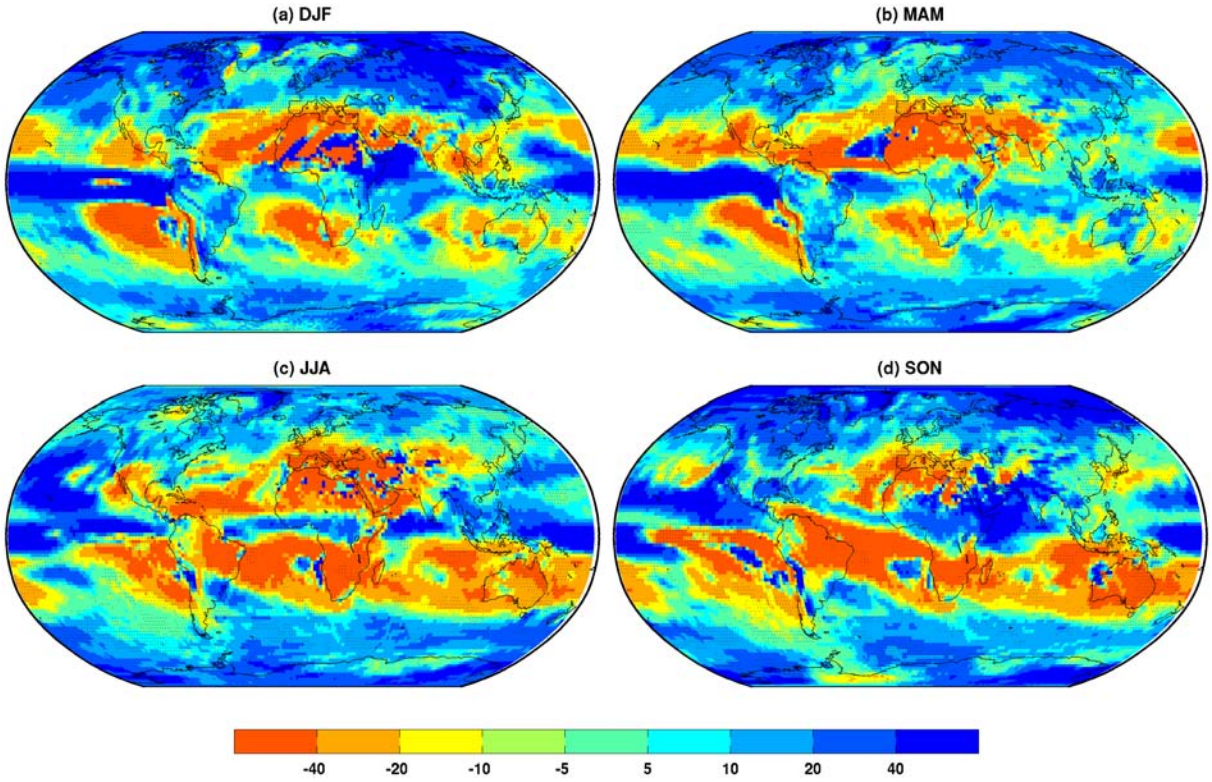

Figure 6.1: A1B ECHAM5 seasonal projections; 2080-2099 relative to 1980-1999. Differences expressed as percentages. Stippling indicates areas where correlation coefficient between observed and nudged simulation (1958-2001) is greater than 0.5.

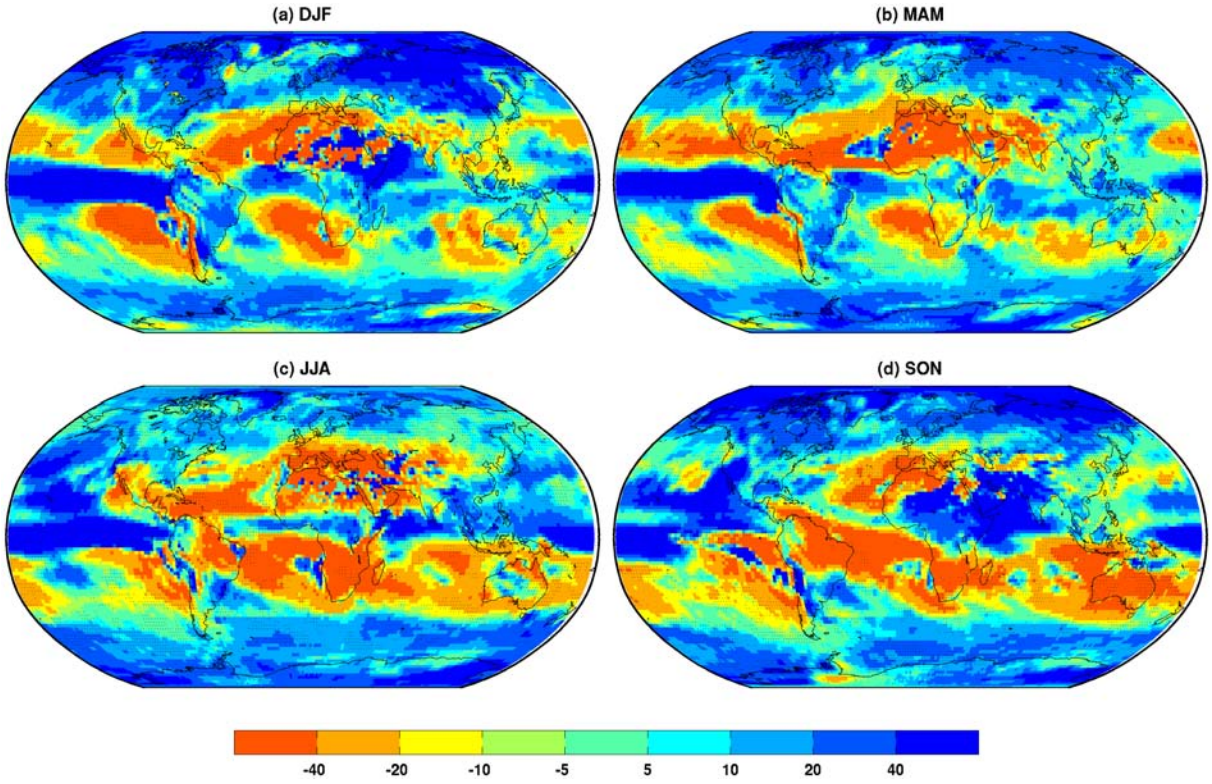

Figure 6.2: As Figure 6.1 but for IPCC SRES A2.

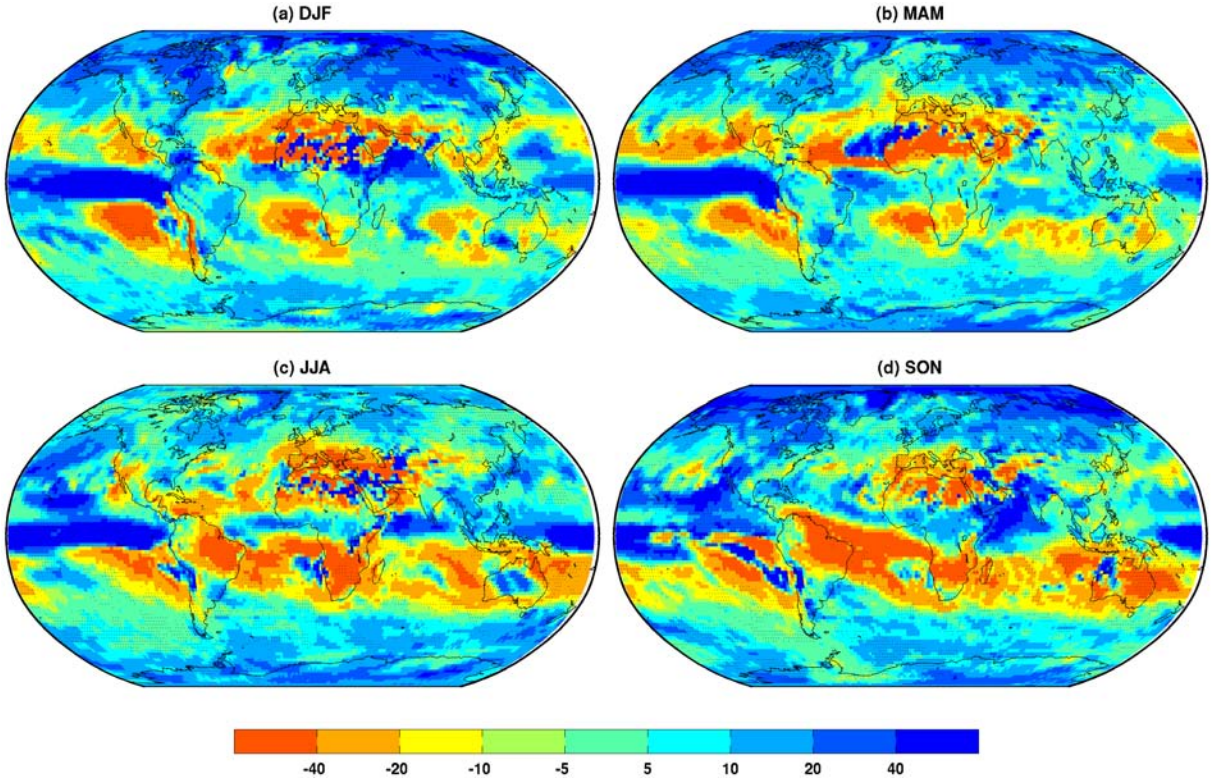

Figure 6.3: As Figure 6.3 but for IPCC SRES B1.

the North Atlantic is also a feature of ECHAM5 which is agreed upon by the multi-model data.

Conversely, a comparison of this nature can also provide extra information on model skill, and thus the degree of uncertainty, in a region where model consensus is lacking. For example, ECHAM5 projects large-scale decrease over much of Australia according to the A1B SRES scenaro (Figure 6.1), particularly during JJA but also during DJF, and skill is good across many central and eastern areas. Results from the IPCC AR4 show that an insufficient number of models agree on the sign of the change across much of Australia, in spite of ECHAM5 projections meriting a seemingly high confidence. Additionally, ECHAM5 shows reasonable skill over the area encompassing north-west United States and western Canada, where a simulated increase in summer precipitation contrasts strongly with the IPCC multi-model projection of drier conditions. During winter, precipitation in Europe also shows good skill in ECHAM5. It is possible to identify with more certainty the north-south gradient in precipitation increase/decrease, and thus the region of zero change which is difficult to pin-point in the IPCC multi-model mean.

GCMs may accurately reproduce large-scale climate features which are spatially 'out of sync' with real world observations. For example, two different GCMs may realistically simulate the magnitude of precipitation events associated with North Atlantic storm tracks, but each simulation may erroneously place the mean track trajectory either north or south of that observed in the real world. Obviously, in a point-to-point comparison of the precipitation fields, these two simulations do not agree with one another

and a multi-model mean would not reflect the relative merits of each simulation's representation of the large-scale features.

6.4 Estimation of regional changes in 21st century precipitation

Focus is now given to downscaled projections of 21st Century precipitation in Europe, North America and Australia, regions where ECHAM5 precipitation is shown to be most skilful. The skill of MOS and Perfect-Prog downscaled models was assessed in detail over these regions in Chapter 4. It is important to understand how different downscaling methods represent actual precipitation characteristics in different areas before analysing projected changes over the next century. Throughout this section, analysis for each region is thus in two parts. Firstly, ECHAM5 projections of seasonal mean precipitation for the period 2080-2099, taken as the mean of a three-member ensemble conducted under the SRES A1B scenario, are compared with downscaled estimates for the same period. Analysis of absolute precipitation quantities permits an assessment of spatial consistency in the actual precipitation features of the raw and downscaled estimates. Secondly, the raw and downscaled projections are assessed in terms of their respective deviation from 1980-1999 precipitation, and with reference to model skill in each case. It is also necessary at this stage to discuss ECHAM5 projections in the context of the results of the multi-model ensemble included in the IPCC AR4, and to briefly summarise the dominant features. Throughout this section, the 'correction' is defined as the downscaled estimate for projected precipitation.

6.4.1 Europe

6.4.1.1 Absolute precipitation

ECHAM5 and downscaled estimates for late 21st Century seasonal mean precipitation are shown in Figure 6.4. The spatial patterns of local precipitation from each downscaling model are generally in good agreement with coarse output from ECHAM5, although offering estimates at far greater resolution and the promising representation of precipitation characteristics not captured by in the initial GCM simulation. The downscaled corrections show a greater concentration of precipitation along the coastline of western Europe during autumn and winter and to a lesser extent during spring, and this pattern extends to the western coastlines of Italy, Greece and the Balkans. In particular, corrected Autumn precipitation is considerably heavier in much of the British Isles, western France and northern Spain.

Spatial variability of precipitation in south-eastern Europe appears poorly represented in ECHAM5. In the corrections, northern and western parts of Turkey show heavier precipitation during autumn and winter whilst the interior is far drier than is suggested by the GCM. Additionally, the arc of the western Caucasus along the Black Sea coast is expected to be considerably wetter than the areas to the north

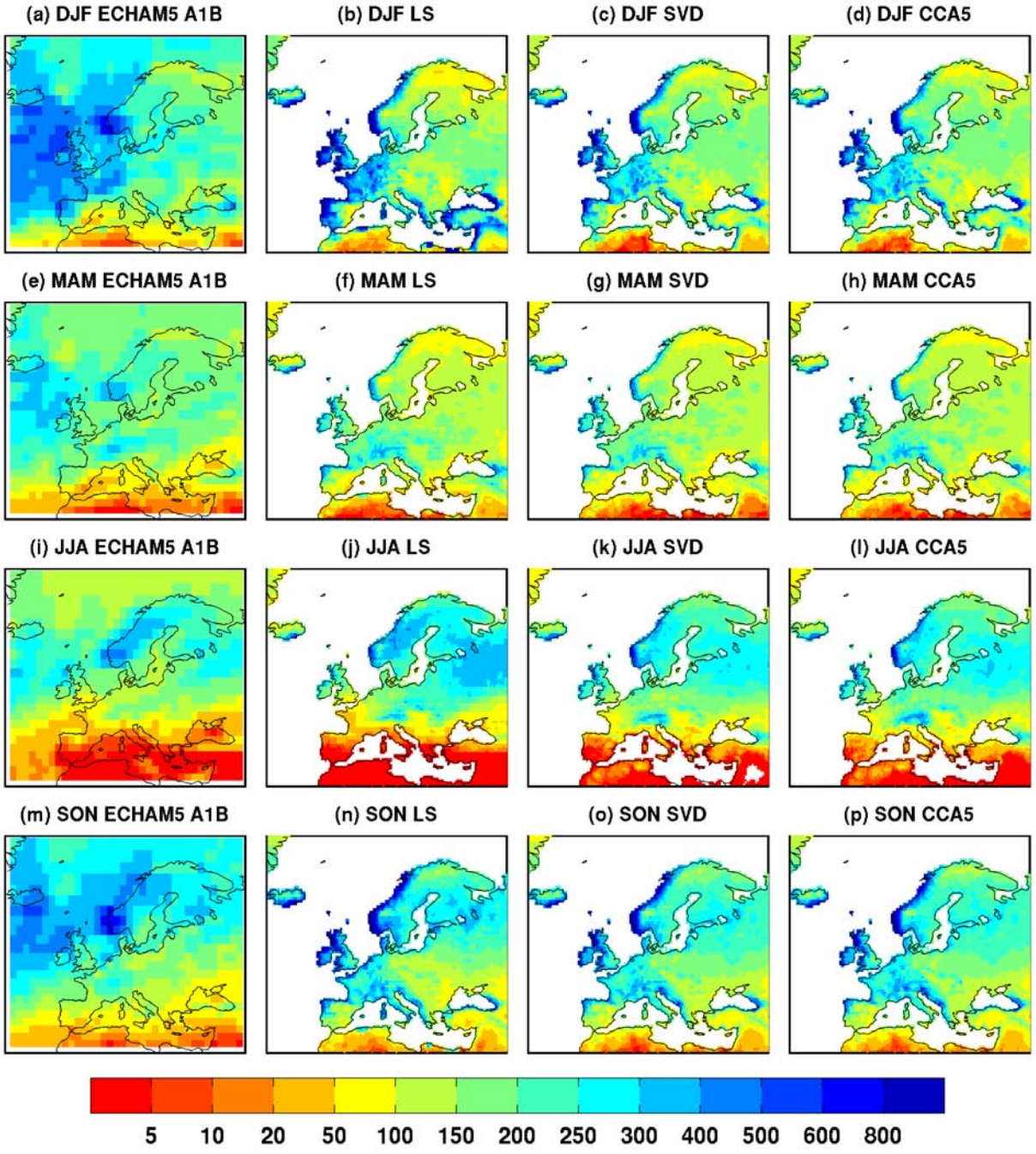

Figure 6.4: Projected ECHAM5 A1B (extreme left panels) seasonal precipitation and downscaled corrections (remaining panels) for the period 2080-2099 over Europe. Precipitation expressed in mm.

and south, a feature reasonaly captured by ECHAM5 during winter and spring but not at other times of year.

As expected, downscaled corrections provide far greater representation of precipitation features over mountainous terrain. Across the Alps, heavier precipitation is not in evidence in ECHAM5 projections during summer, but each correction shows heavier precipitation along the northern and eastern slopes. During the winter months, a clear Alpine rain shadow exists but the region of heaviest precipitation in ECHAM5, which one would associate with orographic rainfall, in the Alps would appear to occur too far north and west. This feature is also in evidence during autumn and spring and suggests that central European topography is insufficiently represented in ECHAM5, which in turn produces an erroneous location of the heaviest precipitation. In the corrections, the spatial distribution of Alpine precipitation is more consistent with topographic influences. Furthermore, heavier precipitation means in the Carpathians and Pyrenees are also captured by the corrections. Over Scandinavia, heaviest corrected precipitation is concentrated along the western coastline during winter, and to a lesser extent during autumn and spring, a feature that ECHAM5 is unable to represent.

The skill of both SVD-RM and CCA-PCR downscaling models is high across most of Europe (Figure 6.5). During summer (JJA), skill is notably poorer in dry parts of southern Europe and the Alps. Overall, there is little difference in the spatial distribution of local-scale precipitation estimated by the three downscaling models. It is notable that the regions of heaviest precipitation appear wetter in the local scaling (LS) estimates than in the other corrections, particularly along the Atlantic coastlines of western Europe and the region of heavy precipitation in north-east Europe spanning Belarus and parts of western Russia. Given that the local scaling model showed a tendency to overestimate months of extreme precipitation (see Chapter 5), caution should be exercised in interpreting these estimates. The non-local models (SVD-RM and CCA-PCR) are similar and many of the small scale features associated with topographic effects are represented by both corrections. The CCA-PCR model estimates slightly wetter conditions in continental mountainous regions, such as the eastern Alps and northern Carpathians during spring and summer.

6.4.1.2 Projected precipitation changes

Projections of seasonal precipitation changes between the late 20th and 21st Centuries are shown in Figure 6.5a-b. Annual precipitation change is expected to follow a north-south contrast according to IPCC ensemble mean, showing a general trend of increases in most of northern Europe and a decrease in the south and Mediterranean (Christensen et al., 2007). In central Europe winter precipitation is likely to increase whereas summer precipitation is likely to decrease. There is also likely to be a shortening of the European snow season and a decrease in snow depth throughout most of Europe (Christensen

et al., 2007). There is a relatively good level of consistency in model simulations compared with other regions but there are still significant uncertainties. Whilst most models support the north-south pattern of precipitation projections, they lack agreement on the magnitude and finer geographical details of future changes. Christensen et al. (2007) point to the particular sensitivity that European climate has to global warming as explanation for this uncertainty.

A stippling overlay in these plots is again used as an indicator of model skill; that is, to highlight regions where the ECHAM5 is able to reproduce temporal variability according to validation tests undertaken in Chapter 3. The north-south contrast in precipitation changes, with an increase (decrease) projected in the north (south), is well-established in GCM climate change simulations (Christensen et al., 2007). Increases are largest in northern and central Europe, and are an effect of the increased atmospheric moisture of a warmer climate (Kendon et al., 2010). Conversely, over the majority of southern Europe, a decreasing relative humidity is likely to lead to a reduction in precipitation occurrence (Rowell and Jones, 2006; Kendon et al., 2010). ECHAM5 is a reputable GCM included in the IPCC multi-model estimates, and its projections concur with these general trends. The stippling in Figure 6.5a-b suggests skill is greatest over northern Europe and during winter. Much of southern Europe and regions of mountainous terrain are associated with poorer skill.

Downscaled corrections of the ECHAM5 A1B projections are shown in Figure 6.5c-f. The local scaling correction is omitted here; as the change is defined as the percentage difference between local (single-point) downscaled corrections of both past and future simulation, the difference itself is identical to that of the raw ECHAM5 output. The SVD-RM and CCA-PCR models are non-local, constructed using information from a surrounding domain and thus the long-term trend (including climate change signal) at a number of locations. The large-scale spatial patterns of changing precipitation are similar to the raw ECHAM5 output, although some important differences are evident. For example, winter precipitation along the western and northern coastline of Scandinavia is not expected to increase as projected by ECHAM5, with the CCA-PCR correction in particular showing little change along much of the coastline even a reduction in the far north. This feature of decreased precipitation is corroborated in RCM and limited area model simulations which are able to better represent the orography of this region and the processes interacting with it (e.g. Ekstrom et al., 2007; Deque et al., 2007). Kendon et al. (2010) suggest that whilst warming is the dominant mechanism driving precipitation increase across the majority of northern Europe, here warming is responsible for changes in circulation which offset the effect of increased moisture and in turn result in an overall decrease. The authors further suggest that the warming gradient in this region, driven largely by an amplification of warming at high latitudes as a result of sea-ice reduction, combines with the local orography and disrupts prevailing flow. It is interesting that, although the CCA-PCR correction does not physically resolve these processes, there is an apparent agreement between independent statistically

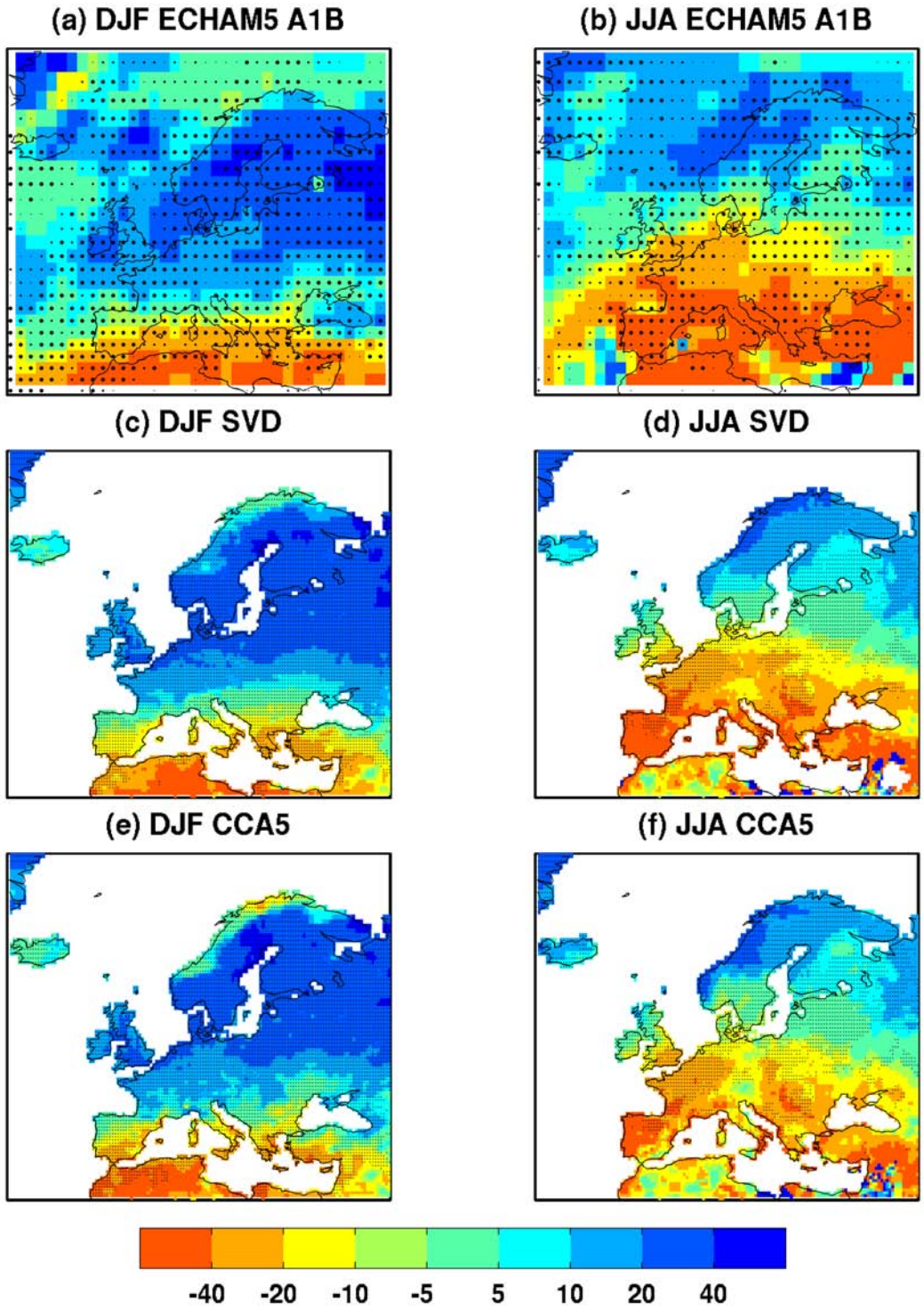

Figure 6.5: Projected (a-b) and downscaled (c-f) seasonal precipitation changes over Europe from the A1B ECHAM5 simulation. Precipitation change between 1980-1999 and 2080-2099 expressed as a percentage. (a-b) Stipples sized to indicate correlation between observed and simulated (nudged ECHAM5 simulation) precipitation; small (r=>0.5), medium (r=>0.6) and large (r=>0.7). (c-f) Stipples indicate areas were correlation between cross-validated downscaled estimate and observation is strong (r=>0.7).

and dynamically downscaled estimates.

Over the Alps, the CCA-PCR correction estimates that the zone of increased winter precipitation will extend further south than is projected by either the SVD-RM correction or ECHAM5 itself. The CCA-PCR model has slightly greater skill then the SVD-RM equivalent in southern and western Alps (section 4.3.1). However, both non-local corrections are in agreement about precipitation across the eastern Alps during summer, projected by ECHAM5 to decrease by 20-40%. SVD-RM suggests a decrease in this region will only be between 10-20%, whilst CCA-PCR estimates between 5-10%. Skill of each downscaling model is reasonable is this region, but not consistently high. A similar difference between the corrections in an area of reasonable skill is over the southern Balkans and central Greece, where the decrease estimated by CCA-PCR is not as severe as that estimated by SVD-RM.

6.4.2 North America

6.4.2.1 Absolute precipitation

Again, downscaled corrections reflect the general spatial patterns of 2080-2099 precipitation projected by ECHAM5 (Figure 6.6). Within the corrections, there exist a number of small-scale features that are not present in the coarse GCM output. Along the north-east Atlantic coast during winter, and to lesser extent during spring, all corrections estimate a concentration of heaviest precipitation along the coast and not reaching as far in land as projected by ECHAM5. The situation is similar in the Pacific Northwest during all seasons, where the precipitation gradient with distance from the coast is gentler and more resolved. Orographic precipitation associated with the Cascade mountains during autumn and winter is captured in each correction. During winter, ECHAM5 projects a low precipitation region to the east of the Canadian Rockies. In the corrections, this driest area exists further south over the Great Plains of the United States, suggesting a spatial discrepancy of this feature in ECHAM5.

The local scaling correction is largely in agreement with both non-local corrections, but does produce estimated precipitation in southern and central United States during spring and summer that is more intense than either the ECHAM5 projection or the subsequent non-local corrections. This is also the case during summer over western Mexico. Furthermore, during winter over the majority of northern and central Canada, precipitation is considerably less than in the projection or the other corrections.

6.4.2.2 Projected precipitation changes

Multi-model projections detailed in the IPCC AR4 indicate a general increase in precipitation over the course of the next century over most of North America, with the exception of the south-western region (Christensen et al., 2007). Consensus amongst the IPCC models is poorer in the southern United States

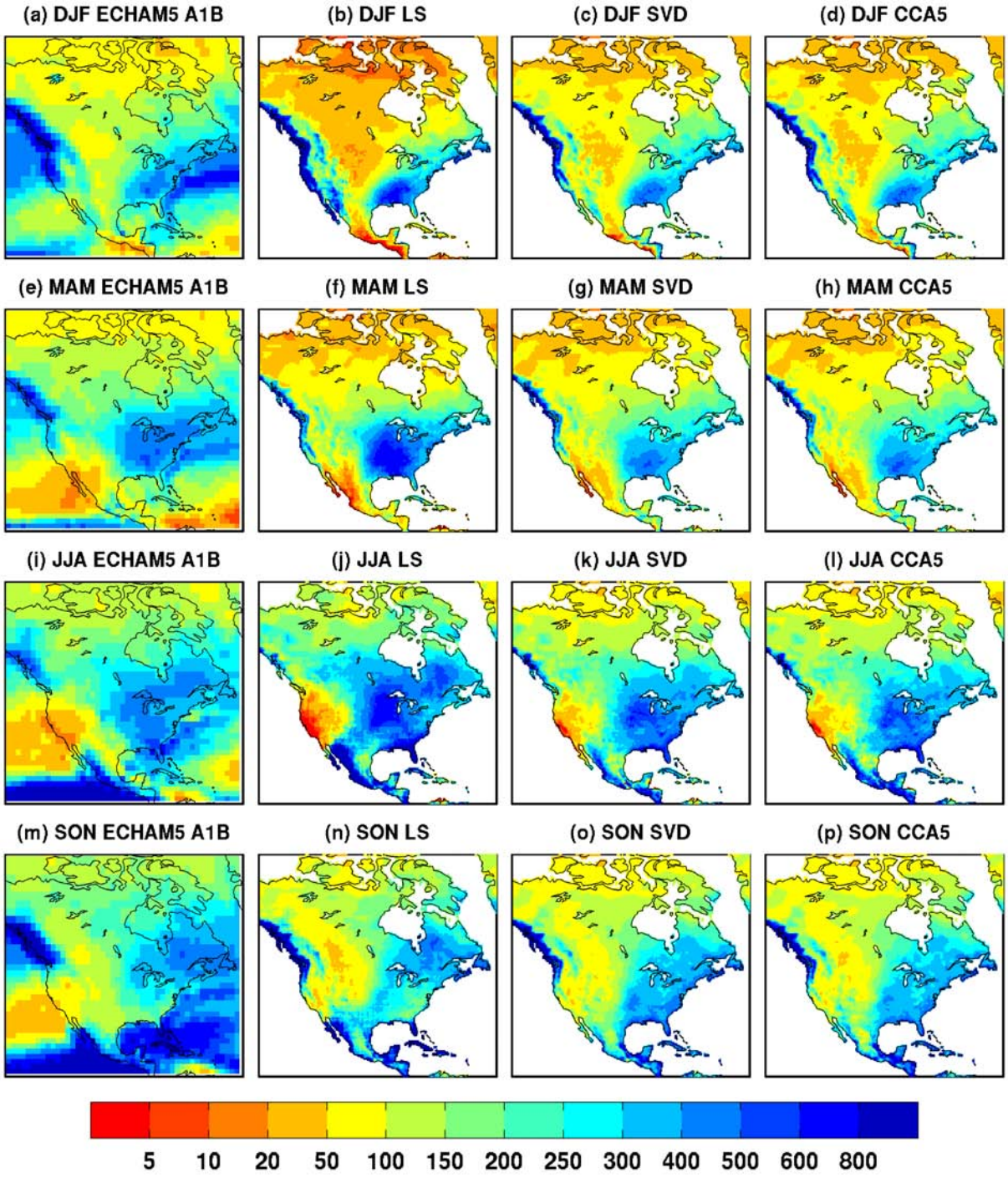

Figure 6.6: Projected ECHAM5 A1B (extreme left panels) seasonal precipitation and downscaled corrections (remaining panels) for the period 2080-2099 over North America. Precipitation expressed in mm.

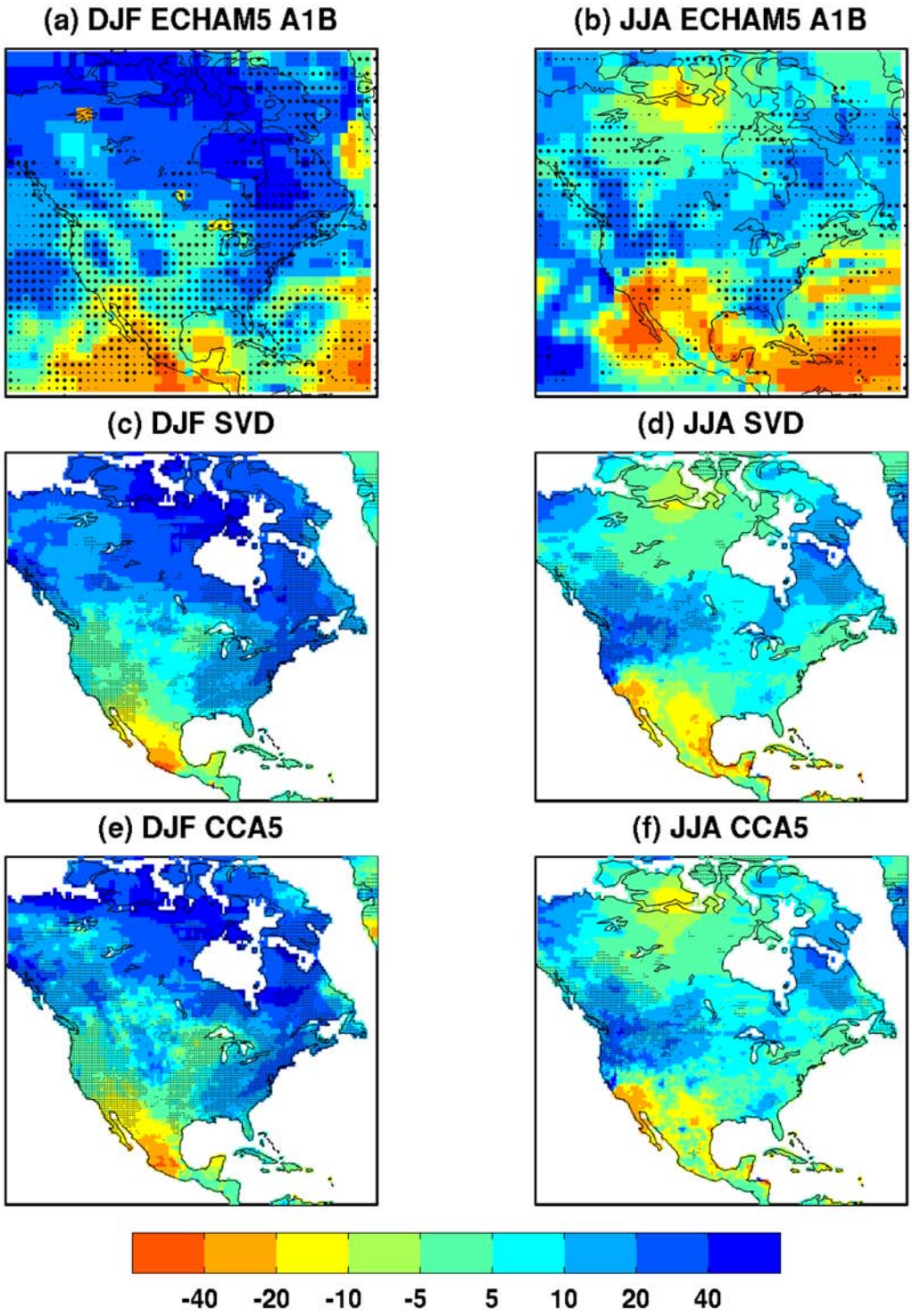

Figure 6.7: Projected (a-b) and downscaled (c-f) seasonal precipitation changes over North America from the A1B ECHAM5 simulation. Precipitation change between 1980-1999 and 2080-2099 expressed as a percentage. (a-b) Stipples sized to indicate correlation between observed and simulated (nudged ECHAM5 simulation) precipitation; small (r=>0.5), medium (r=>0.6) and large (r=>0.7). (c-f) Stipples indicate areas were correlation between cross-validated downscaled estimate and observation is strong (r=>0.7).

and northern Mexico during both winter and summer. Changes in mean precipitation have also been shown to be associated with the frequency and spatial variability of extreme events (e.g. Bell et al., 2004). The far north of Canada is the only region where snow depth and snow season length are not likely to decrease (Christensen et al., 2007).

Similarly to Europe, ECHAM5 precipitation skill is again greater during winter than summer (Figure 6.7a-b). Likewise, skill of the SVD-RM and CCA-PCR downscaling models is greater during winter across large parts of the continent with the exception of the Rocky Mountains range (Figure 6.7c,e). During summer, skill is lacking across central and southern United States (Figure 6.7d,f).

Projected downscaled changes between 2080-2099 and 1980-1999 are, again, largely similar to raw ECHAM5 projected changes. Work using RCM simulations have shown little, if any, extra information at smaller scales (Chen et al., 2003; Plummer et al., 2006; Christensen et al., 2007). Here, the corrections are able to better represent a more concentrated zone of increased winter precipitation along the north-east coast.

6.4.3 Australia

6.4.3.1 Absolute precipitation

During summer (DJF) and autumn (MAM), a narrow band of heaviest precipitation along the east coast is captured by each correction (Figure 6.8). During summer, this band is shown to extend further inland than is suggested by the ECHAM5 projection, particularly in southern Queensland. During autumn, large parts of eastern Australia are associated with heavier rainfall and a more pronounced gradient with distance from the coast. Between autumn and spring, the heavier concentration in southern Victoria is captured in the corrections, as is the west-east precipitation gradient in Tasmania. Precipitation features in the south-west corner of Australia are also represented

The local scaling model has already been shown to exhibit bias in the downscaling of tropical precipitation, and locally scaled estimates of precipitation across the north of the continent are over-estimated during summer (DJF). In some areas of the Northern Territory, the local scaling correction estimates twice as much precipitation as the non-local corrections. The reverse is evident during the drier winter months (JJA) when locally scaled precipitation is lighter than that simulated by ECHAM5. The non-local corrections estimate marginally wetter conditions in the majority of Australia during all seasons and particularly in eastern Australia during winter and spring than is projected by ECHAM5.

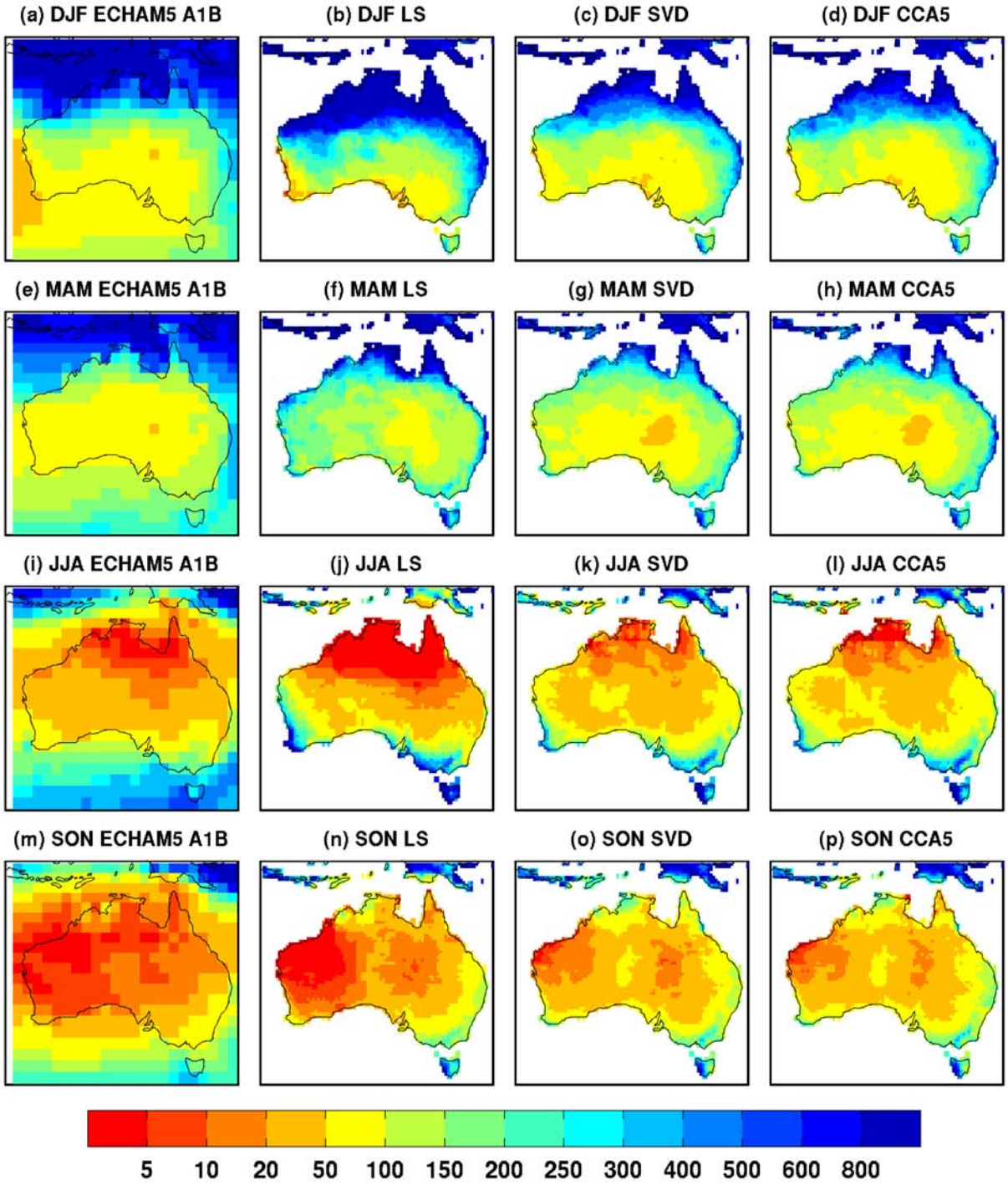

Figure 6.8: Projected ECHAM5 A1B (extreme left panels) seasonal precipitation and downscaled corrections (remaining panels) for the period 2080-2099 over Australia. Precipitation expressed in mm.

6.4.3.2 Projected precipitation changes

Skill of each downscaling model varies greatly by season (Figure 6.9). Good skill is considerably more widespread across the bulk of the continent during the wetter summer months than the drier winter. Much of the eastern part of Australia is associated with good skill during both winter and summer. During summer, the downscaled corrections largely agree with the changes projected in ECHAM5, which can be broadly described as a decrease throughout much of the continental interior and little or no change along the coasts. Only in parts of the south-western region is any sort of increase corroborated in the downscaled corrections (in this case, the CCA-PCR estimate). There is a notable difference in the spatial extent of a 10-20% precipitation decrease in the SVD-RM and CCA-PCR corrections. In the former, this zone extends from the tropical north to South Australian coast whereas the latter is associated with a smaller area of projected decrease in excess of 10% and a steeper gradient of change. In the CCA-PCR correction, the region of a projected decrease of greater than 20% extends into south-eastern Australia and into the Murray-Darling basin, as projected in the raw ECHAM5 output.

Winter (JJA) precipitation is expected to decrease across the majority of the continent. In much of northern Australia, precipitation is expected to decrease by more than 40% according to raw ECHAM5 output. In the far north, this cannot be corroborated by the downscaled corrections due to insufficient model skill. However, CCA-PCR (and to a lesser extent SVD-RM) skill is acceptable along the coast of south-eastern Queensland, and extends southwards over the majority of eastern Australia as already mentioned. Corrections in these parts of Queensland can be considered skilful, and indicate that the projected decrease actually falls within the range of 10-20%, except for the most easterly coastline and in the western lee of the Great Dividing Range.

Also during winter, as mentioned previously, the downscaling models are able to capture the heavier precipitation features in the south-east, and particularly along the south coast of Victoria. The corrected 1980-1999 to 2080-2099 change suggests drying in this region is unlikely to be as severe as is projected by ECHAM5. In the CCA-PCR correction, change in parts of the coastal region is expected to be negligible, although precipitation in the remainder of Victoria is still expected decrease by 10-20%. Additionally, there is a suggestion, supported by good skill, that winter precipitation in Tasmania will be largely unchanged expect in the south/south-west part where a slight increase is anticipated.

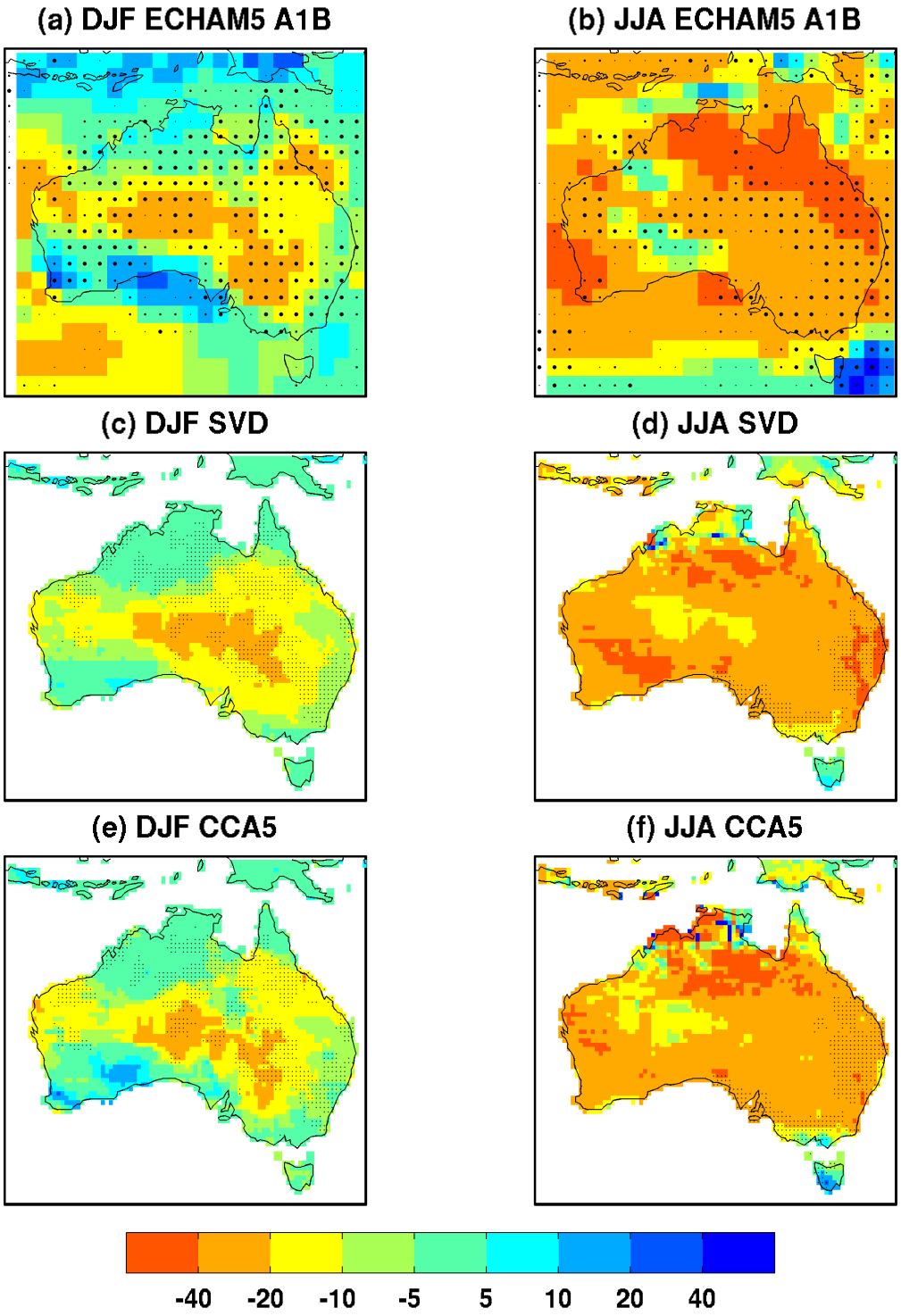

Figure 6.9: Projected (a-b) and downscaled (c-f) seasonal precipitation changes over Australia from the A1B ECHAM5 simulation. Precipitation change between 1980-1999 and 2080-2099 expressed as a percentage. (a-b) Stipples sized to indicate correlation between observed and simulated (nudged ECHAM5 simulation) precipitation; small (r=>0.5), medium (r=>0.6) and large (r=>0.7). (c-f) Stipples indicate areas were correlation between cross-validated downscaled estimate and observation is strong (r=>0.7).

6.4.4 Other regions

Whilst skill in both ECHAM5 precipitation and subsequent MOS corrections is poorer in many other land surface regions, it is important to briefly summarise some of the general features and what, if anything, can be offered by MOS downscaling. In particular, one must consider that in many of these regions the quality, density and completeness of station observational records is poorer than in the areas described in more detail in earlier parts of section 6.4. This, of course, has implications for the reliability of ERA-40 and other reanalysis products upon which a potential nudged simulation is based. Most importantly, the development and, crucially, the validation of MOS corrections is hugely dependent on reliable observational data for the predictand. However, whilst quality data may not be available for long periods in many regions, it is important to consider that global observation is a growing part of climate science, particularly in the last decade or two. With improvements in the observational network, MOS downscaling may still have good potential even in regions where skill has been shown to be relatively poor.

Throughout most of Asia, the IPCC ensemble of GCMs suggests an increase in precipitation with models showing most consistency in North and East Asia (Christensen et al., 2007). These projected increases are expected to apply to both winter and summer precipitation. However, summer precipitation in central Asia shows simulated decreases, particularly in western parts. In broad terms, model simulation of climate change in Asia is consistent, although areas of uncertainty in future precipitation projections tend to be those associated with dramatic topography. The largest precipitation biases from the AR4 are over the Tibetan Plateau where some model simulations of present annual precipitation are 2.5 times more than observations. Christensen et al. (2007) suggested that model assessment is limited by a poor observational network throughout most of Asia and pointed out that little work has been conducted at the regional scale.

The monsoon is the key climatic process over the vast majority of Asia and understanding how the monsoonal flow is driven is crucial in understanding the aspects of climate change. Large-scale tropical circulation and monsoonal flows have been shown to be weaker in simulations of increased global temperatures (Knutson and Manabe, 1995). The strength of the monsoonal flow is known to be related to the phase of ENSO (Pant and Rupa Kumar, 1997) although this connection would most likely be different under globally warmer conditions (Kumar et al., 1999; Sarkar et al., 2004).

Downscaling studies in Asia have been limited to RCM (e.g. Kumar et al., 2006) and high-resolution GCM simulations (e.g. Dairaku and Emori, 2006). MOS downscaling of raw GCM-simulated precipitation may be possible in parts of Asia where the skill of the precipitation field is strong. In ECHAM5, year-round skill is evident in much of east Asia, and MOS methods may compare well with those from RCMs

in this area.

The main expected changes in African precipitation over the course of the next century appear to fit a general decrease in the subtropics with the tropics becoming (slightly) wetter. Thus, the rainfall gradient across the continent is likely to become greater. This is a realistic response to a warmer climate in which increased water vapour transport in the atmosphere is a result of increased evaporation (Christensen et al., 2007). Amongst the most confident projections are a decrease in annual precipitation in Mediterranean Africa and the northern Sahara, a decrease in winter precipitation in southern Africa and a likely increase annual precipitation in East Africa (e.g. Hulme et al., 2001; Ruosteenoja et al., 2003). There are a number of areas where precipitation changes are uncertain including the Sahel, parts of coastal West African and the southern Sahara (Christensen et al., 2007). Particularly in the Sahel, model projections show little agreement (e.g. Hewitson and Crane, 2006) and Christensen et al. (2007) suggested that further research is requirement in the understanding of the different simulations of precipitation responses.

The potential to downscale model output for Africa is unclear and with some exceptions (e.g. Moron et al., 2008) has rarely been a focus for downscaling research. Most recently, greater focus has been given to Africa in the CORDEX framework (Giorgi et al., 2009) but there are still relatively few RCM simulations available so far. Southern Africa is one few areas with a reasonably reliable observational network, and ECHAM5 precipitation does show decent skill at all times of year. Previously, statistical methods have been used to downscale GCM output, with model parameterisation schemes cited as the major factor in disagreement amongst GCMs (Hewitson and Crane, 2006). Whilst the large-scale decrease in southern African rainfall is widely anticipated, decreases at other times of year appear limited to the western coastal areas (Christensen et al., 2007).

New Zealand has been outlined as an area of key importance for downscaling research given the substantial physiographic controls on its precipitation patterns. Some studies have begun to address downscaling in New Zealand (e.g. Kidson and Thompson, 1998; Sansom and Renwick, 2007) but given that IPCC projections anticipate increased precipitation occurrence along the west coast of the South Island, future research may focus on possible changes in the rain shadow of the Southern Alps. It may be interesting to further assess the skill of, particularly, non-local MOS downscaling corrections in regions of complex topography, such as New Zealand and also the European Alps. As discussed in section 6.4.1, MOS corrections show potential in resolving future precipitation changes in the rain shadow of the Scandinavian mountains, and it is important to understand whether this potential extends to other regions with similar physical characteristics.

6.5 Summary and conclusions

Local and non-local MOS downscaling models have been applied to ECHAM5 climate change simulations in order to produce more spatially resolved estimates of future precipitation. The downscaling models were fitted on the precipitation field from the nudged GCM simulation and simultaneous GPCC observations. In application, each model is a correction of the future simulated precipitation field. All downscaled corrections have been made with reference to the skill of that particular downscaling model, determined by comparing a cross-validated reconstruction of historical local-scale precipitation with observations (see Chapter 4).

Currently, confidence in climate change projections is largely based on multi-GCM consensus and this work represents the first attempt to assign local-scale estimates with a quantification of skill for a specific model. It is shown that in many regions a high level of agreement amongst the GCMs used in the IPCC AR4 is matched with good skill in ECHAM5. However, it is also possible to identify regions where multi-GCM consensus is poor but ECHAM5 skill is strong. One may conclude that GCM consensus plots do not distinguish between the different error components of GCM-simulated precipitation, as outlined in Chapter 3 (section 3.2.1). The inclusion of a nudged simulation (potentially for every GCM) in a climate change investigation would offer additional and invaluable information. GCM consensus and calculation of multi-model means is important, but this procedure currently considers all models to have identical skill. A far more robust estimation of future precipitation changes based on GCM output (or downscaling corrections) would be a quantification of skill in a number of GCMs (which is conceptually possible with a set of nudged simulations) and then alternative consensus and multi-model means can be derived in which each GCM is weighted relative to its skill at a given month/season and region.

Downscaled estimates of 21st century regional precipitation made using the local and non-local corrections capture the main features in raw ECHAM5 output but with far greater descriptive resolution. Europe, North America and Australia were identified in Chapter 4 as regions where skill in ECHAM5, and consequently downscaling corrections, is high. The SVD-RM and CCA-PCR (non-local MOS) models show good skill in detailing the precipitation features of mountainous regions, particularly in the European Alps, although subtle differences exist between the two corrections. The local scaling correction has a tendency to over-estimate regions of heavy precipitation, such as along windward coastlines, but is poor in topographically varied areas by comparison.

Chapter 7

MOS downscaling for daily precipitation distributions

7.1 Introduction

The majority of climate change research has focused on understanding long-term changes and variability in mean global climate. In the context of worldwide precipitation, this has primarily involved an analysis of the global distribution of monthly and seasonal mean precipitation (e.g. Xie and Arkin, 1997; Huffman et al., 1997; Adler et al., 2003; Groisman et al., 2005). However, it is also necessary to consider characteristics of precipitation on a daily time scale such as precipitation intensity, frequency of occurrence (including the number of 'wet' days), and the contribution to monthly and seasonal totals of events of differing magnitude. Analysis of these characteristics is arguably increasingly important in a changing and, specifically, warming climate in which the rate of increase in atmospheric moisture content is likely to exceed that of total precipitation, leading to an increase in precipitation intensity but probable decrease in frequency of occurrence (Trenberth et al., 2003; Sun et al., 2006).

Additionally, increased focus in recent years has been given to the impact of a changing climate on extreme events (e.g. Hegerl et al., 2004; Allan and Soden, 2008). In the case of precipitation, it is possible to consider extreme (or at least heavy) precipitation on monthly or seasonal time scales, and this was drawn on briefly in Chapters 3 and 5. However, it is extreme short-term precipitation events that are most relevant for impact applications, and the understanding and quantification of such extremes involves analysing precipitation on a finer temporal resolution than simply seasonal or monthly means. For small-scale studies, particularly those with a hydrological focus, the evolution of specific extreme precipitation events may require an analysis on sub-daily or hourly data. On a regional scale however,

daily precipitation, or more specifically the distribution of daily precipitation, is usually a sufficient basis for study.

Whilst GCM simulation of large-scale features of precipitation is reasonable, there is usually some deficiency in the simulation of daily precipitation characteristics, and in particular the simulation of precipitation extremes. If such deficiencies are consistent, it is possible to develop some form of MOS correction by making a direct comparison between simulated and observed precipitation distributions for some historical period. Previously, MOS downscaling of daily precipitation distributions has been successfully applied to RCM simulations, with specific focus on the downscaling of extremes. There has been less progress in the development of a MOS downscaling correction for GCM-simulated precipitation, which may prove useful in regions where RCM simulations have not been undertaken, or in regions of extremely variable topography where even the highest resolution RCMs are unable to dynamically resolve precipitation processes on sufficiently small-scales (Schmidli et al., 2006; Maraun et al., 2010). However, as discussed in Chapter 3 and in the case of both RCM and GCM simulations, it is problematic to make comparisons between observations and a freely-evolving model simulation which does not mirror real world day-to-day variability.

It is the purpose of this chapter to extend the development of a MOS downscaling approach for European precipitation to daily time scales. As in previous chapters, calibration of downscaling models is performed using the simulated precipitation from the nudged ECHAM5 simulation. As mentioned, it is possible to develop such corrections on a standard (non-nudged) GCM simulation but it is not possible to assess GCM skill in simulating temporal (in this case, day-to-day) variability. Whilst undertaking a MOS approach with a nudged simulation has been shown to be successful in downscaling monthly and seasonal precipitation, the precise benefits of including the additional nudging step in the downscaling process for precipitation distributions is unknown. The work in this chapter can therefore be considered an exploratory study of the potential of this approach.

The analytical approach in the remainder of this chapter is scaled-down but conceptually similar to an amalgamation of that previously undertaken on monthly precipitation in Chapters 3, 4 and 6. Section 7.2 describes the downscaling methods implemented and the observational data used to derive and cross-validate the downscaling models. Section 7.3 focuses on a comparison between the precipitation field of nudged and standard (non-nudged) ECHAM5 simulations. Section 7.4 details the construction and cross-validation of downscaling models based on a correction of the cumulative distribution function (cdf) and a discussion of the potential application of non-local methods (SVD and CCA) described in Chapter 4. In section 7.5, downscaling models developed in the previous section are applied to ECHAM5 21st century simulations according to the SRES A1B scenario. Estimates are made for local-scale changes in seasonal characteristics of daily precipitation with reference to the overall skill of ECHAM5 and skill of

the downscaling correction itself. Conclusions and suggestions for further analysis are made in section 7.6.

7.2 Methods for downscaling daily precipitation

Downscaling models based on quantile mapping are developed for correcting precipitation distributions in the ECHAM5 climate change simulations. Each downscaling model is defined by the relationships between simultaneous simulated and observed daily precipitation totals. As with all work preceding this, modelled precipitation is that simulated by ECHAM5 nudged to ERA-40, as described in Chapter 3. This work represents the first attempt in developing a quantile mapping correction based on a nudged GCM simulation. Focus is also given to the potential application of a non-local MOS downscaling correction for daily precipitation.

7.2.1 Quantile mapping

Quantile mapping, also known as quantile matching, is a commonly used MOS method for downscaling or otherwise correcting model-simulated distribution of daily precipitation (e.g. Hay and Clark, 2003; Dettinger et al., 2004; Deque et al., 2007; Piani et al., 2010). In simple terms, the method is an adjustment of the model-simulated cumulative distribution function (cdf) to provide a better representation of observations. The adjustment itself is based on the derivation of an empirical transformation of precipitation at a number of quantiles in each distribution. Thus, different intensities of precipitation are accounted for and adjusted.

Mapping may be based on quantiles of the empirical cdf, for which a large number of quantile transformations is needed to accurately represent the shape and scale of the distribution. As with all hindcast-derived statistical downscaling models, application to a simulated future climate involves an assumption of stationarity (Trenberth et al., 2003). Piani et al. (2010) argue that this assumption can be better satisfied if the statistical correction is based on as few parameters as possible and thus considered more robust. In the case of quantile mapping, this has been achieved by deriving a transfer between quantiles of gamma distributions fitted to precipitation. A gamma distribution can be defined by two parameters, k and θ, which represent the shape and scale of the distribution respectively. In describing the empirical distributions as a two-parameter gamma distributions, one is able to remove an element of noise which may be detrimental to the statistical correction.

Analysis in this chapter follows the same gamma fitting method, and Figure 7.1 illustrates the fitting of gamma distributions to example empirical distributions of observed and simulated precipitation, and the visual derivation of a transformation for the 90th percentile (p90). The assumption that distributions of

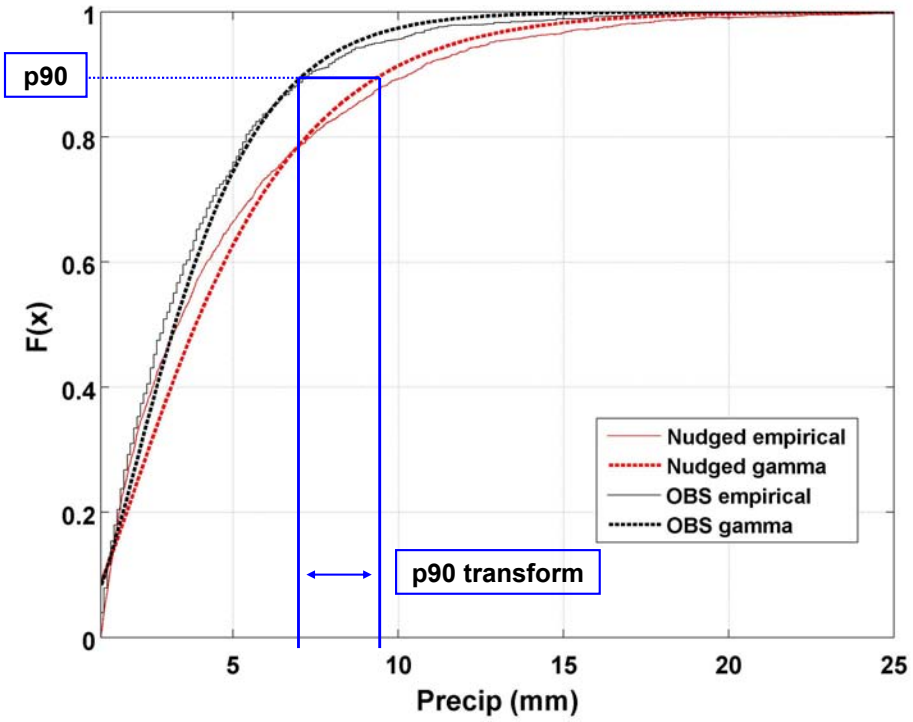

Figure 7.1: Conceptual basis behind fitting a gamma cumulative distribution function to observed and simulated precipitation.

observed daily precipitation can be realistically approximated by a gamma distribution is well-established in the literature (e.g. Katz and Parlange, 1998; Wilks, 2006). However, GCMs often tend to simulate unrealistically small amounts of precipitation in synoptic situations which would otherwise be associated with dry days in the real world (Dettinger et al., 2004). Given this excess 'drizzle' in the GCM, it is useful to identify a threshold to differentiate between dry and wet days. Thus, gamma distributions are fitted on precipitation on wet days only, which are defined in this analysis as days with total precipitation greater than or equal to 1mm. In doing so, it is also necessary to compare the proportion of wet days in each distribution (Piani et al., 2010).

Similarly to the local scaling method described in Chapters 3 and 4, it is possible to construct a gamma-fitted quantile mapping correction using an equilibrium (non-nudged) simulation. The benefits of formulating the correction using precipitation from a nudged simulation is currently unknown.

7.2.2 Potential application of MCA and CCA

In addition to quantile mapping, the potential for downscaling using non-local MOS methods is also discussed. Such methods include regression techniques based on one-dimensional MCA or CCA (which

are referred to in this and all other chapters as SVD-RM and CCA-PCR respectively), which are described in detail in Chapter 4 in terms of their application to downscaling monthly mean precipitation. Here, SVD-RM and CCA-PCR models are developed for downscaling daily precipitation, and are conceptually similar in construction to those developed for downscaling monthly mean precipitation in Chapter 4. Predictor information, in the form of simulated daily precipitation, is taken from an identically-sized spatial domain (20° longitude× 10° latitude). It is currently not known to what extent the success of these non-local methods, particularly CCA, as a downscaling solution for monthly precipitation can be extended to daily time scales.

7.2.3 Cross-validation of methods

Downscaled corrections are made to season-long distribution of precipitation during winter and summer. All downscaling in earlier chapters is performed on monthly mean precipitation at the sufficient grid cell resolution of $0.5° \times 0.5°$ offered by the global GPCC dataset, which permits a worldwide application of each downscaling model irrespective of eventual skill. However, downscaling at the daily time scale obviously requires consideration of precipitation variability on fine temporal and spatial resolutions. Additionally, with a knowledge that ECHAM5 has a generally good capability to realistically simulate precipitation in most of Europe, it is logical to restrict the analysis of downscaling potential to European daily precipitation. Comparisons with observations therefore uses the E-OBS dataset from the ENSEMBLES EU-FP6 project (Haylock et al., 2008). European land-only daily precipitation data is available at a $0.25° \times 0.25°$ grid cell resolution for the period 1950-2006.

All downscaling methods are cross-validated using the same fundamental 'leave-seven-out' approach described in Chapters 3 and 4. This approach requires the daily precipitation distribution for each year to be estimated independently. For example, in attempting to correct daily winter precipitation in year t, the correction model is fitted on data from all other years within the period for which observed data is available (1958-2001 in this instance). In reality, the data used is exclusive of the seven year period centred on year t so as to account for autocorrelation of neighbouring years.

7.3 Using a nudged simulation

It is necessary to assess how nudging the prognostic variables in ECHAM5 impacts upon the simulated daily precipitation. As shown in Chapter 3, the nudging procedure generally does not substantially alter the spatial distribution of monthly and seasonal means, but it is not known how intra-month or intra-season precipitation is altered. An important precipitation characteristic to consider is the percentage of 'wet' days; that is, the proportion of days with total precipitation greater than 1mm. Differences between

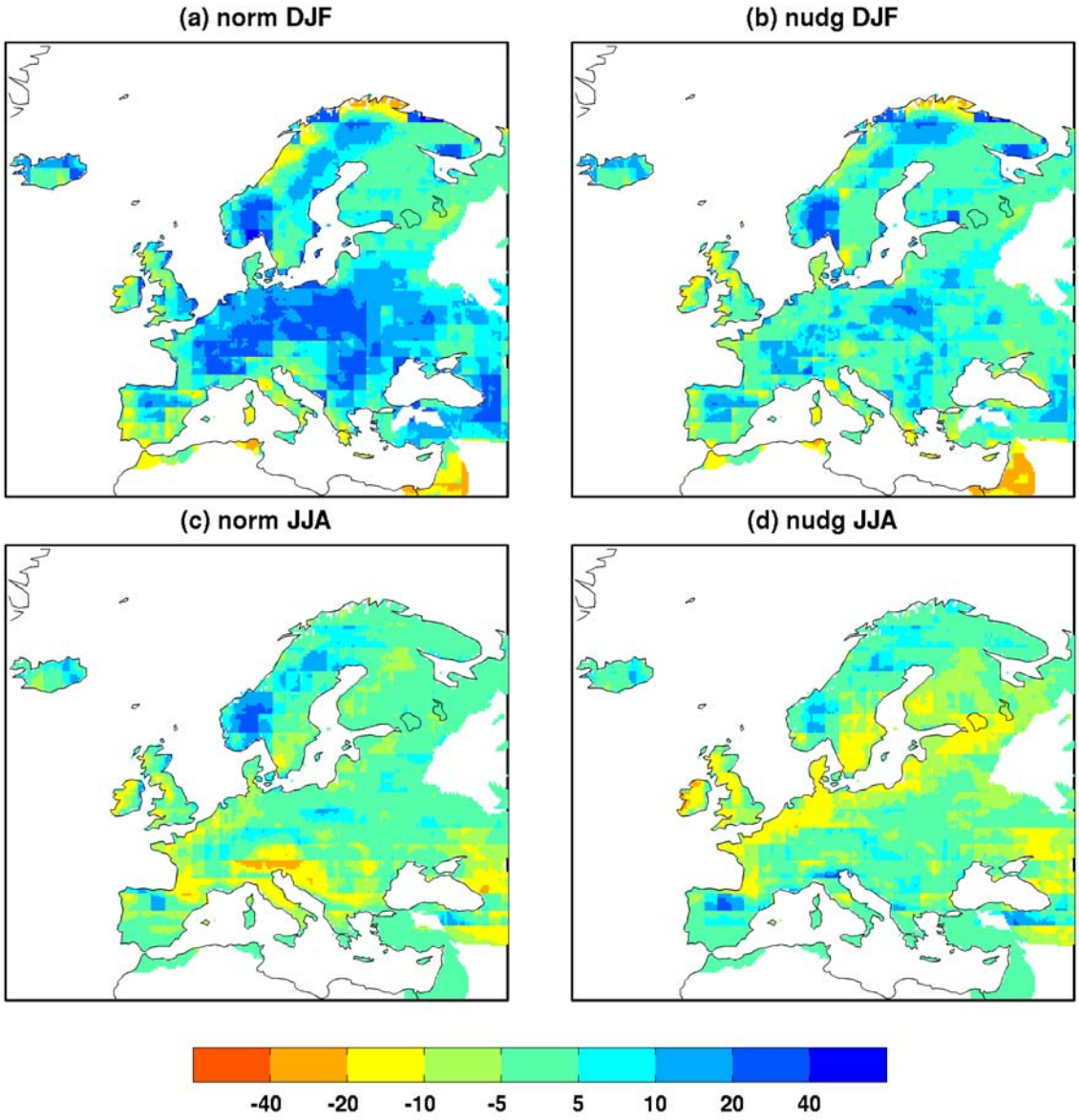

Figure 7.2: Percentage difference in the number of wet days in standard ('norm') and nudged ('nudg') ECHAM5 simulations compared to E-OBS during (a-b) winter and (c-d) summer (1958-2001).

simulated and observed percentages of wet days during winter (DJF) and summer (JJA) are detailed in Figure 7.2. During winter, it is clear that the proportion of wet days is over-estimated in the standard simulation throughout much of continental Europe (Figure 7.2a). ECHAM5 typically simulates between 10-20% more wet days than is observed in the E-OBS dataset. This over-estimation is greatly reduced in the nudged simulation over much of central Europe, particularly in coastal regions (Figure 7.2b). It would therefore appear that these differences are due to errors in circulation (the type 1 and type 2 errors described in Chapter 3) rather than parameterisation errors (type 3). There is little change in southern Scandinavia and parts of central Iberia and central France where a similar degree of over-estimation

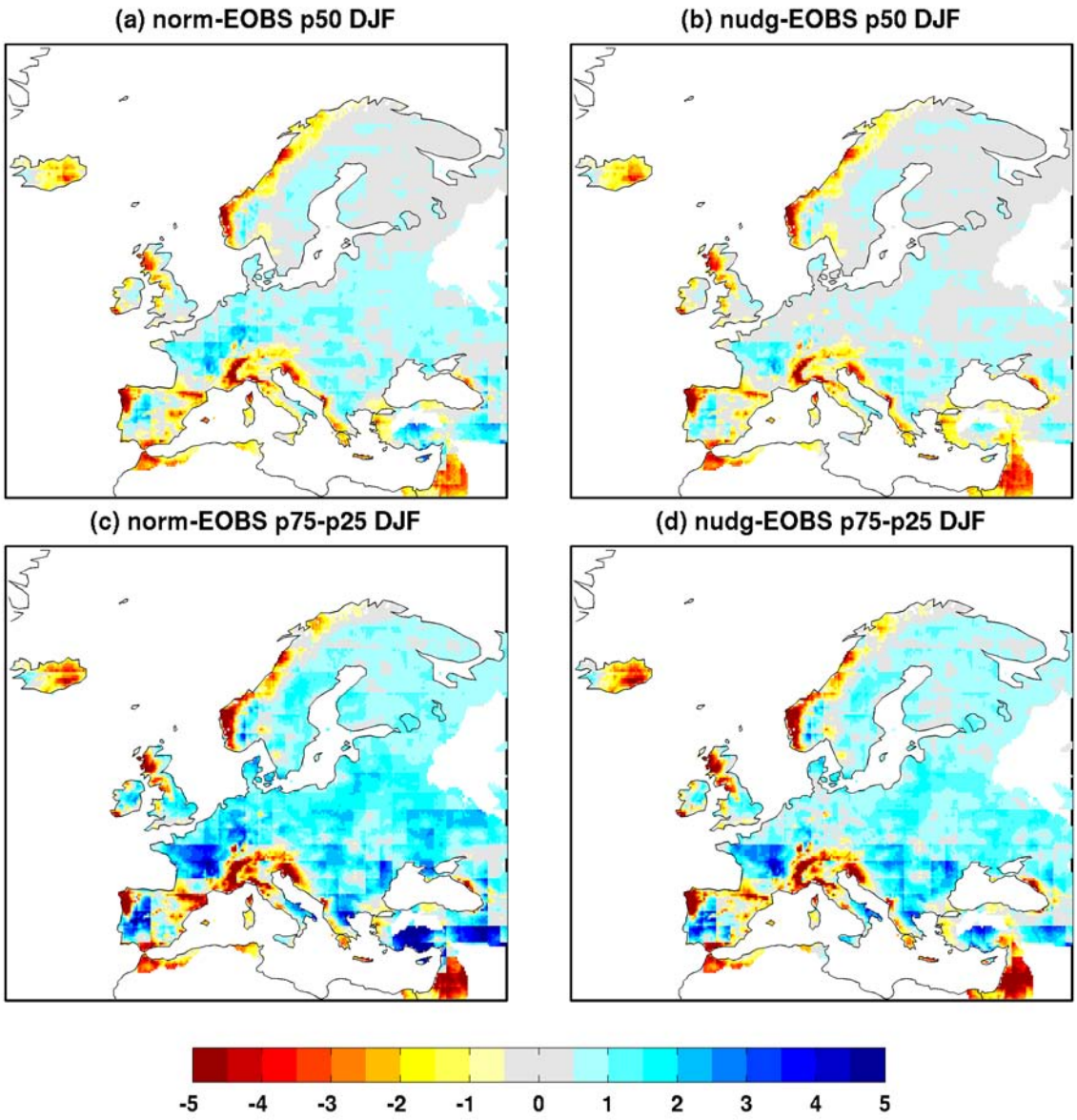

Figure 7.3: Absolute difference (mm) in winter (a-b) median and (b) inter-quartile range statistics between the equilibrium ('norm') and nudged ('nudg') ECHAM5 simulations compared to E-OBS.

remains in the nudged simulation. In some regions, such as western Iberia and western Scandinavia, the extent of under-simulated proportion of wet days is reduced in the nudged simulation, although this is not the case in western parts of British Isles where even fewer wet days are simulated. During summer, there are further inconsistencies in evidence (Figure 7.2c-d). Under-estimation of wet days across much of the Alpine regions is righted to a large extent in the nudged simulation. However, in many windward areas of western and northern Europe, particularly western British Isles, northern France and the Low Countries, extending northwards to Denmark and southern Sweden, a moderate under-estimation in the standard simulation is actually accentuated in the nudged simulation.

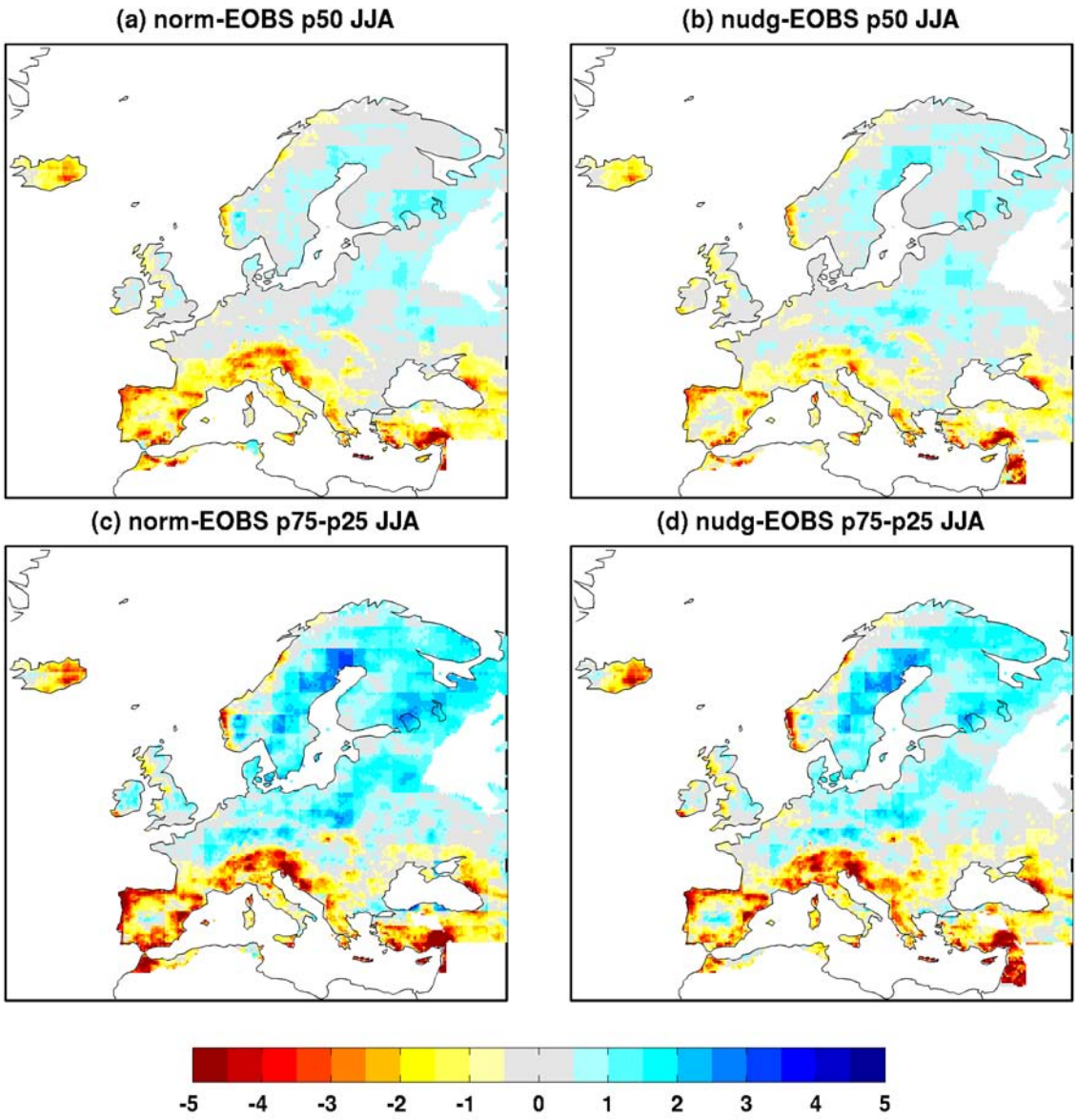

Figure 7.4: Absolute difference (mm) in summer (a-b) median and (b) inter-quartile range statistics between the equilibrium ('norm') and nudged ('nudg') ECHAM5 simulations compared to E-OBS.

Figures 7.3 and 7.4 show differences between simulated and observed values of the median and inter-quartile range during winter and summer respectively. During winter, the spatial distribution of median precipitation quantities is similar in both simulations. In much of the continental interior, both the median and inter-quartile range values are marginally over-estimated in the standard simulation and better represented in the nudged simulation. There is little change in the wettest regions of Europe (western British Isles, north-western Iberia, western Scandinavia and mountainous regions such as the Alps) between each simulation. During summer, the median is very well-simulated across the majority of northern Europe and there is little change between the standard and nudged simulations. Additionally,

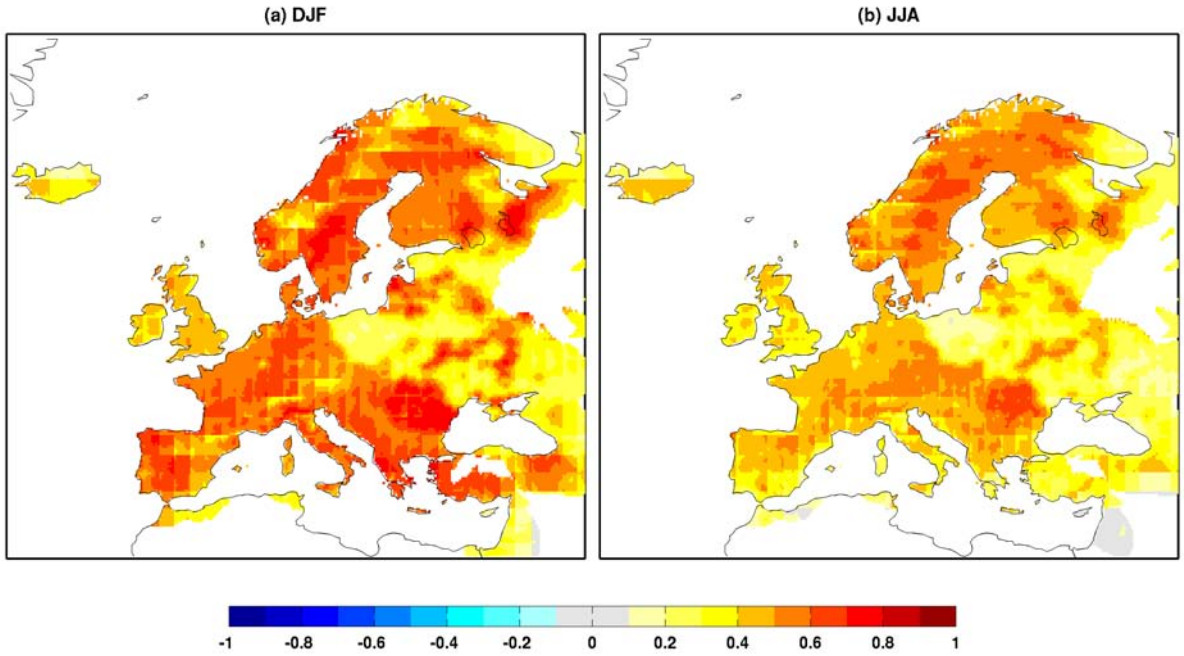

Figure 7.5: Correlation between nudged ECHAM5 and E-OBS daily precipitation.

the extent of the under-estimation of the median in much of southern Europe is reduced in the nudged simulation, and this is particularly apparent across the Alps. Simulation of the inter-quartile range is subject to an over- (under-) estimation in the north (south) and there is little difference in this spatial extent or its magnitude between the standard and nudged simulations. It is known that convective events tend to make a greater contribution to total precipitation during summer, and particularly in southern Europe so it is unsurprising that biases in the median and inter-quartile range are not reduced in the nudged simulation. Generally, biases in median and inter-quartile range precipitation statistics, particularly in the wettest regions of Europe, are similar in both the standard and nudged simulations, suggesting that deficiencies in the precipitation parameterisation, rather than in the simulation of large-scale circulation, are responsible for errors of this nature. Given that the temporal evolution of the ECHAM5 simulation of day-to-day weather is consistent with the observed record, it is possible to compare the simulated and observed precipitation time series (Figure 7.5). This has been performed previously for seasonal and monthly precipitation means in Chapters 3 and 4, and it can be concluded that strong, positive correlation is indicative of good skill in the parameterisation of precipitation.

7.4 Downscaling corrections

7.4.1 Quantile mapping correction

The cross-validated corrections developed using quantile mapping are compared with observations in the historical record. As the fitting of the corrections is based on a nudged simulation, which is in temporal phase with the real world, a direct comparison is possible between the corrected and observed distributions. This permits an analysis of correlation between the corrected (cross-validated) and observed time series in addition to the distribution statistics of each. The latter are analysed first in Figures 7.6 and 7.7, which illustrate median, interquartile range and skewness statistics of the corrected distribution and how closely these match those statistics of the observed distribution. The spatial pattern of precipitation in the correction is physically sound during each season. In winter (DJF), a general west-east gradient in median precipitation is apparent and agreement with observations is generally strong (Figure 7.6). Although heaviest precipitation is estimated along the western coastline fringes and west-facing mountainous regions, precipitation totals in these regions remain moderately under-represented. This rough

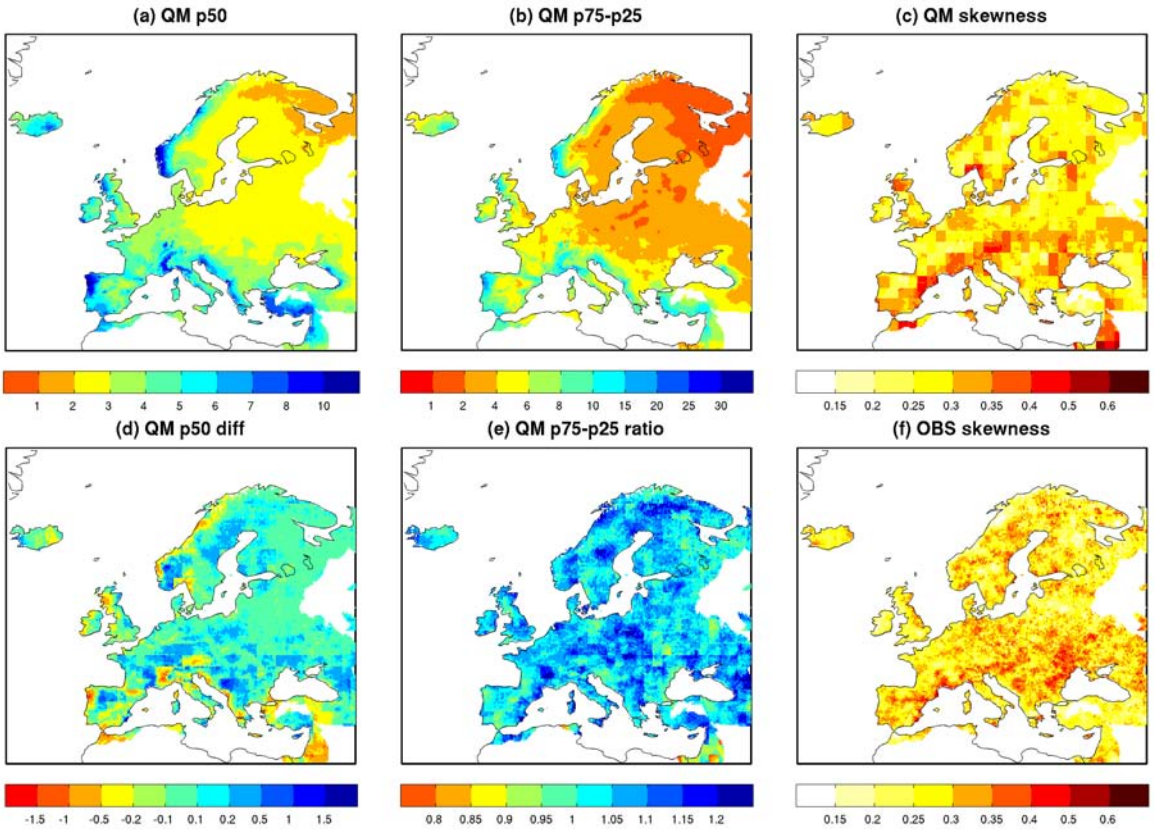

Figure 7.6: Comparative winter precipitation (DJF) statistics in quantile mapping correction and E-OBS observations; (a) median (mm), (b) interquartile range (mm) and (c) measure of skewness in the correction; (d) difference in median and (e) ratio of the interquartile range between the correction and observations; (f) measure of skewness in the observations.

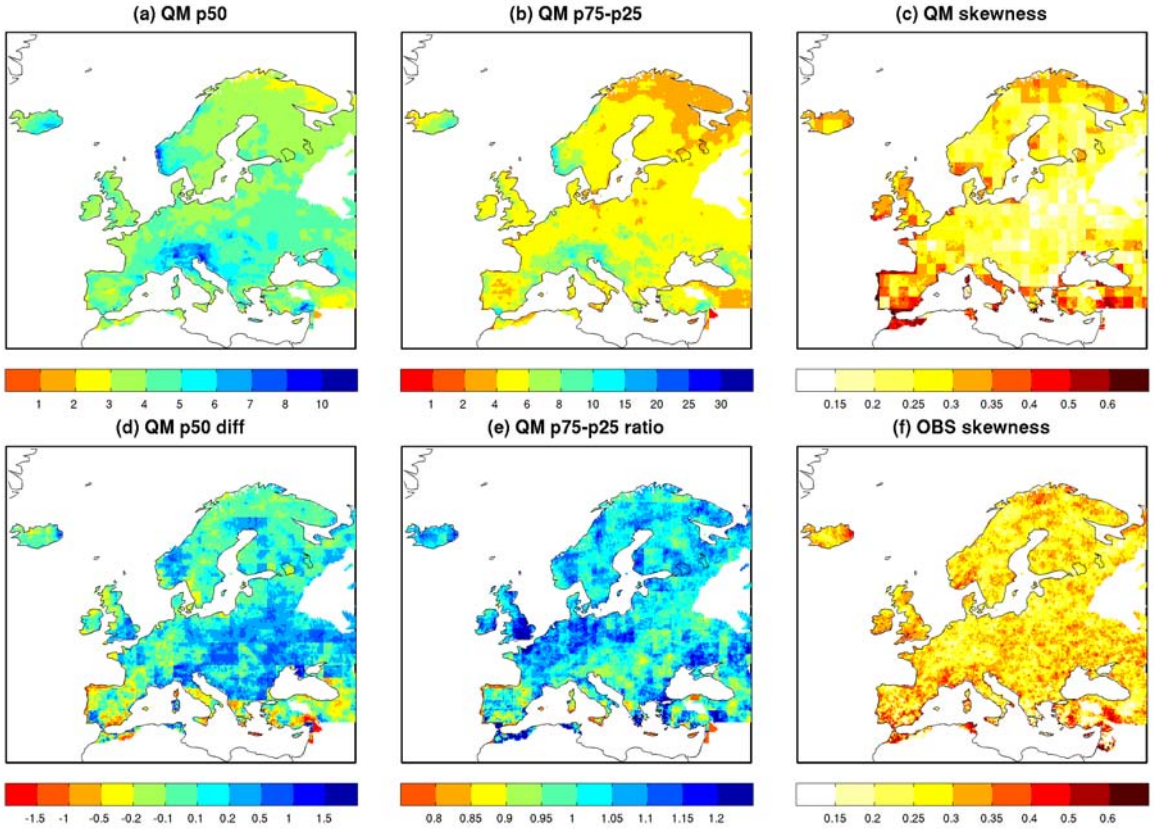

Figure 7.7: As Figure 7.6 but for summer (JJA) precipitation.

pattern is reflected in the interquartile range, although with fewer instances of under-estimation. During summer (JJA), areas of over-estimation of median precipitation are concentrated in central and eastern Europe, with much of coastal Mediterranean regions under-estimated (Figure 7.7). The magnitude of the corrected interquartile range is excessive in parts of north-west Europe, and between 20-40% over-estimated in southern and eastern parts of the British Isles. Positive skewness in both distributions is widespread, although in winter the corrected distribution broadly estimates a greater positive skew in mountainous regions (notably, the Alps and the arc of the Carpathians) than is shown in the observations. A similar pattern is evident along the Mediterranean coast during summer.

As quantile mapping includes only a distribution's wet days (total daily precipitation greater than 1mm), it is of equal importance to analyse the proportion of wet days in the corrected and observed distributions (Figure 7.8). During winter, this proportion is well-reproduced by the correction in the majority of Europe, particularly across continental lowlands. There is some over-estimation within the continent interior, notably in central Iberia and in Scandinavia leeward of the coastal mountains. During summer, estimates for central and eastern Europe are skilful, although much of the coastline of western Europe, with the exception of Scandinavia, is under-estimated. The apparently larger difference in the

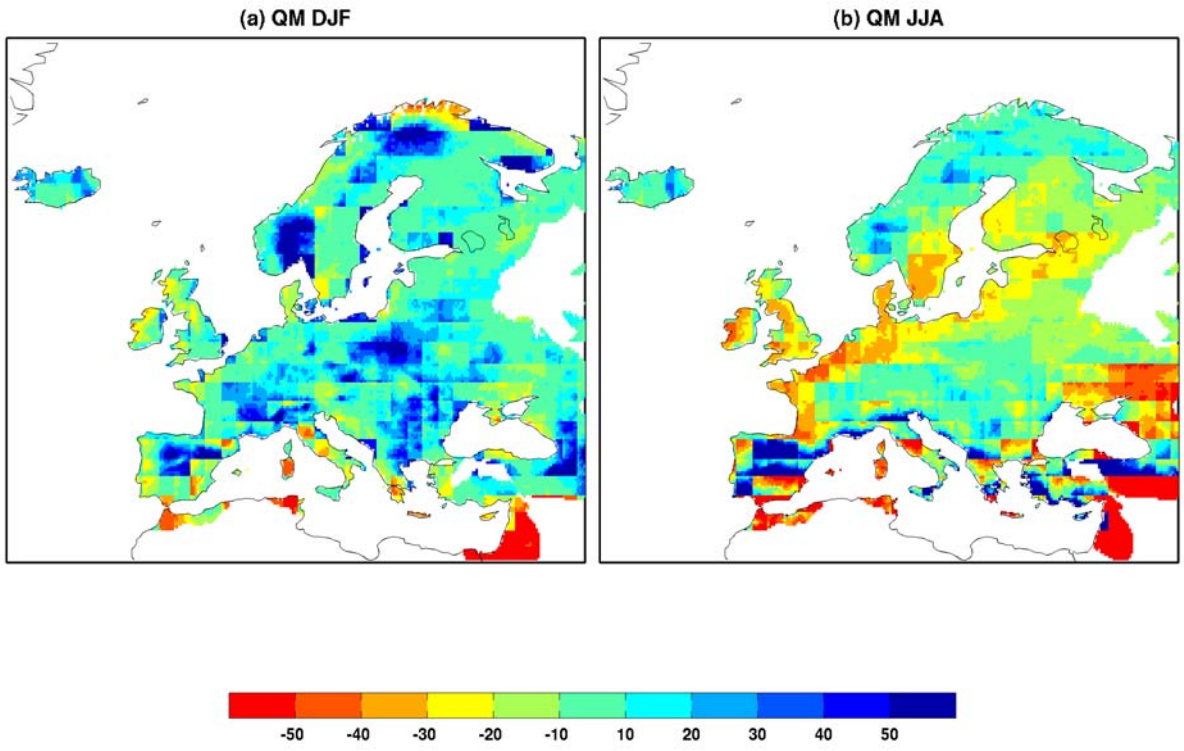

Figure 7.8: Differences in the proportion (%) of wet days between the corrected and observed distributions.

proportion of wet days in southern Europe is difficult to interpret given the low mean precipitation.

Further investigation is given to differences between six corrected and observed quantiles in Figures 7.9 and 7.10. The differences are expressed as a ratio, which is displayed in terms of its natural logarithm to show linear consistency in each sign. The quantiles chosen for analysis include the three quartiles (p25, p50 and p75) and three quantiles at the extreme end of the distribution (p90, p95 and p99). Definitions of 'extreme' precipitation are varied, but a typical measure for heavy precipitation is 90th percentile on wet days (Haylock et al., 2006; Goodess et al., 2010). The 95th and 99th percentiles are included here also, although the extent to which these can be reliably corrected from a fitted gamma distribution is uncertain and, with some exceptions (e.g. Ferro et al., 2005), not usually investigated. Nevertheless, at this stage it is important to gauge an understanding of the limitations of quantile mapping based on a nudged simulation, and key to this is an assessment of the skill of the correction to represent precipitation at the top end of the distribution.

During winter, estimates for the three quartiles (p25, p50 and p75) are closely matched with observations across most of Europe (Figure 7.9a-c). Regions of moderate under- and over-estimation can be largely explained by topographic features. In Scandinavia, the western mountains again have an impor-

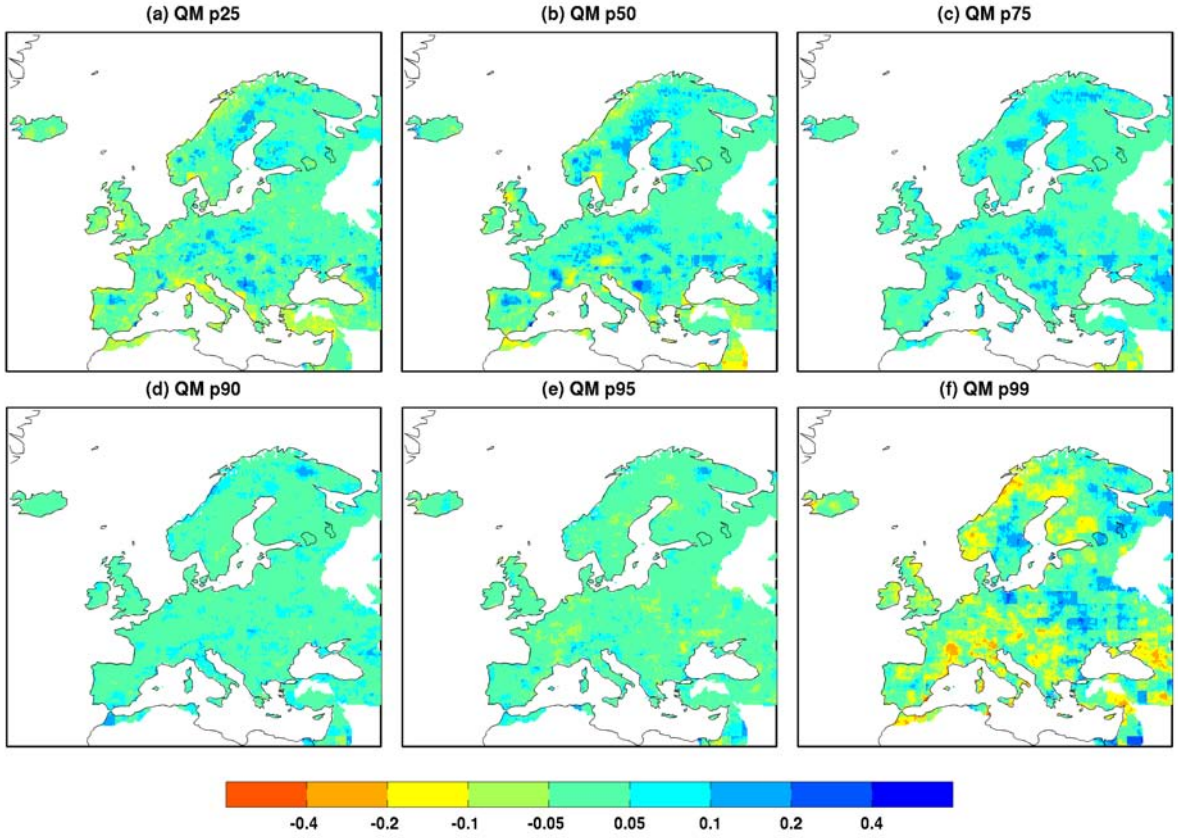

Figure 7.9: Log ratios of six quantiles of winter (DJF) daily total precipitation between the quantile mapping correction and observations (1958-2001).

tant influence on the skill of the correction. Each quantile is consistently over-estimated in the eastern lee of the north-south mountains, roughly consisting of eastern and northern Sweden. This pattern is also evident in the immediate lee of mountains in southern Norway. This over-estimation of precipitation intensity and suggests the influence of the Scandinavian mountains in forming an orographic barrier to western-driven precipitation is under-represented. This conflicts with patterns in other topographically complex regions, such as the western and northern Alps, where it is the windward slopes that experience an over-estimation of each quantile. One must perhaps consider the model's ability to simulate snowfall, particularly in Scandinavia. Similar patterns in the correction of each quartile are evident during summer (Figure 7.10a-c). Additionally, low precipitation totals make these corrections troublesome in southern Europe.

The estimation of quantiles representing heavy precipitation is encouraging. During winter, bias in the estimation of the 90th and 95th percentiles is again very small over the majority of Europe with some isolated exceptions (Figure 7.9d-e). As previously mentioned, quantile mapping based on a fitted gamma distribution is considered ineffective when correcting extreme precipitation, but in this instance there is good model skill in estimating precipitation intensity up to the 95th percentile. This good skill

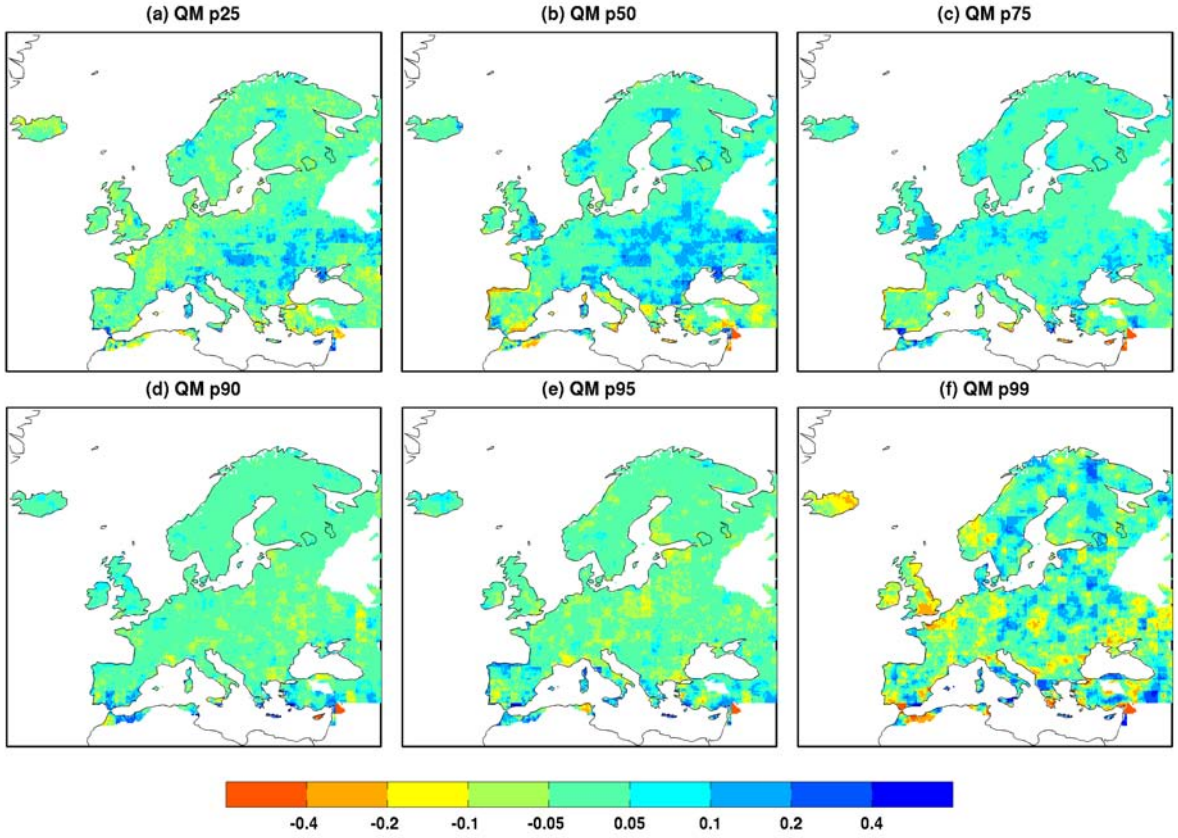

Figure 7.10: As Figure 7.9 but for summer (JJA) daily total precipitation.

does not extend to the 99th percentile, which is consistently under-estimated in the correction across much of central Europe (Figure 7.9f). In spite of this, it is interesting to note that the wettest parts of Europe, occurring on average along the coastlines of north-west Spain, France and the British Isles, are reasonably well-represented. This potential predictability may be owed to the relative regularity in the frequency and magnitude of dominant large-scale frontal precipitation. The 99th percentile is marginally under-estimated along the western coastline of Scandinavia, which again points to model representation of orographic forcing in the Scandinavian mountains. During summer, a similar pattern emerges in the correction the 90th and 95th percentiles, which are well-reproduced over much of Europe (Figure 7.10d-e). An exception is in parts of Mediterranean Europe, where low precipitation totals make heavy precipitation, which consists of infrequent, largely convective events, difficult to estimate. Of interest is the northern coast of Spain, where the 50th and 75th percentiles are under-estimated and the 95th percentile over-estimated. The correction has an apparent tendency to erroneously attribute a greater portion of total precipitation to the most intense precipitation events. There is little to suggest that a skilful correction of the 99th percentile is possible, except in areas where the distribution of daily total precipitation is associated with less variance and a shorter extreme tail.

7.4.2 Skill of downscaled corrections to represent temporal variability

In basing a correction scheme on a simulation in temporal phase with the observed record, it is possible to make a simulation-observation comparison of the time series in addition to the distribution. Correlation maps have been used previously in Chapter 3 and 4 to illustrate where a simulation or downscaling correction is able to reproduce inter-annual variability of monthly mean precipitation. Here, the analysis is extended to both inter-annual and intra-seasonal (essentially, inter-daily) variability.

The non-local regression-based downscaling corrections discussed in section 7.2.2 are impractical for directly estimating actual daily precipitation values due to a large under-estimation of spatial variability as a result of area-averaging. However, there may remain some potential for these methods to provide relevant information about temporal variability. Thus, in this analysis, it is possible to compare all three MOS downscaling methods (quantile mapping, SVD-RM and CCA-PCR). Maps of correlation between daily precipitation in observations and each correction are shown in Figure 7.11. As expected, correlation with observations of the quantile mapping correction is similar to that of raw ECHAM5 output (Figure 7.5) in both distribution and magnitude. During winter, the key features are strong correlations over most of continental Europe, with exception of eastern Europe extending to western Russia. The gradient in correlation strength in eastern Europe does not appear gradual. It is difficult to explain this very sudden difference simply by citing the physical characteristics of the region, whether climatological or topographical, and the sharp differences are likely due to geopolitical influences on the reliability of the observed data. Non-local correlations are stronger and more spatially coherent, with CCA-PCR again showing a marginally better performance than SVD-RM. Both methods provide far greater representation than the quantile mapping correction in mountainous regions, particularly the Alps and the Pyrenees.

Figure 7.12 assesses the gamma distributions simulated, observed and corrected winter (DJF) precipitation at three locations where correlation of the corrected and observed time series are different. Note that these locations were chosen specifically to represent distinct regions of model skill, and not in terms of any differences in physiographic features or long-term climatology. Location (b) is Lodz in central Poland where correlation is distinctly weak. Comparison of the corrected and observed distributions reveals an over-estimation of precipitation falling within the inter-quartile range, an observation also made in Figure 7.6 where each corrected quartile is an increase on the observation.

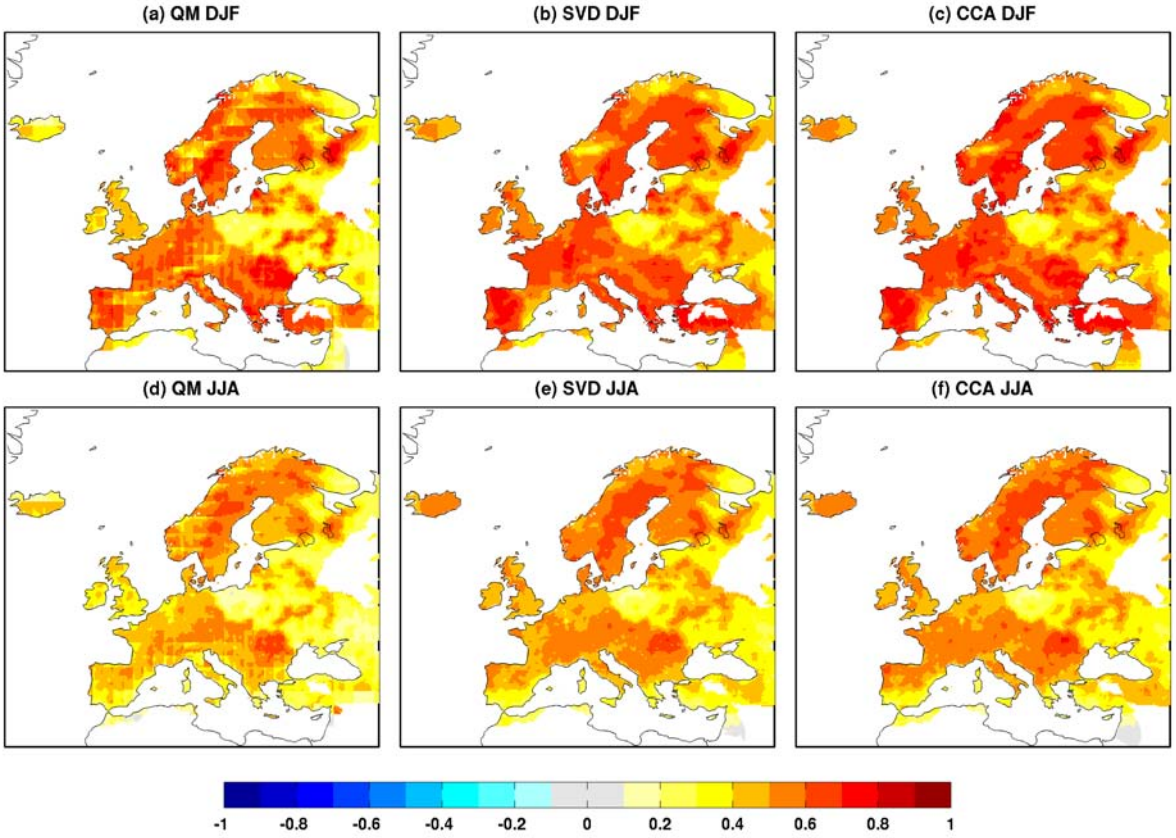

Figure 7.11: Correlation of corrected and observed (E-OBS) daily winter and summer precipitation. Corrections based on quantile mapping (QM), SVD-RM and CCA-PCR.

7.5 Downscaling future changes in daily precipitation distribution

Downscaling corrections derived from quantile mapping were applied to the precipitation field of a future ECHAM5 simulation. Figures 7.13 and 7.14 show raw ECHAM5 output and downscaled corrections of different quantiles for winter (DJF) and summer (JJA) respectively. Simulated and corrected quantiles are calculated from the distribution of all wet days within a three-member ensemble simulation of a control period (1980-1999) and a climate change scenario according to SRES A1B (2080-2099). Analysis focuses on four quantiles: 25th, 50th and 75th percentile indicate the median and the extent of the inter-quartile range, whilst the 90th percentile was chosen as a measure for heavy precipitation (Haylock et al., 2006; Goodess et al., 2010). Many of the key spatial differences between ECHAM5-simulated and downscaled daily precipitation quantiles mirror those in the downscaling of monthly mean precipitation, as shown in Chapter 6 (Section 6.4.1). Therefore, it is unnecessary to restate the key processes behind the obvious patterns that are apparent in downscaled corrections of both monthly and daily precipitation.

During both seasons, the downscaled correction is able represent the key precipitation features cap-

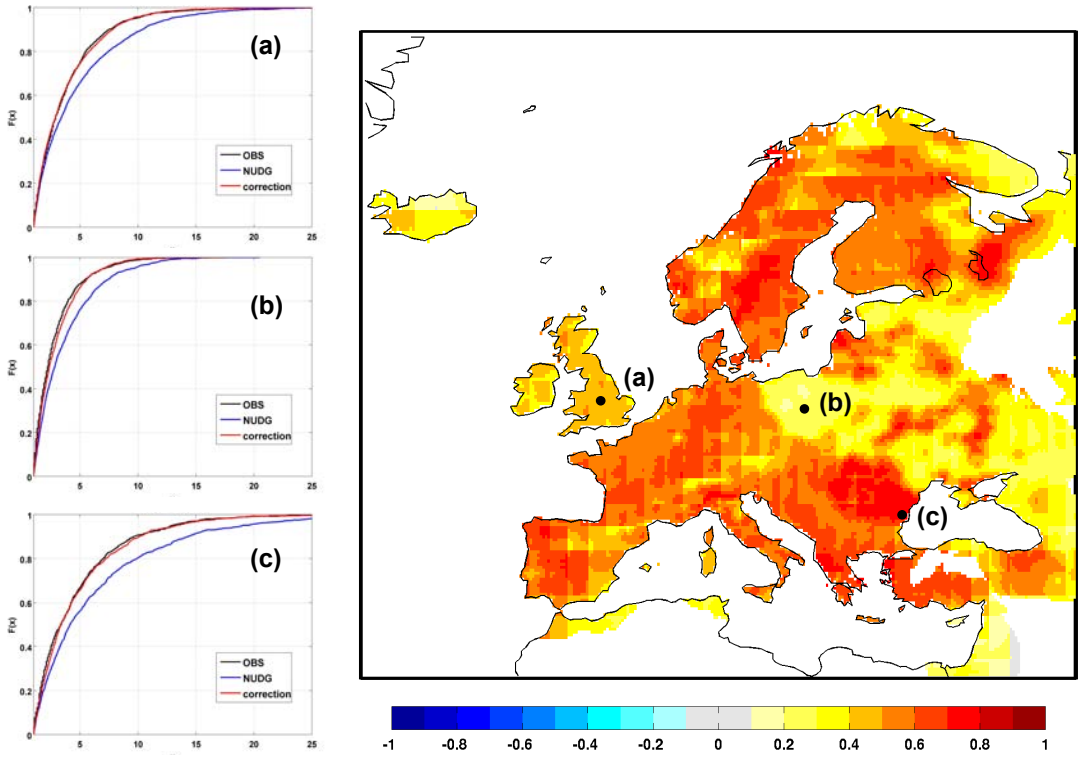

Figure 7.12: Gamma cumulative distribution functions fitted to observed ('OBS'), simulated ('NUDG') and corrected winter (DJF) daily total precipitation (mm) in three example locations with different correlation statistics: (a) Birmingham, United Kingdom; (b) Lodz, Poland; and (c) Constanta, Romania. Note: correlation map taken from Figure fig:plot$_c orrelation_d aily_D JF_J JAa$.

tured in ECHAM5 but at a much greater resolution. With an output resolution of 0.25° × 0.25° (all downscaling on monthly means was constructed to a 0.5° × 0.5° resolution), small-scale variability induced by topography or a sudden change in surface coverage can be represented much more clearly. In most cases it is immediately obvious where the coarser resolution of ECHAM5 is insufficient in resolving precipitation distribution, such as along windward coastlines and across mountain ranges.

During winter, a west-east difference is evident across Europe for the 50th and 75th percentiles (Figure 7.13). To some extent, this pattern is captured in the raw ECHAM5 output but the cross-continent gradient is more pronounced in the correction. Although the prevalence of moisture-laden air masses originating in the Atlantic Ocean would support this connotation, it is uncertain to what extent the relatively poor skill in the downscaling correction in parts of eastern Europe is responsible for the magnitude of the gradient. The fact that the differences are so pronounced and spatially similar for each quantile suggests that model skill is an important factor.

During summer, the corrected quantities of both p50 and p75 are larger over much of Mediterranean Europe, particularly in northern and eastern Spain and the region east of the Adriatic. Given that p25

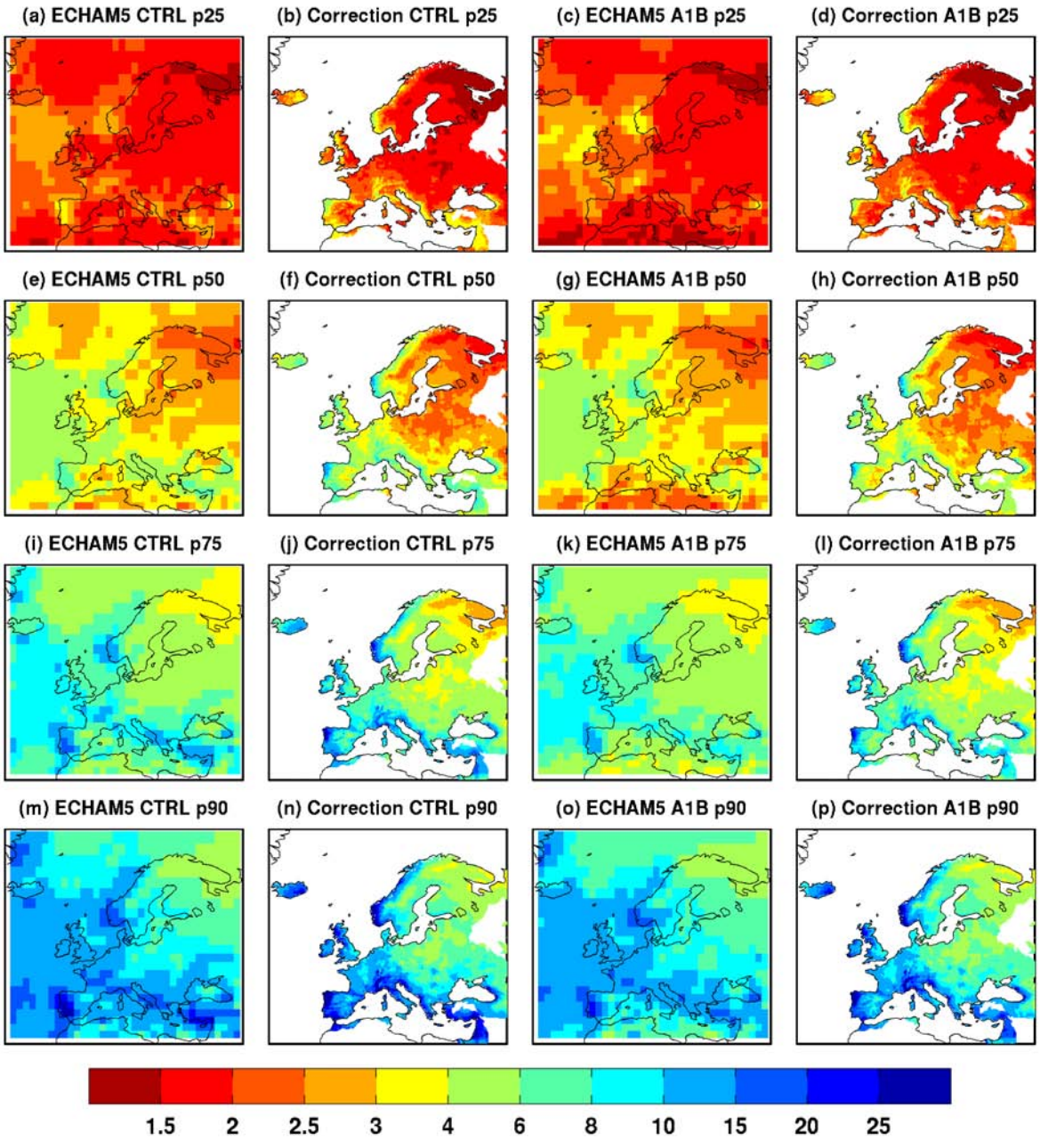

Figure 7.13: ECHAM5 and corrected quantiles of winter (DJF) precipitation for the control (CTRL) period (1980-1999) and SRES A1B scenario (2080-2099). Quantiles expressed in mm.

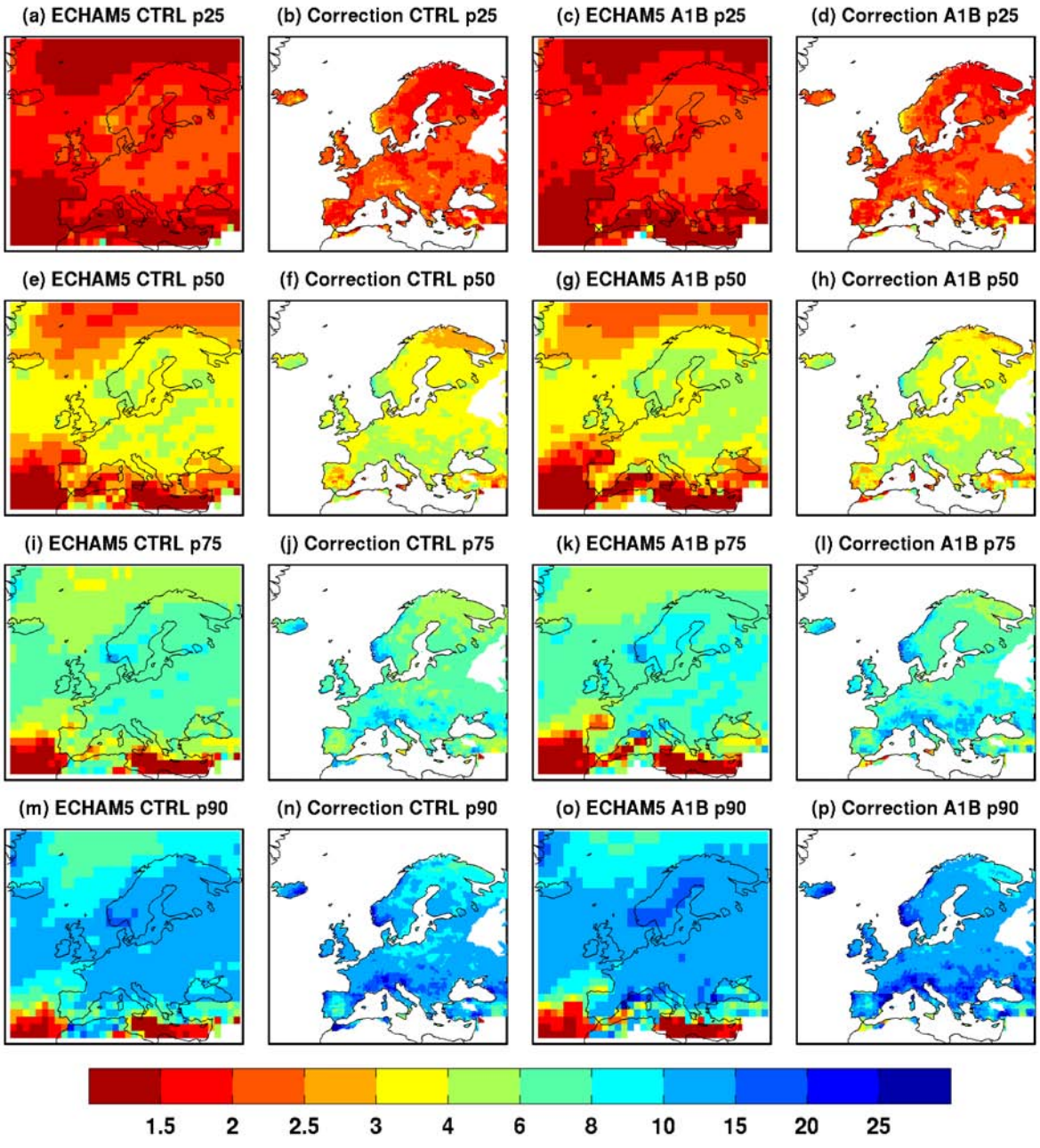

Figure 7.14: ECHAM5 and corrected quantiles of summer (JJA) precipitation for the control (CTRL) period (1980-1999) and SRES A1B scenario (2080-2099). Quantiles expressed in mm.

and p90 are relatively unchanged in the correction, this may be explained by a skew within the simulated inter-quartile range which understates the contribution of heavier precipitation events to the seasonal total. As previously discussed, the tendency of a GCM to simulate 'drizzle' can make it difficult to distinguish between (what in the observed record are referred to as) 'dry' and 'wet' days. The drizzle phenomenon may continue to exist during light precipitation events that are still classified as wet days. In reality, the dominant contribution to total precipitation during summer is from short-term convective events, the magnitude and infrequency of which may not be fully resolved in the GCM. The extent to which this model correction is able to better represent such characteristics of summer precipitation is uncertain. As the skill in the simulation of day-to-day variability is poor in this region (Figure 7.11) caution should be exercised in interpreting these corrections. Additionally, in many regions, low skill can be directly attributed to an extremely small number of precipitation events used to construct the gamma distributions from which corrections are derived.

During both seasons, locations associated with heaviest daily precipitation totals tend to occur along the coastlines of western Europe and in mountainous regions. Whilst the raw ECHAM5 simulation is sufficient in representing heavy coastal precipitation, particularly in north-west Iberia, western British Isles and southern Scandinavia, performance is poorer in resolving heavy continental precipitation induced by orographic effects. This is particularly evident over the Alps in the 75th and 90th percentiles during winter (Figure 7.13i-j,m-n) and to a lesser extent during summer (Figure 7.14i-j,m-n). ECHAM5 better resolves daily precipitation processes at land-sea boundaries than in regions of complex topography, and it is in the latter where downscaling, particularly at this resolution, is able to provide crucial extra information.

During winter, the precipitation gradient in the Scandinavian mountains is clearly evident in corrected precipitation occurring above p50. Although ECHAM5 is able to roughly resolve the intense precipitation in southern Norway, the downscaled correction is required to represent the extension of heavy daily precipitation northwards along the coast. Similarly, the drier eastern lee of the mountain range is more resolutely defined in the correction, particularly at p50, p75 and p90.

In understanding simulated precipitation in a changing climate it is necessary to interpret future simulated precipitation relative to a baseline period. As in Chapter 6, this 'control' (CTRL) period is 1980-1999 as simulated by a three-member ensemble forced by known quantities of greenhouse gas emissions. Figures 7.15 and 7.16 detail the percentage change in four precipitation quantiles between 2080-2099 and 1980-1999 on both ECHAM5 and downscaled grids. Stippling overlays, introduced in Chapter 6, are again used to describe the spatially-varying skill of the quantile mapping correction to reproduce day-to-day variability. All corrections are 'local', in that they are derived from gamma distributions fitted to the empirical cdf of all wet days for a given season at one particular grid cell. As

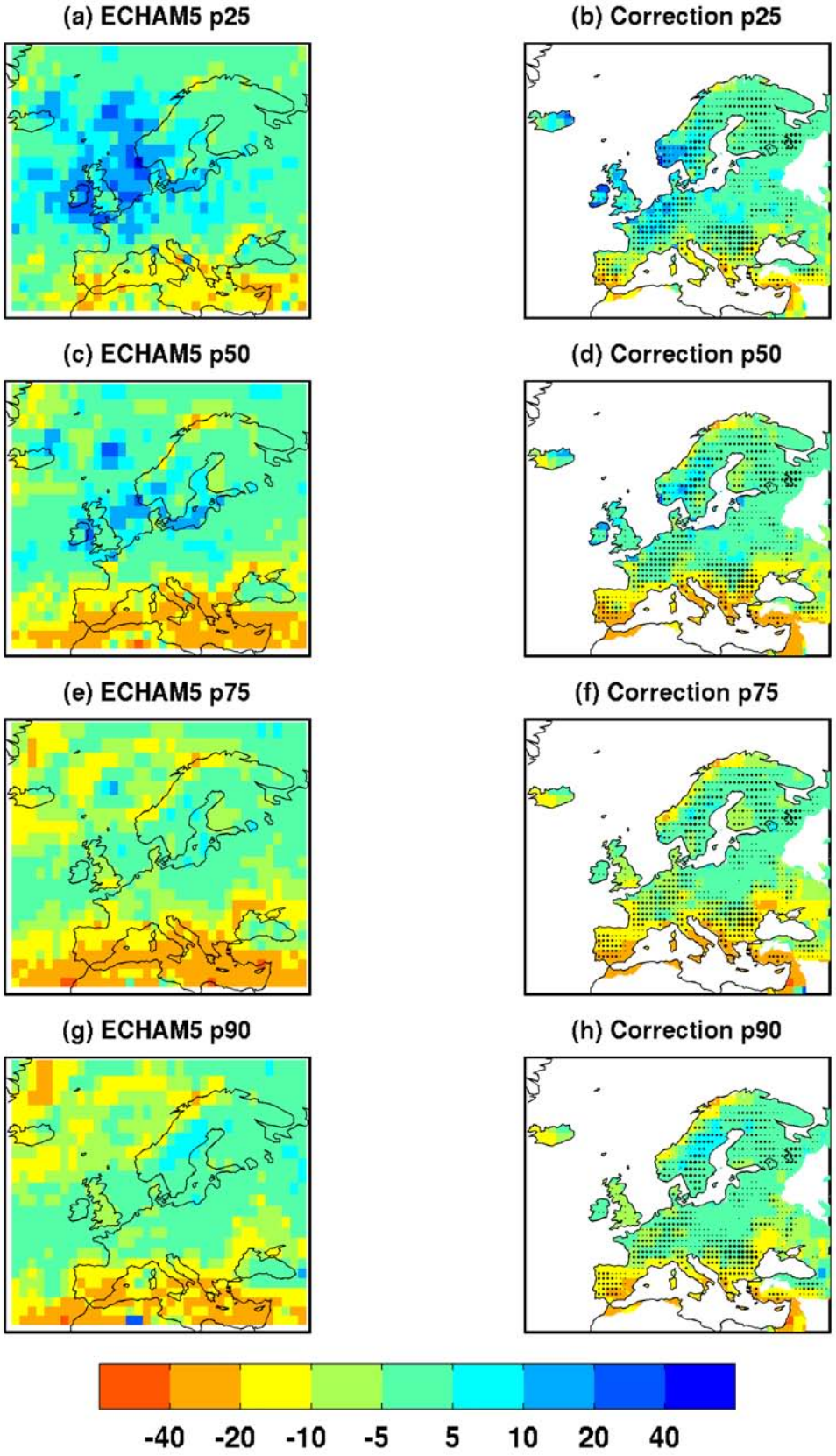

Figure 7.15: Differences in ECHAM5-simulated and corrected winter (DJF) quantiles between the control (CTRL) period (1980-1999) and SRES A1B scenario (2080-2099).

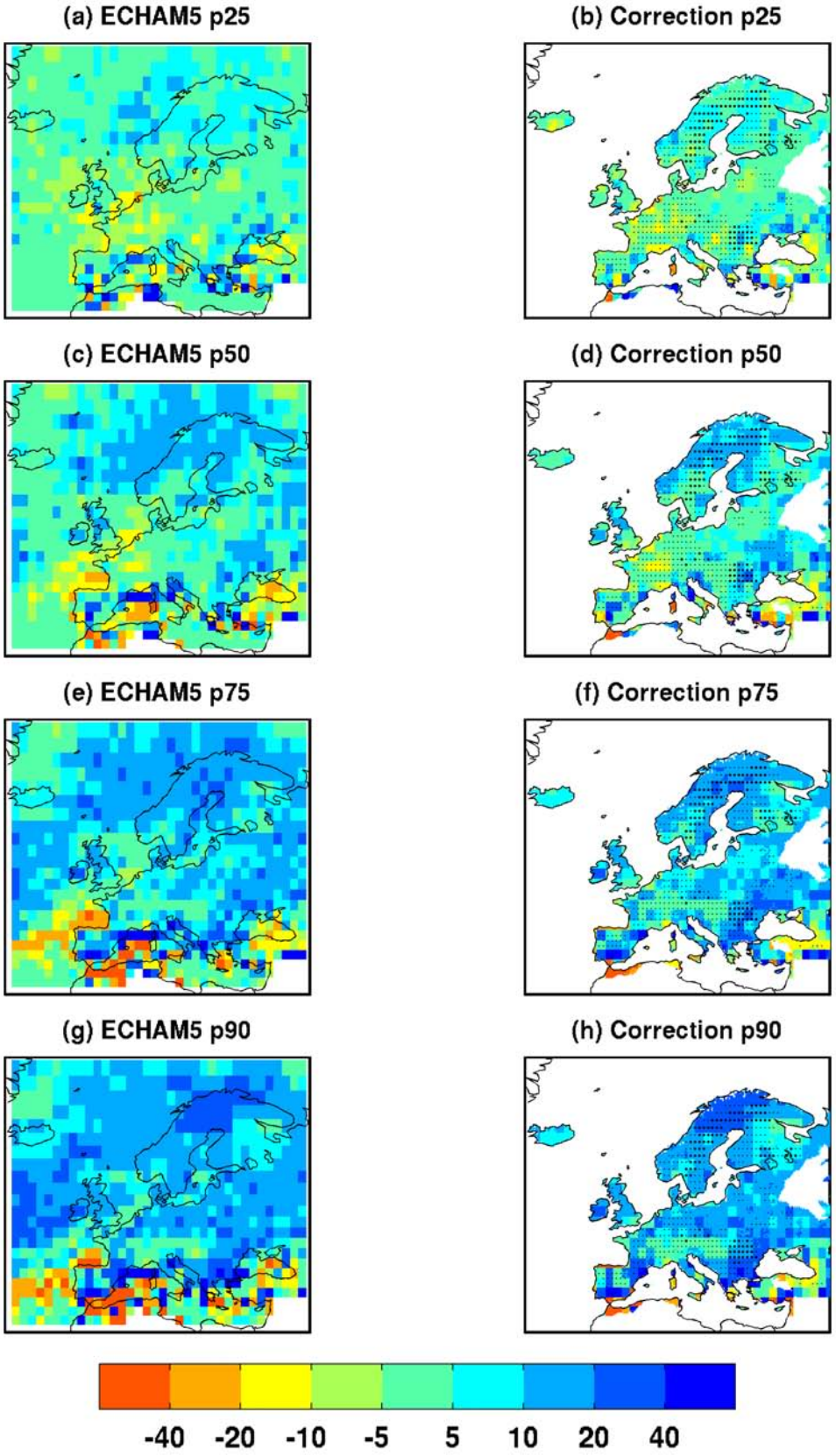

Figure 7.16: Differences in ECHAM5-simulated and corrected summer (JJA) quantiles between the control (CTRL) period (1980-1999) and SRES A1B scenario (2080-2099).

each correction is constructed from information at just one grid cell, the change in corrected precipitation between two time periods is very similar to that which exists in raw ECHAM5 output.

It was noted, particularly during winter, that there were only small differences between GCM simulations and high-resolution RCMs used in the PRUDENCE experiments (Christensen and Christensen, 2007). Furthermore, there was little discernible difference between the PRUDENCE RCMs run at the standard (50×50 km) resolution and those run at a higher resolution (25×25 km) (Christensen and Christensen, 2007). So, the large-scale controls on dynamically-downscaled simulations appear to exert a similar influence on the corrected quantiles shown here.

Before identifying some of the more notable differences, it is first of all necessary to understand the projected changes in daily precipitation distribution as simulated in ECHAM5. During winter, time mean precipitation is shown to increase across northern and central Europe (see Chapter 6, section 6.4.1), and this pattern is generally reflected across the daily distribution (Figure 7.15a,c,e,g). An exception is over the British Isles, where p25 and p50 show little change, and a decrease is estimated in values of both p75 and p90. The lessening contribution of heavy precipitation events to what is expected to be greater mean precipitation totals conflicts with the observed increase in the intensity of winter precipitation in the UK between 1961-1996 (Osborn et al., 2000). However, whilst errors in the parameterisation of precipitation in ECHAM5 are small in terms of reproducing realistic precipitation amounts and temporal variability in this region (see Chapter 3), this apparent skill does not extend to daily time scales as demonstrated by the lack of stippling across the British Isles in Figure 7.15.

Although, the tail of the distribution is not sufficiently modelled in this analysis, the suite of GCMs used in the IPCC AR4 indicate an increase in both the frequency and magnitude of 'extreme' precipitation events (Christensen et al., 2007). During summer (Figure 7.16), differences are less pronounced, although the time mean increase in precipitation in northern Europe is matched by an increase across the daily distribution, which appears stronger at the high quantiles (p75 and p90). Elsewhere, it is difficult to draw reliable conclusions, particularly in the Mediterranean where there is a lack of spatial coherence in the extent (and in some areas even the sign) of the change. The projected decrease in mean precipitation is largely the source of this uncertainty. It has been suggested that whilst an increased water atmospheric moisture content is a simple explanation for increased occurrences of heavy precipitation, a decreased number of wet days may result in a reduced frequency of intense events (Christensen et al., 2007).

The main differences in the long-term shift of precipitation intensities in the downscaled correction are largely concerned with magnitude rather than a change of sign. This is evident in parts of eastern Sweden during winter where the downscaled corrections suggest smaller values for changes in p50, p75 and p90 than are simulated by ECHAM5 (Figure 7.15). This is also a region where the downscaling model is shown to be skilful in reproducing temporal variability, suggesting that high confidence may be

placed in the corrected quantiles. Additionally, the region encompassing southern Finland and the Baltic States shows a similar pattern of reduced magnitude in the projected increase of p75 and p90 although model skill is weaker here.

At high latitudes and mountainous regions, and especially during winter, one must consider the contribution of snowfall to the simulated precipitation total. Changes in snow conditions are not explicitly analysed here, although it is important to consider the implications of differences in inter-GCM parameterisation of rainfall/snowfall. Previous work has shown that a general warming trend is likely to result in a strong decrease in precipitation in areas with consistent seasonal snow cover (Meehl et al., 2007). In Europe, a shortening of the snow season and reduction in snow depth is likely over the course of the next century (Raisanen and Alexandersson, 2003; Rowell, 2005). There are some regions however, such as the area encompassing northern Scandinavia and parts of north-western Russia, where mean temperature and precipitation changes exert less influence on snowfall than regions at lower latitudes (Raisanen and Alexandersson, 2003; Christensen et al., 2007). This also appears to be the case at high altitudes, such as the peaks of the Alps (Beniston et al., 2003).

During summer, additional information about precipitation change is more difficult to pinpoint in the downscaled correction (Figure 7.16). Furthermore, the skill of the quantile mapping method is poorer than that during the winter months. However, in parts of central Europe, the magnitude of the projected decrease in each quantile appears greater in the correction than the raw ECHAM5 simulation. Skill is stronger in this region than in the rest of continental Europe.

7.6 Summary and conclusions

A MOS approach for correcting GCM-simulated European precipitation has been extended for downscaling daily precipitation simulated by ECHAM5. A downscaling method was developed based on quantile mapping of a fitted gamma distribution and cross-validated on its skill to reproduce spatial patterns of observed precipitation quantiles. Downscaling corrections were applied to the ECHAM5 A1B SRES climate change simulation with reference made to regions of Europe where model skill was shown to be strong. It was demonstrated that, in calibrating a quantile mapping correction on precipitation from the nudged simulation, it is possible to produce robust estimates for future distributions of daily precipitation. Furthermore, utilising a nudged simulation permits a comparison of raw GCM-simulated daily precipitation with observations, and an understanding of spatially-varying GCM skill in reproducing temporal variability.

There are some implications for extending the broad concept of using a nudged simulation in the derivation of a correction for daily precipitation distribution. The nudging coefficient, N, of each variable

defined in Chapter 3 (Table 3.1) may not be suitable to reproducing variability on daily time scales. Although the coefficients are designed to reflect the different temporal evolution tendencies of each variable, the relaxation e-folding times ($\tau = 1/N$) are 24 hours for temperature and log(surface pressure) and 48 hours for divergence. So, by definition, the weights of these coefficients are insufficient to capture the evolution of these variables on a sub-daily time scale. Whilst this is not relevant in reproducing temporal variability in monthly means, it is uncertain whether the same nudging coefficients are appropriate for daily precipitation. Altering the coefficient for each variable is no simple procedure as the weighting of the nudging terms must be relative to one another. Ultimately, a physically consistent state amongst the suite of dynamically-resolved variables is required.

Most important for impact studies is understanding future changes in the magnitude and frequency of extreme precipitation events (Trenberth et al., 2003; Allan and Soden, 2008). In this chapter, 'heavy' precipitation events are defined as those above the 90th percentile of wet days (e.g. Haylock et al., 2006; Goodess et al., 2010). Whilst there is some encouraging skill in a quantile mapping correction in the representation of the 90th and 95th percentiles, a gamma distribution is insufficient in resolving events that form the extreme tail. There is potential, however, for MOS corrections (calibrated on a nudged simulation) to be based on a Generalised Extreme Value (GEV) distribution (Maraun et al., 2010), although this is beyond the scope of the current work.

Chapter 8

Conclusion and outlook

8.1 Summary and conclusions

The aim of this research has been to develop a MOS downscaling correction approach for estimating regional precipitation changes for the latest generation of climate change simulations. The MOS approach is based on a statistical correction of GCM-simulated precipitation and is conceptually different to the traditional statistical downscaling approach, which is generally known as Perfect-Prog. Key to the MOS approach developed here is the nudged hindcast GCM simulation upon which the statistical correction between simulated and observed (local-scale) precipitation is calibrated. The freely-evolving nature of a GCM means day-to-day variability of large-scale atmospheric states in a standard, non-nudged hindcast simulation does not match with observations, and so previously, MOS implementation on GCM-simulated precipitation has been limited to a correction of long-term means or distributions. Here, a nudged simulation of the ECHAM5 GCM forces the simulated large-scale fields into temporal phase with the observed record. Such a simulation provides a greater understanding of the value of ECHAM5-simulated precipitation and permits the development of sophisticated MOS corrections that are underpinned by a quantification of their respective skill.

This chapter aims to draw together the key findings of this research and to make recommendations for further study. In the remainder of this section, the objectives of this work outlined in Chapter 1 (section 1.3) are reasserted and discussed in terms of the extent to which each has been achieved. In section 8.2, the limitations of this research are identified and discussed in the context of scope for subsequent research that could potentially follow from what has been conducted here.

8.1.1 Objectives revisited

1. **To conduct a hindcast simulation using the ECHAM5 GCM in which the large-scale circulation and temperature variables are forced to corresponding fields in the ERA-40 reanalysis.** A nudged simulation of ECHAM5 has been conducted by forcing its prognostic fields (divergence, vorticity, temperature, log of surface pressure) to corresponding variables in ERA-40 (Chapter 3). The nudging process was shown to be most effective throughout the extra-tropics in forcing circulation and temperature variables into phase with observations, and least effective across tropical land regions. The nudging coefficients chosen for each variable were based on those used in previous work.

2. **To assess and quantify the skill of ECHAM5 to simulate precipitation given large-scale climatic conditions that are in temporal phase with real world observations.** In Chapter 3, three sources were identified for errors in GCM-simulated precipitation: unrealistic average properties of large-scale atmospheric states (e.g. bias in mean or variance) ('type 1'), the random generation of day-to-day weather and its 'mismatch' with the real world ('type 2'), and the deficiencies in the precipitation parameterisation ('type 3'). It was demonstrated that by conducting a nudged simulation, it is possible to approximately remove the type 1 and type 2 errors, thus isolating the type 3 error and permitting a quantification of the performance of the GCM's parameterisation of precipitation. Using correlation maps, regions were identified where ECHAM5 is able to reproduce interannual variability in monthly or seasonal precipitation given realistic large-scale conditions. It was stressed throughout Chapter 3 and in later chapters that there is a danger in some regions of an additional 'type 4' error induced by the nudging process. This is not considered to be evident where correlation is strong and where the difference in long-term means between the standard and nudged simulations is small. However, an explicit diagnosis of type 4 error is considered beyond the scope of this work.

3. **To develop and validate several statistical downscaling techniques following a MOS approach, in which local-scale monthly mean precipitation is estimated from precipitation as simulated by ECHAM5.** In Chapter 3, a simple local scaling technique (conceptually similar to bias correction) was used to demonstrate the potential for a statistical downscaling correction. This method was fully cross-validated in Chapter 4, along with two additional 'non-local' regression MOS methods based on one-dimensional MCA (SVD-RM) and PC-MLR (CCA-PCR) respectively. MOS downscaling models were generally shown to exhibit good skill in regions where ECHAM5 precipitation was shown to be skilful. This includes much of the northern hemispheric land mass and parts of Australia. Downscaling model performance shows a degree of seasonal de-

pendence, with European and North American precipitation estimates more reliable during winter than during summer. In general, performance of non-local corrections is not substantially greater than local scaling correction, and the latter may be preferable in many areas given its ease of development and implementation. However, non-local corrections do perform strongly across areas of complex topography, such as the European Alps. In general, the PC-MLR correction outperforms the MCA-based correction, particularly in terms of resolving a realistic amplitude of variance.

4. **To develop and validate Perfect-Prog statistical downscaling techniques that have been used successfully in the literature.** The same non-local regression methods used to develop MOS models were also developed under a Perfect-Prog approach. These models used ERA-40 fields at three atmospheric levels as solitary predictors or paired combinations of predictors and were constructed so as to provide a reference in the assessment of MOS downscaling models. Model skill is location- and season-dependent, but generally geopotential height and humidity variables exhibit the strongest predictor performance as solitary predictors. Paired combinations of geopotential height, temperature and specific humidity at lower atmospheric levels also exhibit good skill in outperforming solitary predictors in most regions. Again, the performance of PC-MLR estimates tends to be greater than that of the MCA-based estimates.

5. **To make a comparison between the relative merits of MOS and Perfect-Prog approaches to statistical downscaling.** Comparison in the merits of methods from each approach focused on three regions where downscaling potential is greatest: Europe, North America and Australia. In many areas of each region, MOS shows greater skill than Perfect-Prog techniques in reproducing realistic precipitation amounts, temporal variability and amplitude of variance. Some Perfect-Prog methods, particularly those based on multiple-predictors, perform more strongly than MOS in the driest parts of continental North America and Australia. It is concluded that, in many regions, MOS may act as an alternative to Perfect-Prog methods, or as a supplementary estimate in regions where Perfect-Prog estimates are skilful.

6. **To assess the applicability of a MOS downscaling correction in future climates.** In Chapter 5, each MOS downscaling model was assessed in terms of its stationarity and validity in a perturbed climate. It was shown that model transferability is reasonably strong in Europe and parts of North America where skill in current climate was shown to be strong (Chapter 4). In terms of estimating heavy monthly mean precipitation in the Northern Hemisphere, non-local MOS corrections show less skill, and particularly so during the summer months.

7. **To apply successful MOS corrections to the latest ECHAM5 climate change simulations used in the IPCC AR4 and to make estimates for future changes in local-scale**

precipitation across selected regions of interest. In Chapter 6, each MOS correction was applied to the ECHAM5 A1B SRES simulation and again analysed in greater detail in Europe, North America and Australia. All three downscaling corrections are able to represent small-scale precipitation features not evident in the raw GCM simulation with some degree of skill. Estimates from non-local regression methods based on MCA and PC-MLR tend to exhibit greater agreement, with local-scaling estimates often outlying, particularly in mountainous regions. It was possible to identify some good-skill regions where the extent of precipitation change differs between the downscaled and raw ECHAM5 projections.

8. **To evaluate the potential for a MOS downscaling correction of GCM-simulated daily precipitation distributions across Europe.** In Chapter 7, precipitation in the nudged simulation was shown to better represent key distribution statistics, including the proportion of wet days, than a standard simulation. It was shown that a quantile mapping correction of the simulated daily precipitation distribution (calibrated on the nudged simulation) has reasonable skill in reproducing observations, and able to resolve precipitation statistics up to the 95th percentile.

8.1.2 Overall conclusions

The overarching aim of this research has been broadly achieved. Five overall conclusions are listed below:

- For the first time, the skill of the parameterisation of precipitation in ECHAM5 to reproduce observed temporal variability given a realistic large-scale climatic state was assessed and quantified. In many regions, including much or Europe, and parts of North America and Australia, this skill was good.

- MOS downscaling models, using simulated precipitation as a predictor variable, performed strongly in estimating local-scale monthly mean precipitation across regions where precipitation is known to be well-simulated in ECHAM5. MOS models also showed better performance than Perfect-Prog regression-based downscaling models

- A MOS downscaling model based on PC-MLR was shown to be the strongest performer, particularly across regions of complex topography. A very simple and far less computationally intensive local scaling technique was shown to exhibit decent skill across lowland regions but its potential for application in mountainous regions was shown to be limited.

- In applying MOS downscaling corrections to ECHAM5 climate change simulations, many local-scale precipitation features were well-captured. The MOS model based on PC-MLR estimated

some precipitation changes that were not evident in raw ECHAM5 output but that do appear in RCM simulations.

- The potential for a MOS correction of daily precipitation distribution (based on a nudged simulation) appears promising, although more work is required in this area.

8.2 Limitations and scope for further research

8.2.1 Diagnosis of GCM parameterisation errors

As discussed at length in Chapter 3, the process of nudging to a reanalysis may potentially induce an further source of error in addition to errors in the parameterisation 'type 3' error. This so-called 'type 4' error can occur as a result of an imbalance between dynamically-resolved and parameterised model fields, which is otherwise physically consistent in a standard (non-nudged) simulation. Spurious precipitation may result from modified diabatic heating and a subsequent violation in energy conservation. It is important to reiterate that, for the purposes of this work, interest is only in areas where model bias is small and temporal variability is well-represented. That is, areas where the combined error (constituted of types 3 and 4) is small and it is assumed that spurious precipitation is not evident. However, whilst it is possible to provide an upper estimate for the extent of the 'type 3' error, it is difficult to fully diagnose the extent of this error, and thus fully quantify the skill of the parameterisation of precipitation in ECHAM5, without isolating the 'type 4' nudging error, or at least understanding its spatial extent.

One possibility is to conduct an additional nudged simulation with ECHAM5 in which the same prognostic variables are forced towards temporally corresponding fields from a previous standard (non-nugded) ECHAM5 simulation started from different initial conditions. Comparison of long-term means in each simulation would reveal areas were spurious precipitation may be problematic, although it remains unlikely that such areas would extend beyond the tropics and sub-tropics.

8.2.2 Extending to downscaling of extreme precipitation

It was demonstrated that MOS downscaling corrections of daily precipitation distributions potentially offer greater skill when calibrated on precipitation in a nudged simulation. Whilst a method based on quantile mapping of fitted gamma distributions is sufficient to correct medium-to-heavy precipitation, the technique is insufficient in resolving extreme precipitation. Further work in MOS downscaling may focus solely on extreme events, which should be based on the concepts of extreme value theory (e.g. Katz et al., 2002). As noted by Maraun et al. (2010), there are relatively few examples of such studies that have tackled precipitation downscaling using this approach, although recent examples of model inter-

comparison do exist (e.g. Frei et al., 2006; Haylock et al., 2006; Beniston et al., 2007; Kendon et al., 2008). As mentioned in the summary of Chapter 7, a commonly used method for estimating future occurrence and magnitude of extremes on local-scales involves fitting a generalised extreme value (GEV) distribution to block maxima, which would provide a good starting point for further work.

8.2.3 MOS contribution to downscaling ensembles

Arguably the most obvious limitation in this research is that only one GCM is considered. All climate change projections, including those made in the IPCC reports, are based on the simulations of up to twenty different GCMs. In many cases, a number of simulations with differing initial conditions and external forcings are conducted for each individual GCM. Despite ECHAM5 being one of the more reputable GCMs available, a solitary model is insufficient when making predictions about future climate changes. Likewise, downscaled estimates based on one driving GCM are likely to be far less robust than those based on a GCM ensemble. It is again important to reiterate that the main goal in this work is the 'development' of a correction approach. The vital additional step in the downscaling process of conducting a nudged simulation is very much a computationally-intensive and time-consuming one, particularly as most groups working on downscaling tend to acquire GCM output (on which to drive RCMs or calibrate statistical models) rather than perform the actual simulations. It is arguable, however, that there need only be one nudged simulation per GCM product, if it is assumed that the impact of nudging-induced spurious precipitation ('type 4' error) can be fully quantified and understood.

Additionally, in accordance with recommendations made by STARDEX (Goodess et al., 2010), future downscaling studies may seek to develop local-scale predictions based on an ensemble of downscaled estimates. This is, of course, common practice when dealing with multiple GCM simulations, and also RCM simulations, whether from different models or different perturbations of the same model. Precipitation projections based on estimates from a number of downscaling models with various predictors may offer more robust estimates, at least in areas where there is consistency in the spatial representation of major precipitation features. A more systematic approach than simply relying on a multi-model mean would be to weight each model by its skill according to a rigorous cross-validation.

8.2.4 Towards probabilistic downscaling

The downscaling models developed throughout this work define a statistical relationship between either GCM-simulated precipitation in the case of MOS, or some other large-scale predictors in the case of Perfect-Prog, and the expected value of local scale precipitation. Such models are deterministic and do not account for any noise that cannot be explained by the predictors. A major limitation of a deterministic

approach to statistical downscaling is that, without an additional noise term, it is not possible to fully quantify the extent to which small-scale spatial differences in precipitation are consistent with random variability. There exist examples of more sophisticated techniques that seek to represent the contribution of extreme values and variability to noise; such techniques are collectively referred to as stochastic, or probabilistic, downscaling.

In producing estimates of future precipitation, particularly at the regional scale, the provision of information regarding uncertainty is an important challenge. Climate projections developed in such projects as UKCP09 have been relatively informative on uncertainty pertaining to both the driving GCMs and the downscaling models themselves, and such uncertainty is likely to be better understood as a result of continued model development. However, in a previously cited recent review paper, Maraun et al. (2010) argue that these projections do not adequately address uncertainty associated with random variability. The authors note that probabilistic projections are particularly important on decadal time scales in order to sufficiently account for climate change signals associated with natural variability.

8.3 Improving projections of future regional precipitation

Future climate is likely to be associated with changes in precipitation patterns and understanding the impact of these changes at different scales is an important focus in climate change science. Being able to make reliable estimates of future precipitation changes on regional- to local- scales remains a major challenge. Over the last couple of decades, attempts to achieve this have developed and implemented a breadth of statistical and dynamical downscaling techniques, many implemented as part of coordinated international projects (e.g. PRUDENCE, ENSEMBLES, STARDEX). The limitations of statistical and dynamical downscaling are well understood, but the most obvious advantage of statistical downscaling methodologies is a lesser computational requirement and the potential to provide support and comparison with RCM simulations (W.J. Gutowski, personal communication). The Coordinated Regional Downscaling Experiment (CORDEX), an international project in its infancy but currently under way, has recently called for a greater provision of statistically downscaled regional climate projections to add to a growing number of RCM simulations. Projects such as CORDEX demonstrate the value that statistical approaches continue to offer to the downscaling community, even in the advancement of RCMs.

The application of a MOS correction scheme for other GCM simulations has the potential to add benefit to precipitation projections, both in terms of producing realistic local-scale estimates and of developing an understanding of GCM skill as a basis for placing confidence in such estimates. Despite the necessary additional computational procedure involved with this approach to statistical downscaling, there is scope for its application to extend to other GCMs, finer spatial resolutions and particular regions.

Bibliography

Adler, R. F., et al., 2003: The version-2 global precipitation climatology project (GPCP) monthly precipitation analysis (1979-present). *Journal of Hydrometeorology*, **4 (6)**, 1147–1167.

Allan, R. P. and B. J. Soden, 2008: Atmospheric warming and the amplification of precipitation extremes. *Science*, **321 (5895)**, 1481–1484.

Allen, M. R. and W. J. Ingram, 2002: Constraints on future changes in climate and the hydrologic cycle. *Nature*, **419 (6903)**, 224+, doi:{10.1038/nature01092}.

Bardossy, A. and H. J. Caspary, 1990: Detection of climate change in Europe by analyzing European atmospheric circulation patterns from 1881-1989. *Theoretical and Applied Climatology*, **42 (3)**, 155–167.

Bardossy, A. and E. J. Plate, 1992: Space-Time Model for Daily Rainfall Using Atmospheric Circulation Patterns. *Water Resources Research*, **28 (5)**, 1247–1259.

Beck, C., J. Grieser, and B. Rudolf, 2005: A New Monthly Precipitation Climatology for the Global Land Areas for the Period 1951 to 2000. *DWD, Klimastatusbericht 2004*, 181–190.

Bell, J. L., L. C. Sloan, and M. A. Snyder, 2004: Regional changes in extreme climatic events: A future climate scenario. *Journal of Climate*, **17 (1)**, 81–87.

Beniston, M., F. Keller, B. Koffi, and S. Goyette, 2003: Estimates of snow accumulation and volume in the Swiss Alps under changing climatic conditions. *Theoretical and Applied Climatology*, **76 (3-4)**, 125–140, doi:{10.1007/s00704-003-0016-5}.

Beniston, M., et al., 2007: Future extreme events in European climate: an exploration of regional climate model projections. *Climatic Change*, **81 (Suppl. 1)**, 71–95, doi:{10.1007/s10584-006-9226-z}.

Boe, J., L. Terray, F. Habets, and E. Martin, 2007: Statistical and dynamical downscaling of the Seine basin climate for hydro-meteorological studies. *International Journal of Climatology*, **27 (12)**, 1643–

1655, doi:{10.1002/joc.1602}, General Assembly of the European-Geosciences-Union, Vienna, Austria, April 2006.

Bosilovich, M. G., J. Y. Chen, F. R. Robertson, and R. F. Adler, 2008: Evaluation of global precipitation in reanalyses. *Journal of Applied Meteorology and Climatology*, **47 (9)**, 2279–2299.

Bretherton, C. S., C. Smith, and J. M. Wallace, 1992: An Intercomparison of Methods for Finding Coupled Patterns in Climate Data. *Journal of Climate*, **5 (6)**, 541–560.

Buonomo, E., 2010: Assessment of daily extremes fom the ENSEMBLES set of regional climate models driven by ERA-40 reanalysis (in preparation). *Climatic Research*.

Buonomo, E., R. Jones, C. Huntingford, and J. Hannaford, 2007: On the robustness of changes in extreme precipitation over Europe from two high resolution climate change simulations. *Quarterly Journal of the Royal Meteorological Society*, **133 (622, Part A)**, 65–81, doi:{10.1002/qj.13}.

Burger, G., 2002: Selected precipitation scenarios across Europe. *Journal of Hydrology*, **262 (1-4)**, 99–110.

Busuioc, A., D. L. Chen, and C. Hellstrom, 2001: Performance of statistical downscaling models in GCM validation and regional climate change estimates: Application for Swedish precipitation. *International Journal of Climatology*, **21 (5)**, 557–578.

Busuioc, A., F. Giorgi, X. Bi, and M. Ionita, 2006: Comparison of regional climate model and statistical downscaling simulations of different winter precipitation change scenarios over Romania. *Theoretical and Applied Climatology*, **86 (1-4)**, 101–123, doi:{10.1007/s00704-005-0210-8}.

Carnell, R. E. and C. A. Senior, 1998: Changes in mid latitude variability due to increasing greenhouse gases and sulphate aerosols. *Climate Dynamics*, **14 (5)**, 369–383.

Carter, G. M., J. P. Dallavalle, and H. R. Glahn, 1989: Statistical forecasts based on the National Meteorological Center's Numerical Weather Prediction System. *Weather and Forecasting*, **4**, 401–412.

Cavazos, T., 1999: Large-scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas. *Journal of Climate*, **12 (5)**, 1506–1523.

Cavazos, T. and B. C. Hewitson, 2005: Performance of NCEP-NCAR reanalysis variables in statistical downscaling of daily precipitation. *Climate Research*, **28**, 95–107.

Charles, S. P., B. C. Bates, I. N. Smith, and J. P. Hughes, 2004: Statistical downscaling of daily precipitation from observed and modelled atmospheric fields. *Hydrological Processes*, **18 (8)**, 1373–1394.

Charles, S. P., B. C. Bates, P. H. Whetton, and J. P. Hughes, 1999: Validation of downscaling models for changed climate conditions: case study of southwestern Australia. *Climate Research*, **12 (1)**, 1–14.

Chen, M., D. Pollard, and E. J. Barron, 2003: Comparison of future climate change over North America simulated by two regional models. *Journal of Geophysical Research-Atmospheres*, **108 (D12)**, doi:{10.1029/2002JD002738}.

Cherry, S., 1996: Singular value decomposition analysis and canonical correlation analysis. *Journal of Climate*, **9 (9)**, 2003–2009.

Christensen, J. H., T. R. Carter, M. Rummukainen, and G. Amanatidis, 2007: Evaluating the performance and utility of regional climate models: the PRUDENCE project. *Climatic Change*, **81 (Suppl. 1)**, 1–6, doi:{10.1007/s10584-006-9211-6}.

Christensen, J. H. and O. B. Christensen, 2007: A summary of the PRUDENCE model projections of changes in European climate by the end of this century. *Climatic Change*, **81 (Suppl. 1)**, 7–30, doi:{10.1007/s10584-006-9210-7}.

Christensen, J. H., et al., 2007: *Regional Climate Projections. In: Climate Change 2007: The Physical Scientific Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M. and Miller, H. L. (eds).* Cambridge University Press.

Conway, D. and P. D. Jones, 1998: The use of weather types and air flow indices for GCM downscaling. *Journal of Hydrology*, **213 (1-4)**, 348–361.

Conway, D., R. L. Wilby, and P. D. Jones, 1996: Precipitation and air flow indices over the British Isles. *Climate Research*, **7 (2)**, 169–183.

Corte-Real, J., H. Xu, and B. D. Qian, 1999: A weather generator for obtaining daily precipitation scenarios based on circulation patterns. *Climate Research*, **13 (1)**, 61–75.

Crane, R. G. and B. C. Hewitson, 1998: Doubled carbon dioxide precipitation changes for the Susquehanna basin: Down-scaling from the genesis general circulation model. *International Journal of Climatology*, **18 (1)**, 65–76.

Dairaku, K. and S. Emori, 2006: Dynamic and thermodynamic influences on intensified daily rainfall during the Asian summer monsoon under doubled atmospheric carbon dioxide conditions. *Geophysical Research Letters*, **33 (1)**, doi:{10.1029/2005GL024754}.

DelSole, T. and P. Chang, 2003: Predictable component analysis, canonical correlation analysis, and autoregressive models. *Journal of the Atmospheric Sciences*, **60 (2)**, 409–416.

Deque, M., et al., 2007: An intercomparison of regional climate simulations for Europe: assessing uncertainties in model projections. *Climatic Change*, **81 (Suppl. 1)**, 53–70, doi:{10.1007/s10584-006-9228-x}.

Dettinger, M. D., D. R. Cayan, M. Meyer, and A. E. Jeton, 2004: Simulated hydrologic responses to climate variations and change in the Merced, Carson, and American River basins, Sierra Nevada, California, 1900-2099. *Climatic Change*, **62 (1-3)**, 283–317.

Douville, H., F. Chauvin, S. Planton, J. F. Royer, D. Salas-Melia, and S. Tyteca, 2002: Sensitivity of the hydrological cycle to increasing amounts of greenhouse gases and aerosols. *Climate Dynamics*, **20 (1)**, 45–68, doi:{10.1007/s00382-002-0259-3}.

Durman, C. F., J. M. Gregory, D. C. Hassell, R. G. Jones, and J. M. Murphy, 2001: A comparison of extreme European daily precipitation simulated by a global and a regional climate model for present and future climates. *Quarterly Journal of the Royal Meteorological Society*, **127 (573, Part A)**, 1005–1015.

Ekstrom, M., P. D. Jones, H. J. Fowler, G. Lenderink, T. A. Buishand, and D. Conway, 2007: Regional climate model data used within the SWURVE project 1: projected changes in seasonal patterns and estimation of PET. *Hydrology and Earth System Sciences*, **11 (3)**, 1069–1083.

Engen-Skaugen, T., 2007: Refinement of dynamically downscaled precipitation and temperature scenarios. *Climatic Change*, **84 (3-4)**, 365–382, doi:{10.1007/s10584-007-9251-6}.

Feddersen, H. and U. Andersen, 2005: A method for statistical downscaling of seasonal ensemble predictions. *Tellus Series A-Dynamic Meteorology and Oceanography*, **57 (3)**, 398–408.

Ferro, C. A. T., A. Hannachi, and D. B. Stephenson, 2005: Simple nonparametric techniques for exploring changing probability distributions of weather. *Journal of Climate*, **18 (21)**, 4344–4354.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: recent advances in downscaling techniques for hydrological modelling. *International Journal of Climatology*, **27 (12)**, 1547–1578.

Fowler, H. J., C. G. Kilsby, and P. E. O'Connell, 2000: A stochastic rainfall model for the assessment of regional water resource systems under changed climatic conditions. *Hydrology and Earth System Sciences*, **4 (2)**, 263–282.

Fowler, H. J., C. G. Kilsby, P. E. O'Connell, and A. Burton, 2005: A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. *Journal of Hydrology*, **308 (1-4)**, 50–66, doi:10.1016/j.jhydrol.2004.10.021.

Frei, C., J. H. Christensen, M. Deque, D. Jacob, R. G. Jones, and P. L. Vidale, 2003: Daily precipitation statistics in regional climate models: Evaluation and intercomparison for the European Alps. *Journal of Geophysical Research-Atmospheres*, **108 (D3)**, doi:{10.1029/2002JD002287}.

Frei, C., R. Scholl, S. Fukutome, J. Schmidli, and P. L. Vidale, 2006: Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. *Journal of Geophysical Research-Atmospheres*, **111 (D6)**, doi:{10.1029/2005JD005965}.

Frias, M. D., E. Zorita, J. Fernandez, and C. Rodriguez-Puebla, 2006: Testing statistical downscaling methods in simulated climates. *Geophysical Research Letters*, **33 (19)**, doi:{10.1029/2006GL027453}.

Genthon, C., G. Krinner, and E. Cosme, 2002: Free and laterally nudged antarctic climate of an atmospheric general circulation model. *Monthly Weather Review*, **130 (6)**, 1601–1616.

Giorgi, F. and R. Francisco, 2000: Evaluating uncertainties in the prediction of regional climate change. *Geophysical Research Letters*, **27 (9)**, 1295–1298.

Giorgi, F., C. Jones, and G. R. Asrar, 2009: Addressing climate information needs at the regional level: the CORDEX framework. *WMO Bulletin*, **58 (3)**, 175–183.

Giorgi, F. and M. R. Marinucci, 1996: Improvements in the simulation of surface climatology over the European region with a nested modeling system. *Geophysical Research Letters*, **23 (3)**, 273–276.

Giorgi, F., et al., 2001: *Regional Climate Information – Evaluation and Projections. In: Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K. and Johnson, C.A.* Cambridge University Press.

Glahn, H. R., 1968: Canonical correlation and its relationship to discriminant analysis and multiple regression. *Journal of the Atmospheric Sciences*, **25 (1)**, 23.

Glahn, H. R. and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, **11**, 1203–1211.

Godfrey, R. A., 1982: An Application of Model Output Statistics to the Development of a Local Wind Regime Forecast Procedure. *Journal of Applied Meteorology*, **21 (12)**, 1786–1791.

Goodess, C. M. and P. D. Jones, 2002: Links between circulation and changes in the characteristics of Iberian rainfall. *International Journal of Climatology*, **22 (13)**, 1593–1615.

Goodess, C. M. and J. P. Palutikof, 1998: Development of daily rainfall scenarios for southeast Spain using a circulation-type approach to downscaling. *International Journal of Climatology*, **18 (10)**, 1051–1083.

Goodess, C. M., et al., 2010: An intercomparison of statistical downscaling methods for Europe and European regions – assessing their performance with respect to extreme weather events and the implications for climate change applications (submitted). *Climatic Change*.

Graham, L. P., J. Andreasson, and B. Carlsson, 2007: Assessing climate change impacts on hydrology from an ensemble of regional climate models, model scales and linking methods - a case study on the Lule River basin. *Climatic Change*, **81 (Suppl. 1)**, 293–307, doi:{10.1007/s10584-006-9215-2}.

Groisman, P. Y., R. W. Knight, D. R. Easterling, T. R. Karl, G. C. Hegerl, and V. A. N. Razuvaev, 2005: Trends in intense precipitation in the climate record. *Journal of Climate*, **18 (9)**, 1326–1350.

Gruber, A., X. J. Su, M. Kanamitsu, and J. Schemm, 2000: The comparison of two merged rain gauge-satellite precipitation datasets. *Bulletin of the American Meteorological Society*, **81 (11)**, 2631–2644.

Gutowski, W. J., Y. B. Chen, and Z. Otles, 1997: Atmospheric water vapor transport in NCEP-NCAR reanalyses: Comparison with river discharge in the central United States. *Bulletin of the American Meteorological Society*, **78 (9)**, 1957–1969.

Hannachi, A., I. T. Jolliffe, and D. B. Stephenson, 2007: Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, **27 (9)**, 1119–1152, doi:{10.1002/joc.1499}.

Hanssen-Bauer, I., E. J. Forland, J. E. Haugen, and O. E. Tveito, 2003: Temperature and precipitation scenarios for Norway: comparison of results from dynamical and empirical downscaling. *Climate Research*, **25 (1)**, 15–27.

Hanssen-Bauer, I., O. E. Tveito, and E. J. Forland, 2001: *Temperature scenarios for Norway. Empirical downscaling from ECHAM4/OPYC3. Klima Report 24/00, Norwegian Meterological Institute, Oslo.*

Hay, L. E. and M. P. Clark, 2003: Use of statistically and dynamically downscaled atmospheric model output for hydrologic simulations in three mountainous basins in the western United States. *Journal of Hydrology*, **282 (1-4)**, 56–75, doi:{10.1016/S0022-1694(03)00252-X}.

Hay, L. E., G. J. McCabe, D. M. Wolock, and M. A. Ayers, 1992: Use of Weather Types to Disaggregate General-Circulation Model Predictions. *Journal of Geophysical Research-Atmospheres*, **97 (D3)**, 2781–2790.

Haylock, M. R., G. C. Cawley, C. Harpham, R. L. Wilby, and C. M. Goodess, 2006: Downscaling heavy precipitation over the United Kingdom: A comparison of dynamical and statistical methods and their future scenarios. *International Journal of Climatology*, **26 (10)**, 1397–1415.

Haylock, M. R., N. Hofstra, A. M. G. K. Tank, E. J. Klok, P. D. Jones, and M. New, 2008: A European daily high-resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research-Atmospheres*, **113 (D20)**, doi:{10.1029/2008JD010201}.

Hegerl, G. C., F. W. Zwiers, P. A. Stott, and V. V. Kharin, 2004: Detectability of anthropogenic changes in annual temperature and precipitation extremes. *Journal of Climate*, **17 (19)**, 3683–3700.

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

Hellstrom, C. and D. L. Chen, 2003: Statistical downscaling based on dynamically downscaled predictors: Application to monthly precipitation in Sweden. *Advances in Atmospheric Sciences*, **20 (6)**, 951–958.

Hellstrom, C., D. L. Chen, C. Achberger, and J. Raisanen, 2001: Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. *Climate Research*, **19 (1)**, 45–55.

Hewitson, B. C. and R. G. Crane, 1996: Climate downscaling: Techniques and application. *Climate Research*, **7 (2)**, 85–95.

Hewitson, B. C. and R. G. Crane, 2002: Self-organizing maps: applications to synoptic climatology. *Climate Research*, **22 (1)**, 13–26.

Hewitson, B. C. and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. *International Journal of Climatology*, **26 (10)**, 1315–1337, doi:10.1002/joc.1314.

Higgins, R. W., K. C. Mo, and S. D. Schubert, 1996: The moisture budget of the central United States in spring as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. *Monthly Weather Review*, **124 (5)**, 939–963.

Hohenegger, C., P. Brockhaus, and C. Schaer, 2008: Towards climate simulations at cloud-resolving scales. *Meteorologische Zeitschrift*, **17 (4, Sp. Iss. SI)**, 383–394, doi:{10.1127/0941-2948/2008/0303}.

Hoke, J. E. and R. A. Anthes, 1976: Initialization of numerical-models by a dynamic initialization technique. *Monthly Weather Review*, **104 (12)**, 1551–1556.

Huffman, G. J., et al., 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bulletin of the American Meteorological Society*, **78 (1)**, 5–20.

Hulme, M., R. Doherty, T. Ngara, M. New, and D. Lister, 2001: African climate change: 1900-2100. *Climate Research*, **17 (2, Sp. Iss. SI)**, 145–168.

Hurrell, J. W., 1995: Decadal trends in the North-Atlantic Oscillation - regional temperatures and precipitation. *Science*, **269 (5224)**, 676–679.

Huth, R., 1999: Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Climate Research*, **13 (2)**, 91–101.

Huth, R., 2000: A circulation classification scheme applicable in GCM studies. *Theoretical and Applied Climatology*, **67 (1-2)**, 1–18.

IPCC, 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 [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, M. Tignor and H. L. Miller (eds.)].* Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 996pp.

Jacob, D., et al., 2007: An inter-comparison of regional climate models for Europe: model performance in present-day climate. *Climatic Change*, **81 (Suppl. 1)**, 31–52, doi:{10.1007/s10584-006-9213-4}.

Janowiak, J. E., A. Gruber, C. R. Kondragunta, R. E. Livezey, and G. J. Huffman, 1998: A comparison of the NCEP-NCAR reanalysis precipitation and the GPCP rain gauge-satellite combined dataset with observational error considerations. *Journal of Climate*, **11 (11)**, 2960–2979.

Jansen, E., et al., 2007: *Palaeoclimate. In: Climate Change 2007: The Physical Scientific Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (eds).* Cambridge University Press.

Jeuken, A. B. M., P. C. Siegmund, L. C. Heijboer, J. Feichter, and L. Bengtsson, 1996: On the potential of assimilating meteorological analyses in a global climate model for the purpose of model validation. *Journal of Geophysical Research-Atmospheres*, **101 (D12)**, 16 939–16 950.

Jimoh, O. D. and P. Webster, 1999: Stochastic modelling of daily rainfall in Nigeria: intra-annual variation of model parameters. *Journal of Hydrology*, **222 (1-4)**, 1–17.

Jones, P., C. Harpham, C. Kilsby, V. Glenis, and A. Burton, 2009: UK Climate Projections science report: Projections of future daily climate for the UK from the Weather Generator. Tech. rep., University of Newcastle, UK.

Jones, P. D. and D. Conway, 1997: Precipitation in the British Isles: An analysis of area-average data updated to 1995. *International Journal of Climatology*, **17 (4)**, 427–438.

Jones, P. D., M. Hulme, and K. R. Briffa, 1993: A comparison of Lamb circulation types with an objective classification scheme. *International Journal of Climatology*, **13 (6)**, 655–663.

Jones, R. G., J. M. Murphy, and M. Noguer, 1995: Simulation of climate-change over Europe using a nested regiona-climate model. 1. Assessment of control climate, including sensitivity to location of lateral boundaries. *Quarterly Journal of the Royal Meteorological Society*, **121 (526, Part B)**, 1413–1449.

Kaas, E., et al., 2000: Final report of the POTENTIALS project (Project On Tendency Evaluations using New Techniques to Improve Atmospheric Long-term Simulations). Available from the EU-Commision (DGXII). 93 pp.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge, 341pp.

Kalnay, E., et al., 1996: The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, **77 (3)**, 437–471.

Karl, T. R., 1979: Potential Application of Model Output Statistics (Mos) to Forecasts of Surface Ozone Concentrations. *Journal of Applied Meteorology*, **18 (3)**, 254–265.

Karl, T. R., W. C. Wang, M. E. Schlesinger, R. W. Knight, and D. Portman, 1990: A method of relating general-circulation model simulated climate to the observed local climate. 1. Seasonal statistics. *Journal of Climate*, **3 (10)**, 1053–1079.

Katz, R. W., 1996: Use of conditional stochastic models to generate climate change scenarios. *Climatic Change*, **32 (3)**, 237–255.

Katz, R. W. and M. B. Parlange, 1998: Overdispersion phenomenon in stochastic modeling of precipitation. *Journal of Climate*, **11 (4)**, 591–601.

Katz, R. W., M. B. Parlange, and P. Naveau, 2002: Statistics of extremes in hydrology. *Advances in Water Resources*, **25 (8-12)**, 1287–1304.

Kendon, E. J., D. P. Rowell, and R. G. Jones, 2010: Mechanisms and reliability of future projected changes in daily precipitation. *Climate Dynamics*, **35 (2-3)**, 489–509, doi:{10.1007/s00382-009-0639-z}.

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}.

Kidson, J. W., 2000: An analysis of New Zealand synoptic types and their use in defining weather regimes. *International Journal of Climatology*, **20 (3)**, 299–316.

Kidson, J. W. and C. S. Thompson, 1998: A comparison of statistical and model-based downscaling techniques for estimating local climate variations. *Journal of Climate*, **11 (4)**, 735–753.

Kilsby, C. G., P. S. P. Cowpertwait, P. E. O'Connell, and P. D. Jones, 1998: Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *International Journal of Climatology*, **18 (5)**, 523–539.

Kilsby, C. G., et al., 2007: A daily weather generator for use in climate change studies. *Environmental Modelling & Software*, **22 (12)**, 1705–1719, doi:{10.1016/j.envsoft.2007.02.005}.

Kistler, R., et al., 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bulletin of the American Meteorological Society*, **82 (2)**, 247–267.

Klein, W. H. and H. R. Glahn, 1974: Forecasting Local Weather by Means of Model Output Statistics. *Bulletin of the American Meteorological Society*, **55 (10)**, 1217–1227.

Klein, W. H., B. M. Lewis, and I. Enger, 1959: Objective prediction of 5-day mean temperature during Winter. *Journal of Meteorology*, **16 (6)**, 672–682.

Knutson, T. R. and S. Manabe, 1995: Time-mean response over the tropical Pacific to increased CO2 in a couple ocean-atmosphere model. *Journal of Climate*, **8 (9)**, 2181–2199.

Kostopoulou, E., C. Giannakopoulos, C. Anagnostopoulou, K. Tolika, P. Maheras, M. Vafiadis, and D. Founda, 2007: Simulating maximum and minimum temperature over Greece: a comparison of three downscaling techniques. *Theoretical and Applied Climatology*, **90 (1-2)**, 65–82.

Krishnamurti, T. N., J. S. Xue, H. S. Bedi, K. Ingles, and D. Oosterhof, 1991: Physical initialization for numerical weather prediction over the tropics. *Tellus Series A-Dynamic Meteorology and Oceanography.*, **43 (4)**, 53–81.

Kumar, K. K., B. Rajagopalan, and M. A. Cane, 1999: On the weakening relationship between the Indian monsoon and ENSO. *Science*, **284 (5423)**, 2156–2159.

Kumar, K. R., A. K. Sahai, K. K. Kumar, S. K. Patwardhan, P. K. Mishra, J. V. Revadekar, K. Kamala, and G. Pant, 2006: High-resolution climate change scenarios for India for the 21st century. *Current Science*, **90 (3)**, 334–345.

Lamb, H. H., 1972: *British Isles Weather Types and a Register of Daily Sequence of Circulation Patterns, 1861-1971*. Geophysical Memoir, 116 HMSO, London, 85pp.

Landman, W. A. and L. Goddard, 2002: Statistical recalibration of GCM forecasts over southern Africa using model output statistics. *Journal of Climate*, **15 (15)**, 2038–2055.

Leander, R. and T. A. Buishand, 2007: Resampling of regional climate model output for the simulation of extreme river flows. *Journal of Hydrology*, **332 (3-4)**, 487–496, doi:{10.1016/j.jhydrol.2006.08.006}.

LeBlanc, M. J., P. Tregoning, G. Ramillien, S. O. Tweed, and A. Fakes, 2009: Basin-scale, integrated observations of the early 21st century multiyear drought in southeast Australia. *Water Resources Research*, **45**, W04 408 10 pp, doi:{10.1029/2008WR007333}.

Lenderink, G., A. Buishand, and W. van Deursen, 2007: Estimates of future discharges of the river Rhine using two scenario methodologies: direct versus delta approach. *Hydrology and Earth System Sciences*, **11 (3)**, 1143–1159.

Lenderink, G. and E. van Meijgaard, 2008: Increase in hourly precipitation extremes beyond expectations from temperature changes. *Nature Geoscience*, **1 (8)**, 511–514, doi:{10.1038/ngeo262}.

Leung, L. R. and S. J. Ghan, 1999: Pacific Northwest climate sensitivity simulated by a regional climate model driven by a GCM. Part I: Control simulations. *Journal of Climate*, **12 (7)**, 2010–2030.

Lyne, W. H., R. Swinbank, and N. T. Birch, 1982: A data assimilation experiment and the global circulation during the FGGE special observing periods. *Quarterly Journal of the Royal Meteorological Society*, **108 (457)**, 575–594.

Maheras, P., K. Tolika, C. Anagnostopoulou, M. Vafiadis, I. Patrikas, and H. Flocas, 2004: On the relationships between circulation types and changes in rainfall variability in Greece. *International Journal of Climatology*, **24 (13)**, 1695–1712, doi:10.1002/joc.1088.

Maraun, D., H. W. Rust, and T. J. Osborn, 2010: The influence of synoptic airflow on UK daily precipitation extremes. Part I: Observed spatio-temporal relationships (in press). *Climate Dynamics*.

Maraun, D., et al., 2010: Precipitation downscaling under climate change: Recent developments to bridge the gap between dynamical models and the end user. *Reviews of Geophysics*, **48**, RG3003, 34 PP., doi:{10.1029/2009RG000314}.

Matyasovszky, I. and I. Bogardi, 1996: Downscaling two versions of a general circulation model to estimate local hydroclimatic factors under climate change. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, **41 (1)**, 117–129.

Maurer, E. P., A. W. Wood, J. C. Adam, D. P. Lettenmaier, and B. Nijssen, 2002: A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States. *Journal of Climate*, **15 (22)**, 3237–3251.

McCutchan, M. H., 1978: Model for Predicting Synoptic Weather Types Based on Model Output Statistics. *Journal of Applied Meteorology*, **17 (10)**, 1466–1475.

Mearns, L. O., I. Bogardi, F. Giorgi, I. Matyasovszky, and M. Palecki, 1999: Comparison of climate change scenarios generated from regional climate model experiments and statistical downscaling. *Journal of Geophysical Research-Atmospheres*, **104 (D6)**, 6603–6621.

Mearns, L. O., F. Giorgi, P. Whetton, D. Pabon, M. Hulme, and M. Lal, 2003: Guidelines for Use of Climate Scenarios Developed from Regional Climate Model Experiments. Data Distribution Centre of the Intergovernmental Panel on Climate Change.

Mearns, L. O., W. Gutowski, R. Jones, R. Leung, S. McGinnis, A. Nunes, and Y. Qian, 2009: A Regional Climate Change Assessment Program for North America. *Eos, Transactions American Geophysical Union*, **90 (36)**, 311, doi:doi:10.1029/2009EO360002.

Meehl, G. A., et al., 2007: *Regional Climate Projections. In: Climate Change 2007: The Physical Scientific Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (eds).* Cambridge University Press.

Mo, K. C., M. Chelliah, M. L. Carrera, R. W. Higgins, and W. Ebisuzaki, 2005: Atmospheric moisture transport over the United States and Mexico as evaluated in the NCEP regional reanalysis. *Journal of Hydrometeorology*, **6 (5)**, 710–728.

Mo, K. C. and R. W. Higgins, 1996: Large-scale atmospheric moisture transport as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. *Journal of Climate*, **9 (7)**, 1531–1545.

Moron, V., A. W. Robertson, M. N. Ward, and O. Ndiaye, 2008: Weather types and rainfall over Senegal. part I: Observational analysis. *Journal of Climate*, **21 (2)**, 266–287.

Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *Journal of Climate*, **12 (8)**, 2256–2284.

Murphy, J., 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *International Journal of Climatology*, **20 (5)**, 489–501.

Nakicenovic, N. and R. Swart, (Eds.) , 2000: *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 599pp.

Neelin, J. D., M. Munnich, H. Su, J. E. Meyerson, and C. E. Holloway, 2006: Tropical drying trends in global warming models and observations. *Proceedings of the National Academy of Sciences of the United States of America*, **103 (16)**, 6110–6115, doi:10.1073/pnas.0601798103.

Noguer, M., R. Jones, and J. Murphy, 1998: Sources of systematic errors in the climatology of a regional climate model over Europe. *Climate Dynamics*, **14 (10)**, 691–712.

Osborn, T. J., D. Conway, M. Hulme, J. M. Gregory, and P. D. Jones, 1999: Air flow influences on local climate: observed and simulated mean relationships for the United Kingdom. *Climate Research*, **13 (3)**, 173–191.

Osborn, T. J., M. Hulme, P. D. Jones, and T. A. Basnett, 2000: Observed trends in the daily intensity of United Kingdom precipitation. *International Journal of Climatology*, **20 (4)**, 347–364.

Palutikof, J. P., C. M. Goodess, S. J. Watkins, and T. Holt, 2002: Generating rainfall and temperature scenarios at multiple sites: Examples from the Mediterranean. *Journal of Climate*, **15 (24)**, 3529–3548.

Pant, G. B. and K. Rupa Kumar, 1997: *Climates of South Asia*. John Wiley & Sons, Chichester, 320pp.

Perica, S. and E. FoufoulaGeorgiou, 1996: Model for multiscale disaggregation of spatial rainfall based on coupling meteorological and scaling descriptions. *Journal of Geophysical Research-Atmospheres*, **101 (D21)**, 26 347–26 361.

Piani, C., J. O. Haerter, and E. Coppola, 2010: Statistical bias correction for daily precipitation in regional climate models over Europe. *Theoretical and Applied Climatology*, **99 (1-2)**, 187–192, doi: {10.1007/s00704-009-0134-9}.

Plummer, D. A., et al., 2006: Climate and climate change over North America as simulated by the Canadian RCM. *Journal of Climate*, **19 (13)**, 3112–3132.

Poccard, I., S. Janicot, and P. Camberlin, 2000: Comparison of rainfall structures between NCEP/NCAR reanalyses and observed data over tropical Africa. *Climate Dynamics*, **16 (12)**, 897–915.

Raisanen, J. and H. Alexandersson, 2003: A probabilistic view on recent and near future climate change in Sweden. *Tellus Series A-Dynamic Meteorology and Oceanography*, **55 (2)**, 113–125.

Ramamurthy, M. K. and F. H. Carr, 1987: 4-Dimensional Data Assimilation In The Monsoon Region .1. Experiments With Wind Data. *Monthly Weather Review*, **115 (8)**, 1678–1706.

Randall, D. A., et al., 2007: *Cilmate Models and Their Evaluation. In: Climate Change 2007: The Physical Scientific Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (eds).* Cambridge University Press.

Reid, P. A., P. D. Jones, O. Brown, C. M. Goodess, and T. D. Davies, 2001: Assessments of the reliability of NCEP circulation data and relationships with surface climate by direct comparisons with station based data. *Climate Research*, **17 (3)**, 247–261.

Risbey, J. S. and P. H. Stone, 1996: A case study of the adequacy of GCM simulations for input to regional climate change assessments. *Journal of Climate*, **9 (7)**, 1441–1467.

Robertson, A. W., S. Kirshner, and P. Smyth, 2004: Downscaling of daily rainfall occurrence over northeast Brazil using a hidden Markov model. *Journal of Climate*, **17 (22)**, 4407–4424.

Roeckner, E., et al., 2006: Sensitivity of simulated climate to horizontal and vertical resolution in the ECHAM5 atmosphere model. *Journal of Climate*, **19 (16)**, 3771–3791.

Rowell, D. P., 2005: A scenario of European climate change for the late twenty-first century: seasonal means and interannual variability. *Climate Dynamics*, **25 (7-8)**, 837–849, doi:{10.1007/s00382-005-0068-6}.

Rowell, D. P. and R. G. Jones, 2006: Causes and uncertainty of future summer drying over Europe. *Climate Dynamics*, **27 (2-3)**, 281–299, doi:{10.1007/s00382-006-0125-9}.

Rudolf, B., H. Hauschild, W. Rueth, and U. Schneider, 1994: Terrestrial Precipitation Analysis: Operational Method and Required Density of Point Measurements. *NATO ASI I/26, Global Precipitations and Climate Change (Ed. M. Desbois and F. Desalmand)*, Springer Verlag Berlin, 173–186.

Rudolf, B. and U. Schneider, 2005: Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. *2nd Workshop of the International Precipitation Working Group IPWG, Monterey, October 2004*, 231–247.

Rummukainen, M., 1997: Methods of statistical downscaling of GCM simulations. Tech. Rep. 80, Swedish Meteorological and Hydrological Institute, Norrkping, Sweden.

Rummukainen, M., S. Bergstrom, G. Persson, J. Rodhe, and M. Tjernstrom, 2004: The Swedish Regional Climate Modelling Programme, SWECLIM: A review. *AMBIO*, **33 (4-5)**, 176–182.

Ruosteenoja, K., T. Carter, K. Jylha, and H. Tuomenvirta, 2003: *Future Climate in World Regions: And Intercomparison of Model-Based Projections for the New IPCC Emissions Scenarios*. Finnish Environment Institute, Helsinki, 83pp.

Salathe, E. P., 2003: Comparison of various precipitation downscaling methods for the simulation of streamflow in a rainshadow river basin. *International Journal of Climatology*, **23 (8)**, 887–901.

Sansom, J. and J. A. Renwick, 2007: Climate change scenarios for New Zealand rainfall. *Journal of Applied Meteorology and Climatology*, **46 (5)**, 573–590.

Sarkar, S., R. P. Singh, and M. Kafatos, 2004: Further evidences for the weakening relationship of Indian rainfall and ENSO over India. *Geophysical Research Letters*, **31 (13)**, doi:10.1029/2004GL020259.

Schmidli, J., C. Frei, and P. L. Vidale, 2006: Downscaling from GCM precipitation: A benchmark for dynamical and statistical downscaling methods. *International Journal of Climatology*, **26 (5)**, 679–689.

Schmidli, J., C. M. Goodess, C. Frei, M. R. Haylock, Y. Hundecha, J. Ribalaygua, and T. Schmith, 2007: Statistical and dynamical downscaling of precipitation: An evaluation and comparison of scenarios for the European Alps. *Journal of Geophysical Research-Atmospheres*, **112 (D4)**.

Schneider, U., T. Fuchs, A. Meyer-Christoffer, and B. Rudolf, 2008: Global Precipitation Analysis Products of the GPCC. Tech. rep., Global Precipitation Climatology Centre (GPCC), Deutscher Wetterdienst, Offenbach a. M., Germany.

Schoof, J. T. and S. C. Pryor, 2001: Downscaling temperature and precipitation: A comparison of regression-based methods and artificial neural networks. *International Journal of Climatology*, **21 (7)**, 773–790.

Sheffield, J., A. D. Ziegler, E. F. Wood, and Y. B. Chen, 2004: Correction of the high-latitude rain day anomaly in the NCEP-NCAR reanalysis for land surface hydrological modeling. *Journal of Climate*, **17 (19)**, 3814–3828.

Shongwe, M. E., W. A. Landman, and S. J. Mason, 2006: Performance of recalibration systems for GCM forecasts for southern Africa. *International Journal of Climatology*, **26 (12)**, 1567–1585.

Sohn, E., 2007: The Big Dry. *Science News*, **172**, 17.

Spak, S., T. Holloway, B. Lynn, and R. Goldberg, 2007: A comparison of statistical and dynamical downscaling for surface temperature in North America. *Journal of Geophysical Research-Atmospheres*, **112 (D8)**.

Stauffer, D. R. and N. L. Seaman, 1990: Use of 4-dimensional data assimilation in a limited-area mesoscale model. 1. Experiments with synoptic-scale data. *Monthly Weather Review*, **118 (6)**, 1250–1277.

Sun, Y., S. Solomon, A. Dai, and R. W. Portmann, 2006: How often does it rain? *Journal of Climate*, **19 (6)**, 916–934.

Taschetto, A. S. and M. H. England, 2009: An analysis of late twentieth century trends in Australian rainfall. *International Journal of Climatology*, **29 (6)**, 791–807, doi:{10.1002/joc.1736}.

Taylor, K. E., 2001: Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research-Atmospheres*, **106 (D7)**, 7183–7192.

Themessl, M. J., A. Gobiet, and A. Leuprecht, 2011: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models (in press). *International Journal of Climatology*.

Thompson, D. W. J. and J. M. Wallace, 1998: The Arctic Oscillation signature in the wintertime geopotential height and temperature fields. *Geophysical Research Letters*, **25 (9)**, 1297–1300.

Timmreck, C., H. F. Graf, and J. Feichter, 1999: Simulation of Mt. Pinatubo volcanic aerosol with the Hamburg climate model ECHAM4. *Theoretical and Applied Climatology*, **62 (3-4)**, 85–108.

Timmreck, C. and M. Schulz, 2004: Significant dust simulation differences in nudged and climatological operation mode of the AGCM ECHAM. *Journal of Geophysical Research-Atmospheres*, **109 (D13)**.

Tippett, M. K., T. DelSole, S. J. Mason, and A. G. Barnston, 2008: Regression-based methods for finding coupled patterns. *Journal of Climate*, **21 (17)**, 4384–4398, doi:10.1175/2008JCLI2150.1.

Tolika, K., P. Maheras, H. A. Flocas, and A. Arseni-Papadimitrioua, 2006: An evaluation of a general circulation model (GCM) and the NCEP-NCAR reanalysis data for winter precipitation in Greece. *International Journal of Climatology*, **26 (7)**, 935–955.

Tolika, K., P. Maheras, M. Vafiadis, H. A. Flocasc, and A. Arseni-Papadimitriou, 2007: Simulation of seasonal precipitation and raindays over Greece: a statistical downscaling technique based on artificial neural networks (ANNs). *International Journal of Climatology*, **27 (7)**, 861–881.

Trenberth, K. E., A. G. Dai, R. M. Rasmussen, and D. B. Parsons, 2003: The changing character of precipitation. *Bulletin of the American Meteorological Society*, **84 (9)**, 1205, doi:10.1175/BAMS-84-9-1205.

Trenberth, K. E. and D. J. Shea, 2005: Relationships between precipitation and surface temperature. *Geophysical Research Letters*, **32 (14)**, doi:10.1029/2005GL022760.

Trenberth, K. E., et al., 2007: *Observations: Surface and Atmospheric Climate Change. In: Climate Change 2007: The Physical Scientific Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M. and Miller, H. L. (eds).* Cambridge University Press.

Trigo, R. M. and J. P. Palutikof, 2001: Precipitation scenarios over Iberia: A comparison between direct GCM output and different downscaling techniques. *Journal of Climate*, **14 (23)**, 4422–4446.

Ummenhofer, C. C., et al., 2011: Indian and Pacific Ocean influences on Southeast Australian drought and hydrology (in press). *Journal of Climate*.

Uppala, S. M., et al., 2005: The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, **131 (612, Part B)**, 2961–3012, doi:10.1256/qj.04.176.

van der Linden, P. and J. F. B. Mitchell, 2009: ENSEMBLES: Climate Change and its Impacts: Summary of research and results from the ENSEMBLES project. Tech. rep., Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.

von Storch, H., 1995: Inconsistencies at the interface of climate impact studies and global climate research. *Meteorologische Zeitschrift*, **4**, 72–80.

von Storch, H., 1999: On the use of "inflation" in statistical downscaling. *Journal of Climate*, **12 (12)**, 3505–3506.

Vrac, M., M. L. Stein, K. Hayhoe, and X. Z. Liang, 2007: A general method for validating statistical downscaling methods under future climate change. *Geophysical Research Letters*, **34 (18)**, doi:{10.1029/2007GL030295}.

Wallace, J. M., Y. Zhang, and J. A. Renwick, 1995: Dynamic contribution to hemispheric mean temperature trends. *Science*, **270 (5237)**, 780–783.

Wang, G. L., 2005: Agricultural drought in a future climate: results from 15 global climate models participating in the IPCC 4th assessment. *Climate Dynamics*, **25 (7-8)**, 739–753, doi:{10.1007/s00382-005-0057-9}.

Widmann, M., 2005: One-dimensional CCA and SVD, and their relationship to regression maps. *Journal of Climate*, **18 (14)**, 2785–2792.

Widmann, M. and C. S. Bretherton, 2000: Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *Journal of Climate*, **13 (11)**, 1936–1950.

Widmann, M., C. S. Bretherton, and E. P. Salathe, 2003: Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. *Journal of Climate*, **16 (5)**, 799–816.

Wigley, T. M. L. and P. D. Jones, 1987: England and Wales precipitation - a discussion of recent changes in variability and an update to 1985. *Journal of Climatology*, **7 (3)**, 231–246.

Wilby, R. L., 1994: Stochastic weather-type simulation for regional climate-change impact assessment. *Water Resources Research*, **30 (12)**, 3395–3403.

Wilby, R. L., S. P. Charles, E. Zorita, B. Timbal, P. Whetton, and L. O. Mearns, 2004: *Guidelines for Use of Climate Scenarios Developed from Statistical Downscaling Methods. IPCC Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA).*

Wilby, R. L., D. Conway, and P. D. Jones, 2002: Prospects for downscaling seasonal precipitation variability using conditioned weather generator parameters. *Hydrological Processes*, **16 (6, Sp. Iss. SI)**, 1215–1234, doi:10.1002/hyp.1058.

Wilby, R. L. and T. M. L. Wigley, 1997: Downscaling general circulation model output: a review of methods and limitations. *Progress in Physical Geography*, **21 (4)**, 530–548.

Wilby, R. L. and T. M. L. Wigley, 2000: Precipitation predictors for downscaling: Observed and general circulation model relationships. *International Journal of Climatology*, **20 (6)**, 641–661.

Wilby, R. L., T. M. L. Wigley, D. Conway, P. D. Jones, B. C. Hewitson, J. Main, and D. S. Wilks, 1998: Statistical downscaling of general circulation model output: A comparison of methods. *Water Resources Research*, **34 (11)**, 2995–3008.

Wilks, D. S., 1992: Adapting Stochastic Weather Generation Algorithms for Climate Change Studies. *Climatic Change*, **22 (1)**, 67–84.

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences.* 2d ed., Elsevier, London, 627pp.

Wilks, D. S. and R. L. Wilby, 1999: The weather generation game: a review of stochastic weather models. *Progress in Physical Geography*, **23 (3)**, 329–357.

Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62 (1-3)**, 189–216.

Woodhouse, C. A., 1997: Winter climate and atmospheric circulation patterns in the Sonoran Desert region, USA. *International Journal of Climatology*, **17 (8)**, 859–873.

Wu, R. G. and S. P. Xie, 2003: On equatorial Pacific surface wind changes around 1977: NCEP-NCAR reanalysis versus COADS observations. *Journal of Climate*, **16 (1)**, 167–173.

Xie, P. P. and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bulletin of the American Meteorological Society*, **78 (11)**, 2539–2558.

Yang, C., R. E. Chandler, V. S. Isham, and H. S. Wheater, 2005: Spatial-temporal rainfall simulation using generalized linear models. *Water Resources Research*, **41 (11)**, doi:{10.1029/2004WR003739}.

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–U4, doi:{10.1038/nature06025}.

Zorita, E., J. P. Hughes, D. P. Lettemaier, and H. von Storch, 1995: Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation. *Journal of Climate*, **8 (5, Part 1)**, 1023–1042.

Zorita, E. and H. von Storch, 1999: The analog method as a simple statistical downscaling technique: Comparison with more complicated methods. *Journal of Climate*, **12 (8)**, 2474–2489.

Zou, X., I. M. Navon, and F. X. Ledimet, 1992: An optimal nudging data assimilation scheme using parameter-estimation. *Quarterly Journal of the Royal Meteorological Society*, **118 (508)**, 1163–1186.

Zwiers, F. W. and H. Von Storch, 2004: On the role of statistics in climate research. *International Journal of Climatology*, **24 (6)**, 665–680.