APPROACHES TO QUANTIFYING AND REDUCING UNCERTAINTY IN GCMS OVER SOUTHERN AFRICA

Suzanne Carter

Thesis presented for the degree

DOCTOR OF PHILOSOPHY

in the Department of Environmental and Geographical Science, Science Faculty

UNIVERSITY OF CAPE TOWN

August 2008

PLAGARISM DECLARATION

I declare that this dissertation is my own, unaided work,
and that is has not been submitted previously
as a dissertation or thesis for any degree
at any other University

Name: Suzanne Carter

Signed: Signed by candidate Signature removed
Date: January 2009

PREFACE

Climate change modelling using General Circulation Models (GCMs) is at present the most widely used tool for analysing what will happen to the Earth under future warming scenarios. The high dependency on these models requires that rigorous model validation is done to ensure that they are simulating a realistic baseline climate. However, model validation is often focused on the areas where the model was developed and is not necessarily as skilful in other regions of the world. As there are no GCMs developed on the African continent, a technique of evaluating model bias and potentially correcting for this bias is desirable. In the past this has been done on a regional scale – producing a single value of model suitability for the entire region. In this work a sub regional scale analysis is conducted with the aim of identifying synoptic level biases and generating a baseline mean from which models can be compared for skill.

Chapter 1 provides an overview of the state of the knowledge regarding climate change science, with particular focus on modelling future climate. In addition it summarises currently used techniques to address model uncertainty, providing details as to why this study is novel in its approach and how it tries to address issues that are not adequately managed in previous studies. **Chapter 2** relates the details of the methodology and provides information on the data used in the study.

The diagnostic analysis in **Chapter 3** identifies sources of model bias relating to frequency and sequencing of synoptic dynamics for each model as compared to observed climate. The individual synoptic states are then assessed in **Chapter 4** whereby each synoptic state is compared between a model's control run and observations. This method of validating the model data has been used in previous studies.

Chapter 5 generates a weight for each model based upon the overall frequency difference between the model's control and the observed data. The weight is used to modify each model individually and to produce a multi-model solution for control and future projections. The technique is advanced by adjusting the baseline to the observed pattern as well as weighting the climate change anomaly to produce a multi-model corrected future projection. **Chapter 6** applies a similar method to produce a multi model solution for each synoptic state by weighting the climate change anomaly and adjusting the baseline.

Lastly, **Chapter 7** summarises the finding of this study, outlining shortcomings and caveats as well as providing recommendations for future work on this subject. Due to the large

number of figures generated in this study, only some are included in the main text. Most of the rest are included in the four appendices, one for each of the analysis chapters.

Acknowledgements

I wish to thank all my colleagues at the Climate Systems Analysis Group for answering many 'little' questions, brainstorming new ideas and generally providing a positive research environment. In particular I wish to thank **Jeremy Main** and **Chris Jack** for helping me come to grips with C programming.

I am also grateful for the financial support provided by the **National Research Foundation** and the **Climate Systems Analysis Group** during my studies.

I would like to thank my supervisor **Bruce Hewitson** for suggesting that I continue studying climatology, for the many opportunities he has given me over the years to travel and take part in collaborative research projects and finally for always giving encouragement when it was most needed.

Finally, to all my friends and family who have been patiently waiting for the day that they can introduce me as "Dr Carter", many thanks for all your encouragement, understanding and love.

T*he most exciting phrase to hear in science, the one that heralds new discoveries, is not Eureka! (I found it!) but rather, "hmm.... that's funny...."* Isaac Asimov

ABSTRACT

Approaches to quantifying and reducing uncertainty in GCMs over southern Africa

Global Circulation Models (GCMs) are the primary tool for simulating future climate changes. These models by necessity make use of various assumptions and simplifications due to computational constraints, and in so doing introduce biases and systematic error. Along with other sources of uncertainty regarding our understanding of the climate system and given the quasi-chaotic nature of the climate, climate projections differ between models whose climate simulation skill is poorly quantified.

A new methodology is presented to assess the regional biases in GCMs and to, in part, compensate for some aspects of these biases. The study will focus on the Southern African region but could be replicated for other regions. Using Self-Organising Maps (SOMs), synoptic archetypal patterns are identified and the distribution and frequency of these patterns assessed. The use of synoptic archetypes to quantify model metrics presents a novel approach with many benefits over standard metrics, such as errors and means per variable. SOMs add a spatial and multi-variable dimension to the analysis as each metric is calculated based on its synoptic circulation pattern and associated to a **set** of atmospheric variables.

Comparison between the control runs of three GCMs and NCEP reanalysis (as a proxy for current observation) data quantify the model biases in simulating daily circulation dynamics. From this analysis a weighting structure for daily data was devised, which uses the difference between the model and observed frequency of occurrence as a metric for weighting. The weight is applied to control and future model projections for these GCMs at both an aggregate scale (the multi-model mean) and for specific synoptic states. The model weights are a measure of the realism of the various synoptic patterns, apportioning greater weights to the patterns that were best reproduced in the model.

Some persistent biases in the models are notable based on comparisons between the NCEP and GCM SOM node mapping, such as an overall cool bias in the models and a shift of the dominant high pressure cells and thus the westerly wave to the south. The weighting techniques provide insight into how much of the model bias is contributed by differences in synoptic frequency and what part is attributable to systematic biases in the models which result in a different mean state for a given synoptic process. The frequency correction enabled a correction of up to 25% of the difference between model and reanalysis data, but

in most cases the change was far smaller than this. The differences in mean conditions remained the largest component of the bias. To correct for this the weighting was applied to the climate change anomaly (difference between future and control projections) per synoptic process to create a multi-model climate change component that is added to the NCEP baseline. This provides the most accurate depiction of future climate from the data provided.

The models generally have different strengths, therefore the weighted multi-model solution allows models to give a greater contribution where they are skilful and less where they do not match the observed dynamics. Comparison of the magnitude of the climate change signal showed that winter states in the weighted multi-model composite had a smaller temperature increase and reduced rainfall compared to the unweighted results. In summer states there is greater warming and increased rainfall, especially over the oceans. This suggests the models are over estimating changes in temperature in winter and underestimating the increases in summer. Synoptic events are the primary driver of climate change impacts. Therefore errors in synoptic state will have a notable influence on the climate change projections and need to be fully considered in any climate change impact study. The use of the weighting technique helped to identify and reduce uncertainties in the climate change projections over Southern Africa.

CONTENTS

CHAPTER ONE

Introduction and Literature Review

1. INTRODUCTION

Climate change is widely considered one of the key challenges of this century. Global temperatures are projected to rise over the next century due to the high concentrations of carbon dioxide and other greenhouse gases present in the atmosphere. Since the industrial revolution in the late 1800s, humans have been adding unprecedented quantities of carbon dioxide into the atmosphere by burning fossil fuels to generate power and run intensive industries. Already there has been an increase in mean global temperature of 0.74°C from 1906 – 2005 (Trenberth *et al*., 2007). It is now almost certain that these anthropogenic emissions of greenhouse gases have changed the composition of the atmosphere and that the effects of these changes are unavoidable (IPCC, 2007).

Global warming is a result of the increased efficiency of the greenhouse effect. This natural process keeps the Earth warmer than would be expected merely from direct solar energy. Carbon dioxide is a long lived greenhouse gas and a key greenhouse gas emitted by anthropogenic use of fossil fuels. Natural recycling mechanisms are currently able to absorb half of the emissions produced annually over 30 years, with the next 30% removed within a few centuries and the remaining 20% residing in the atmosphere for many thousands of years (Denman *et al*., 2007). There are different response times for different anthropogenic forcings. For example oceans have much longer residence times than the atmosphere resulting in the persistence of warming in the climate system, even without new emissions it will be impossible to avoid some climate changes.

Methane is another greenhouse gas that has a much shorter residence time in the atmosphere but has a much larger global warming potential, making it the second most important greenhouse gas. Large amounts of methane are released from intensive agriculture production, especially from rice and cattle farming. Currently six greenhouse gases are regulated under the Kyoto Protocol, all of which human activity has increased the production of in the last century. However, not all major emitting countries are taking on these targets; therefore the current reduction targets for 2008-2012 are relatively low given the reductions needed according to the latest science predictions (IPCC, 2007).

As we produce more and more carbon dioxide and other greenhouse gases the Earth's greenhouse effect is increasingly leading to the phenomenon referred to as global warming. This additional warming changes the heat distribution in the atmosphere and oceans and

results in large scale changes to weather systems, leading to impacts such as rising sea levels due to thermal expansion, melting of glaciers and ice sheets due to raised ambient temperatures, and increasingly warmer annual mean temperatures. However, the regional impacts are varied for different climate variables (e.g. rainfall, cloud cover, winds). Climate is so much a part of our daily lives that it is often difficult to communicate the enormity of the implications of seemingly slight variations in climate parameters. A single degree increase in global mean annual temperature has far reaching regional implications for almost all sectors of the economy (Stern Review, 2006). Examples include increased heat stress for organic life, changing distribution of vector diseases such as malaria, reduced rains, shifts in agricultural suitability, and rising sea levels due to melting glaciers, which can affect coastal zones. The term climate change is used to describe all of these processes related to global warming.

The projected unavoidable impacts of a changed climate will have some impact on almost every aspect of life, and in the first instance require adaptation measures. The projected impacts include: rising temperatures with greater warming at the poles; rising sea levels worldwide; changes in weather patterns; changes in amount, seasonality and intensity of rainfall; increased likelihood of extreme weather events; and loss of plant and animal species as habitats are altered (IPCC, 2007). Urgent action is required to curb further emissions that may lead to dangerous climate change. Nonetheless, while a clear case exists for mitigation efforts to reduce greenhouse gasses in the long term, the regional scale uncertainty of the projected change constrains the global and regional adaptation responses.

Recent events in international politics have highlighted the overarching importance of taking measured responses in both mitigation and adaptation, with numerous meetings between 2007 and 2009, taking up the issue of climate change at the highest levels of government. The adoption of the Fourth Assessment Report (AR4) from the Intergovernmental Panel on Climate Change (IPCC) in 2007 marked the acceptance by all governments that climate change is being driven by human activity and that countries need to act in order for the most dangerous climate impacts to be avoided. Notable recognition of this step came from the Noble peace prize jointly awarded to the IPCC and Al Gore *"for their efforts to build up and disseminate greater knowledge about man-made climate change, and to lay the foundations for the measures that are needed to counteract such change*[1]*".*

In this context, it is the regional projections of climate change that have become increasingly important (Christensen *et al*., 2007) in order for regional response measures to be

[1] http://nobelprize.org/nobel_prizes/peace/laureates/2007/press.html

effectively identified and executed. To this end, this thesis explores a technique that seeks to reduce inconsistencies or biases of projected regional climate change in a region and provide a first order correction to obtain more robust information.

1.1 REGIONAL PROJECTIONS AND UNCERTAINTY

Almost all regional projections are derived from simulations by Global Circulation Models (GCMs) which simulate the climate interactions using a set of physics equations describing the basic laws of motion and thermodynamics. These models have been developed exclusively in developed countries which have the resources to conduct such time and computing intensive projects. At present there are roughly 25 groups around the world developing coupled Atmosphere and Ocean GCMs. Models are run using various emission scenarios, the most widely used being the IPCC Special Report on Emission Scenarios (SRES) which outline different future emission pathways based upon socioeconomic predictions of future world trends (see section 1.3).

In the AR4, according to various climate model projections and scenarios, global mean temperature will increase between 1.1-6.4°C by the year 2100 relative to 1980-1999 (IPCC, 2007). Sea level rise is estimated at 18-59cm from the lowest to highest ranges (Meehl *et al*., 2007). Local effects can be greater as this is a global average. Even with the large differences between the concentrations of the various greenhouse gases in these scenarios, the end results of sea level change and temperature change remain relatively similar in the near term. This is in part due to the inertia of the climate system in response to the added greenhouse gases of global warming. From the AR4 results, the projections for 2020-2029 show it is very probable that there will be an increase in global temperature of ~0.6°C above 1980-1999 levels (Meehl et al., 2007), with little difference between scenarios. Over Southern Africa the estimates are a little higher, ranging from 0.5-1.5°C depending on location, with coastal projections lower than over interior. The AR4 was particularly careful in communicating uncertainty, depicting both the level of confidence in scientific understanding and the likelihoods of specific occurrences.

The large spread of possible warming has made it difficult for policy makers to commit to adaptation strategies. Further efforts to quantify local responses are required to facilitate decision-making processes by national governments. Choices made now regarding mitigation of greenhouse gas emissions will not have an immediate effect on stabilising temperatures, but will produce benefits by the end of the century and make significant differences to the total warming experienced. Although the changes in concentrations, radiative forcing and sea level rise will be similar globally, the impacts will differ regionally dependent on local vulnerability and coping capacity to these changes.

From the IPCC Third Assessment Report (IPCC, 2001) results from regional future projections over Southern Africa showed greater than average warming but of inconsistent magnitudes and for rainfall small decreases with model results conflicting not only in terms of magnitude but in direction of change (Giorgi et al., 2001). In the IPCC AR4 (IPCC, 2007) the Southern African regional multi-model projections for 2080-2099 showed the median temperature increase lies between 3°C and 4°C in all seasons (Christensen *et al*., 2007). For precipitation projections, robust results indicate that there is likely to be drying during the winter period, especially in the south west. The south west is the only part of the subcontinent that receives winter rainfall, however, local orographic forcings may result in different changes at the local scale (Christensen et al., 2007; Midgley et al., 2005).

The reason for the large envelope of future model projection results stems from the numerous uncertainties pertaining to climate modelling and the use of different GCMs and scenarios. Uncertainties result from:

- wide range of possible future world trends in population, economy, technology, energy, agriculture and land use (largely compensated for by using multiple scenarios);
- difficulties in projecting gas emissions, concentrations and radiative forcing due to incomplete knowledge of the climate system interactions, especially feedbacks;
- difficulty in modelling climate responses to changes in radiative forcing;
- uncertainty in model parameterizations;
- natural variability;
- incomplete knowledge of physical processes;
- converting model responses into suitable material for an impact study.

Another consideration is the uncertainty generated from the models themselves and will be elaborated on in the next section.

1.2 GENERAL CIRCULATION MODEL DYNAMICS

The GCMs used in this analysis are coupled atmosphere-ocean models (or AOGCMs). They have two primary model components; one represents the atmosphere and the other the ocean (along with additional sub-models for ice, land surfaces processes, etc). The transfers of heat, momentum and matter connect these components. GCMs are three dimensional representations of the earth's natural systems that are run at reduced complexity in order to run simulations within a reasonable time period. The spatial resolution of GCMs is too coarse to resolve all processes, and thus requires the use of parameterizations – approximations to represent sub-grid-scale processes. GCMs are built upon physics equations, which are in turn based upon the general laws of conservation of

mass, momentum and energy and the ideal gas law (McGuffie and Henderson-Sellers, 1997) to represent the behaviour of the atmosphere and other components of the Earth system. Parameterizations schemes use simplified physics equations which are needed to represent the processes at the temporal and spatial resolution required for the purpose of the model (Peixoto and Oort, 1992).

Each model has different combinations of parameterizations and has differing degrees of complexity in representing the various sub components such as vegetation cover, ice and snow cover and topography. For example the sea ice component is represented in most models by a thermodynamic ice model often with 'free drift' dynamics (Randall *et al*., 2007). Most models have categories of sea ice thickness; however, terrestrial snow processes still vary widely. Additionally the choice of which aerosols (if any) are represented in the model will influence the cooling effects in the model. GCMs also differ depending upon the use of various modelling techniques to represent processes, the choice of grid spacing and boundary conditions, available computing power and the assumptions made about future conditions (Cracknell, 2001; Randall *et al*., 2007). Other differences between the models include their ability to capture modes of variability and the expression of natural variability cycles such as the El Nino Southern Oscillation.

The climate sensitivity of the model is a measure of the amount of global average surface warming following a doubling of carbon dioxide concentrations and this varies between models from 2 to 4.5°C in the models assessed in the AR4 (Randall *et al*., 2007). Feedbacks, especially water vapour feedbacks, were identified as contributing three times more uncertainty than radiative forcing to the differences in climate sensitivity (Randall *et al*., 2007). Feedbacks are therefore an important area of further research but are not addressed further in this work.

Stationarity is another issue to consider, i.e. whether statistical relationships are valid under future climate regimes. This assumption could be a source of error as in the future dynamics may not function in the same way they have in the observed record. Non-stationarity has been shown to result in an underestimation of the magnitude of the change (Hewitson and Crane, 2006).

As climate models are formulated similarly, there is a worry that the reasons for the models producing similar results has less to do with the models converging on the true answer and more to do with the similarities in the experiment design.

As a result of these and other uncertainties, the models give a spread of results in their climate projections. Some of this spread is structural (how they represent processes), some is error (some of their processes are wrong), and some due to natural variability. This results in an envelope of uncertainty, which is in part due to the range of future possibilities and in part due to model error. To differentiate between the types of bias, this thesis refers to all structural biases as process specific bias and where the models are wrong as persistent bias, which is independent of synoptic process, e.g. in general, models tend to produce too many days with weak precipitation and too little precipitation overall in intense events (Christensen *et al*., 2007).

1.3 SRES SCENARIOS – CAPTURING A RANGE OF POSSIBILITIES

The IPCC commissioned a report on emission scenarios in 1996. The final product was published as the Special Report on Emission Scenarios (Nakic´enovic´ *et al*., 2000). The scenarios are modelled on socio-political projections of possible future trends, based upon a number of expert opinions and from the scientific literature. Scientists from many nations put a wide range of scenarios forward. There are 35 official SRES scenarios (IPCC, 2001) all of which follow the four marker storyline assumptions, namely A1, A2, B1 and B2, to different degrees. Each of them represents different combinations of various demographic, social, economic, technical and environmental developments in the future. The driving forces of population growth, economic growth, technological advances, energy use, land use and agricultural practices affect each of the scenarios to different degrees.

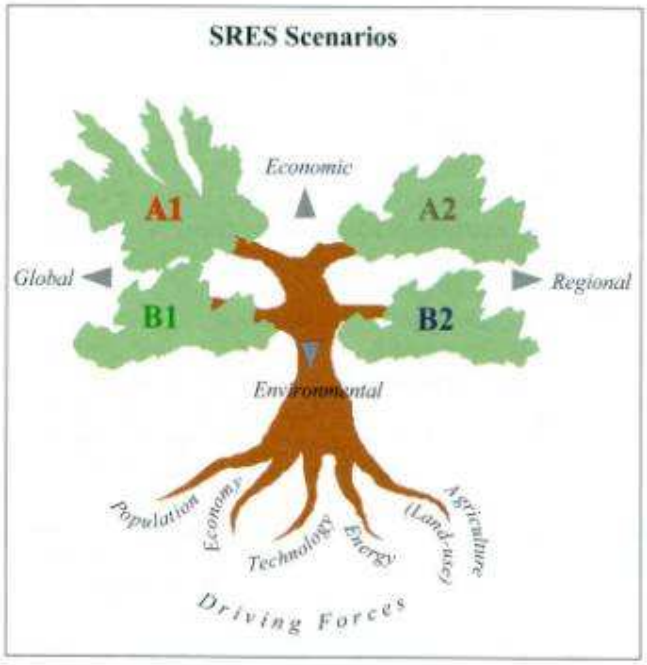


Figure 1.1: SRES scenarios family tree diagram, outlining the driving forces which in different combinations result in the four marker scenarios (from: IPCC Special Report on Emission Scenarios, 2000)

There are two principal axes along which the four marker scenarios are orientated – these are Economic/Environmental and Global/Regional. Figure 1.1 shows this diagrammatically. All four storylines start from the same initial conditions and gradually diverge. From the four marker storylines 6 general scenarios are derived. There is one each for A2, B1 and B2, and another 3 scenarios are produced through sub-setting the A1 storyline to consider 3 different energy sources (Fossil fuels (A1FI), non-fossil fuels (A1T) or a balance between the two (A1B)). The values for the various forcings and projected values of certain variables, for example radiative forcing and emissions, were produced by expert analysis of various types of projection models.

Uncertainties identified in the SRES scenarios are: choice of storyline parameter combinations, author's interpretation of storylines, links between driving forces and qualitative inputs, methodological differences, differing source data and inherent uncertainties (Nakic´enovic´ *et al*., 2000).

All scenarios are meant to have equal probability; however, the scenarios A2 and B2 have been favoured in current research. The Data Distribution Centre (DDC[2]) of the IPCC (Intergovernmental Panel on Climate Change) is the primary dissemination body for data of GCM simulations based on SRES forcing. At the time of data collection for this study, most GCMs only had A2 and B2 forced simulations available. This is due to the high computational demands required to run each model/scenario combination. To minimise costs the two middle-of-the-road scenarios were used by most modelling groups. Hence, only the A2 and B2 scenarios have been widely utilised. More recently the IPCC has shifted to the use of A2 and B1 scenarios, which are available through the Program for Climate Model Diagnosis and Intercomparison (PCMDI)[3]. For sea level rise and temperature change the A2 scenario has larger increases than the B2 scenario.

In the near term projections for the next 40 years, the choice of emission pathways does not alter the model response much (IPCC, 2007) but has a larger impact at the end of 21st century. Given this, only one of the emission scenarios is used in this study, SRES A2, as this scenario was universally available for all the models used in this analysis and the time period of interest was mid 21st century projections.

1.4 STATISTICAL MEASURES OF CLIMATE CHANGE UNCERTAINTY

There has been some headway into quantifying the range of uncertainty, if not yet in outright reducing it. The needs of decision and policy makers for regional assessments with the lowest uncertainty for evaluating local options for both climate change mitigation and

[2] http://www.ipcc-data.org/
[3] http://www-pcmdi.llnl.gov/

adaptation, has resulted in much effort being devoted to try and provide greater certainty. However, there is a range of opinion on how best to go about assessing and reducing uncertainty.

Some would argue that there is not sufficient accuracy in GCM simulations to match observed climate to allow for the weighting of individual models to produce a multi-model ensemble and because the models cannot be meaningfully calibrated (as they are simulating a never before experienced state of the climate system). *"The lack of any ability to produce useful model weights, and even to define the space of future possible models, rules out the possibility of producing meaningful PDFs for future climate based simply on combining results from multi-model or perturbed physics ensembles, or emulators thereof* (Stainforth *et al*., 2007)*."* As a result, they treat the current ensemble of climate projections as an envelope which provides the lower bound on the maximum range of uncertainty, which they argue is still a useful guide to decision and policy makers and does not result in misleading projections which may later be proved invalid.

In contrast to this perspective, others are choosing to assign probabilities to the range of climate model projections (e.g. UKCIP08, 2008[4]; Greene et al., 2005; Tebaldi et al., 2005) and seek to derive probability density functions (PDFs) based on a measure of skill for the different models. Underlying this approach is the assumption that model simulations may be ascribed a weighting factor based on a metric of performance to create a multi-model ensemble solution. Typically, however, these approaches use metrics based on time and space averaged values, and do not address the question of skill on daily and synoptic time scales – the scales at which local climate is directly forced and which averaging may readily obscure. This question of the synoptic daily time scale is at the heart of this thesis; exploring model weighting using a metric based on the frequency of daily synoptic processes.

Below is a review of current tools being employed to address issues of GCM uncertainty. These range from using single model projections to a large envelope of results from many models to complex probability based approaches.

The most simplistic approach is to take the output of one or more climate models that are shown to be indicative of current climate and use this models future projections as the most realistic simulation of future climate (UKCIP02, 2002[5] and (Hayhoe et al., 2006). The problem with this method is that it does not capture the range of climate sensitivities and is therefore unlikely to generate the most complete assessment. There is also the danger of

[4] http://www.ukcip.org.uk/index.php?option=com_content&task=view&id=163&Itemid=293
[5] http://www.ukcip.org.uk/index.php?option=com_content&task=view&id=161&Itemid=291

providing single values to policy makers which they take as being the absolute answer and can result in strategies that are not robust to the full range of climate possibilities.

(New and Hulme, 2000) identified a simplistic method, referred to as a *Monte Carlo approach*, for scenario uncertainties. It postulates the idea that the wider the range of models, scenarios and ensemble members, the better the range of uncertainty is captured. Therefore, the end product is a wide 'envelope' of results. The key to this approach is to consider as wide a range of possibilities as is feasible.

Others have taken ensemble analysis further, weighting models according to 'model credibility'. This measure compares model ensembles, assessed based on how well each model realistically reproduced selected statistics of historical climate (Brekke et al., 2008). Comparison of model bias, based on error, means and other metrics have tried to identify optimal subsets of variables for specific applications (Glecker et al., 2008).

Currently, the UK Met Office's Hadley GCM is being used to extrapolate a range of future climates by running a large range of model parameterisations and scenarios. The rationale behind this is that for many variables *"the main uncertainty in multi-decade climate prediction is not in the initial state nor in the external driving, but in the climate system response* (Allen and Stainforth, 2002)*."* To address this particular aspect of model uncertainty a large number of simulations are required. The project is run through www.climateprediction.net, making use of personal computers during their dormant periods from volunteers from the general public, in a similar fashion to the SETI project[6]. Each computer runs a separate model version (set of parameterisations and initial conditions) to produce an ensemble of ensembles. Roughly 290 000 simulations had already been completed in June 2008. This wide envelope of ensembles is constrained by the fact that the Hadley climate model (HadSM3) is not a perfect simulator of current climate. Therefore whatever results are derived will carry through the biases of that particular model. On the other hand it does provide invaluable information on what set of parameterisations should be combined to get the most accurate results. When assessing the climate sensitivity of each of the model versions, they found a wide range of response, from 2-11°K, and argued that the wide range of possible behaviour within a GCM show that high sensitivities cannot be ignored (Stainforth *et al*., 2005).

Giorgi and Mearns (2002) qualitatively measure uncertainty using their *Reliability Ensemble Averaging* (REA) methodology that looks at regional monthly means and calculates for each region: the average, the uncertainty range and two reliability indices:

[6] http://setiathome.berkeley.edu/

- *Performance criterion* - how well the model simulates current climate (or model bias). Models that represent current climate well are given higher reliability.
- *Convergence criterion* – measures the convergence of simulations from different models giving higher reliability to those with greater convergence. This assumes that greater convergence of models is indicative of a better model. This assumption is not necessarily true as outlier projections may be more likely in a future climate if the climate system dynamics were to change.

It was seen that the performance criterion had consistently lower reliability than the convergence criterion, making convergence a more dominant determinant of skill. Variables analysed were temperature and precipitation. The uncertainty could not be reduced by eliminating poorer models, as there was "*no model that ubiquitously performed better or worse than all the others… over all regions* (Giorgi and Mearns, 2002)." Over Southern Africa it was seen that increases in temperature from the REA analysis showed increases in both A2 and B2 scenarios of roughly 2-5ºC and indicated some drying. However, there was little agreement in precipitation changes especially in summer (DJF). Another consideration is that it used monthly averages only and therefore temporal resolution was poor. Although the results gave some broad scale indications of regional reliability of models, sub regional spatial differences and gradients were not explored.

Statistical Science has been using Bayes theorem to explain probability for more than 100 years. Many statistical climatologists have been using Bayesian methods to explore the uncertainty related to climate change projections (Berliner *et al*., 2000; Tebaldi *et al*., 2005). There have been limitations to these methods due to the high dimensionality of climate data and arguments over the choice of prior and posterior distributions which make objective assessment difficult. More recently the REA method has been extended, using a Bayesian approach (Tebaldi *et al*., 2005), which they found to be a useful method in synthesising the information from an ensemble of simulations. In addition, due to the flexible nature of any Bayesian formulation (choice of prior and posterior distributions) the method allows for the exploration of the assumptions that generate probabilistic forecasts.

Multi-model ensemble forecasts are used by (Krishnamurti et al., 2000) to weight model ensembles using a simple linear regression. Regression (using least squares) analysis is performed for a training period of current climate for all models. There has been a lot of debate as to the merits of this technique. It seems to have gained some popularity as a methodology as it was used in two papers presented at 9th International Meeting of Statistical Climatologists in 2005.

Another approach taken by (Stott et al., 2006) is to apply a regional scaling technique to model predictions, according to their ability to represent decadal mean observed trends for

the 20th century in the region of interest. The scaled values produce a probabilistic prediction of continental-scale temperature changes. The method differs from other probabilistic methods in that it uses observed trends attributable to past anthropogenic influence on the climate system to constrain the forecast instead of simulations of climatology.

The UK Climate Impacts Programme has recently released a high resolution (25 x 25 km) probabilistic climate projections for the UK, in December 2008. In total 31 parameters were varied to provide around 300 model versions, creating a large ensemble of projections of future global climate, all using the Hadley Centre climate model (HadCM3). The model versions were then weighted according to their ability to simulate observed climate, which makes the assumption that those models that are best at simulating observed climate will also be better at simulating future climate. Finally, single model results from other IPCC climate models were incorporated to better reflect the modelling uncertainty associated with future climate projections. Using this method, probabilistic climate projections were produced for three future greenhouse gas emissions scenarios. Other products include a weather generator and marine projections, in addition to observational data sets.

Collectively these approaches reflect the growing need to derive regional information from GCMs, and the different perspectives of approaches being employed. This is an evolving area and is likely to be an increasing priority focus in years to come. This thesis seeks to contribute to the work by exploring an alternative approach using daily resolution data from the GCMs.

1.5 THESIS OBJECTIVES

In this work, uncertainties relating to the GCM spread of projected climate are explored using the daily temporal resolution of synoptic patterns and a method based on Self Organizing Maps (SOMs) to produce a weighted multi-model solution. SOMs were developed more than a decade ago (Kohonen *et al*., 1995) as a variant of artificial neural nets, and are the principle tool used in this analysis, extending methods presented by Hewitson and Crane (2002). The SOM is a pattern recognition tool which characterises data into a predetermined number of states. The first application of SOMs to atmospheric data was by (Main, 1997). Since then, SOMs have proved valuable in synoptic circulation pattern mapping and their use in validating GCM projections has been proven to be quite robust. (Hope et al., 2006) used SOMs to link synoptic states with observed rainfall characteristics, which were then used in a subsequent paper as a basis for future projections of rainfall based on future GCM synoptic states (Hope, 2006). They have also been used in the study of climate change and variability (Cavazos, 1999; 2000; Cavazos *et al*., 2002; Hewitson and

Crane, 2002; Tennant and Hewitson, 2002; Tennant and Reason, 2005; Tadross *et al*., 2006; Thomas *et al*., 2007), identification of climatic regions (Malgrem and Winter, 1999) and both up- and downscaling of atmospheric data (Crane and Hewitson, 2003; Gutiérrez *et al*., 2005; Hewitson and Crane, 2006).

Most of the methodologies described in section 1.4 were based upon monthly average climate data and often assign a single value for an entire region. Moreover, the variables used are usually the derivative attributes of the underlying circulation forcing, and tend to focus on surface air temperature and precipitation. In contrast, this methodology chooses instead to examine daily data. Additionally, variables used reflect the synoptic patterns of the large scale circulation features, which are then subsequently related to the more societally relevant variables of temperature and precipitation.

By looking at daily data from GCMs, it is possible to determine the inherent biases that these models have in simulating synoptic scale climate. A selected number of GCM simulations of the present climate is compared with "observed" historical circulation as represented by the National Centre for Environmental Prediction (NCEP) reanalysis data (Kalnay *et al*., 1996) to identify the circulation modes that the models simulate with greater or lesser skill, and so in turn, help to determine the model biases. The methodology is fully described in Chapter 2.

Biases in frequency of synoptic modes are identified in Chapter 3, according to three indices: the average mapping error for each synoptic state, the frequency at which the model simulates each synoptic state, and the temporal evolution and transitions between states. A diagnostic analysis of the climate models and further determination of regional model bias will be covered in Chapter 4. A weighting structure, determined by frequency bias, is then applied to the daily results, dependent on atmospheric state, for the mean in Chapter 5 and per synoptic state in 6. Combining synoptic specific future weighted model projections is shown to result in a better forecast of future climate than using an individual model. The forecast is 'better' because a measure of noise has been removed from the information, improving the understanding of the forecast. However this does not necessarily improve predictive accuracy.

The following research questions guide the study:

1. *How can synoptic scale circulation, as an integrator of processes, be used to identify synoptic specific and persistent model bias?*
2. *What synoptic features do the models fail to capture and what features do they simulate accurately?*

3. *How does the projected regional climate change when using a metric based on the realism of the synoptic circulation, and does this reduce the uncertainty envelope of future regional projections?*

These questions address gaps in current scientific knowledge regarding the nature of model bias at a sub regional scale, provides insight into the dynamic basis for the biases identified, and distinguishes between persistent biases (those seen in all models and in all states) and those inherent to a specific synoptic process or model. This study is a first step in finding simple bias correction techniques for GCMs at the daily time scale and as such there will be areas that can be expanded and elaborated in future work.

CHAPTER TWO

Data and Methodology

As outlined in Chapter One, Global Circulation Models (GCMs) provide a reduced complexity simulation of the physical world. As such, GCMs have inherent bias due to various limitations in our understanding of the physics of the atmosphere and the restrictions imposed by a computer's capacity to solve the physics equations describing these processes in a reasonable time frame. The methodology developed in this thesis attempts to apply a bias adjustment in a regionally specific manner.

2.1 INTRODUCTION TO SELF-ORGANIZING MAPS

Synoptic climatology studies have commonly sought to reduce data to more manageable proportions by applying a synoptic classification tool that assigns data into different categories, types, or synoptic states. This is particularly useful in process studies to determine the relationships between circulation and local environmental parameters (Crane, 1978; Yarnal, 1984; Wigley and Jones, 1987). Self Organizing Maps (SOMs) present a method of analysing climate change in terms of the pattern and frequency response of weather events. SOMs have a wide range of applications and have proven valuable when used to evaluate and validate climate simulations by GCMs (Main, 1997; Hope, 2006; Hope *et al*., 2006). In this study SOMs are used to assess daily synoptic circulation patterns in the data, which in turn allows for the biases in the model output to be explored, and to develop a means of correcting for some of this.

A SOM is a type of Artificial Neural Net (ANN) developed by (Kohonen, 1995) which can recognise commonly occurring patterns within multi-dimensional data sets. A SOM treats all data inputs as a continuum, making no assumptions about the statistical properties of the data. A SOM will cover the data space of the input data, clustering where there are more data points in the multi-dimensional space. This means that the SOM is able to represent non-linear relationships, which other visually based synoptic classification tools struggle to do. The map is trained in an iterative manner, so that each node or "type" is weighted by the input data that most closely match that node. Each node is also weighted by adjacent nodes in the SOM space, treating the data as a continuum rather that separate classes (Hewitson and Crane, 2002). The SOM arranges the patterns so that the most dissimilar patterns are widely separated and transitional states are found inbetween these extremes. States that are located close together have greater similarity. The SOM provides the user with a 2D representation of the multi-dimensional data distribution.

In atmospheric applications each node represents a different synoptic state (for an individual/group of atmospheric variable/s) and one can find the frequency at which these patterns occur, in which season they are predominant and if there is more than one variable in the data set, which data reflect which type of synoptic states. This last characteristic is what is most useful for validating GCMs. An observed data set is compared to the GCM using the same SOM training. The differences between the observed and GCM are determined for various parameters to give an indication of model bias.

2.2 DATA

In this work, data from the National Centre for Environmental Prediction (NCEP) (Kalnay *et al*., 1996) is used to generate the observed climatology. NCEP data is a GCM reanalysis based upon multiple observed data sources, including weather stations and satellites. This observed data is integrated using a global model to create a gridded data format that is comparable to the output from a GCM. The NCEP data set is only reliable post 1979 (Kalnay *et al*., 1996) due to poor spatial coverage in many areas, especially in the Southern Hemisphere in the pre-satellite era; therefore the period 1979-2000 will be used in this study. NCEP reanalysis data is not true observed data, and will have it's own inherent errors. However, it does present an observationally constrained reference whose temporal sequencing of synoptic events does reflect reality.

The GCMs output used in the analysis are all from three models from the IPCC's Fourth Assessment Report. These models were selected as they all had variables for winds, specific humidity, temperature and precipitation at a given geopotential heights and the surface whereas other models use some sigma levels for winds and geopotential heights for other variables. With the following models all analysis is based on geopotential heights to avoid confusion in interpreting between the two kinds of heights. A brief description of each is given below:

- The **Geophysical Fluid Dynamics Laboratory** CM 2.1 (GFDL) developed at Princeton University(Delworth et al., 2006) and (Gnanadesikan, 2004)
- The **Model for Interdisciplinary Research on Climate** (MIROC) developed by the Centre for Climate System Research (CCSR), National Institute for Environmental Studies (NIES) and Frontier Research Centre for Global Change (FRCGC). The simulation used was the medium resolution run (K1 model developers, 2004)
- **Meteorological Research Institute Coupled GCM** 2.3.2 (MRI CGCM) developed by the Japan Meteorological Agency (Yukimoto et al, 2001)

In Table 2.1 the main model specifications are outlined. GFDL has the lowest resolution of the three and the simplest sea ice and land surface scheme.

Model	Atmospheric	Ocean	Sea Ice	Land
GFDL	Spherical (R30) 2.3 x 3.8 14 levels	1.9 x 2.3 18 levels	Free drift, no leads	Bucket (single layer), routing
MIROC 2 (med res)	2.8 x 2.8 20 levels	0.5-1.4 x 1.4 43 levels	Rheology, leads	Layers, canopy, routing
MRI CGCM	Spherical (T42) 2.8 x 2.8 30 levels	0.5-2.0 x 2.5 23 levels	Free drift, leads	Layers, canopy, routing

Table 2.1: GCM specifications for the model components (adapted from: PCMDI model appraisal, 2005; Randall et al 2007)

GFDL and MRI are spectral models whereas MIROC is a grid point model. MIROC has a much higher resolution ocean model and has a complex sea ice and land surface component. GFDL model is the simplest model of the three and therefore may have some of the largest biases, due to its lower resolution and less complicated feedback cycles. MRI has most of the same feedback complexity seen in MIROC, but with lower oceanic resolution compensated for with more levels in the atmospheric component. All these models are run using the SRES A2 emissions scenario for control (1979-2000) and future (2045-2064).

The data used here are daily averages of lower troposphere winds and specific humidity for the domain 18-42°S and 9-42°E. Figure 2.1 shows the domain, indicating countries, oceans and regions of interest.

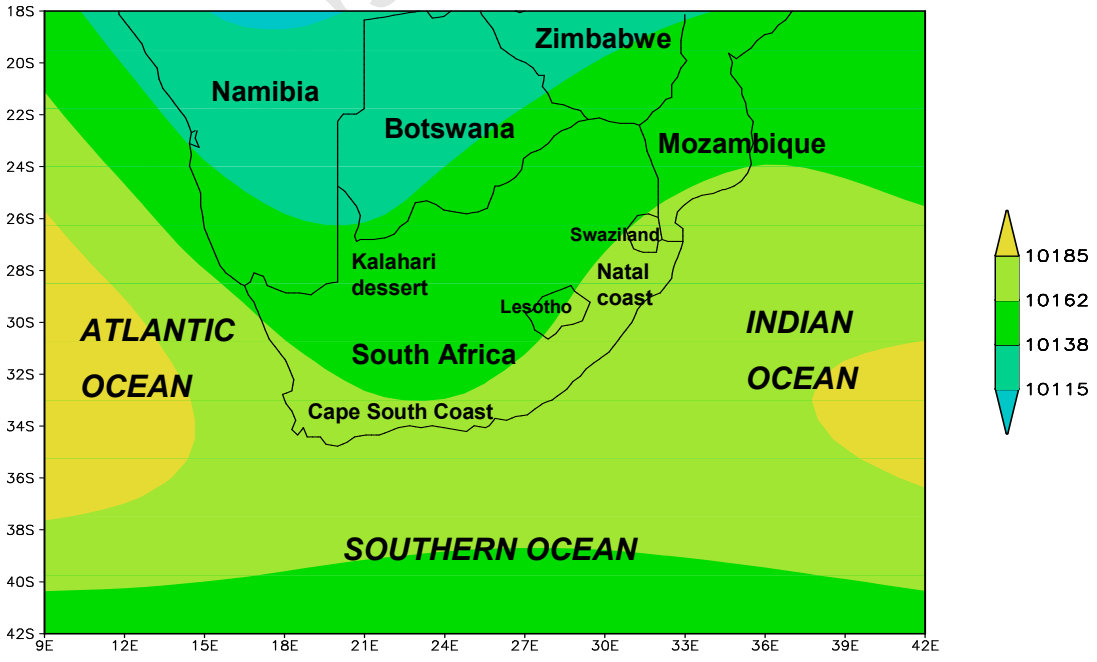


Figure 2.1: Domain used in analysis displaying NCEP (1979-2000) average sea level pressure in colour shading

The input variables used in the study depict lower atmospheric winds at the surface and 700hPa and specific humidity at 700hPa. These inputs are used to produce an array of synoptic patterns that describe present day large scale synoptic circulation.

All data was re-grided to a common 3°x3° (latitude by longitude) grid and normalised (i.e. standardized by subtracting its own mean and dividing by the model's standard deviation). This technique of standardising removes systematic bias and allows models to be more easily compared, by expressing the data as a difference from the models own mean for each variable. All model and NCEP means and standard deviations are presented in Table 2.2.

Model	U wind surface ms^{-1}	V wind surface ms^{-1}	U wind 700 ms^{-1}	V wind 700 ms^{-1}	Specific humidity 700 g/m^3
NCEP mean	0.379	0.781	5.841	-0.520	0.00302
Std dev	4.563	4.350	7.724	4.761	0.00174
GFDL mean	1.039	0.468	7.429	-0.537	0.00269
Std dev	4.409	3.743	8.113	5.005	0.00167
MIROC mean	1.053	0.133	7.584	-0.551	0.00266
Std dev	4.691	3.857	8.919	4.975	0.00193
MRI mean	-0.0043	0.666	5.959	-0.185	0.00258
Std dev	3.760	3.729	7.338	4.664	0.00158

Table 2.2: Means and standard deviations for the input circulation variables, used to standardise the data

From these statistics there are some points worth noting. GFDL and MIROC have larger u wind means on both levels than NCEP, whereas MRI means are closer to NCEP, but the standard deviation is considerably smaller. In the v wind component, MRI is closer in the surface field with GFDL underestimated and MIROC overestimated. All models standard deviations are smaller than NCEP by roughly the same margin. At the upper level, GFDL and MIROC are very close to the NCEP mean while the MRI mean is overestimated. All models have a lower specific humidity mean than NCEP.

2.3 SOM PROCESS

The software package used is SOM_PAK version 3.2 (Kohonen *et al*., 1996). The first operation of the SOM routine defines a random or linear distribution of nodes in the data space. Each node is defined by a reference vector of weighting coefficients, where each coefficient is associated with a particular input variable (Hewitson and Crane, 2002). The SOM reference vectors are adjusted through the iterative process until they span the data space. Each node represents a position approximating the mean of the surrounding data space and is referred to as an archetype. The data entry does not become associated with the group at this stage, but rather it is used to adjust the location of the node in the data space.

The user determines the size of the SOM in the x and y direction (e.g. a 5 x 3 spatial grid results in 15 nodes). As it is an arbitrary choice to a large extent, using a varied number of nodes or classes is advisable, and then choosing the optimal number where there are not too few to result in generalisation, neither too many so as to make the analysis overly complex. The size of the SOM has been shown to have little effect on the results, with only the degree of generalisation affected (Crane and Hewitson, 2003). However, the number of data points that map to each node will decrease with a larger SOM and if there are very few points for a node, the interpretation of the results may become difficult. In this analysis the optimal size SOM was 5 x 3, after testing a range of domains as large as 9 x 7 and as small as 3 x 3. A smaller domain did not produce sufficient variety to differentiate synoptic processes whereas larger domains proved difficult to group as variations became less discernable in adjacent nodes. The chosen SOM domain offered enough variation to pick out the large scale processes clearly.

The geographical extent of the domain was chose to cover southern Africa, including oceans which influence the region but not including the tropics as the synoptic drivers there are distinctly different.

Other user inputs specify the degree of similarity between the reference vector and the input vectors (learning rate) and the influence one node exerts on those around it (radius). In the first training these were kept high while running through the data roughly 10 times whereas in the second training the radius and learning rate are reduced and runs through the data only once or twice. The SOM cannot be overtrained, therefore the input data is presented to the SOM multiple times until there is no change in node assignment (Hewitson and Crane, 2006). The training for this experiment used a linear initialisation (instead of random) with an initial radius of 3 nodes. Random initialisation proved to be less effective than the linear initialisation. This could possibly be due to the large number of input variables used.

The second run, which fine tunes the spatial sequencing, was run for 100 000 iterations with a radius of 1 node. Each data entry was then assigned to the node which most accurately represents that data entry. The error value is a measure of how far the Euclidean distance is between the data entry and the archetype/node it is assigned to.

2.3.1 Creating a SOM for the observed climatology

In this study, NCEP data is used as the input data for the SOM, which depicts 15 (5 X 3) different synoptic states (nodes). The SOM_PAK software reads in the data and by an iterative learning process orders the initially random structure of the data. The SOM reads through the data and determines the similarity (in terms of Euclidean distance) between the data entry and the reference vectors of the randomised nodes. The reference vector of the

best matching node is then modified so as to reduce the difference between input vectors and the learning rate. A second, longer training, with a smaller radius and slower learning rate fine tunes the distribution of the patterns. Patterns that are similar to one another are close together and patterns that are most different will be placed on opposite sides of the SOM space.

The data is arranged in n-dimensional data space, where n is the number of inputs. A Sammon map (figure 2.2) is generated, showing a 2D approximation of the relative Euclidean distances between nodes in the data space. A reasonable Sammon map has a relatively uniform distribution such as in figure 2.2, with all nodes almost equidistant from one another. Unsuitable trainings will have folds in the Sammon maps which would indicate that the same area of data space has been oversampled, leading to too much influence given to a small set of nodes in a certain area and neglect of other areas.

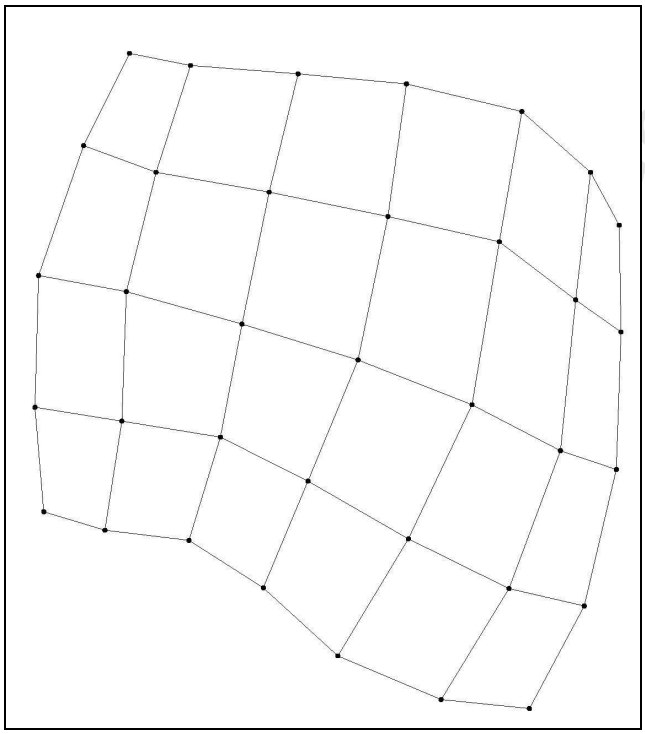


Figure 2.2: Example of a 7x5 SOMs Sammon map with well spaced nodes

Each data entry (in this case a daily average) is assigned to the node that most accurately matches that entry or has the closest reference vectors. The relative frequency (i.e. how many dates were assigned to a node) will show which synoptic states are most common. All frequencies are expressed as percentages of the total data file so that data files of different sizes can be compared. Another valuable statistic produced by the SOM is the error (i.e. how closely the node represents each date mapping to that node on average), which is the

Euclidean distance of each data entry from the reference vectors of the node it is assigned to. Both the frequency and error values will be used in Chapter 3 to determine biases.

Using the node assignment for each date, other variables that were not inputs to the SOM can be analysed. The node assignment for each time step is taken for a new variable and new nodes are generated by averaging all the dates that are assigned to a particular node for this new variable. For example, a mean temperature map can be generated for each node by averaging temperature for all the dates that map to that node. The same technique can be used for any other variable. The results of the diagnostic approach to determine model bias will be further scrutinised in Chapter Three.

2.3.2 Applying GCM data to the SOM

The GCM control and future data for the same fields are then run through the master SOM structure. The SOM field is still based on the observed values of NCEP, the model merely maps its own data entries to the node which most closely represents that date. From the basic statistics of error and frequency, one can compare the differences between the NCEP 'observed reality' and the models simulation of current (control) and future climate. Deviation in frequency will indicate what the model is failing to capture or its 'bias'. A mindmap of the data processing is shown in figure 2.3 below.

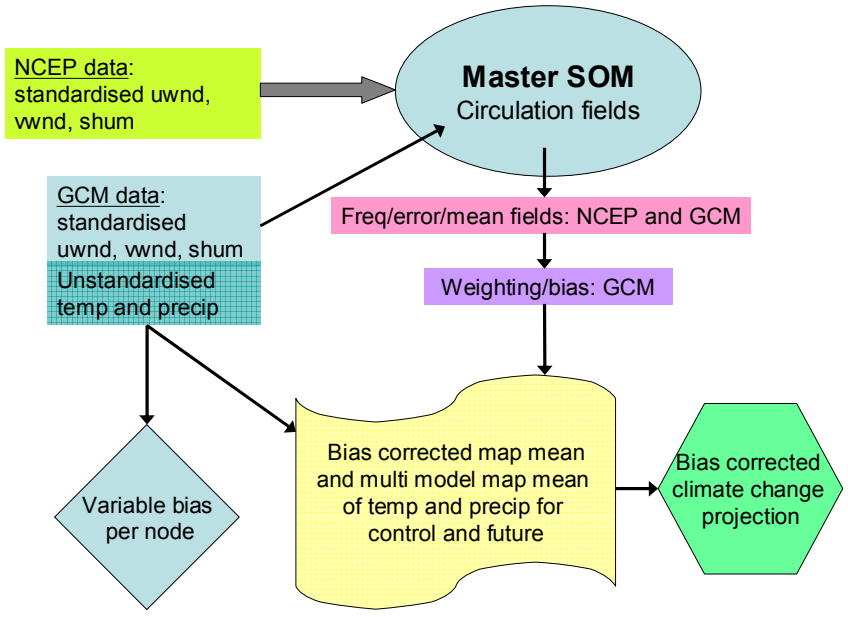


Figure 2.3: Mind map of data processing to weight model outputs

The frequency difference between the GCM control run and the NCEP data is calculated for each node. This statistic can be used to weight the models on the premise that those

models with greater similarity in frequency to the observations are most accurate and that they should be weighted higher than those models with less similarity. Each node map, which represents a synoptic state, is then weighted according to its accuracy in representing the correct frequency of that state.

The same technique is then applied to the future data, using the same weight (model control – NCEP difference over 100). The benefit of using the control weight is that it keeps the climate change signal (difference between control and future) undisturbed while correcting for the frequency bias inherent in the control. The weighting function is described in more detail in section 2.4.

2.4 MULTI-MODEL SOLUTIONS

The weighting of models requires the user to understand the data that is being weighted, in order for the accurate determination of the corrections needed. In Chapter 3, the basic statistics of error and frequency and the sequencing of synoptic states helps to determine how the observed and modelled systems differ from one another. In Chapter 4, the internal dynamics within the node are examined to determine the magnitude and position of the corrections that need to be made. This provides the context for the correction techniques demonstrated in Chapters 5 and 6.

In Chapter 5.1, a first order frequency correction is demonstrated for individual models. A summated mean of each node is produced by taking the observed frequency of NCEP for each node and multiplying this by the mean node value for each model. This simple technique allows for the exploration of how the model means behave given an observed frequency and helps to establish how large an impact frequency differences have on the weighting.

2.4.1 Creating model weights

The second attempt at weighting uses the difference between observed and model frequency as a metric of accuracy. The rationale is that if the models are simply depicting patterns at the wrong frequency (e.g. model has too many cold fronts) then the adjusted control projection should match the observed patterns. For each model, the frequency anomaly between the GCM's control frequency (Cf) and NCEP's frequency (Nf) is made into an absolute value and divided by 100 (to insure that the number is smaller than 1), thereby creating a weight (W) for each node (i) with a value between 0 and 1, shown in Equation 1. If using a SOM of a very large number of nodes (not recommended as it would be difficult to interpret synoptic characterisation) the denominator might need to be increased.

(E.1) $$Wi = 1 - \frac{|Cfi - Nfi|}{100}$$ per model, per node

This is done to eliminate problems with multiplying and dividing by negative weights and allows for the summation of weight totals. Each model will have a different node weight and these can be compared to see which dynamics are well resolved in different models and if there are certain dynamics that are universally well/poorly represented.

In Chapter 5.2, weighting is done on an aggregated level, therefore each of the node frequency anomalies are summated first and then differenced from 1 to determine a single weight (Wm) for the whole model, shown in Equation 2. This technique should help identify overall stronger or weaker models.

(E.2) $$Wi = 1 - \sum_{i=1}^{n} \frac{|Cfi - Nfi|}{100}$$

Each grid cell of the average mean GCM control data for a variable is then multiplied by the model weight. The summation of the models is divided by the summated weights to produces a single map average, shown in Equation 3. The same can be done for the future average ($Pave$). In this work the variables used are temperature and precipitation.

(E.3) $$MutiM = \frac{\sum_{m=1}^{3} Cave \times Wm}{\sum_{m=1}^{3} Wm}$$

To further correct for the systematic bias, in addition to the frequency bias, the difference between future and control map averages for each model are weighted using the model weight (Wm). These are then summated for all three models and divided by the summated model weights. This produces a weighted climate change anomaly (correcting for frequency bias) and is added to the NCEP baseline (thereby correcting for systematic bias) as illustrated in Equation 4.

(E.4) $$Mfutr = \frac{\sum_{i=1}^{n} Wm(Pave - Cave)}{\sum_{m=1}^{3} Wm} + Nave$$

2.4.2 Weighting individual synoptic states

In Chapter 6, similar equations are used to produce node specific multi-model solutions. The benefit of this is that the performance of each node is determined and the strongest model for that synoptic pattern is used instead of assuming that the model has equal skill for all

processes as in above. Node specific weights from Equation 5 are used to weight node average fields for each model. The weight is expressed as a proportion of the summated weights from all models thereby making a new weight (*Wi'*) which conveys the amount of "correctness" contributed by that model.

$$(E.5) \quad Wi' = \frac{Wi\,|}{\sum_{m=1}^{3} Wi}$$

Once all model node averages are weighted the three models are summated for each node and divided by the summated model weights for that node. The weighted multi-model solution (*MultiM*) is shown in Equation 6.

$$(E.6) \quad MultiMi = \frac{\sum_{m=1}^{3} Cave \times Wi}{\sum_{m=1}^{3} Wi} \qquad \text{per node}$$

Finally, to create a future weighted node specific multi-model solution (*Mfutr*), the difference between future and control node averages for each model are weighted using the node specific model weight (*Wi)* and divided by the summated weights for the three model for that node. The NCEP mean for that node (*Ni*) is added to the weighted climate change anomaly to produce a node specific weighted future projection, as shown in Equation 7.

$$(E.7) \quad Mfutri = \frac{\sum_{m=1}^{3} Wi'(Pi - Ci)}{\sum_{m=1}^{3} Wi'} + Ni$$

In comparing the average mean with the multi-model mean, an improvement in the skill of the mode is shown in section 6.2. The temperature range is reduced slightly and the positioning of the warming is better understood. Comparison by node shows a great deal of variation even with the weighting of nodes within each model. The use of the NCEP baseline to correct for error is one approach, which assumes that models biases have the same magnitude and spatial location under future conditions (i.e. stationarity). However, it is defensible to work on this assumption given the absence of knowledge of how these relationships might change. These outputs will be fully discussed in Chapter 5 and Chapter 6, regarding future corrected projections.

CHAPTER THREE

Bias Identification

In this chapter the results of the SOM technique's characterisation of the data are explored. The large scale characterization sets the scene for the rest of the analysis. To determine which models are performing well and for which processes, we compare the model control (current climate) values to the observed values for each node. Bias is the quantification of how closely the model matches the observed data and can be measured in different ways. The first order model biases in this analysis are determined in terms of:

- the average error (how alike each daily entry is to the node) for each synoptic state;
- the frequency at which each model simulates certain synoptic states;
- the rate and direction of transitions between states (i.e. the evolution of synoptic systems).

Each of these metrics is used to evaluate the SOM and determine the bias each model has for each process in terms of whether the processes in the model are similar to the observations, how often the model depicts a process, and the rate and direction of transitions between states. This allows for the identification of the processes that are simulated well by the models and those features which are not well represented. All three metrics could potentially be used to create a weighting function. However, the frequency bias has much greater variability than the other two metrics and becomes the focus of the weighting function in subsequent chapters. These components form a discussion on the diagnostic components of the experiment.

Each master SOMs is generated from NCEP data in the process described in Chapter 2, using the full annual data set to produce an annual training that will be discussed first (to provide an overall context) and monthly subsets for January, April, July and October representing the four seasons (to provide greater intra-annual detail).

3.1 SOUTHERN AFRICAN CIRCULATION DYNAMICS

To provide some context for the analysis, the baseline climate of the region is described briefly in this section. Southern African climate is driven by a few large scale features that are present in all months but have different influences and predominance in each season. The subtropical anticyclones (high pressure cells) are always dominant features. The two dominant cells are named after the oceans they are located over: the South Atlantic High Pressure (SAHP) and South Indian High Pressure (SIHP). These high pressure cells move northwards in winter in response to the seasonal shift in maximum sunlight which moves the inter tropical convergence zone north and all associated large scale features with it. When

anticyclones ridge in over the continent from the Atlantic Ocean, it results in stable conditions as the high pressure hinder uplift and convection. The increased intensity of these features in winter explains why most of Southern Africa experiences little winter rainfall.

The exception to this is the Cape south coast which has winter rains associated with the cold sector of frontal systems that are attached to the westerly wave. In the winter the subtropical anticyclones move further north allowing for the westerly wave to shift north too and results in the frontal systems moving over the southern coast. The pressure gradients are much stronger in the winter, which has steeper gradients than in any other month in the analysis. As winds are driven by strong pressure gradients, this also explains the gale strength winds common to the Cape in winter. Cut-off lows result in intense precipitation events over the southern coast and are experienced a few times a year, often resulting in much flooding and damage.

Summer dynamics are more commonly driven by tropical troughs extending down from the tropics and bringing moist air further south. These troughs are at times connected to the westerly wave to form squall lines that bring intense precipitation. Due to the summer heat over the continent, heat lows are easy to form, which result in much thunderstorm activity over the eastern highlands. All of these features result in heavy summer rains over the eastern half of the continent. The subtropical highs are less dominant and weaker in summer and are also further south, blocking most westerly wave fronts and resulting in a dry south coast.

In both spring and autumn the subtropical anticyclones are dominant, bringing stable conditions to most of the southern continent. Weak troughs form over the continent in both seasons, sometimes linked to the westerly wave. These features extend from the tropics and are fuelled by the moist air originating there.

3.1.1 Circulation characterization by NCEP

All SOM trainings in this experiment used normalised NCEP data to derive 15 synoptic archetypes (5 in the x-axis and 3 in the y-axis). The numbering of the nodes is shown in figure 3.1 and this numbering will be used throughout the analysis. The input data, as described in Chapter 2, comprises of lower troposphere winds at the surface and 700 hPa and specific humidity at 700 hPa, to describe large-scale tropospheric circulation dynamics. The reason for choosing fields depicting large scale feature is that the GCMs tend to be most skilful at this scale (Randall et al., 2007).

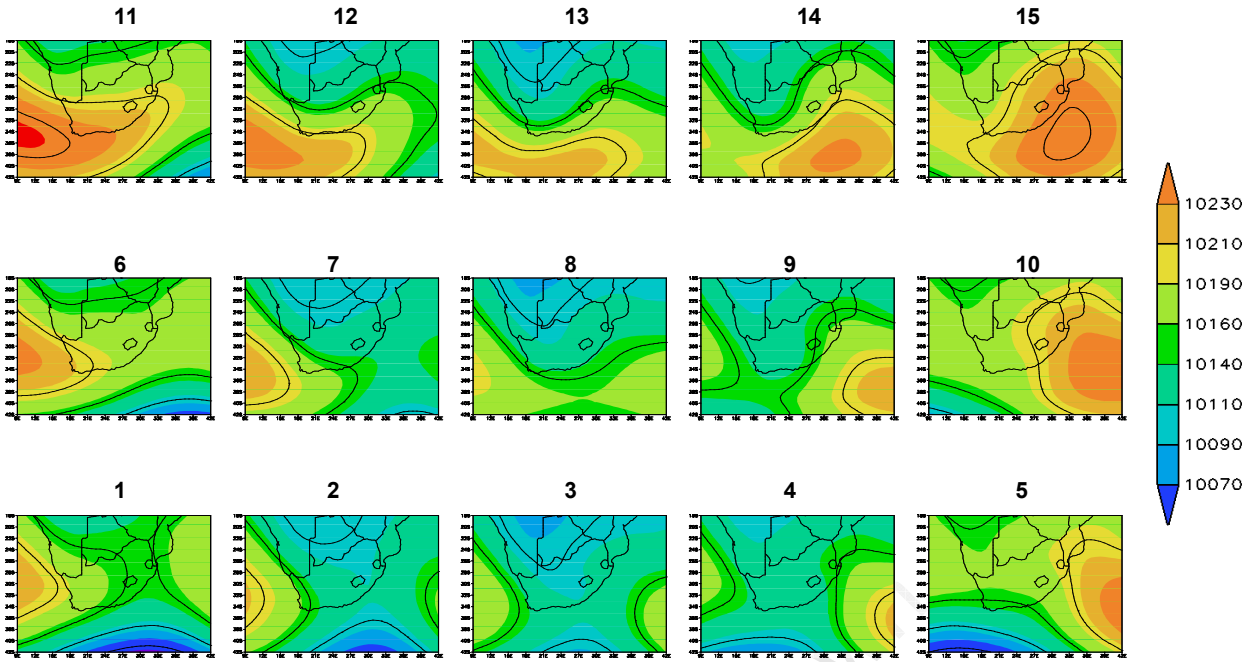


Figure 3.1: Layout reference guide for 15 node SOM, nodes depicting annual SOM sea level pressure from NCEP (1979-2000).

The trained annual SOM has a range of synoptic states, from states with a strong westerly jet to states with dominant ridging high pressure cells over the continent. Figure 3.2 provides a rough guide to the dominant circulation types in the SOM, using the terminology from Tyson and Preston-Whyte (2000). The classification was based on the large scale synoptics of average sea level pressure fields (which was not a direct input in the SOM training but an average derived for each node based upon the average of the data entries mapping to that node), together with the wind fields which are further described in Chapter 4.

The two outer columns of the annual SOM (figure 3.2) depict typically winter states, with strong anticyclones and in the bottom left and right corner cold front systems. The three middle columns depict typically summer states including ridging anticyclones, heat lows and troughs.

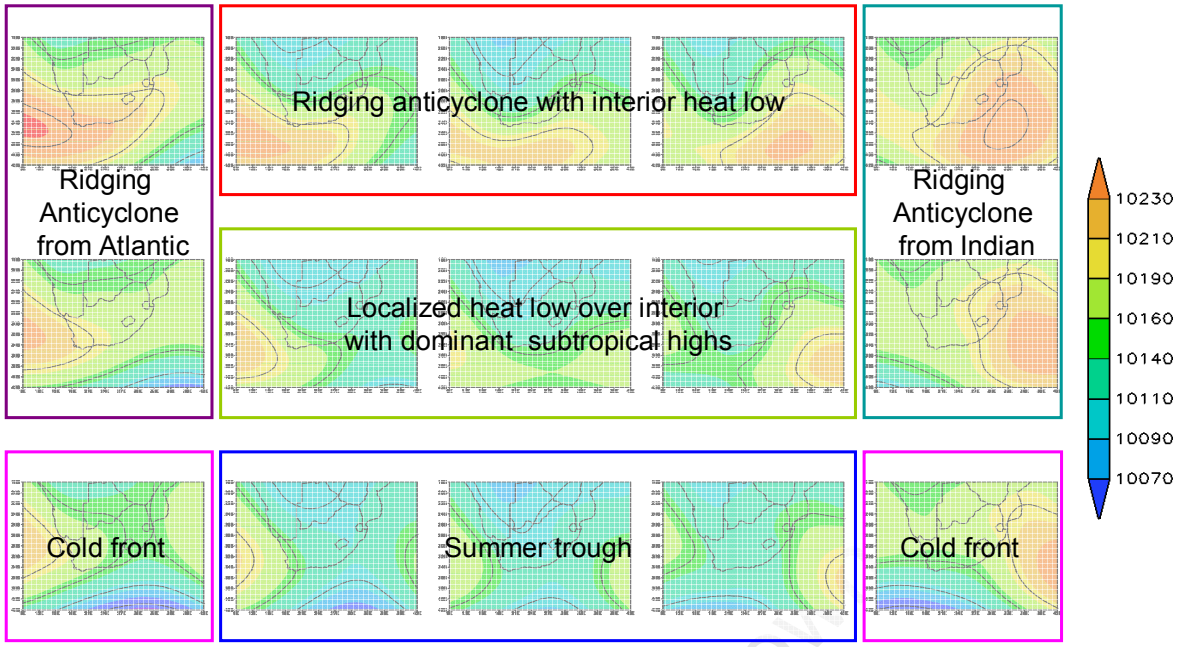


Figure 3.2: SOM mean composite of NCEP (1979-2000) annual sea level pressure, describing the synoptic characteristics in colour textboxes

In figure 3.3 the 700hPa winds are displayed as vector arrows and the colour of the arrows indicates the 700hPa winds multiplied by specific humidity to show moisture transport (red/orange for high moisture, purple/blue for low moisture). Tropical moisture sources play a dominant role in specific humidity characterization, as air from the south is generally cool and dry in comparison to more humid tropical air.

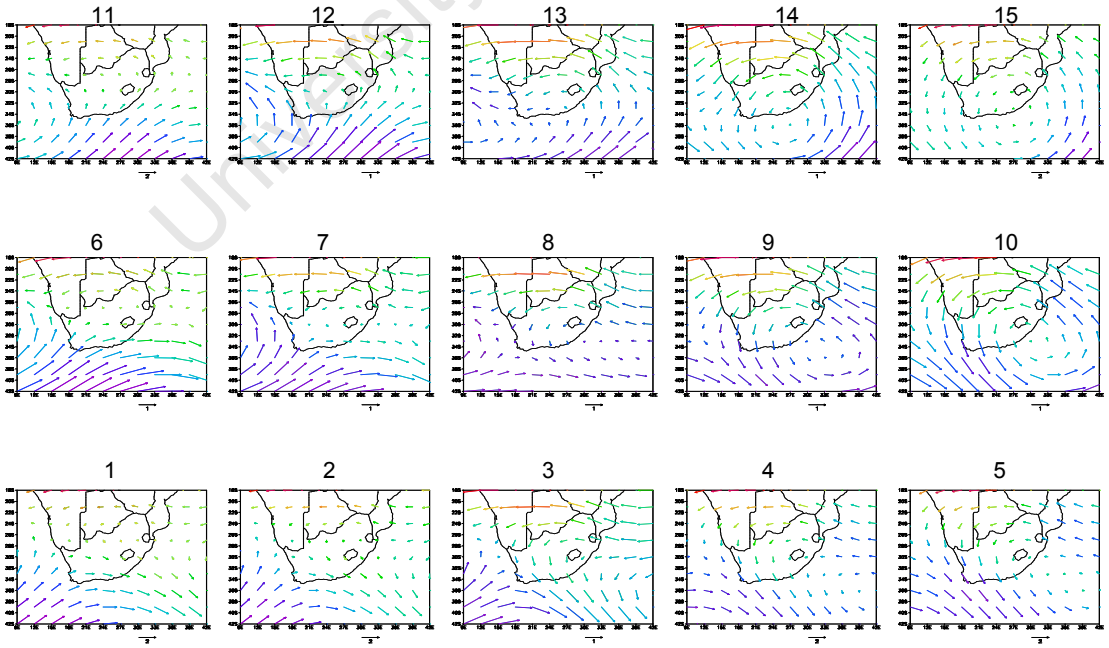


Figure 3.3: NCEP input data for annual SOM training for 700hPa level winds depicted by vector direction and arrow length, moisture transport is represented by colour of the arrow (blue low moisture, red high moisture content)

To provide an understanding of the inter-annual variation in bias and to unpack the annual signals, monthly subsets from the annual data were trained into SOM distributions. Four months were chosen to depict each of the seasons: January (summer), April (autumn), July (winter) and October (spring). Each month was subsetted from the annual data to produce monthly data series for NCEP and each of the GCMs. Each NCEP monthly data series was then trained in a separate SOM to describe that months (season) dynamics. This provides a snapshot of the different mechanisms working at different times of the year, without having to make large datasets for each season which are then more difficult to get detailed information from, due to the effects of averaging. Below is a brief description of each month's circulation dynamics in terms of the positioning of the monthly SOM distribution, to allow for easy referencing in following chapters. These descriptions will help to identify in future sections and chapters what circulation characteristics have biases that are persistent between all subsets and those that are specific to a certain process or time of the year.

January: Figure 3.4 shows the observed SLP composites obtained from the SOM patterns. Most of the circulation patterns have strong anticyclones ridging in over the east (left side of SOM) or west (right side of SOM) coast, with tropical troughs extending down from the equator. On the right side there are tropical lows that are connected to the westerly wave. Figure 3.4 provides a visual orientation to the January circulation fields.

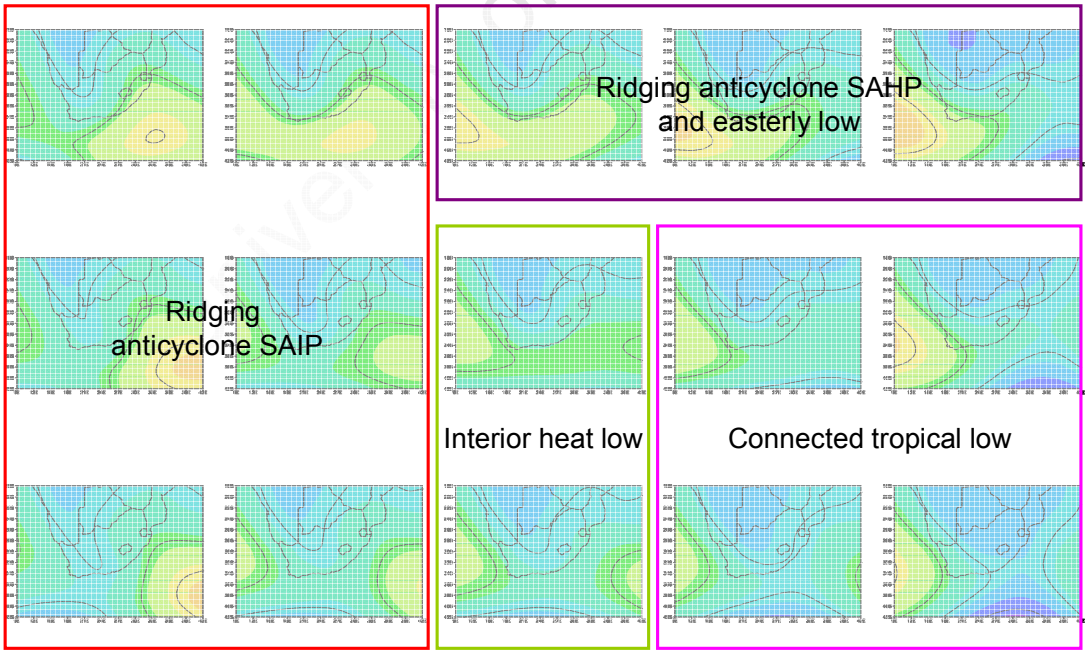


Figure 3.4: SOM composite of NCEP January (1979-2000) of sea level pressure, describing the synoptic characteristics in colour textboxes

April: As seen in figure 3.5, ridging anticyclones dominate the left side of the SOM space with dominant South Indian High Pressures in the top row and strong and ridging South Atlantic

High Pressure cells in the bottom two rows. Weak trough features are on the right side of the SOM with the bottom right two nodes depicting the westerly wave connecting up with a weak interior trough.

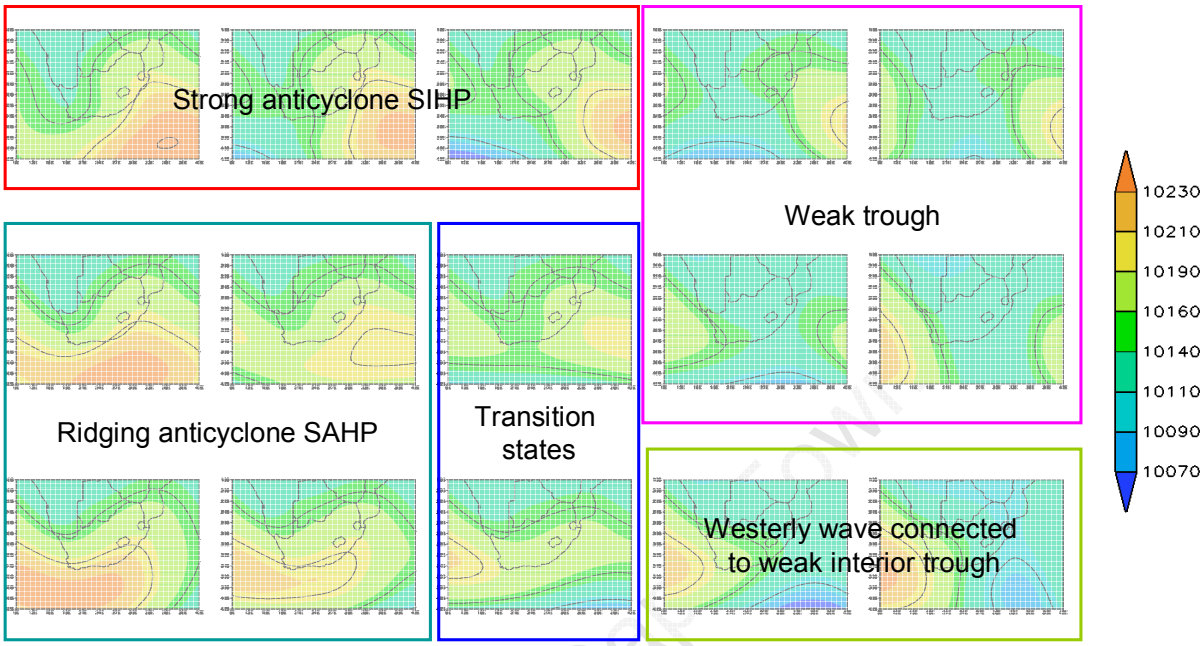


Figure 3.5: SOM composite of NCEP April (1979-2000) of sea level pressure, describing the synoptic characteristics in colour textboxes

July: The dominant circulation types in July (figure 3.6) are strong ridging highs from the Atlantic Ocean (top two rows) and strong high pressure over the Indian Ocean (bottom row) on the left side of the SOM. In contrast, the strong westerly wave with associated frontal systems dominates the right side of the SOM.

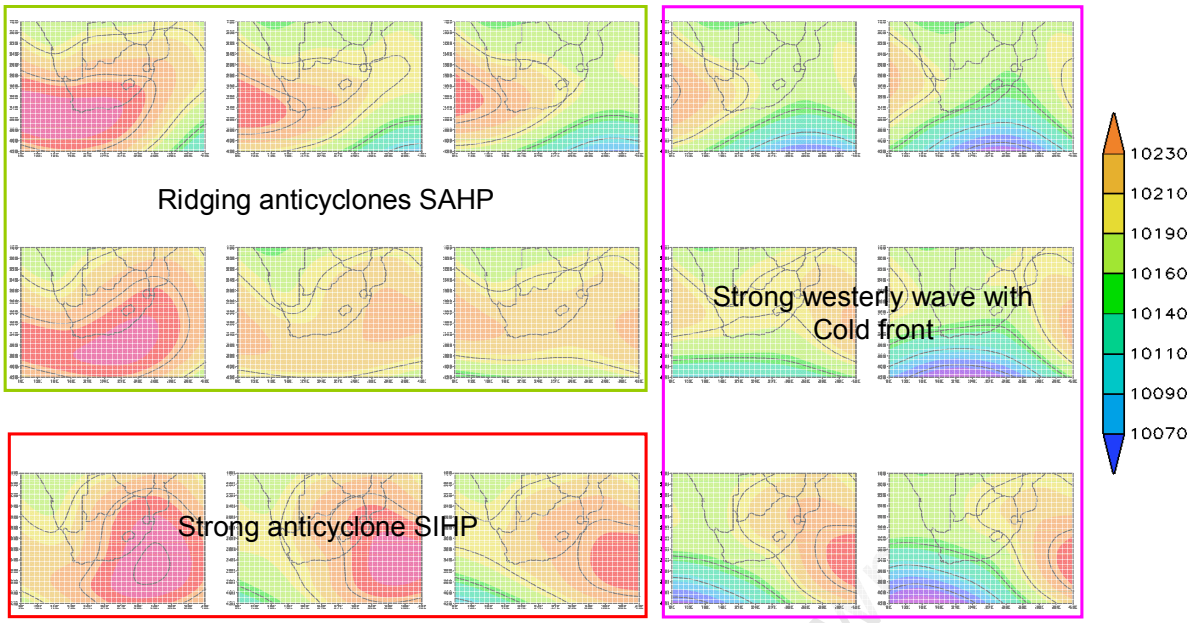


Figure 3.6: SOM composite of NCEP July (1979-2000) of sea level pressure, describing the synoptic characteristics in colour textboxes

October: Ridging anticyclones occupy the bottom left side of the October SOM (figure 3.7) with weak troughs on the top right side. The bottom right nodes depict an interior heat low with a strong anticyclone over the Atlantic. The bottom rows on the left side depict a high ridging over the coast, which bring stable conditions. Top left side has strong high pressure systems over the South Indian Ocean.

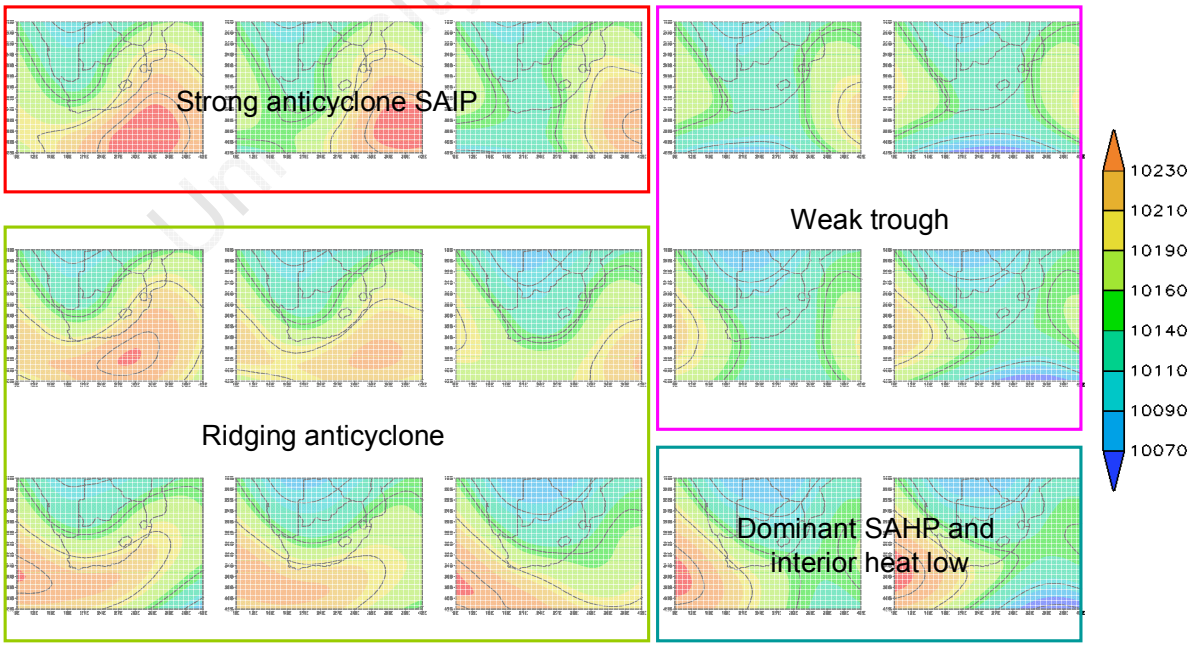


Figure 3.7: SOM composite of NCEP October (1979-2000) of sea level pressure, describing the synoptic characteristics in colour textboxes

From the SOM maps in this section, the large scale dynamics were assessed and related to the overall circulation characteristics of the region. These maps will be referred to throughout the analysis using the numerical position within the SOM followed by with a description of the dynamic state represented in brackets. In Chapter 4 the GCM equivalents of the above SOM fields will be compared to NCEP, to determine differences inside each node for all input variables as well as temperature and precipitation.

3.1.2 NCEP error and frequency

Average error values for each node give an indication of how well the SOM characterizations match the actual data. Each day of data has an error value attached to it from which an average error value per node is determined. The error values are in fact a measure of the average Euclidean distance each data entry has from the node reference vectors. If the error values are very large then the experiment is not capturing all the dynamics accurately and the SOM domain is likely to be too small. Low error confirms that the distribution of nodes adequately reflects the data space. If single inputs are used the size of the error is based upon the single metric and it is easy to determine what percentage of the total Euclidean distances the error has for a given node. With multi-dimensional data, as in this experiment, it is not as easy to determine what proportion of the data space the error makes up. If the reference vectors are examined it can be determined what the full span of the data distribution is. The SOM space usually has similar error values for all nodes, but in this analysis corner nodes tend to have slightly higher values. The corner states represent the most extreme differences within the data and therefore the outlier data map to these nodes. For this reason slightly elevated error readings are expected as states that are less similar are forced to take on the same node given a lack of alternative choices. Middle nodes tend to have lower error values as these states are transition states. In figure 3.8.a the error values for the annual NCEP SOM ranges from 14.0 – 15.8, which is quiet a narrow range of error, giving us confidence that the SOM is not too tightly constrained and that there is a good distribution of the data space (i.e. the SOM size is appropriate). Error maps for NCEP monthly data have a small error range for most months and in some months the error values are lower than the annual data.

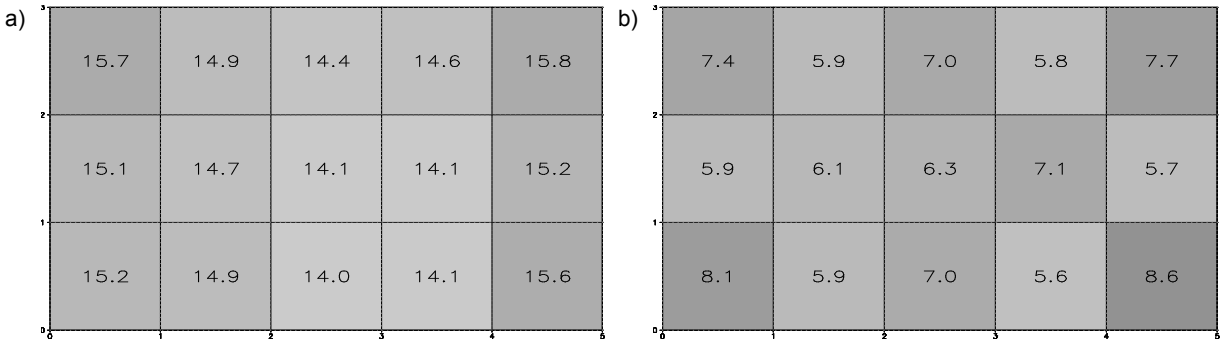


Figure 3.8: NCEP annual SOM training a) error map (Euclidean distance) and b) frequency (%) for the 1979-2000 period.

The frequency value represents the percentage of the data that maps to a particular node. The NCEP annual frequency distribution varies between 5.8 – 8.6 %, with slightly higher frequencies for corner nodes (see figure 3.8.b). This illustrates that the most commonly occurring nodes are in the corners. However, none of the states have very low values, indicating that the distribution within the SOM space is almost even. In general corner nodes generally have higher frequency values, but in some cases another node may have the highest value. The monthly NCEP error and frequency plots will be fully analysed in conjunction with the GCM plots in section 3.2.

3.1.3 NCEP Trajectories

The use of trajectories to follow the evolution of synoptic states was first demonstrated in SOM mapping by (Main, 1997). In this analysis it is used to determine whether the models follow the same evolutionary pattern as observed data. If the trajectories in the observed and model data are similar then the models are mimicking the cycle well. Trajectories are determined by the sequence of node selection or day to day progression of synoptic states. For each node the average subsequent node is determined and this average transition is depicted by an arrow showing the direction of change and the arrow length depicting the magnitude of the transition.

The NCEP annual average trajectories are plotted for each node in figure 3.9. The length of the arrow depicts the average number of states moved in the subsequent time step or the transition rate. Middle nodes tend to appear stationary, which is logical as these nodes represent transitory weather patterns which interlink all states. This explains why the magnitude and direction are very low as the average transition is well dispersed amongst the other nodes. Large magnitude trajectories are found in all the corner nodes, as these are extreme states.

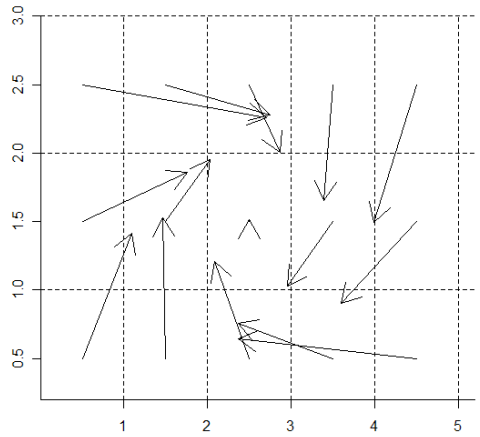


Figure 3.9: NCEP annual SOM training trajectories, length of arrow indicates average transition between states

From this we can see that nodes track down the right hand side of the SOM and then skip across to the bottom left corner and track up the left side in a clockwise motion. This evolution of states depicts the progression from a westerly wave with an associated front moving eastward, followed by ridging high pressure cells moving in over the continent. Monthly NCEP trajectories will be discussed in section 3.2.2 in comparison to the model trajectories, determining bias for this metric.

3.2 GCM CHARACTERISATION

In this section the model values for the three metrics are compared with the observed data to determine bias. The common trends between the five SOM trainings will be discussed in section 3.3. The analysis of the monthly SOMs highlight some interesting results, however, only a few of the plots will be shown to keep this section concise. All of the figures not in the text can be found in Appendix I.

3.2.1 GCM error and frequency

When looking at model error, it is important to remember that the values represent the average Euclidean distance that the data entries are from the node which is representing them. The models tend to have higher error values than NCEP, particularly in the outer nodes (as seen in figure 3.10). This is to be expected given that the model data was not incorporated in the SOM training data and therefore the NCEP data should be closer to the node representatives as the patterns are drawn from NCEP. This is also indicative of the fact that the SOM space mapped by NCEP does not fully encompass the range of control climate states (see table 3.1). Interestingly the central node error values in the model are often lower than NCEP which is likely related to the fact these are transition states and the models tend to have lower variability for these modes.

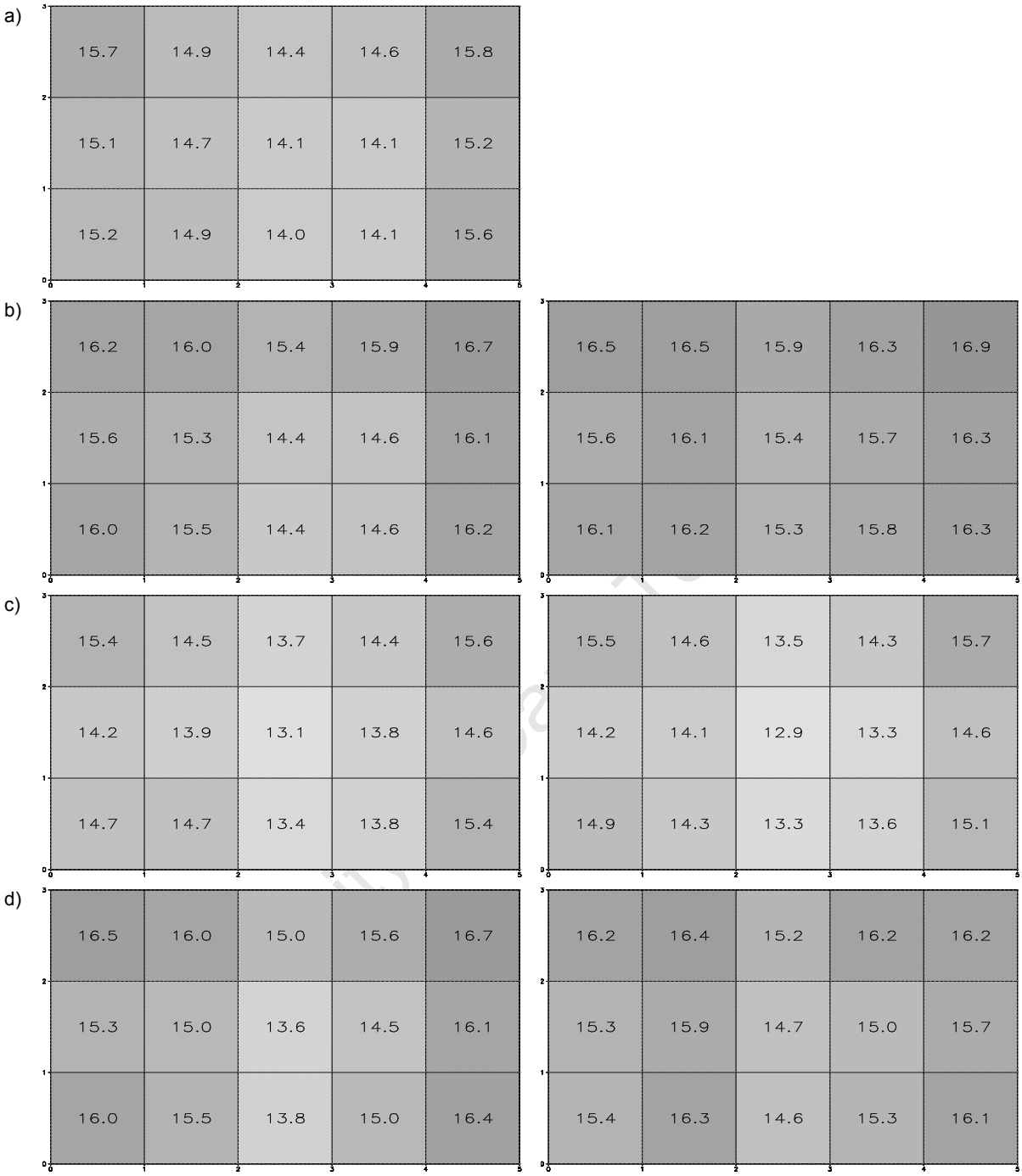


Figure 3.10: Error maps of annual SOM training, expressed as Euclidean distances for a) NCEP, b) GFDL, c) MIROC and d) MRI control years on the left and projections years (future) to the right

In table 3.1 the range of difference in error between each GCM and NCEP for all nodes is expressed as a percentage of total NCEP error. This measure gives us a guide as to how closely the overall SOM space from NCEP matches the model. The maximum error differences are found in GFDL July control projection where error is no less than 30% and up to 65%, however the future projection for the same model and month is very much lower indicating that the future conditions occupy the SOM space better than the control. This same pattern of lower future errors is seen in all of MIROC projections, which even has many

error values that are lower than NCEP. However this is not consistent in the other two models.

Model	Annual	January	April	July	October
GFDL control	2.5 → 8.5	18 → 39	6 → 20	30 → 65	18 → 36
Future	4 → 11	12 → 30	14 → 34	10 → 26	5 → 35
MIROC control	-7 → - 1.5	10 → 24	-4 → 16	-4 → 12	5 → 40
Future	-8 → -1	3 → 27	0 → 12	-8 → 10	-20 → 30
MRI control	-3 → 7	10 → 28	0 → 16	2 → 20	18 → 38
Future	2 → 10	10 → 23	10 → 32	2 → 22	0 → 30

Table 3.1: Indicative ranges for the difference in error values between model control and NCEP expressed as a % of the NCEP error (Negative numbers indicate that the NCEP error is larger than model error)

Future error values (right column of figure 3.10) are largely consistent in magnitude with the control, indicating that the circulation patterns are of a similar type in both future and control. But as noted above the future errors can be smaller than the control. As the models dynamics do not alter with time (i.e. the equations governing the simulation), one would hope that the control and future runs of a model would be consistent for this metric. A large difference in error would indicate that a significant shift in circulation patterns in the region. Given this constraint in all models, this metric is unlikely to be useful in a weighting function and strengthens the idea that a frequency bias plays a large role in overall model biases.

Looking at the differences between the model and observed frequencies in figure 3.11, model control frequencies show a similar distribution to NCEP. GFDL has higher values in the first column than the other models (ridging South Atlantic anticyclone with strong westerly wave), indicating that the model may over represent these states. Future frequencies for all models show a drop off on the right side of the SOM (dominant South Indian high), indicating that these states are occurring less frequently in the future. If there are less SIHP to block low pressures, an increase in east coast rainfall is likely to result.

All the models tend to favour the central node (strong heat low over the South African interior with ridging anticyclones linking over the tip of the continent). GFDL and MRI both have the central column (interior heat low) showing large increases in the future, which could indicate an increase in total rainfall. Corner nodes' frequencies decrease in the future in both GFDL and MRI. MIROC, by contrast, has increasing frequency for the left bottom corner (strong westerly wave and dominant South Atlantic high) and all other corner nodes maintaining similar frequencies to the control. The MIROC bias when seen in conjunction with the precipitation values shown in Chapter 4, shows that the model is tending towards a drier climate as the corner nodes all have low specific humidity and precipitation values.

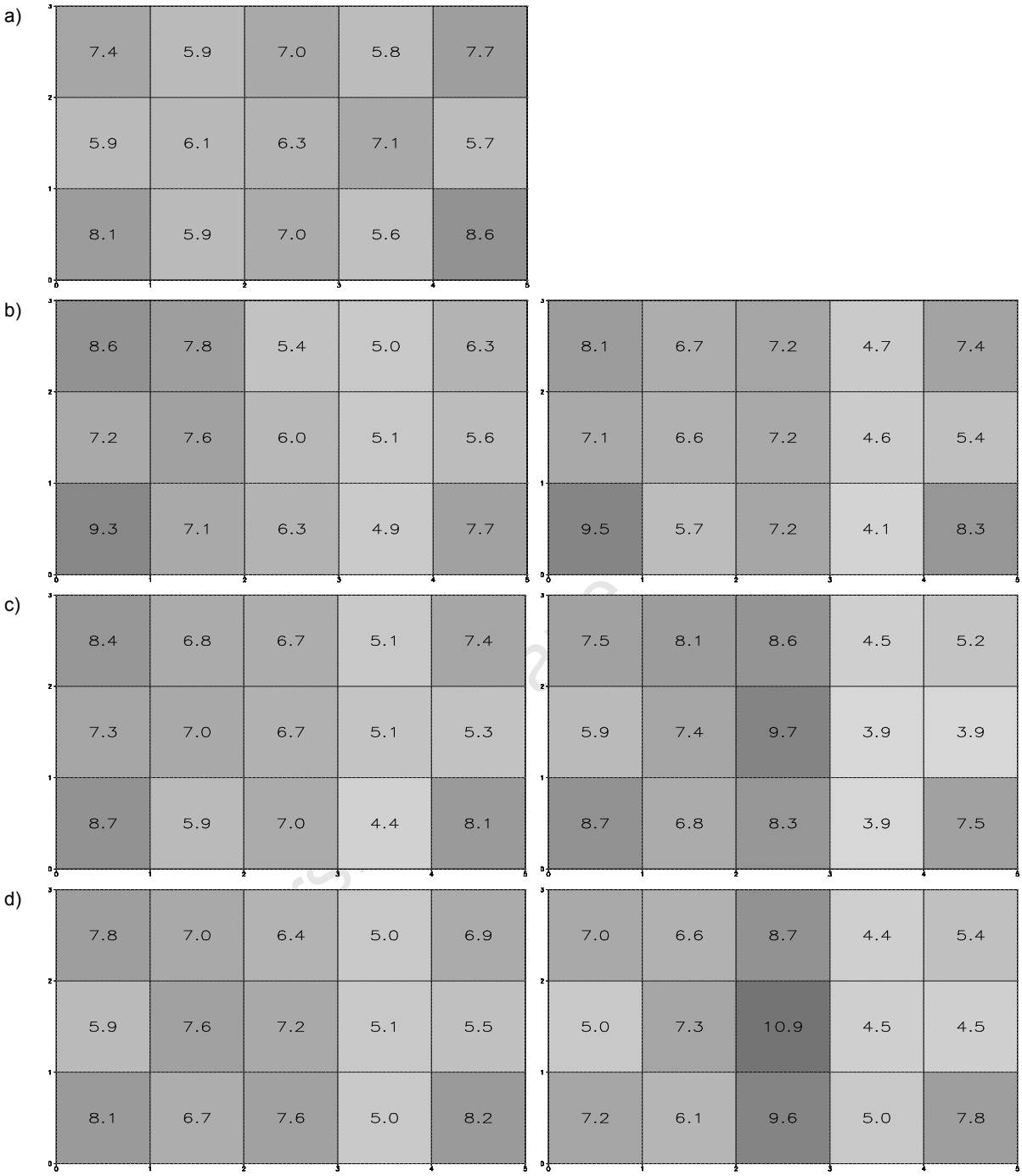


Figure 3.11: GCM frequency maps for control years on the left and projection years (future) on the right, expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

From these basic statistics, we observe that the frequency differences are more variable than the error values and could provide a useful metric for weighting models.

Monthly subsets

Error values are generally larger in monthly SOMs than in the annual SOM. This could be due to the smaller data sample (i.e. there are fewer data points, therefore the average is

more likely to be skewed by outlier values than if there were a large sample size). Some nodes have very low frequencies making these outliers more prominent. Most of the monthly error and frequency maps are in Appendix I.

January: Error values (Appendix I, figure A2) tend to be larger in the models than NCEP, particularly in GFDL, which has higher error for the left top corner node (ridging anticyclone from Indian Ocean). In the GFDL future projections, the same node decrease in error. All models have low error for the central node (interior heat low) in both control and future. Corner nodes tend to have the largest errors, as these are the extreme states. Future errors in corner nodes are mostly less than the control, except in MIROC which shows increases in the first column (ridging anticyclones) but less error for most other nodes. This indicates that the future data fits into the SOM space occupied by NCEP better than the control data for some nodes.

Each model has a different bias for its highest and lowest frequency node (Appendix I, figure A3). GFDL favours the far right column (connected tropical low with strong South Atlantic high), MIROC the right bottom corner node (connected tropical low) and top central column node (ridging South Atlantic high), with a very low weighting for node 7 (weak South Indian anticyclone). MRI favours the node 3 (interior heat low). Looking at the model anomalies from NCEP (figure 3.12) all models have much lower values node 1 (ridging South Indian high with trough to west). Generally, dominant South Atlantic highs are over represented and dominant South Indian highs are under represented in all the models.

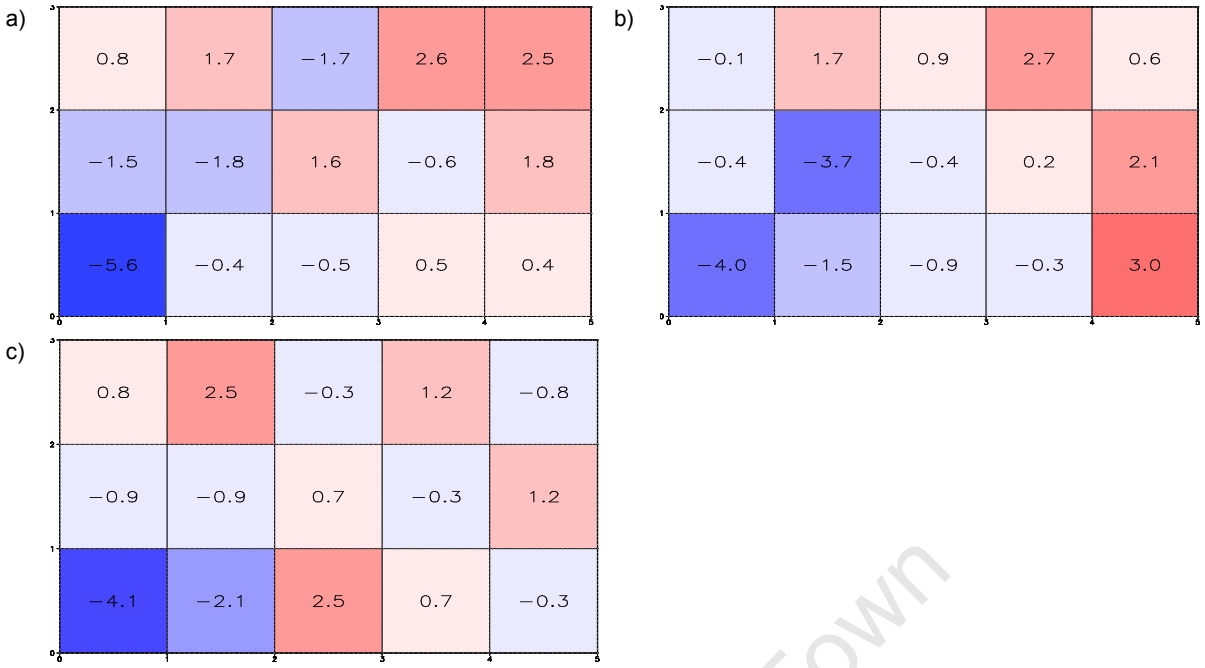


Figure 3.12: Anomaly frequency of January GCM control from NCEP a) GFDL, b) MIROC and c) MRI. Red indicates where the model frequency is higher than NCEP and blue where it is lower.

In the future frequencies there are some more consistent trends with all models favouring node 5 (connected tropical low), supporting literature that the summer rainfall region will have increased precipitation (Midgley et al., 2005). Dominant South Indian systems show little change between future and control, but as these states are under represented in the control these biases are important.

April: Error is highest in the right corner nodes of the NCEP plot (weak interior trough connected to westerly wave) – see Appendix I, figure A4. All of the models also have high error in these nodes. In GFDL and MRI future plots, the error increases in all nodes especially the corner nodes. This could indicate that some extreme states can be expected in the future that do not correspond to the NCEP archetypes; however the error values are still not large enough to indicate completely different dynamics but it is plausible that intensity of pressure gradients may increase for example. In Chapter four this will be further explored.

NCEP frequencies favour the corner nodes and the models have mostly lower frequencies in the top row (ridging South Indian highs and weak troughs) and higher frequencies for the bottom row (ridging south Atlantic high and connected easterly trough) as seen in figure 3.13. All models have very low frequencies for the node 12 (ridging South Indian high). GFDL and MRI favours node 9 (weak trough) whereas MIROC favours node 1 (ridging south Atlantic high). MIROC has a more even distribution amongst the weak trough states in comparison to the other models which favour a single node. All indicate that the models tend to over

simulate these weak trough features and under simulate ridging Indian highs, which could lead to a slight increase in precipitation.

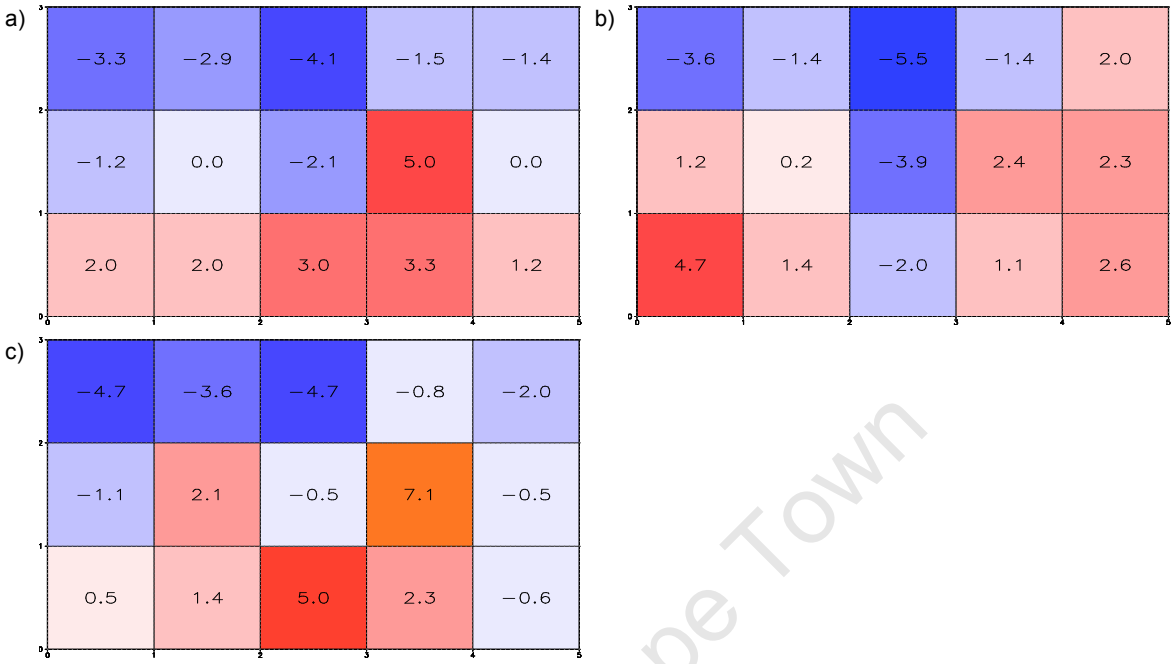


Figure 3.13: ***Anomaly frequency for April****: GCM control from NCEP a) GFDL, b) MIROC and c) MRI. Red indicates where the model frequency is higher than NCEP and blue where it is lower.*

Future frequencies show a large reduction in node 3 (weak South Atlantic anticyclone ridging over south). In MIROC there is an increase in the weak trough states in the right corner. GFDL and MIROC both have increases in frequency of the under represented South Indian anticyclones which could indicate that this feature may become even more dominant in reality. MRI favours node 1 (ridging South Atlantic high) and the previously favoured node 9 (weak trough).

July: NCEP error is fairly evenly spread amongst nodes, while MIROC and MRI have slightly higher than NCEP error especially in the bottom row (Appendix I, figure A6). MIROC has the largest errors coming from ridging highs, however, most of these decreases in the future. In the MRI control plot, most of the bottom row (ridging South Indian high and strong westerly wave to right) has high error. This changes position slightly to favour node 6 (ridging Atlantic anticyclones), with an overall tendency for increasing error in the future. The error values in the GFDL control plots are much higher than other models, although these do decrease in the future plot. Reduction in error is greatest for westerly wave features connected to fronts on the left side of the SOM.

In the frequency plots (Appendix I, figure A7), GFDL has large frequencies for nodes 7, 8 and 9 which are all well above NCEP (see figure 3.14 a) and low frequencies in both first (very

deep anticyclones) and last (strong westerly wave) columns. In future anomaly plot (figure 3.14.b) there is an increase in the top row (ridging South Atlantic anticyclone and strong westerly wave) with large decreases in the central nodes (ridging anticyclones).

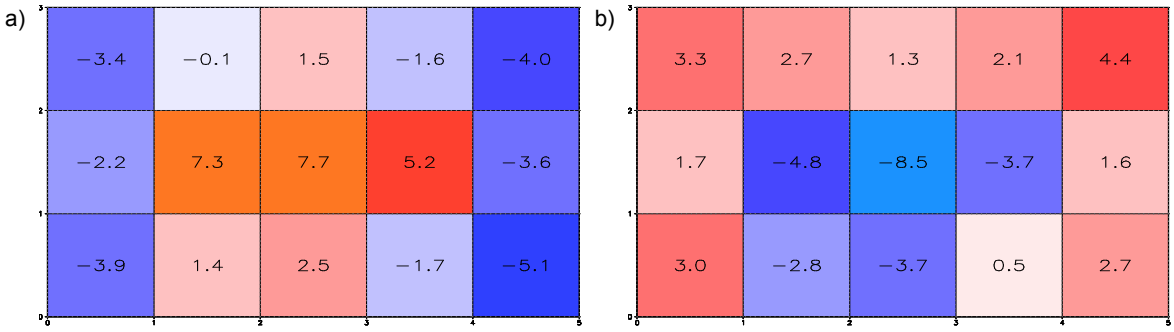


Figure 3.14: Anomaly frequency plot for July for a) GFDL – NCEP and b) GFDL future – control. Red indicates where the model frequency is higher than NCEP/control and blue where it is lower.

MIROC control plot shows a preference for dominant anticyclones and this is continued in the future plots with an increase in node 4 (a continental high over the interior and the westerly wave bringing frontal systems to the south coast). MRI has a bias for ridging South Indian highs which decrease in the future plots indicating even greater reductions are likely. It also favours node 7 (both subtropical highs sitting over the continent), which has a very low frequency in NCEP. This node's frequency decreases in the future and there is a relative increase along the top row (ridging anticyclones from the South Atlantic). This would support the idea that the western half of the subcontinent will have increased drying in the winter season (Hewitson and Crane, 2005; Midgley et al., 2005).

October: NCEP's error map indicates that it had low error for the stable conditions of the bottom row, especially node 3, where the South Atlantic high is dominant with an interior low (see Appendix I, figure A8). All models have higher error values for the control in the outer nodes. Future plots show a general decrease in error compared to respective control plots (figure 3.15) especially for the outer nodes.

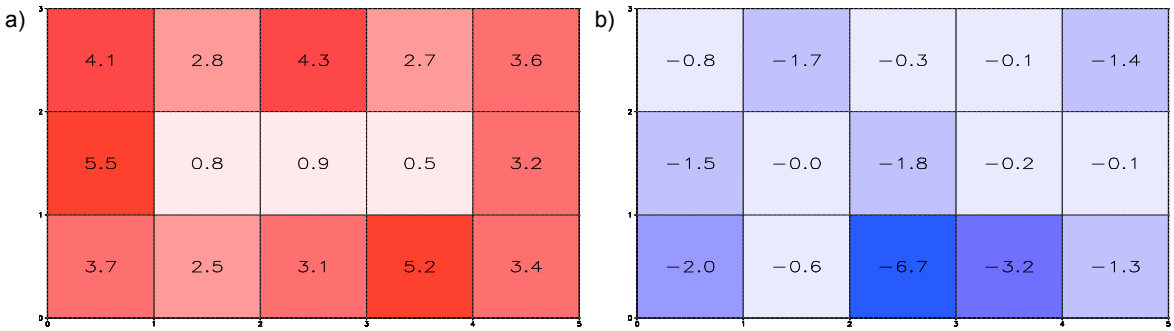


Figure 3.15: Anomaly error plot for October for a) MIROC – NCEP and b) MIROC future – control. Red indicates where the model error is higher than NCEP and blue where it is lower than the control in future projections.

NCEP frequency is highest for bottom row corner nodes depicting a ridging South Atlantic high and interior heat low (Appendix I, figure A9). All models favour the last column (trough features) and future plots show small increases and decreases dependant on model for these states. A consistent signal is for an increase in frequency for node 1 (ridging South Atlantic high) and node 8 (interior low with both subtropical highs present).

The frequency differences have illuminated some of the processes for which the models are failing to capture the correct number of days in a particular state, either under or over simulating a process. No model is a clear winner in terms of matching observed frequencies and therefore a weighted multi-model solution should help to reduce the incorrect frequency aspects and generate more accurate model results.

3.2.2 GCM Trajectories

In figure 3.16, model trajectories are very similar to NCEP in the control projections in terms of direction of trajectories, with some slight variations in magnitude of transition, tending to be slightly larger in the models in corner nodes. The direction and magnitude of transitions remains consistent between the future and control runs of all models in the annual SOM. This indicates that the evolution of the climate system remains unchanged in the model. Differences between models show that the magnitude of transitions can differ but that the general direction is consistent, which shows that the models are following the same large scale synoptic sequences but that the rate of evolution of these sequences can differ. This gives us confidence that the technique used later in this analysis (Chapter 5) will be useful as the correction is based on the frequency of synoptic states.

a)

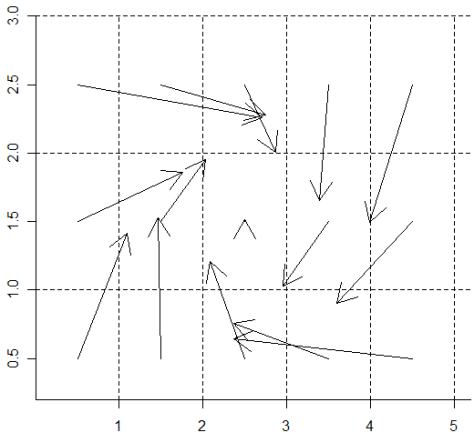
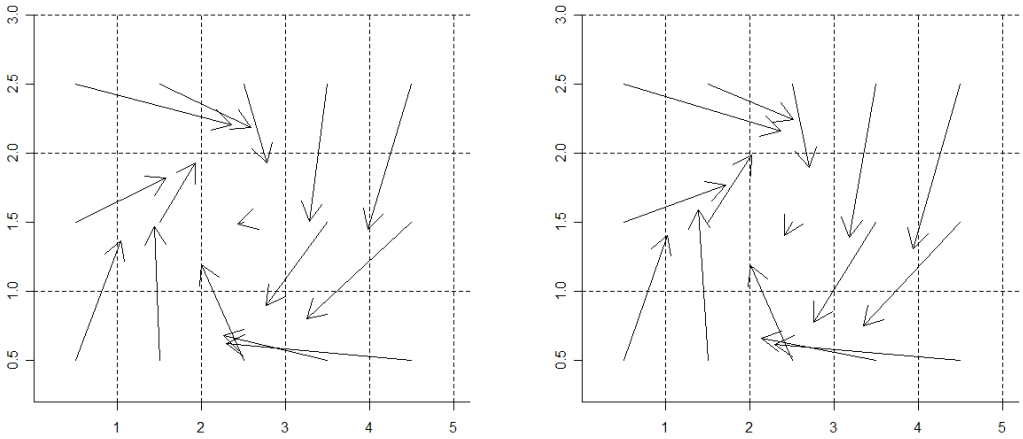
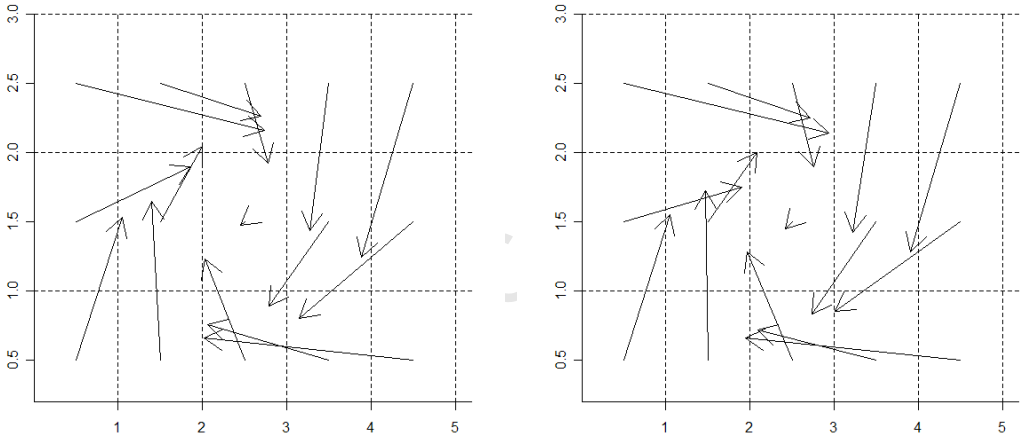


Figure 3.16 a): *Trajectory map of annual SOM training for control projection, arrows direction indicates the average transition between nodes and the arrow length indicates the speed of transitions for a) NCEP*

b)



c)



d)

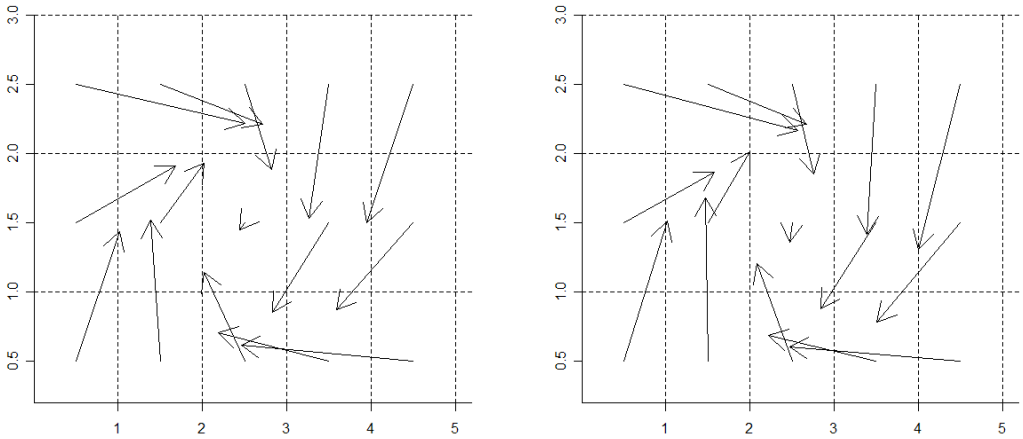


Figure 3.16 b) – d Trajectory map of annual SOM training for control projection, arrows direction indicates the average transition between nodes and the arrow length indicates the speed of transition between successive days for NCEP observed and GCM control (left) and future (right), for b) GFDL, c) MIROC and d) MRI

January: NCEP trajectories depict anticlockwise progression of transitions between states. The direction of transitions remains consistent between models and NCEP and of similar magnitude, except for node 1 (trough with dominant South Indian high) which has much larger transition magnitude to interior heat low nodes, particularly in MIROC. In the future trajectories, magnitudes increase in node 11 (dominant South Indian high) in all models towards the middle of the SOM field (interior heat low). Other transitions are unchanged in direction although magnitudes vary slightly for each node. Movement towards heat low states appears to be faster in the models than in observed.

April: Trajectories for most nodes are in clockwise direction, with outer nodes in the top left side and bottom right side having the largest magnitude transitions. Trajectories for the top left nodes (ridging South Indian anticyclones) move towards the east (weak troughs) and bottom right (westerly wave and connected trough) to the west (ridging Atlantic highs). The bottom left side nodes (ridging South Atlantic highs) and central nodes (both subtropical highs dominant) are the most stable conditions with slowest transition rates, indicating that these are intermediate features from which transitions to many different kinds of state ate possible. MIROC does deviate slightly in node 6 (southerly high pressure cell) moving in stronger more easterly direction towards a weaker ridging high from South Indian Ocean.

All three models have very similar transition rates and almost identical direction of trajectories to each other. However, all models have node 7's direction more easterly than NCEP which has a weaker north east orientation. This suggests the models tend to favour the more stable central conditions, whereas the observational data favours transitions to a deeper high with the westerly wave moving in from the Atlantic.

July: Transitions are in a clockwise direction with larger transitions rates in outer nodes. All models are broadly consistent with NCEP. MIROC shows nodes 3 and 4 (strong South Indian highs with a low to west) to be decreasing in transition rate but maintaining transition direction, indicating greater persistence of these features or that they are states from which transitions are variable as to which kind of state it progresses to. GFDL shows larger transitions in the future for (ridging South Atlantic anticyclones) to central nodes which have weaker subtropical highs. GFDL also has a stronger transition rate in node 10 (strong westerly wave). MRI has a future direction change for node 4 (strong South Indian highs with low to west) which has a northeast orientation, towards stronger westerly wave patterns.

October: NCEP trajectories are clockwise, as are all the models. GFDL has a greater transition rate for node 7 (southerly high pressure cell) towards the central node. MIROC has increased transition speed in nodes 7 and 11 (ridging South Indian highs) and all three models have increased transition speed from node 6 (also a ridging South Indian high). From

this it would seem that the models move through the South Indian High Pressure systems a little faster than occurs in reality and thus would explain why they are under represented in the frequency plots.

Future plots of GFDL and MIROC show increases in magnitude for nodes 4 (interior heat low with dominant South Atlantic high) moving towards ridging South Atlantic highs. MIROC has a decreased rate of transition in node 10 (weak trough) and MRI has node 11 (ridging South Indian high) weakening and moving towards central stable conditions.

Broadly, trajectories are consistent between NCEP and models and between future and control indicating that there is little bias in the sequencing of synoptic states. All models tend to have similar transitional directions indicating evolution in models is similar and accurately reflects reality as they are mostly consistent with NCEP. This reveals that although the trajectory metric is useful for understanding the underlying data, it will not be suitable for a weighting metric.

3.3 COMMON BIASES

From the two sections above, some patterns have emerged that are further clarified here, based upon all five SOM trainings and highlighting the similar biases identified so far.

High error values indicated were there are dynamics in the model that may not be accurately represented by the NCEP master SOM. For the annual SOM (figure 3.10) the error values from both the control and future projections were a little higher than NCEP, especially in the outer nodes, but error was at times lower in the models, particularly in the centre. Monthly SOMs have larger differences between NCEP and models but as stated before this should be expected given the smaller sample size of the monthly data. As in the annual plots, error is highest in the outer nodes and often decreases in the future projections, especially in the MIROC model. Errors do not seem to be higher for any particular process but are generally lower in the transient states that occupy the centre of the SOM.

As with the error plots, the outer nodes tend to have higher frequencies in both annual (figures 3.11) and monthly plots (Appendix II figures A3, A5, A7 and A9). The annual plots show a great deal of similarity between the models choices for high and low values and, although these are roughly the same as NCEP, there is greater similarity between models than an individual model and NCEP. Some months have consistent biases such as April and October, where all models consistently favour two or three nodes and have similarly low weights for other features. Less consistent signals are seen in the other seasons.

Some general biases **from the baseline period (1979 - 2000)** are:

- Strong South Indian anticyclones are under represented by the models in all seasons;
- Trough features are increasing in dominance in October, perhaps resulting in overestimated rainfall;
- January has an increased number of continental lows, which could result in overestimation of rainfall in this season. This will have to be confirmed by the results in Chapter 4.

In all the trajectory plots the models capture the same basic evolutionary cycle with slight variance in the average direction and magnitude of change, but overall maintaining much the same pattern. The models seem to behave more similarly to each other than to NCEP, although deviations are slight. There is some indication that South Indian High Pressure cells may become more persistent, although the sequence of evolution remains the same.

In the Chapter 4, each of the circulation input variables, temperature and precipitation will be assessed to see how the dynamics in the NCEP master SOM match up to the average GCM field for each variable. From this we can see whether the biases identified are the result of misplaced synoptic features, inaccurate gradients or over and under estimated magnitudes. Future changes in the variables will also be discussed. From this individual node biases for each model are identified, leading up to the bias weighting function to be discussed in Chapter 5.

CHAPTER FOUR

Differences in GCMs representations of climate variability

The ability of a model to simulate present climate variability is an important pre-requisite in order to have confidence in the model, as this defines the baseline against which the future climate projections are compared. Model validation is therefore a central necessity in any model analysis. The use of SOM states allows for comparison of not just mean and global model values but also for a specific domain, as well as allowing for a disaggregation into the various synoptic states. In this way, skill is tested for the sub region and for a number of circulation dynamics, highlighting each model's ability to depict the correct position and magnitude of large scale synoptic features. This gives insight into how the model represents circulation features and provides a guide as to what biases a correction function should be reducing.

The NCEP reanalysis data set (1979-2000) was used as the baseline data for the SOM. Although NCEP is a good proxy for observed data, it should not be considered as reality (Kalnay et al., 1996; Tennant and Hewitson, 2002), due to the constraints mentioned in section 2.2. For the purposes of this analysis however it is assumed that this data is "reality" in the absence of a more accurate gridded observation data set at the time of the study.

The NCEP SOM archetypes shown in the previous chapter are used as the reference state for observed climate. In addition, a composite node mean of climate variables for each model is generated by calculating the mean for each grid cell of all the dates mapping to that node. The end result is a disaggregation of regional climate patterns as a function of synoptic state for each model, which allows for comparison with NCEP.

This chapter thus uses SOMs as a means to focus on the synoptically based differences between models and NCEP for key atmosphere and surface variables. It is recognized that frequencies of these states vary between models, but that is the focus of Chapter 3.

The SOM input variables used, which describe the circulation dynamics, are lower troposphere winds and specific humidity. The circulation data was standardized using a normalization technique, which expresses each variable as a departure from its own mean over the standard deviation. As the variables ranges in magnitude are large, the normalised data brings all variables to a baseline from which they all add the same influence to the SOM training. However, as normalised data are already anomaly fields, comparison by differencing the fields would not be meaningful.

The normalisation was therefore inverted (by multiplying by the standard deviation and adding the mean back to the node values) in the anomalies shown in the following section - to see the real differences in winds and specific humidity per node. The inverted normalisation figures are shown in greyscale at the head of every section to allow for easy reference to anomaly conditions.

The sequence of data manipulation is thus; original data → normalised → training for SOM → reverse normalising per node → SOM composite anomalies

Composite maps of temperature and precipitation data were generated by taking all the days that mapped to the same node and averaging them. However, when looking at these composite node fields, remember that the temperature and precipitation data was not standardized, and therefore anomalies represent real numerical differences.

The monthly SOM fields will also be discussed. However, as there are many plots generated in this analysis, but not all are necessary to illustrate the biases, most of these figures can be found in Appendix II. All plots use a red/white/blue colour palette where, red indicates a positive bias and blue a negative bias, and white areas are close to zero.

The implications of the biases described for each variable will then be synthesised and summarised in a section at the end of the chapter, outlining the possible dynamical explanations for the biases. This will help inform the appropriateness of the bias correction results as it can be determined whether the correction techniques are addressing the biases identified.

In the rest of this work, greyscale figures are used to represent a single variable field which has values that are all in the positive range, and colour maps are used where the single variable has both positive and negative values or is an anomaly.

4.1 SURFACE WINDS

The surface and mid level winds are split into their vector components and each model's composite node anomaly from the NCEP SOM is analysed to see which synoptic states are being accurately simulated and which the models struggle to depict. For the u wind component (east-west winds), blue is an easterly (east to west) wind and red is a westerly (west to east) wind. For the v wind component (north-south winds), blue indicates a northerly (from north to south) wind and red a southerly (south to north) wind.

East – West winds:

The NCEP mean conditions for the surface u wind component (east-west winds) shown in figure 4.1 show that westerly winds dominate in the southern Ocean (associated with the westerly wave in the upper air) while over the continent and adjacent oceans the winds tend to be easterly. The westerly wave is dominant over the subcontinent in the outer columns (typically winter states) and is weaker in the three central columns (typically summer states).

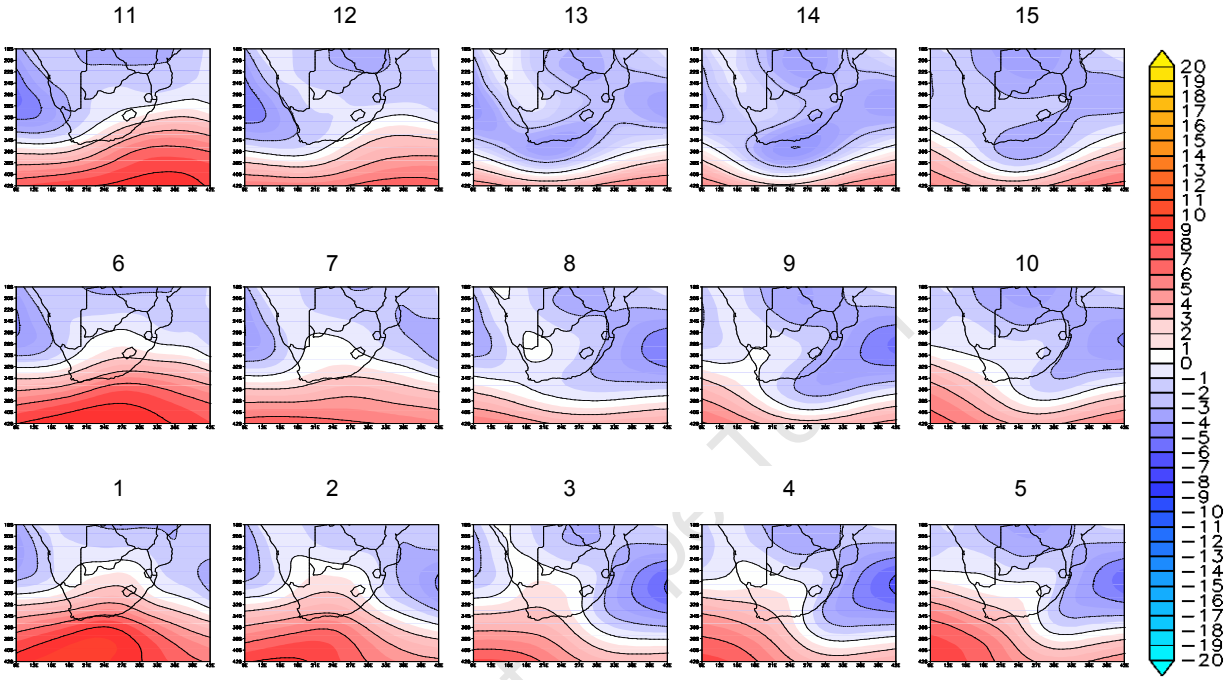


Figure 4.1: NCEP master SOM training showing composite map of annual average surface winds (reversed normalisation) per node in u component (east-west), contour interval 2ms^{-1}

For the u wind component anomalies between NCEP and the GCM, blue indicates decreased strength in prevailing wind and red an increased strength of prevailing wind in the model. MIROC (figure 4.2) and GFDL (Appendix II, figure B1) anomalies mostly depict an increased westerly over the southern Ocean and a persistent reduced easterly in all nodes off Mozambique's north coast. MIROC has stronger easterly winds in the strong ridging anticyclone states in the first and last columns (typically winter states) over the southern Ocean and has a strongly reduced easterly wind over Namibia in the top row (ridging anticyclones with interior low). For MRI (Appendix II, figure B2) most of the domain has a reduced westerly wind with a weakly increased easterly wind over Zimbabwe, Botswana and Namibia. Over the Indian Ocean there are also stronger easterly winds, which are strongest when coincident with dominant South Indian high.

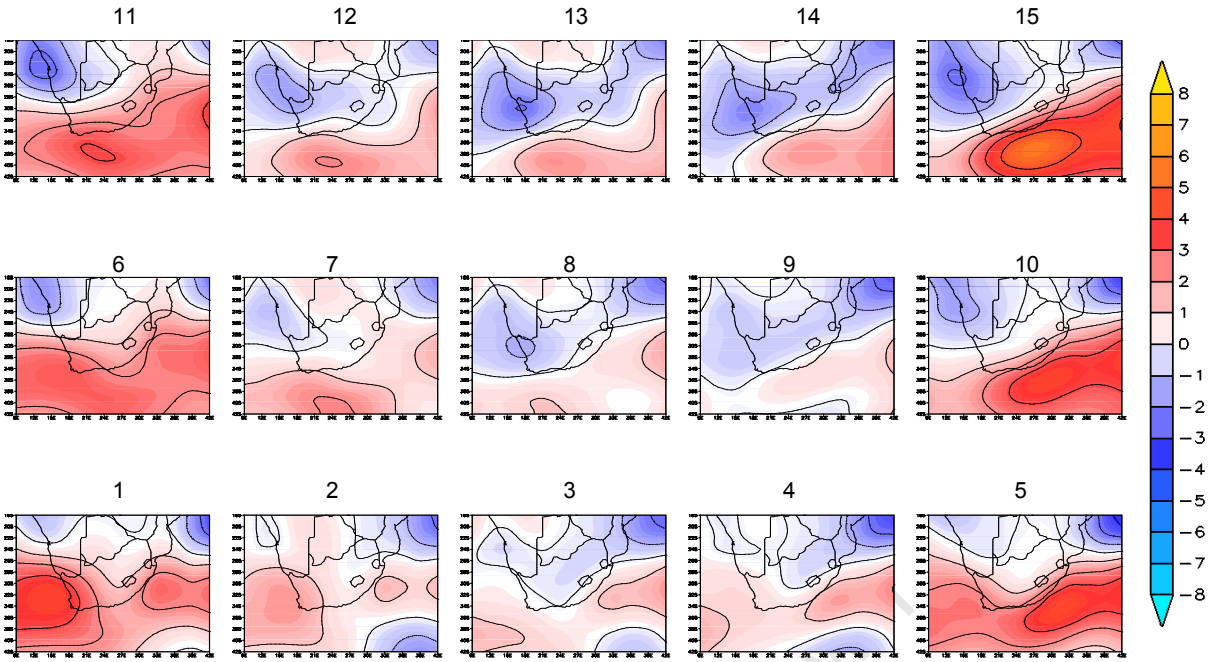


Figure 4.2: Anomaly of MIROC-NCEP for each SOM node for annual average u winds at the surface, contour interval 1ms^{-1}

North – South winds:

There is a persistent southerly wind on the Atlantic coast. Where there are summer trough features (nodes 2, 3, 4, 6, 7 and 8) the east coast and parts of southern Ocean have strong northerly flow.

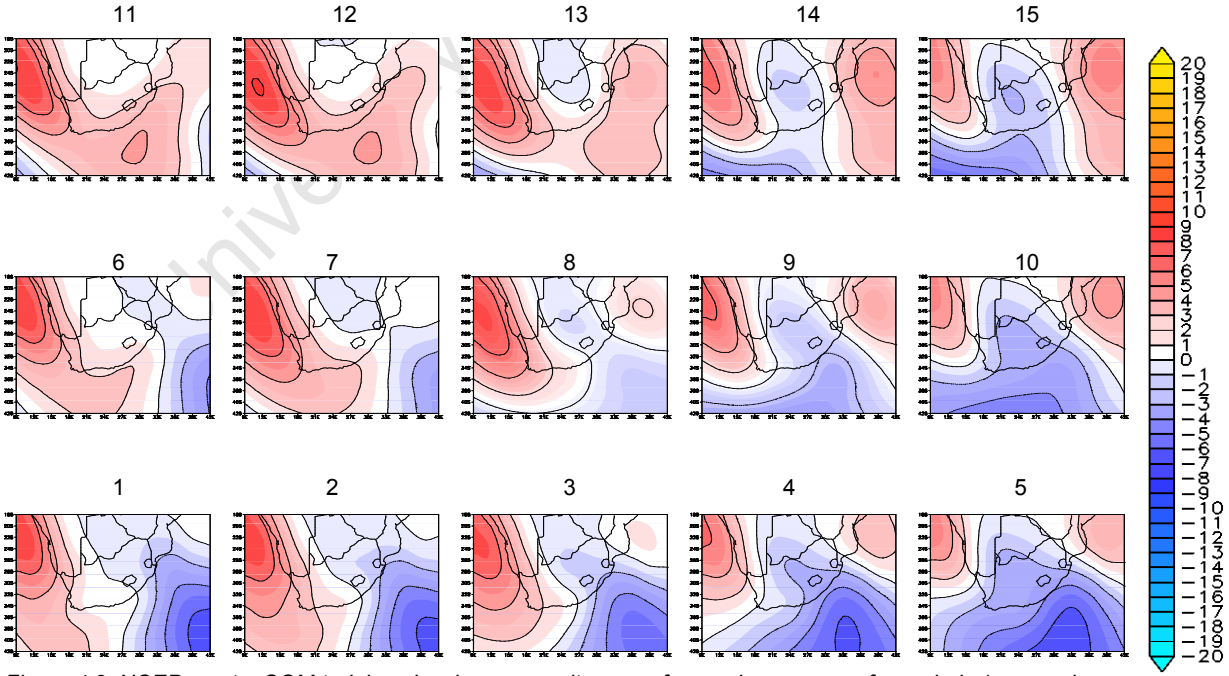


Figure 4.3: NCEP master SOM training showing composite map of annual average surface winds (reversed normalisation) per node in v component (north-south), contour interval 2ms^{-1}

For the v wind component (north-south winds) anomalies between NCEP and the GCM, blue indicates a reduced strength in prevailing wind and red an increase in prevailing wind

strength in the model. All models in the v wind component (north-south) have a similar pattern of bias. In the top row especially (ridging anticyclones with interior low) there is a reduced southerly flow over the Indian Ocean and Natal coast while over the Atlantic Ocean and western half of the continent, there is a increase in northerly winds. MIROC (Appendix II, figure B3) has a greater reduction in southerly flow over the Atlantic Ocean than the other models especially in the top row (strong ridging anticyclones). For the same states, over the Indian Ocean, all the models have an increase in southerly winds associated with the ridging anticyclone. GFDL (figure 4.4) has a slightly reduced northerly flow over the Indian Ocean (strongest reduction in bottom right corner node of figure 4.4 for example). MRI patterns are very similar to GFDL.

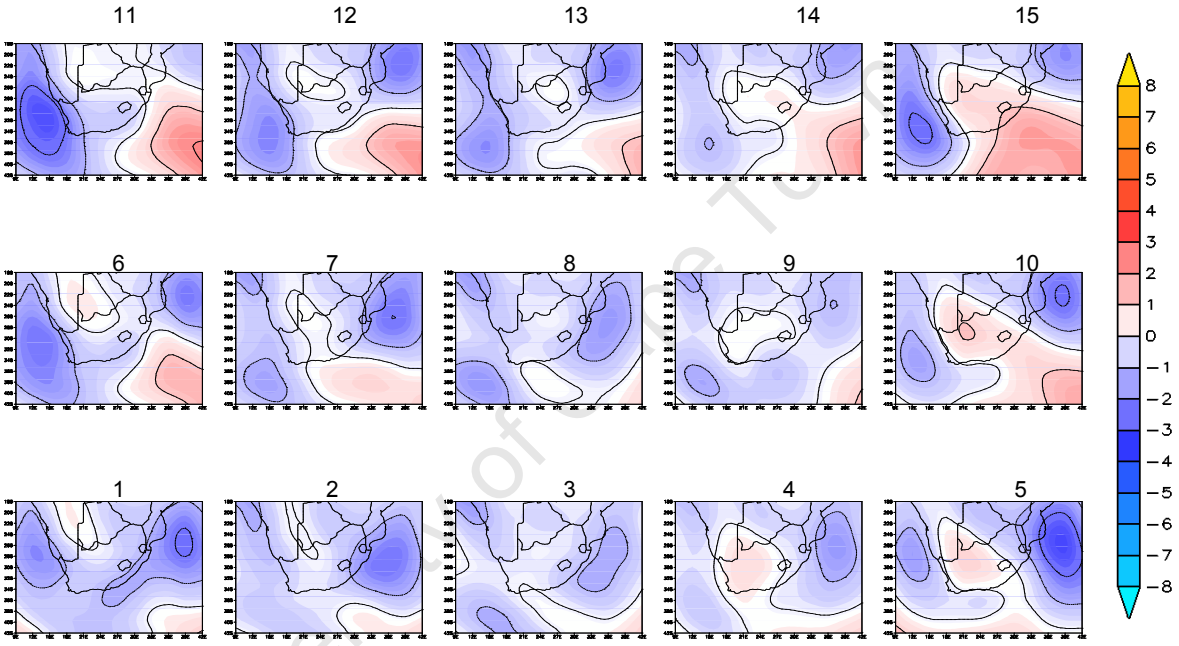


Figure 4.4: Anomaly of GFDL-NCEP for each SOM node for annual average v winds at the surface, contour interval 1ms^{-1}

4.2 MID LEVEL WINDS

Analysis of the 700hPa level reveals information about large scale formations such as the westerly wave and deep pressure cells.

East – West winds:

In the upper air flow, the influence of the westerly wave is much stronger and extends further into the continent from the southern Ocean. Much larger wind speeds are seen over the southern Ocean with the largest values in the outer nodes (typically winter conditions). A weaker eastward flow is seen over northern Nambia, Botswana and Zimbabwe.

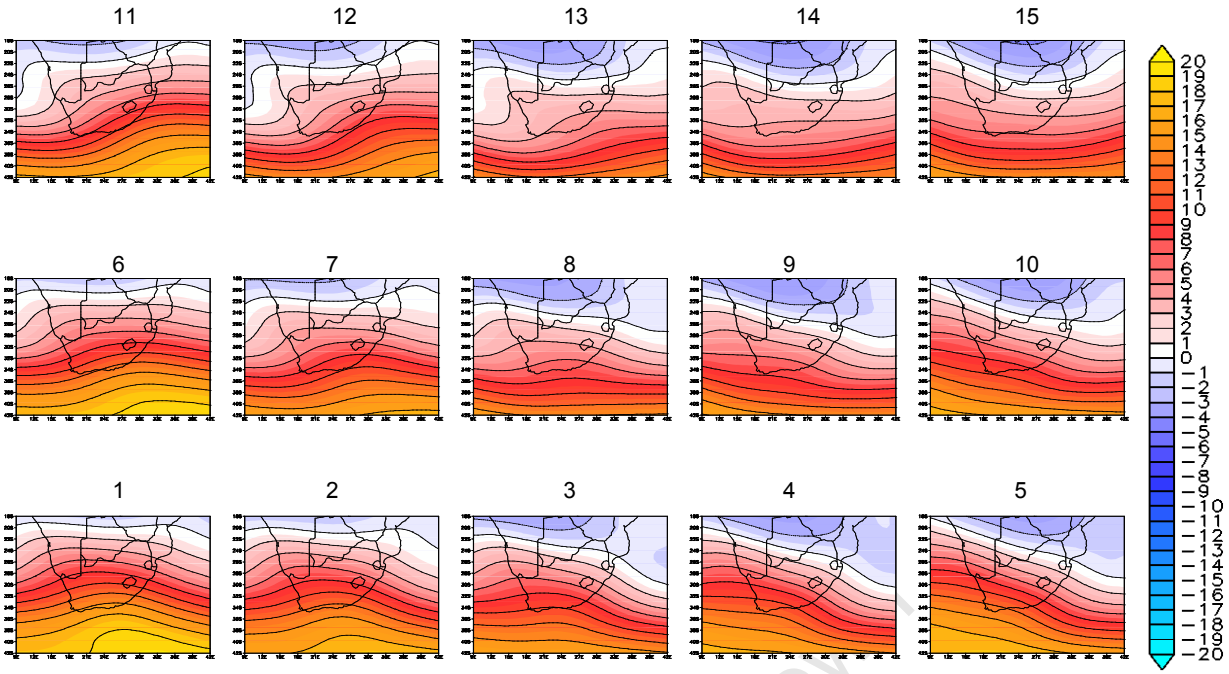


Figure 4.5: NCEP master SOM training showing composite map of annual average mid level winds (reversed normalisation) per node in u component (east-west), contour interval 2ms^{-1}

In the model anomalies of the u wind component (east-west), GFDL (Appendix II, figure B5) and MIROC (figure 4.6) anomalies from NCEP show mostly increased westerly winds, especially in the first and last rows (typically winter conditions) over the Southern Ocean and adjacent interior. In most nodes there is a reduced easterly flow in the top right corner off the coast of Mozambique. MRI (Appendix II, figure B6) has a reduced westerly flow over the Southern Ocean for most nodes, which is strongest in the top row (ridging anticyclones with interior low) and to the north of the domain the nodes have increased easterly flow. Where the westerly wave is dominant (e.g. node 1 and 5) there is stronger westerly flow over South Africa, indicating that the model may over-strengthen the westerly wave.

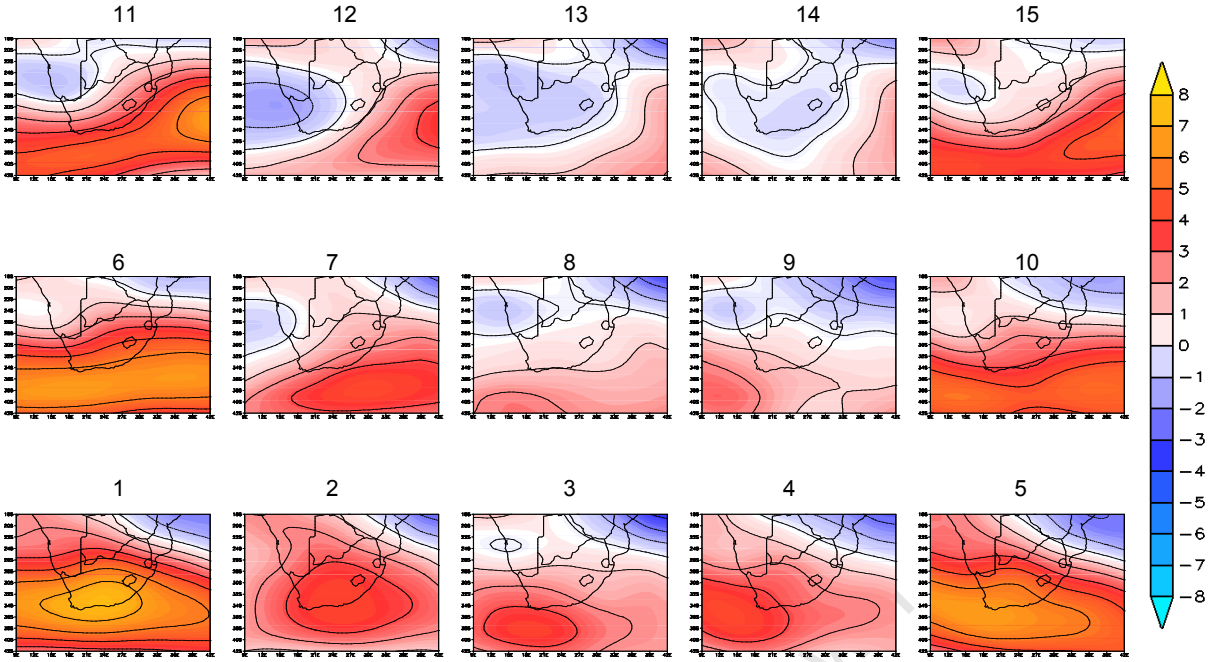


Figure 4.6: Anomaly of MIROC-NCEP for each SOM node for annual average u winds at 700 hPa, contour interval 1ms^{-1}

North – South winds:

The v wind component is much weaker in the upper atmosphere. Where the Indian Ocean anticyclones are dominant (last two columns on right side of SOM) there is a stronger southerly flow over the Indian Ocean and northerly flow on the west coast. Where Atlantic anticyclones are ridging the inverse is true, strengthening southerly flow over the Atlantic. Summer trough features (such as nodes 2, 3 and 4) are associated with strong northerly flow.

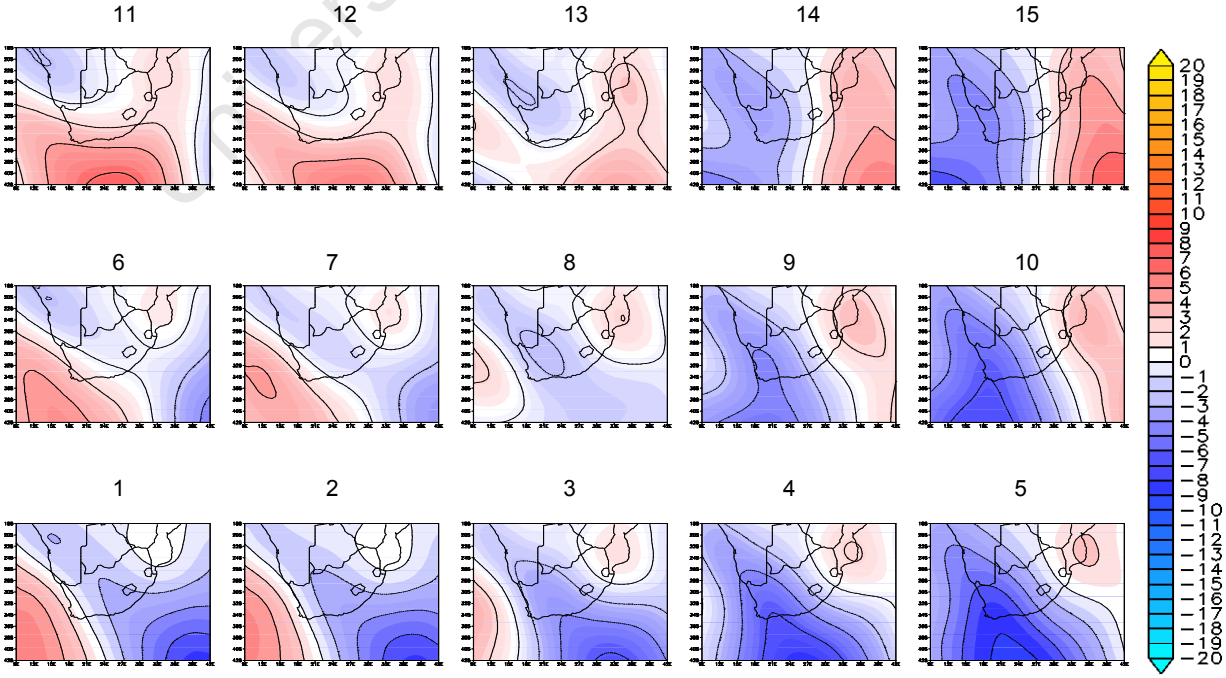


Figure 4.7: NCEP master SOM training showing composite map of annual average mid level winds (reversed normalisation) per node in u component (east-west), contour interval 2ms^{-1}

In the v wind component (north-south), all models show the bottom row (cold fronts and summer troughs) as having strongly reduced northerly flow over the southern Indian Ocean. Figure 4.8 shows that GFDL states with dominant and ridging anticyclones (top row) all have increased southerly flow over the oceans east of the continent, whereas on the Atlantic coast there is a reduction in prevailing winds. In states where the south Indian high is dominant (nodes 9, 10, 14 and15), the reduced northerly flow over the Atlantic is strongest. This indicates that Atlantic highs may be positioned too far south. In cold front states (nodes 1 and 5) there is a strong reduction in northerly flow, suggesting the westerly wave is indeed positioned too far south. These patterns are seen in both MIROC and MRI as well (Appendix II figures B7 and B8).

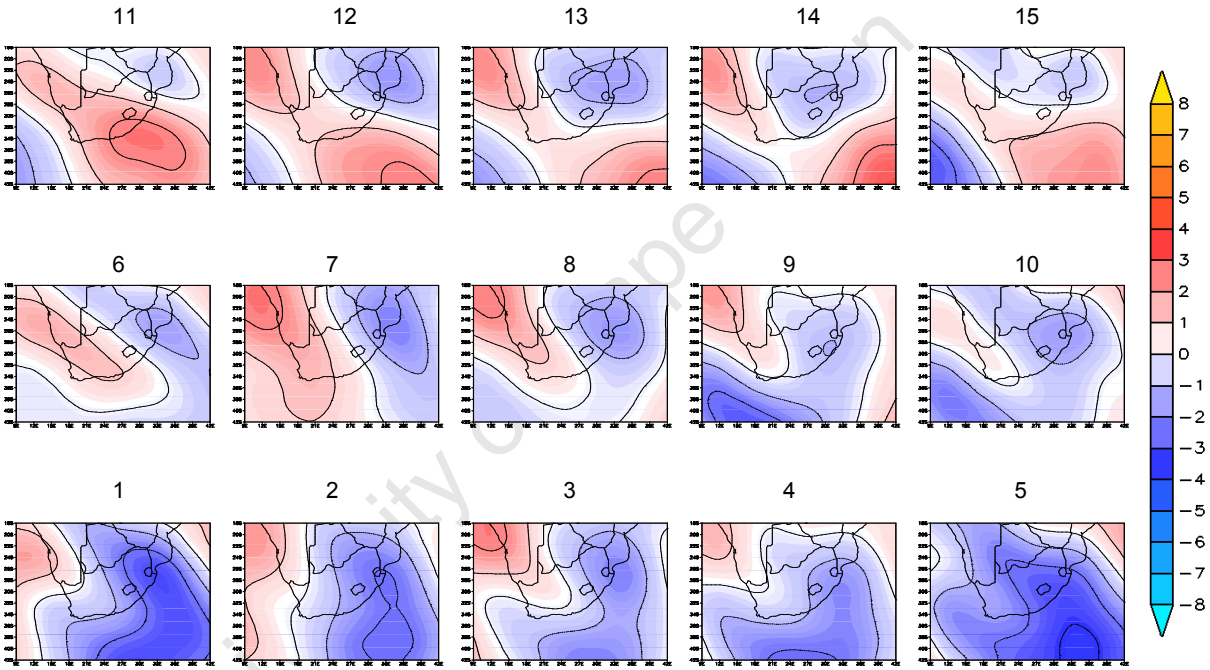


Figure 4.8: Anomaly of GFDL-NCEP for each SOM node for annual average v winds at 700 hPa, contour interval 1ms^{-1}

4.3 SPECIFIC HUMIDITY

The NCEP SOM of specific humidity (figure 4.9) shows that outer nodes (winter states) have low specific humidity values over the continent and in the three middle columns (summer states) there is higher moisture over the continent with lower values to the south. Highest moisture concentration occurs in the middle column over the continent, extending south from the tropics. This indicates that moisture is highest when there is an interior low to feed moisture from the tropics and lowest when the anticyclones are dominant. Interestingly the cold front conditions (nodes 1 and 5) have lower than expected moisture values, which could result in reduced frontal rainfall in the model predictions.

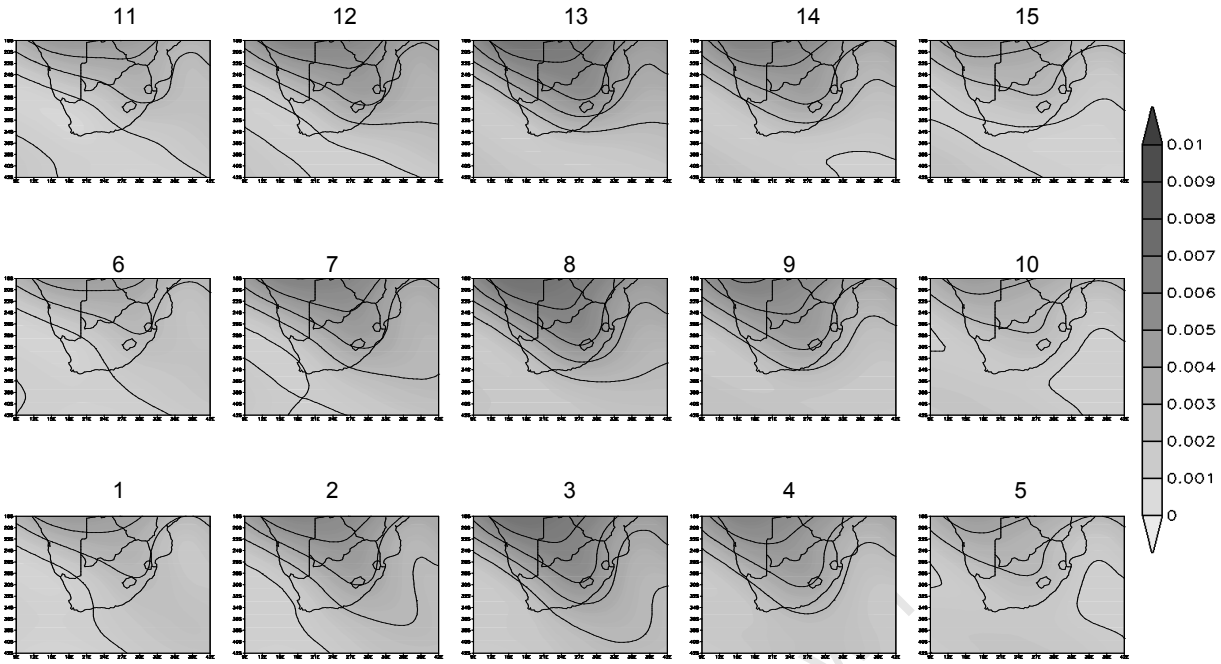


Figure 4.9: NCEP master SOM training showing composite map of annual average specific humidity at 700hPa per node, contour interval 0.001g/m^3

In GFDL (figure 4.10), central nodes have positive anomalies over the west of southern Africa, indicating that models may precipitate too much in the west, while outer nodes have negative anomalies extending from the east coast over the continent with very small positive anomalies over the Indian Ocean. MIROC (Appendix II, figure B9) has the largest positive anomalies which are in the three central columns (summer states) over the whole continent, extending further south in states with strong westerly wave. The rest of the continent and ocean is negative with largest anomalies to the Namibian northeast coast in the outer columns. Negative anomalies are greatest in the outer columns (winter states). This indicates that the models are under predicting humidity in the winter and over predicting summer humidity on the east coast. MRI (Appendix II, figure B10) positive anomalies are much smaller and constrained to northern Mozambique with strong negative anomalies in all nodes off the Namibian coast, extending further east over the continent in the outer nodes. The persistent bias in all models off the Namibian coast indicates that the models underestimate the amount of moisture, which may be related to most of this moisture being in the form of fog which is a difficult precipitating state for the models to depict.

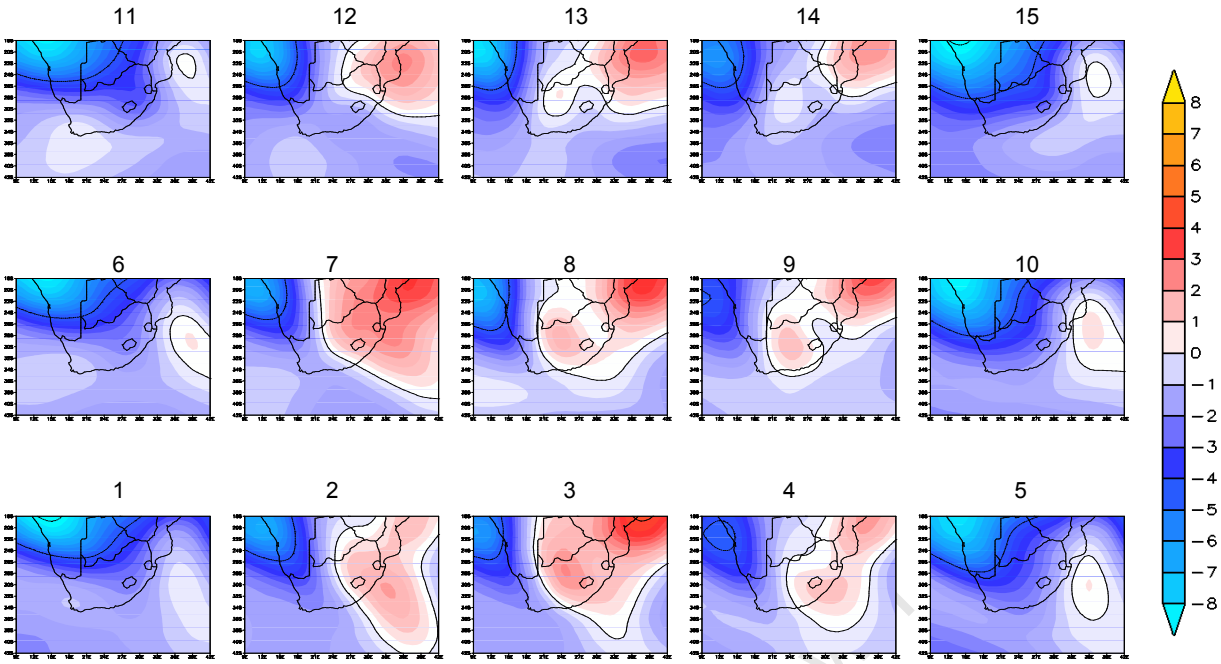


Figure 4.10: Anomaly of GFDL-NCEP for each SOM node for annual average specific humidity, contour interval 1 g/m^3.

From this, all models underestimate specific humidity especially in the two outer columns (strong ridging anticyclones) over northern Namibia, indicating that rains in this region are likely underestimated. Models all overestimate, to varying degrees, the east coast specific humidity in the three central columns (interior low in all nodes). MRI is closest to NCEP and MIROC has the largest positive bias, and therefore probably over simulates rains associated with interior lows.

4.4 TEMPERATURE MEAN

The temperature and precipitation variables are not part of the SOM training and therefore they were extracted separately. Each day of temperature or precipitation data is assigned to the same SOM node that the date was assigned to in the circulation variables. Mean composite maps are generated by averaging all days mapping to that node.

The composite temperature field associated with each SOM node for annual NCEP temperature is shown in figure 4.11, where the outer columns depict cooler states, with the coolest temperatures centred over the Free State province of South Africa. The three central columns all have high temperatures over the Indian Ocean off the coast from Mozambique and over the northern interior.

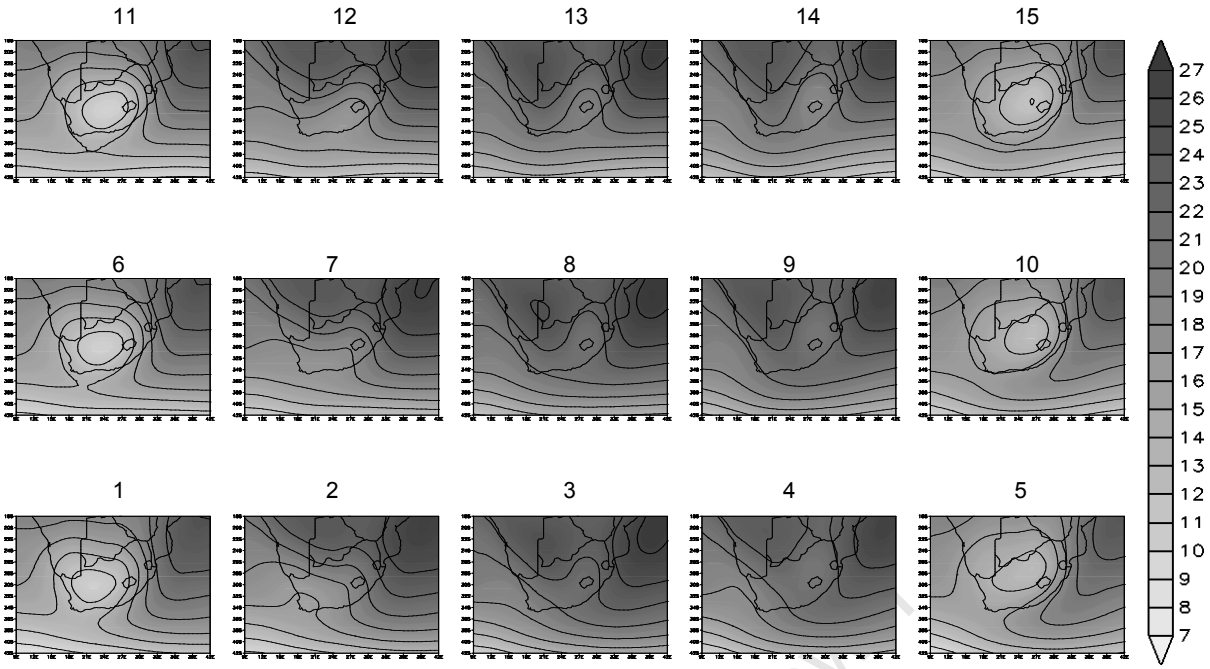


Figure 4.11: NCEP master SOM showing composite map of annual average surface air temperature per node, contour interval of 2°C

MRI (Appendix II, figure B12) generally has larger temperature values than NCEP over the continent and smaller values over the oceans. Central nodes show a negative anomaly over the middle of the continent, which could indicate that the model is failing to represent heat lows in stable conditions. GFDL (Appendix II, figure B11) has predominantly negative anomalies over the whole domain; however, there is a persistent positive bias on the west coast, which in outer nodes spreads over Namibia and Botswana. MIROC (see figure 4.12) also has a persistent positive bias that remains similar in size and intensity in all nodes along the west coast of Namibia, indicating the model is overestimating temperature there and a strong negative bias in the southern Ocean in all nodes, indicating that models underestimate the temperature over this region. The elevated temperatures off Namibia could be associated with the models struggling to accurately represent the cold ocean currents found here, which have a localised cooling effect.

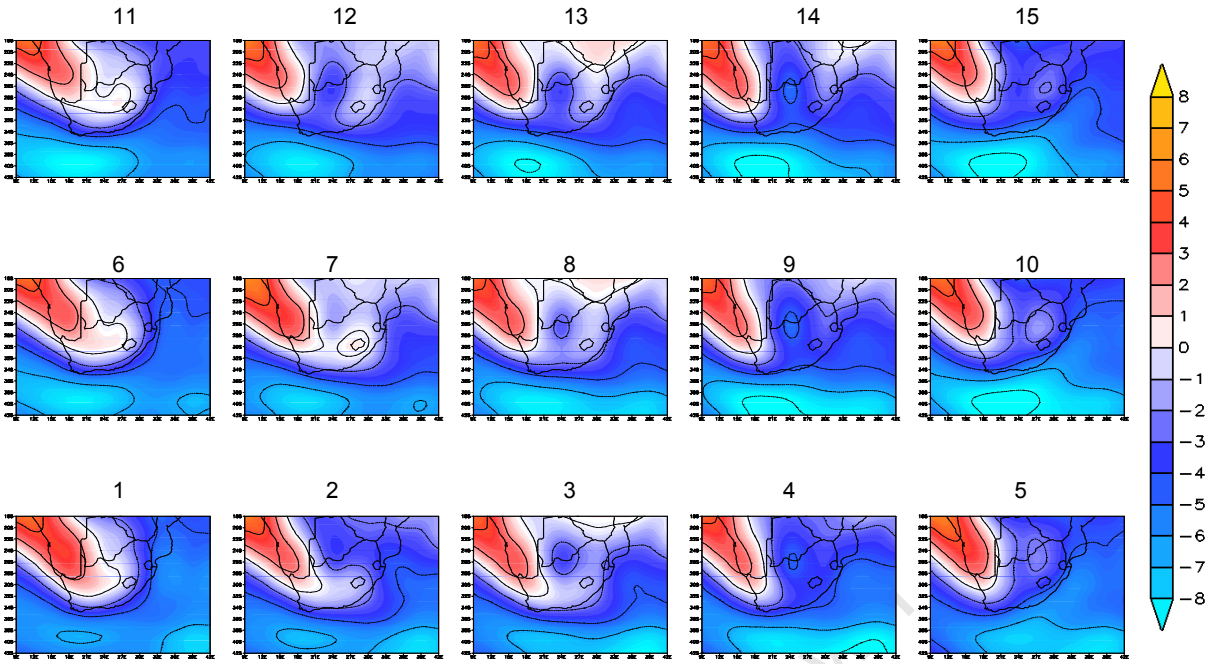


Figure 4.12: Anomaly of MIROC-NCEP for each SOM node for annual average surface air temperature, contour interval 1°C

As the variability within the annual SOM was low and dominated by persistent biases irrespective of season, a closer look at the monthly SOM fields for temperature is discussed below.

Monthly temperatures:

January: The January NCEP SOM (Appendix II, figure B13) has the highest temperatures over the eastern half of the continent, centred over southern Namibia and the Kalahari, as well as over the Indian Ocean off the coast of Mozambique. Warmer temperatures extend further south in the ridging anticyclone states to the left side of the SOM and on the right are more restricted to southern Botswana and Namibia in the trough states. Anomalies are mostly negative and become more so the further south in the domain. Slight positive anomalies over Zimbabwe and in MRI (Appendix II, figure B16) over Zimbabwe and Namibia, may indicate that warming here in the future projections may be slightly elevated. Over the Southern Ocean, the bias is as much as 6°C below NCEP values in MIROC (Appendix II, figure B15) and is present in both the other models too. This shows that all the models make the Southern Ocean and western interior too cold. All nodes for all models had a slightly higher bias over the western interior of South Africa, with generally lower bias to the east for the same latitude. Otherwise the gradation of bias has a linear character, with increased negative bias when moving southwards.

April: The NCEP SOM (Appendix II, figure B17) shows little variation in the temperature patterns for April, with highest temperatures concentrated over the Indian Ocean in all nodes. April anomalies are positive over the western interior especially in MRI (Appendix II,

figure B20) where the bias is 4°C above NCEP. GFDL (Appendix II, figure B18) has the smallest anomalies between the three models. The largest node anomalies are seen on the left side of the SOM for ridging anticyclones. In the weak trough states on the right side, the bias is smaller but spreads into the eastern interior as well. Over the southern Ocean all models have a negative bias, which is strongest in the weak trough states over the Atlantic Ocean particularly. This persistent bias is less in April than it was in January.

July: In all July SOM states (Appendix II, figure B21) the coldest average temperatures are found over the centre of South Africa, with lowest temperatures over the Indian Ocean off the coast of Mozambique. July has the largest positive anomalies, which are mostly over the northern part of the domain centred on Botswana. MIROC (figure 4.13) centres the highest anomalies west of where GFDL (Appendix II, figure B22) and MRI (Appendix II, figure B23) have large anomalies. GFDL's bias is largest in the strong westerly wave states on the right side whereas MIROC and MRI have the greatest positive bias for nodes 3, 7, 9 and 13, all of which have a strong westerly wave to the south with dominant anticyclones. Small negative biases over the Southern Ocean remain, constrained to the west in the strong westerly wave states on the right and over the whole ocean in the dominant anticyclone states to the left. MIROC has the greatest variance within a node. All models fail to capture the cooler temperatures over the central interior. This is probably due to the diurnal cycle not being accurately depicted and therefore the cold night temperatures are not adequately represented in the models.

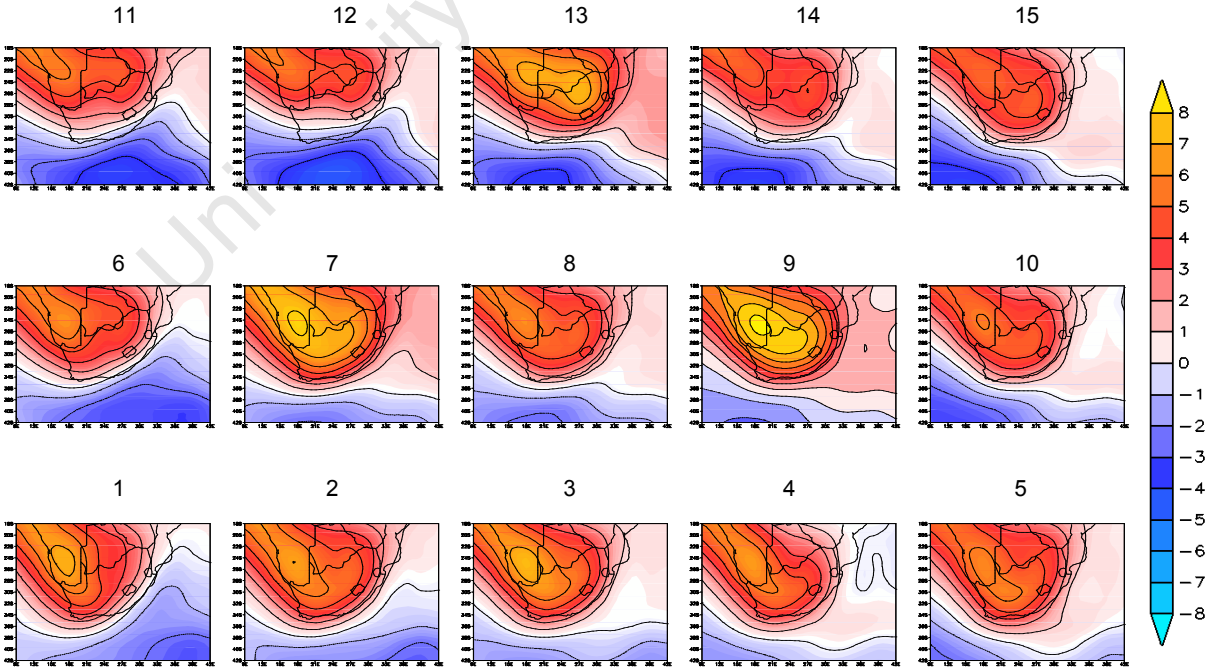


Figure 4.13: Anomaly of MIROC-NCEP for each SOM node for July average surface air temperature, contour interval 1°C

October: The NCEP SOM for October (Appendix II, figure B24) has the highest temperatures over northern Namibia and Botswana with a general decrease in temperature moving southwards. Ridging anticyclones from the Indian Ocean have the highest temperatures. October anomalies are strongly negative for all models. NCEP temperatures are consistently warmer than all models, which suggest that the models are not capturing the warmer temperatures of this season. MIROC (Appendix II, figure B26) has a small positive bias in all nodes off the Namibian west coast. Anomalies are largest over the central interior, with Southern Botswana and the Northern Kalahari, South Africa, as the focal point. MRI (figure 4.14) and GFDL (Appendix II, figure B25) have small negative anomalies over the Southern ocean.

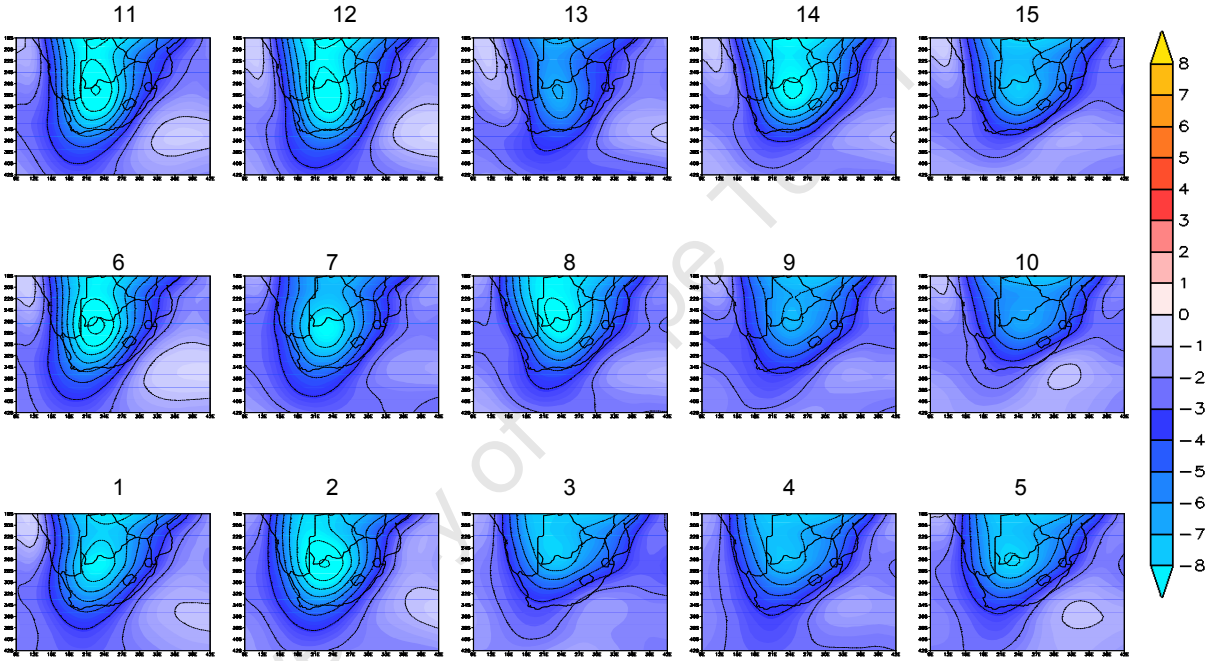


Figure 4.14: Anomaly of MRI - NCEP for each SOM node for October average surface air temperature, contour interval 1°C

Across all time periods for all nodes and models, there is a bias to underestimate temperature in the Southern Ocean. Dependent on season, the models all either under or overestimate temperatures over the continent. This implies that the models may not accurately capture the seasonal variability and boundaries and therefore overestimate the winter minimum temperature and underestimate the summer high temperatures.

4.5 PRECIPITATION MEAN

NCEP annual precipitation follows the specific humidity pattern, with most precipitation occurring over the continent in the three central columns (see figure 4.15). These summer rainfall nodes all have interior heat lows with links to the westerly wave in the bottom row. Anticyclones keep precipitation away and therefore the outer nodes which favour these features (node 6 and 11 favouring strong Atlantic high and nodes 10 and 15 strong Indian

highs) have low precipitation values. Rains to the south coast are associated with the westerly wave and frontal systems connected to them. These states are seen in nodes 1 and 5 and show high rainfall in the south.

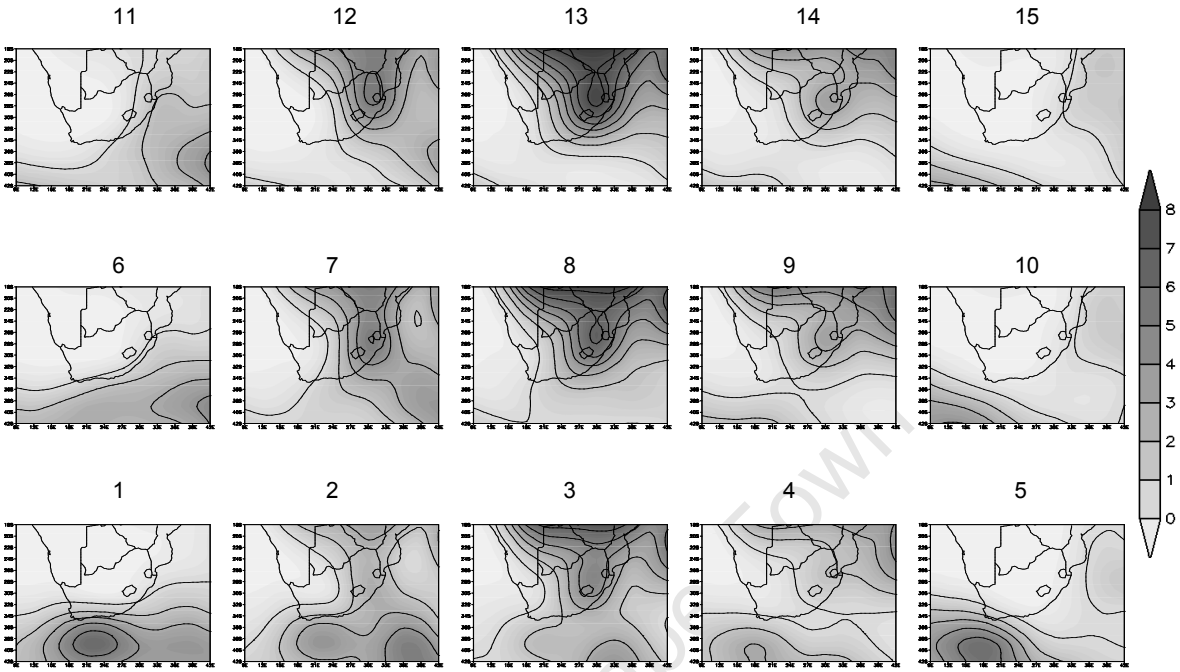


Figure 4.15: NCEP master SOM showing composite map of annual average precipitation per node, contour interval 1mm

MIROC (figure 4.16) has a predominantly positive bias for precipitation except in central nodes where there is a negative bias over Zimbabwe. All nodes have a negative bias over the Indian Ocean off the Mozambique coast. In the bottom row MIROC overestimates precipitation over Namibia, South Africa and Botswana and there is an area of underestimation to the south of Cape Town over the Southern Ocean. This might indicate that cold fronts are slightly underestimated in this model.

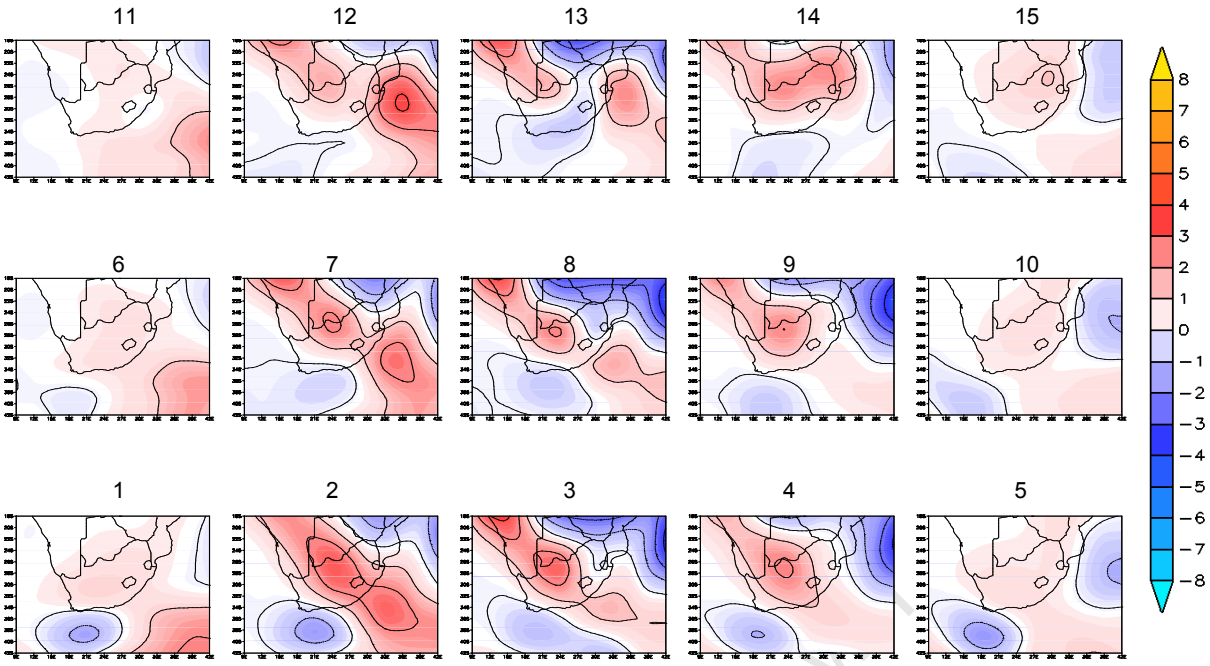


Figure 4.16: Anomaly of MIROC-NCEP for each SOM node for annual average precipitation, contour interval 1mm

MRI (Appendix II, figure B28) has a negative bias in central nodes on the east coast of the continent. In some outer nodes a positive bias is located in the South Indian Ocean, with a dipole negative bias below Cape Town. GFDL (Appendix II, figure B27) has mostly positive bias especially over the south Indian Ocean and in the bottom row displays a similar dipole pattern to MRI and MIROC. In the last two columns a negative bias to the north of the domain is prevalent. The over-simulation of precipitation in the Indian Ocean is not surprising given the same bias identified in specific humidity.

In the next section the monthly precipitation composite anomalies will be explored to find seasonal variations in the precipitation response.

Monthly precipitation:

January: NCEP data (Appendix II, figure B29) shows that the January rains fall mostly on the east coast and in the tropics, while the west coast receives almost no rains at all. January anomalies are negative over most of the continent, indicating that rains are under represented in the GCMs especially over the summer rainfall region to the east. In the connected low states to the right of the SOM the models underestimates precipitation over Zimbabwe and the Northern Province of South Africa, while overestimating rain off the Natal coast. To the left side of the SOM, where there are strong South Indian highs, rains are underestimated over northern Namibia and Botswana as well.

April: In April the NCEP SOM (Appendix II, figure B33) has large precipitation values for the connected trough states off the Natal coast in the bottom right corner. Most other states that

have the westerly wave present have a band of rain to the south west, and another cell off the Mozambique coast. April anomalies are mostly positive and located over the continent. MIROC (Appendix II, figure B34) has the largest positive anomalies and MRI (Appendix II, figure B35) the smallest. In weak trough states on the right side of the SOM there is a large negative anomaly over the southern Ocean. In ridging anticyclones states on the left side of the SOM the negative anomaly is stronger off the east coast, indicating that the models are under estimating rains where there are strong South Indian highs, as seen below in GFDL (figure 4.17).

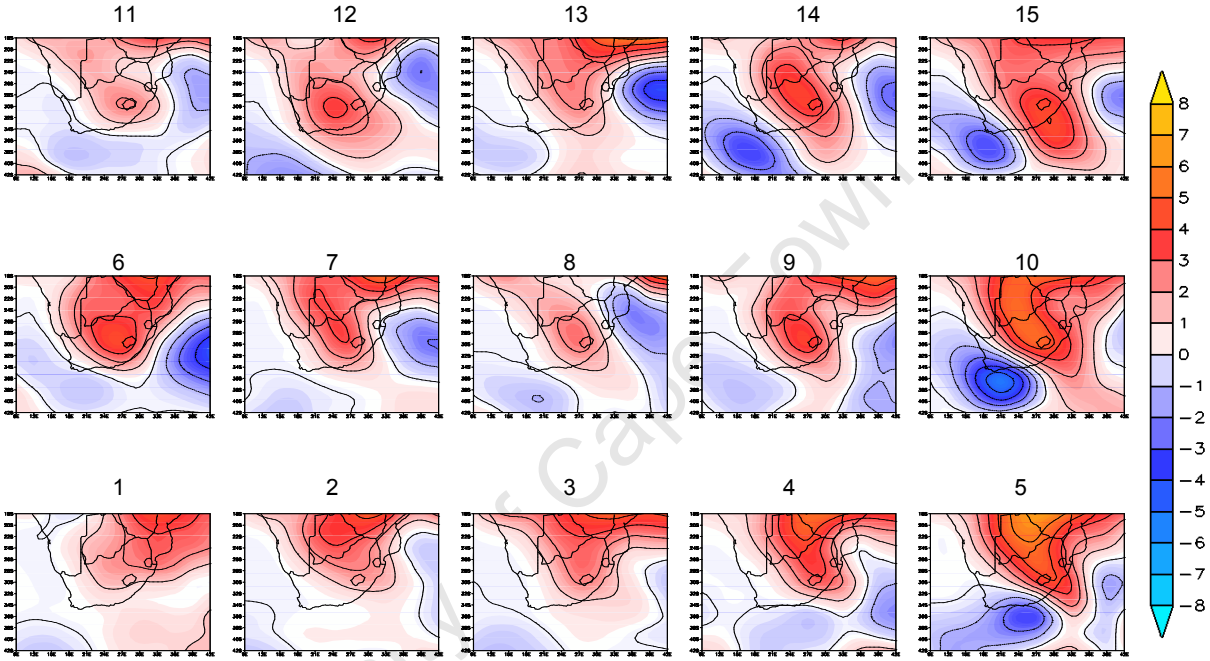


Figure 4.17: Anomaly of GFDL-NCEP for each SOM node for April average composite precipitation, contour interval 1mm

July: The NCEP SOM (Appendix II, figure B36) shows that the right side (strong westerly wave with cold front) has the large rainfall values concentrated on the Cape south coast. The ridging anticyclones to the left of the SOM have drier conditions. Where there are strong South Indian highs, there is less intense rain over the Indian Ocean. For July anomalies, GFDL (Appendix II, figure B37) has generally positive anomalies, strongest off the Natal coast. All models have a consistent bias for the last two columns (strong westerly wave with associated front) which all have a dipole anomaly pattern where models underestimate rain off the Cape south coast and overestimate off the Natal coast (in GFDL and MIROC (Appendix II, figure B38) this positive anomaly extends into the interior). The anomaly could indicate that the models do not get all the orographic rain associated with the Cape coast as they flatten the topography with the large grid size. Over the Drakensberg, the models overestimate rain in states with a summer thermal trough. MRI anomalies are much smaller than either of the other two models for all nodes.

October: The NCEP SOM (Appendix II, figure B40) has the highest precipitating states on the right side, particularly in states with strong interior troughs. Positive anomalies are largest over the South Indian Ocean, with the largest anomalies are found in nodes with troughs or interior lows. The largest anomaly is 9mm for node 3. Over the continent there are small negative anomalies in the top row states with strong South Indian highs, as seen in figure 4.18 for MIROC. This implies that models slightly underestimate rainfall associated with ridging Indian highs and overestimates rain off the Natal coast when there is a trough extending south from the tropics.

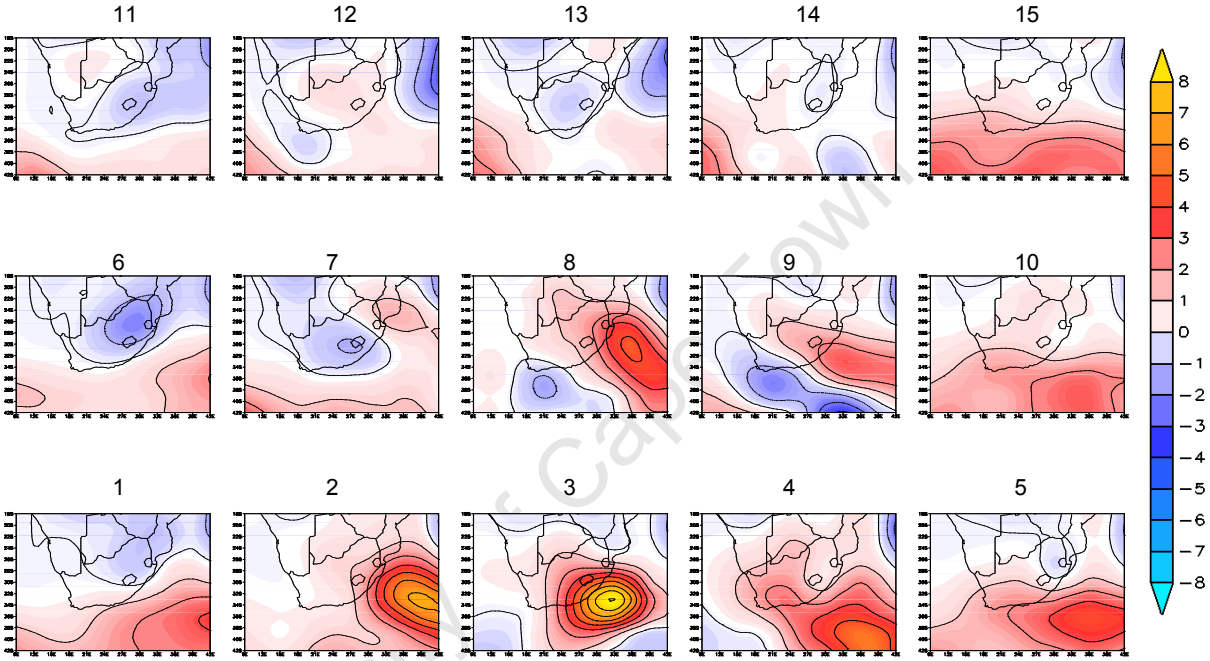


Figure 4.18: Anomaly of MIROC -NCEP for each SOM node for October average composite precipitation, contour interval 1mm

4.6 POSSIBLE DYNAMIC EXPLANATION FOR BIASES

In all the anomaly plots, it seems apparent that the models are more similar to each other than to NCEP, with some persistent biases identified across models, especially in the monthly temperature SOMs. Estimation of surface level variables is generally more complex than simulating upper air features as topographic effects come into play.

Circulation differences identified:

- increased westerly flow over the Southern Ocean in surface and mid level winds, which is strongest over South Indian Ocean when South Indian high present;
- persistently reduced easterly winds in all nodes, all models off Mozambique's north coast in surface and mid level winds;
- stronger westerly wave in models;
- reduced southerly flow in the Atlantic Ocean and at times the adjacent western half of continent, especially when South Atlantic high dominant in surface winds;

- reduced southerly flow over the Indian Ocean and Natal coast in surface winds;
- increased northerly winds over the Atlantic Ocean and western half of the continent in surface level winds (results in southwards displacement);
- westerly wave positioned too far south, especially in association with cold fronts in mid level winds;
- underestimate specific humidity in strong ridging anticyclone states over northern Namibia;
- overestimate specific humidity along the west coast for interior lows.

All the models have a bias towards south westerly flow, indicating that the high pressure cells are shifted further south than they should be and the westerly wave is stronger in the models that in observations. A persistent anomaly over the Mozambique coast to under represent the easterly flow could be linked to the shift south of the high pressures, thus reducing onshore easterlies associated with the anticyclonic flow around a high. This would also explain the surface level increases in easterly flow just below Mozambique which is strongest when Indian high is dominant over this region. With regards to specific humidity, the models overestimate along the west coast and underestimate on the east coast leading to larger anomalies for precipitation on the west and lower values to the east. The moisture that is brought on shore with the Indian highs circulation would therefore be displaced southwards and can explain the reduced precipitation over this area in almost all nodes in the annual SOM.

Temperature differences identified thus far:

- Southern Ocean surface air temperature is too cold in all models for all nodes in all seasons;
- under estimate the summer high temperatures; and
- over estimate winter minimum temperature.

This indicates the models have a cooling bias in general. This is important for future projections as the increased temperatures are likely to be even greater due to the inherent cool bias.

Precipitation differences identified thus far:

- over simulation of precipitation in the Indian Ocean;
- under estimate summer rainfall over eastern half of continent;
- slightly underestimate rainfall associated with ridging highs; and
- over estimate rain off the Natal coast in trough states extending south from the tropics.

From these trends, it seems most of the models struggle to accurately reflect the dynamics over the eastern half of the continent and South Indian Ocean. This area receives mostly convective rainfall therefore the convection scheme used may not be the most appropriate. However as the specific humidity values were overestimated in the model over this area other dynamics could also play a role.

At this point it may be useful to note some of the main differences between the three models themselves. MIROC and GFDL surface and mid level circulation are similar, with most states depicting an increased westerly flow over the southern Ocean and a persistently reduced easterly in all nodes off Mozambique's north coast. MRI westerly surface flow is weaker, with stronger easterly winds for dominant South Indian high states. In the mid level, however, MRI has a more dominant westerly wave over South Africa, possibly over-strengthened in mid level and therefore too weak at surface. The north-south component is consistent between models; however, MIROC has a greater reduction in southerly flow over the Atlantic Ocean than the other models especially for strong ridging anticyclone states.

GFDL precipitates too much in the west in summer states, while winter states have reduced humidity extending from the east coast over the continent. When the South Indian high is dominant, GFDL tends to underestimate rains to the north of the domain. MIROC's summer states are the most humid over the whole continent, extending further south in states with a strong westerly wave. Precipitation is overestimated over Namibia, South Africa and Botswana and there is an area of underestimation to the south of Cape Town over the Southern Ocean, possibly indicating cold fronts are slightly underestimated in this model. MRI has lower humidity values, constrained to northern Mozambique with less rainfall in all nodes off the Namibian coast, extending further east over the continent in winter states. In summer states MRI underestimates rainfall on the east coast.

MRI is the warmest model with the highest temperature values over the continent and smaller values over the oceans. Summer states have lower temperatures over the middle of the continent, which could indicate that the model is failing to represent heat lows in stable conditions. GFDL has lower temperatures over the whole domain, and both GFDL and MIROC are colder over the Southern Ocean. Both models are overestimating temperature on the west coast, which in winter states spreads over Namibia and Botswana. This could be associated with the models struggling to accurately represent the cooling effect of the cold ocean currents offshore.

In the next Chapter, the frequency bias correction function will be applied. From these results we can see how these biases are reduced by using the weighting technique.

CHAPTER FIVE

Mean multi-model climate projections

In the previous two chapters, biases were identified for each model in terms of the frequency at which the model represents particular synoptic states, transitions between states, the regional biases in positioning of features and the differences in magnitude for a given variable and state. In this chapter, three techniques are employed to adjust the mean baseline using the metric of frequency in different formulations. The objective is to see what impact the frequency weighting will have on the results and whether the weighted results are closer to the observed for control simulations, thereby correcting for biases identified in the previous chapters. Multi-model results for some formulations are also demonstrated, whereby the relative strengths of each model are combined. Future projections are considered in conjunction with an adjusted mean baseline.

The first technique uses the observed (NCEP) data frequency per node and applies the observed node frequency to each composite model node for the variables of temperature and precipitation and summated to produce a model mean, which is then compared to NCEP. This technique results in an adjustment of the baseline, but it is shown that it cannot be used alone to correct the projections due to the large differences in the node composites.

The second technique attempts to rank the overall performance of each model and the resultant model weight is then applied to the mean of the data set. The models are combined to produce an average multi-model mean that reflects each model proportionately to the weighting. The correction metric used here is the frequency anomaly between the model control frequency (*Cfr*) and observed NCEP frequency (*Nfr*) for each node and summates these differences to create a single model weight. The weight represents the general accuracy the model has in representing all synoptic processes. The weighting for each model is then applied to both future and control data for the variables of temperature and precipitation. The weighted daily data are summated for all models and divided by summated model weights, producing a single “multi-model map mean” for all models of control and future data variables. The idea behind this is to try compensate for differential bias and to allow for comparison between models independent of their individual systematic biases. This technique also gives models with a higher weight a larger component in the final multi-model mean. The multi-model map mean is also compared to the unweighted multi-model average to assess the improvement in skill.

The final method explores the difficulty of correcting the baseline state and therefore uses the NCEP data as a control baseline, adding a weighted multi-model future anomaly. This

attempts to remove the systematic bias entirely by using the observed baseline. A weighting of the future change component is then added to this, which should produce more realistic approximations of the future state. The metric used is the synoptic level anomaly difference between the model control frequency (Cfr) and observed NCEP frequency (Nfr) for each node and produces a weighted multi-model mean for each synoptic state. The weighting for each model is applied only to the mean climate change signal (future – control) and then added to the NCEP (observed) baseline. The result is then compared to the result of using the unweighted multi-model average added to the NCEP baseline.

5.1 OBSERVED FREQUENCY ADJUSTMENT OF MODEL MEANS

In this method, we see how large an influence frequency has on the mean state. The premise is that changing model frequencies to observed frequency (making the number of observations the same) will result in a more accurate model means for individual variables. This will help to determine how much of the model bias is related to mismatched frequency of occurrence and how much is dependant on the model mean itself. The weighting functions used in subsequent sections all use the metric of frequency in some way; therefore it is important to understand the influence of this factor.

The model mean for each node is multiplied by the NCEP frequency for that node. All nodes are summated to produce a map mean for each model that uses the NCEP data as a reference frequency to allow for bias adjustment, while retaining the models individual dynamics. In this way the bias in the model mean is compensated for against a reference of the NCEP ("observed") data. In the annual data the effects of this technique would be expected to be less apparent as frequency differences between models and observations are generally smaller than for the monthly data. However, the resultant anomaly pattern for temperature and precipitation show that the result is still much closer to the original model mean than NCEP. As the models composite node means for individual variables of temperature and precipitation were not normalized and were not input fields in the SOM training, this is not wholly unexpected. Therefore, we should still expect large differences between the observed and resultant composite fields. The differences identified in Chapter Four between observed and model node means are therefore largely retained, adjusting the map mean by frequency but not altering the individual node's positioning of features. An approximation of the future data can be generated by adding the delta change value onto the corrected base line. However this is not shown here, given the limited effectiveness of this method.

5.1.1 Temperature

The differences shown in Chapter Four between NCEP and GCM composite means for each node were as high as 7°C for some geographic locations and synoptic modes.

However, this section looks at a variable's mean anomaly for the whole data set, which was not explored in previous chapters. To provide an idea of the scale of correction we are hoping to achieve, the difference between NCEP and model averages is shown in figure 5.1. The maximum temperature anomaly between NCEP and the models is around 3.5°C. All these biases were explored in Chapter Four and therefore, only the annual anomaly is shown here. However, the positioning of the anomaly is interesting; with all models having a negative bias (NCEP values are warmer) over the Southern Ocean, indicating a cold bias which is greatest in MIROC. GFDL and MIROC both have negative anomalies over the continent and all three have a positive bias off the Namibian coast of varying strength, which as discussed in Chapter 4, probably indicates that the cold ocean currents here are not accurately simulated. MRI has a positive bias over the east half of the continent not shown in the other two models.

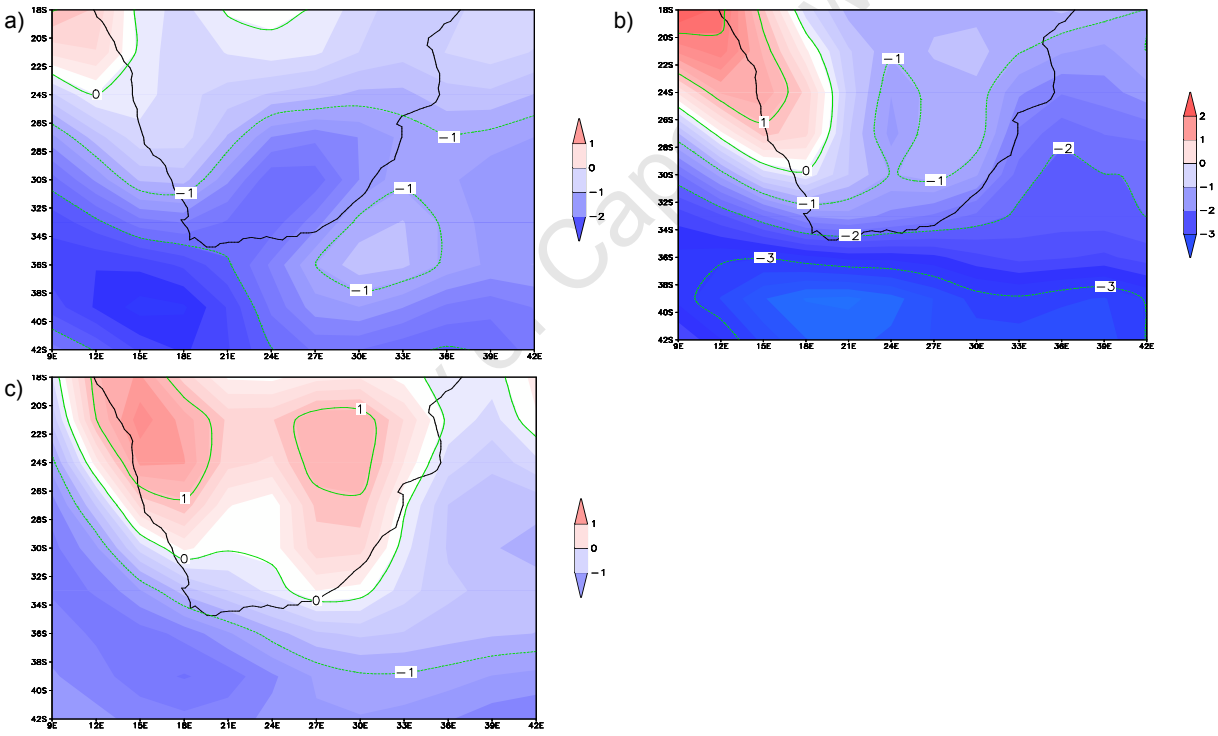
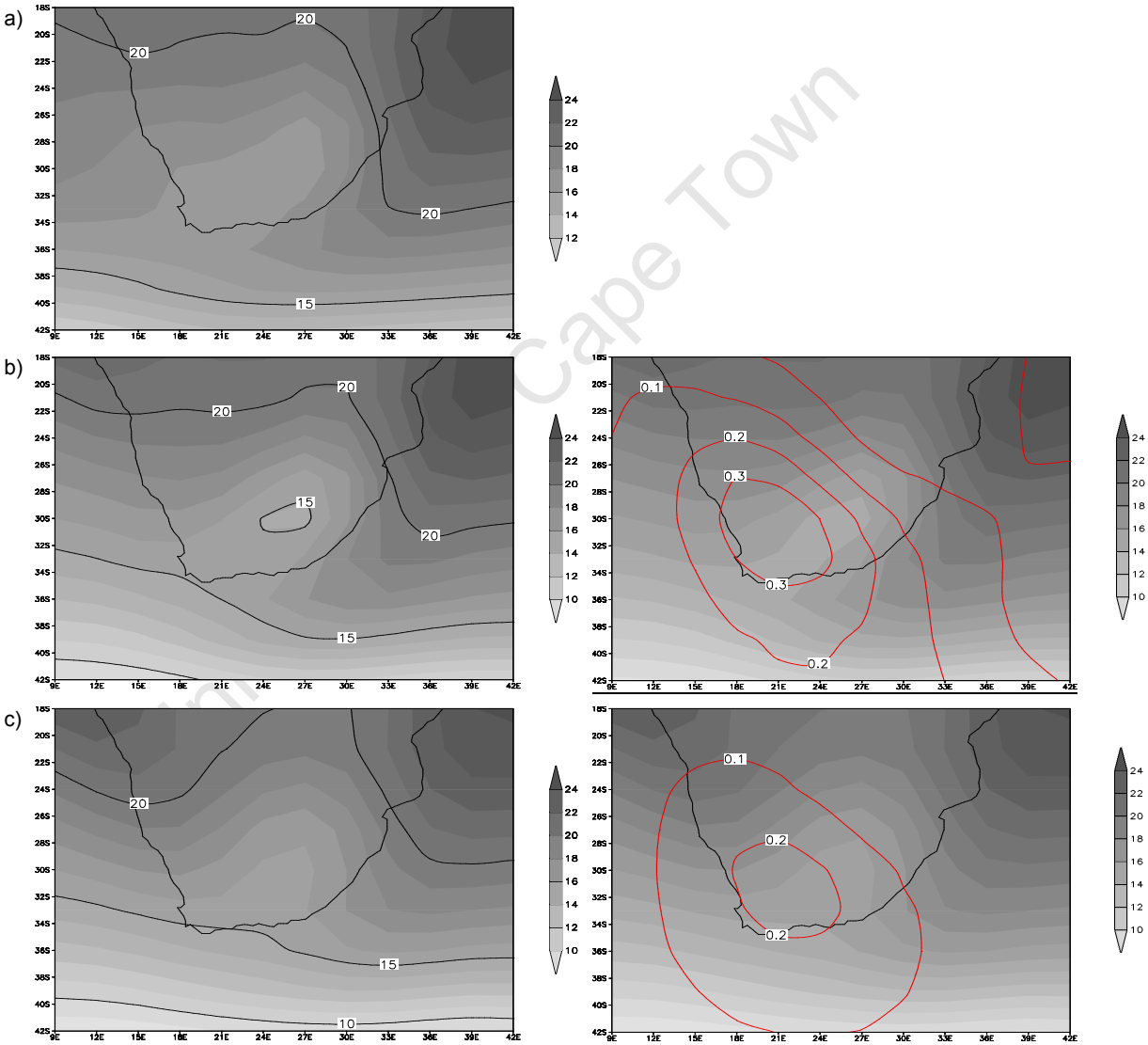


Figure 5.1: Anomalies of NCEP annual (1979-2000) mean temperature in °C from a) GFDL, b) MIROC, c) MRI

The results in figure 5.2 depict NCEP mean temperature in the left column's first block and below it the three models adjusted temperature, using the NCEP frequencies instead of the models. The variables mean for each node is assigned NCEPs frequency. All nodes are then summated and divided by the total number of nodes to produce an adjusted variable mean. The right side shows the original average temperature values in grey scale, overlaid with contour lines indicating the anomaly between weighted and the original values.

There is generally little change between NCEP and corrected data as apposed to NCEP and control. The annual adjustment is of the order of up to about 0.5°C, compared to the 1.5-3.5°C range required to bring the model data in line with observations. Therefore, the correction is accounting for roughly only a tenth of the difference between the NCEP and GCM. However, this should be seen as a first step towards a better solution, in which we see how large a role frequency has in determining the baseline state. MIROC and GFDL have a positive correction over the South Atlantic Ocean and Western Cape that is in line with increasing these values to the observed values. Similarly, MRI has a negative anomaly over the east continent to correct for the overestimation previously noted there.



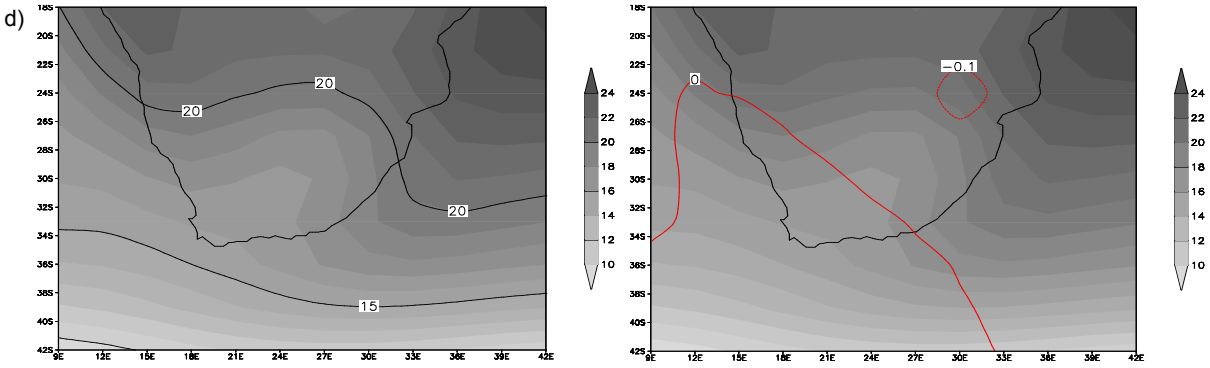


Figure 5.2: ***Frequency corrected annual temperature:*** *a) NCEP annual mean temperature in °C compared to annual control frequency corrected temperature (left) and original control mean temperature in shading with difference between corrected and mean in overlaid red contours (right) for b) GFDL, c) MIROC, d) MRI*

This simple correction tool has moved the models mean closer to the NCEP observed climate. The adjustment is still quiet small but the positioning of the gradients is brought closer to the observed distribution.

5.1.2 Precipitation

The non-corrected average anomalies shown in figure 5.3 once again show that different models have different gradients of intensity in their biases although there are some commonalities. All have underestimated precipitation to the north of the domain and GFDL and MIROC have overestimated rains over the South Indian Ocean and over South Africa. MRI has stronger negative biases in the South Atlantic and over Zimbabwe and Mozambique.

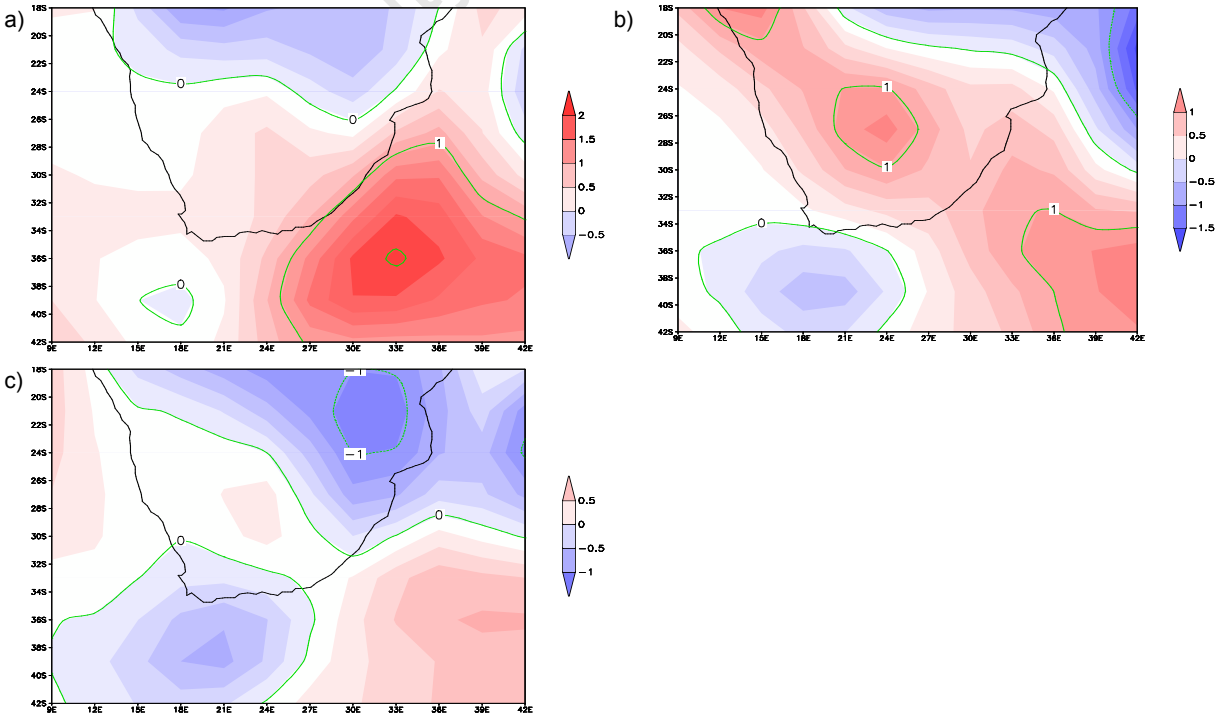


Figure 5.3: *Anomalies of NCEP annual mean precipitation (mm/d) from a) GFDL, b) MIROC, c) MRI*

The same frequency correction technique is applied to the precipitation data as shown in figure 5.4, with similar results. The adjustment is up to 0.3mm which is about a fifth of the anomaly differences. All three models are adjusted to lower values to varying degrees over the South Indian Ocean, relative to their bias (i.e. GFDL has largest correction as it had largest bias).

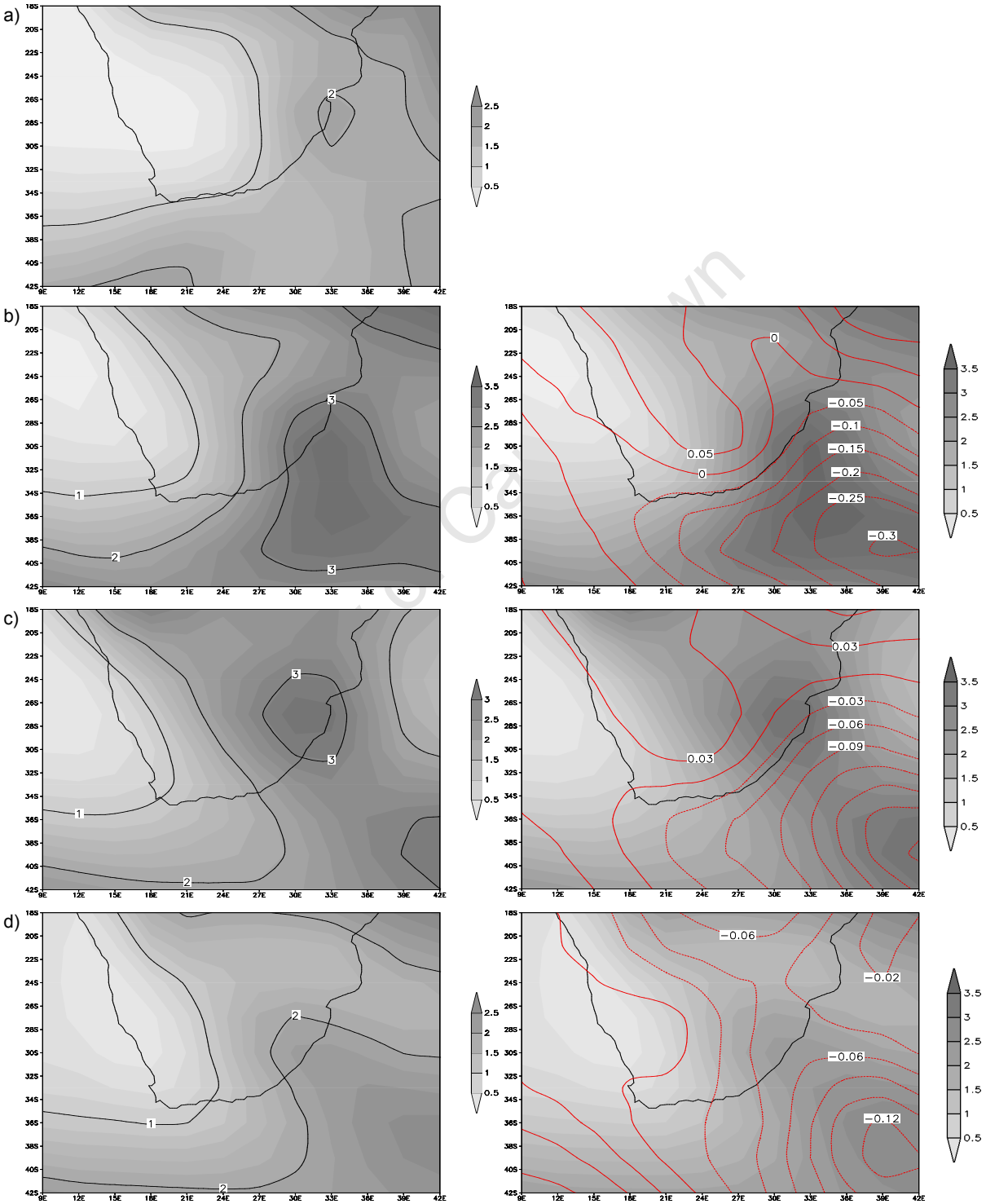


Figure 5.4: ***Frequency corrected annual precipitation:*** *a) NCEP annual mean precipitation (mm/d) compared to annual control frequency corrected precipitation (left) and original control mean precipitation in shading with difference between corrected and mean in overlaid red contours (right) for b) GFDL, c) MIROC, d) MRI*

Monthly frequency corrected plots are shown in Appendix III, figures C1-C8. The adjustments are generally of similar magnitude and in some instances negligible, such as July temperatures in GFDL (figure 5.5), where the overall pattern of the map means is indistinguishable from NCEP as the contour interval in the plot is 1°C and the differences are all at the fractional level. This implies that although the frequencies were different, the mean dynamics were not substantially different.

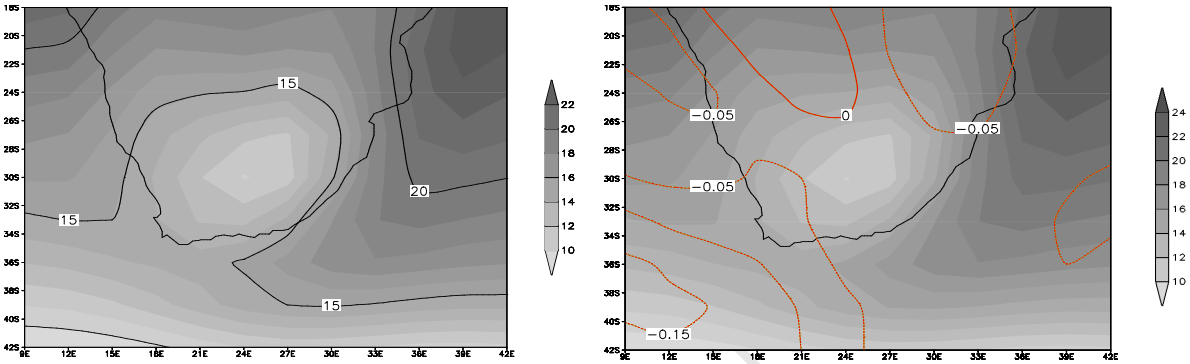


Figure 5.5: GFDL corrected July temperature in °C (left) and GFDL unweighted July temperature overlaid with the difference between weighted and unweighted model results in red (right)

In some instances the results show substantial differences, such as October precipitation in MIROC (figure 5.6) where differences are as high as 0.8mm over an area where anomaly from NCEP was up to 2mm, therefore correcting for roughly a third of the bias.

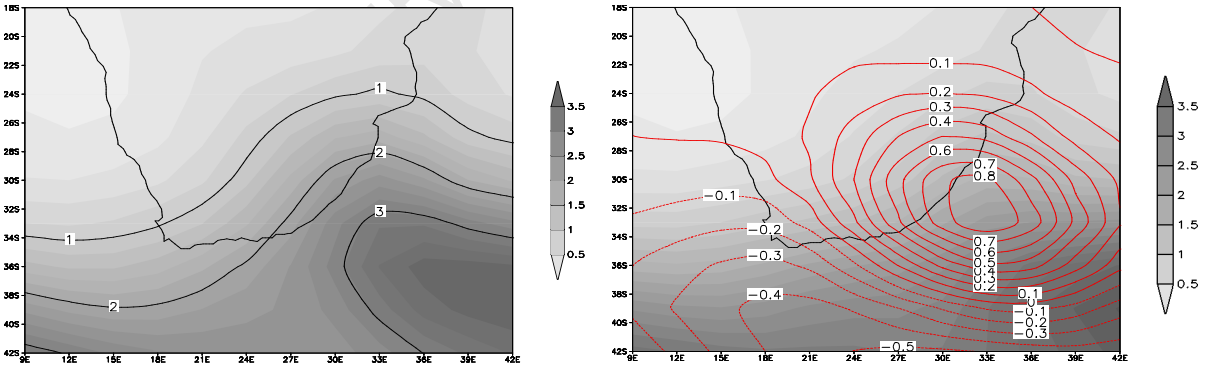


Figure 5.6: MIROC corrected October precipitation in mm/d (left) and MIROC unweighted October precipitation overlaid with the difference between weighted and unweighted model results in red (right)

The frequency adjustment is producing only subtle deviations from the unweighted data. The limitation of the frequency adjustment technique is in most cases the mean model composite for each node is distinctly different enough from the observed mean that this difference eclipses the small correction from frequency Therefore, although the frequency adjusted map mean is a closer match to NCEP, it cannot be corrected further as the gradients within each node are not synchronous, as depicted in the sometimes large

anomalies depicted in figures 4.4 and 4.5. For this reason the technique is not sufficient on its own to correct the map mean.

The future projections cannot be adjusted using this method and therefore the techniques usefulness is limited to adjusting the control frequencies. What can be done is that the anomaly between future and control can be added to the adjusted control which should improve the accuracy of the individual models mean average. This idea is not perused here as in the next sections better tools are identified.

5.2 WEIGHTED MULTI-MODEL MEANS

This method looks at overall model performance in the form of a single weight for each model. The idea is to see if one model is consistently better overall than all the others for each of the 5 time series analysed. The model with the largest weight has the most influence on the multi-model mean solution in this method.

A single model weight is derived for each model for each of the five SOM mappings (annual, January, April, July, October); using the frequency anomaly between the model control frequency (Cfr) and observed NCEP frequency (Nfr) divided by 100, for each node and summating these differences to create a single model weight.

NCEP's frequency is differenced from each model to produce an absolute value of difference per node. The absolute value is used as we are interested in the magnitude of the anomaly but not the direction (i.e. if the value isn't the same as NCEPs it gets a lower weight whether it overestimates or underestimates). Each nodes frequency difference between NCEP and the model is divided by 100, so that the difference is always less than one. All these node anomalies are summated and subtracted from one to produce a single weight for each model that approaches one when most correct and approaches zero when most incorrect. The weights are created using equation E.2 from Chapter 2 as shown below.

$$\text{(E.2)} \quad Wi = 1 - \sum_{i=1}^{n} \frac{|Cfi - Nfi|}{100}$$

Table 5.1 shows each of the model's weights for the different time periods. As some of the frequency differences are larger in monthly circulation data, there are some weights that are much smaller than those in the annual weights (e.g. GFDL July weight).

	GFDL	MIROC	MRI
Annual	0.83	0.89	0.89
January	0.77	0.76	0.81
April	0.67	0.65	0.63
July	0.49	0.78	0.75
October	0.74	0.46	0.68

Table 5.1: Weights (rounded off) for each model derived from the fractional difference between observed and model frequencies for each SOM node, combined into a single overall model weight. Values approaching 1 have the greatest accuracy.

In the annual weights there is little difference between models and all three perform well at the aggregate level. In the monthly weights there are some interesting variations. January has the highest monthly weights (thus making it the closest to reality) and April has consistent values although accuracy is not as high. In July and October, however, there are obvious winners and losers. GFDL has the lowest weight in July but performs best of the three in October, and *vice versa* in MIROC. MRI performs consistently well amongst all seasons but is only the best in January. This confirms pervious studies findings that no one model is more accurate than all the others all the time (Giorgi and Mearns, 2002). Using a multi-model solution is therefore advisable to maximize the strengths of all the models used.

The time series average for each grid cell is multiplied by its respective model weight. When the three models are summated and divided by the summated weights, they produce a single frequency corrected map mean of temperature or precipitation as shown in E.3 below.

$$\text{(E.3)} \quad MutiM = \frac{\sum_{m=1}^{3} Cave \times Wm}{\sum_{m=1}^{3} Wm}$$

Each model is thus corrected for its own overall frequency bias and its contribution to the multi-model mean is scaled by its relative accuracy. The future and control corrected map mean for each model is then compared and the anomaly between the corrected and uncorrected averages assessed.

When weights are similar amongst models there will be little improvement in the multi-model solution, as in the annual data for example. This is because the similar size weights create a map that is very similar to the average map mean (dividing by a common denominator) and therefore no benefit is achieved. In these instances the observed frequency technique from section 5.1 is more effective.

5.2.1 Temperature

In the annual data, the correction function is very weak as the model weights are so close that control and future weighted map means are virtually indistinguishable from the average. In figure 5.7 the shaded image represents the corrected function, overlaid in contours by the correction anomaly (difference between corrected and uncorrected). Corrections are of the order of one hundredth whereas the corrections needed are three times greater.

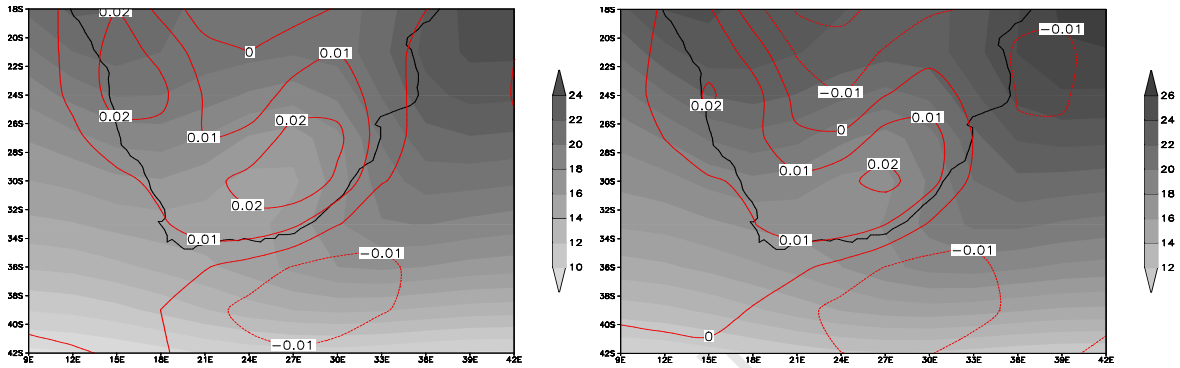


Figure 5.7: ***Multi-model annual temperature****: created by weighting each models average control (left) and future (right) temperature in °C overlaid with the anomaly between weighted and average values*

In figure 5.8 the monthly weighted map means are shown. In some cases the mean is closer to the observed distribution. However, in other examples especially where all of the models weights are similar (such as in January and April), the multi-model mean is unable to improve on the average of all three models. In July and October the weighted multi-model mean shows an improvement on the average and will be discussed further below.

The July weighted map mean increases temperatures over the continent, especially over Namibia and the Free State province of South Africa. Decreases (indicating weighted mean has lower values) are over the South Indian Ocean and Northern Botswana. Future projections show a similar pattern over the continent with a slightly larger increase in temperature but over the southern Ocean and Northern Botswana the decreases are smaller.

For October there are small decreases off the Namibian coast, with increases over the Southern Ocean and central and northern parts of the continent. The same pattern is seen in the future projections with stronger positive anomalies over Botswana indicating increased temperatures.

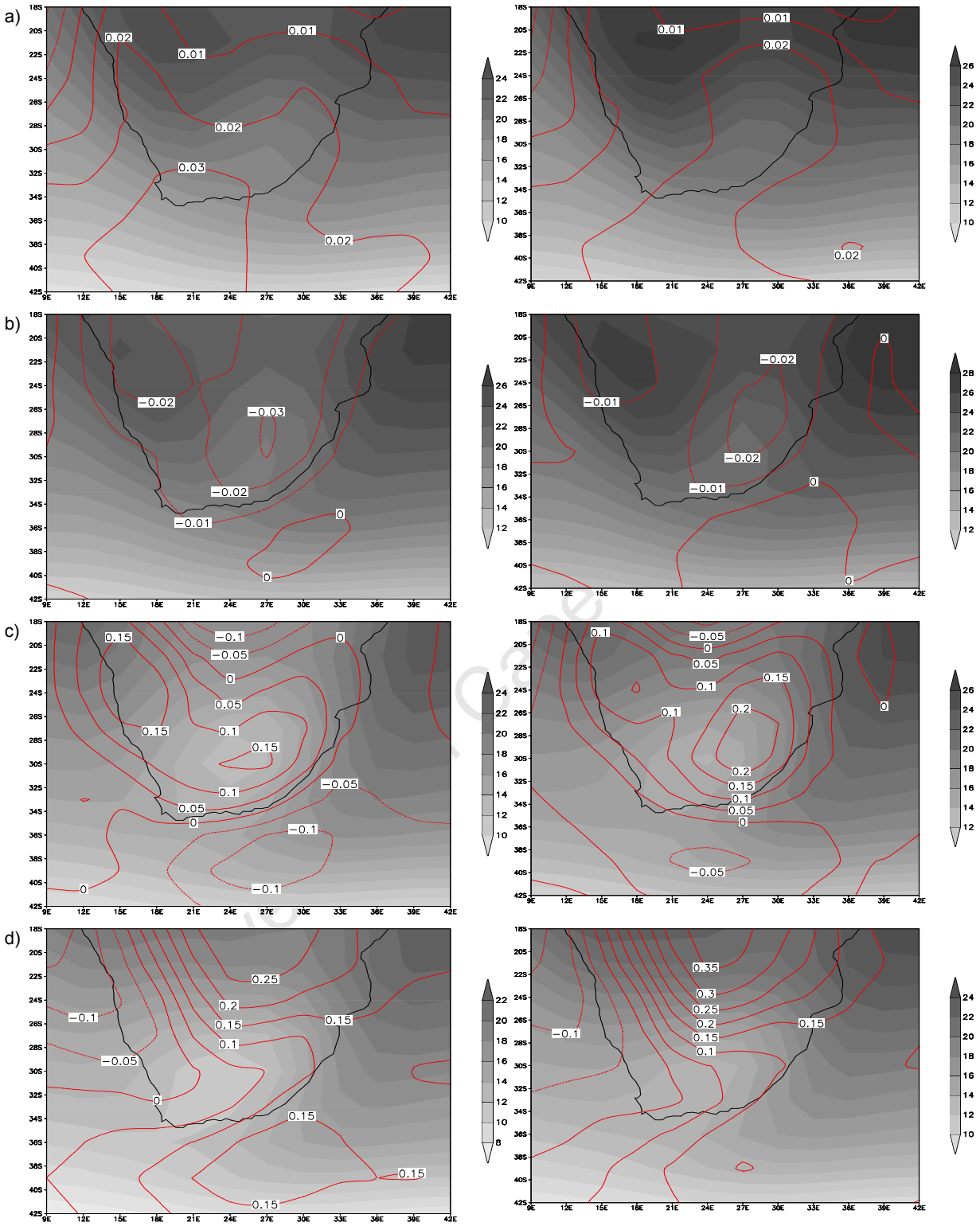


Figure 5.8: ***Multi-model temperature****: created by combining all the models weighted average control (left) and future (right) temperature in °C overlaid with the anomaly between weighted and average values for a) January, b) April, c) July and d) October*

The order of magnitude of the correction is on the same scale as that for the observed frequency correction (section 5.1), which is insufficient to bring the control projections to the observed baseline. To allow for better comparison of the relative improvement in skill, figure

5.9 and 5.10 depict the anomaly of the observed average (NCEP) from the model control for all three models and lastly compare the anomaly of the corrected multi-model solution to NCEP.

From the July anomalies, all models are seen to have a large positive bias over the continent and a negative bias over the Southern Ocean. The weighted multi-model is still too close to an average of all three models (which is not necessarily the most accurate) to really correct for the biases identified in Chapter 4. Even though GFDL had the lowest weight, its map average anomaly is actually the most accurate. This indicates that even though the GFDL model may have large biases for specific nodes, overall the average is closer to the observed than other models. This brings into doubt the usefulness of this technique when used merely as a frequency adjustment without considering the baseline conditions. The underlying assumption that model's projections were influenced the most by frequency biases has been shown to be only a part contributor to model error, and that the systematic biases remain a strong component of overall model error.

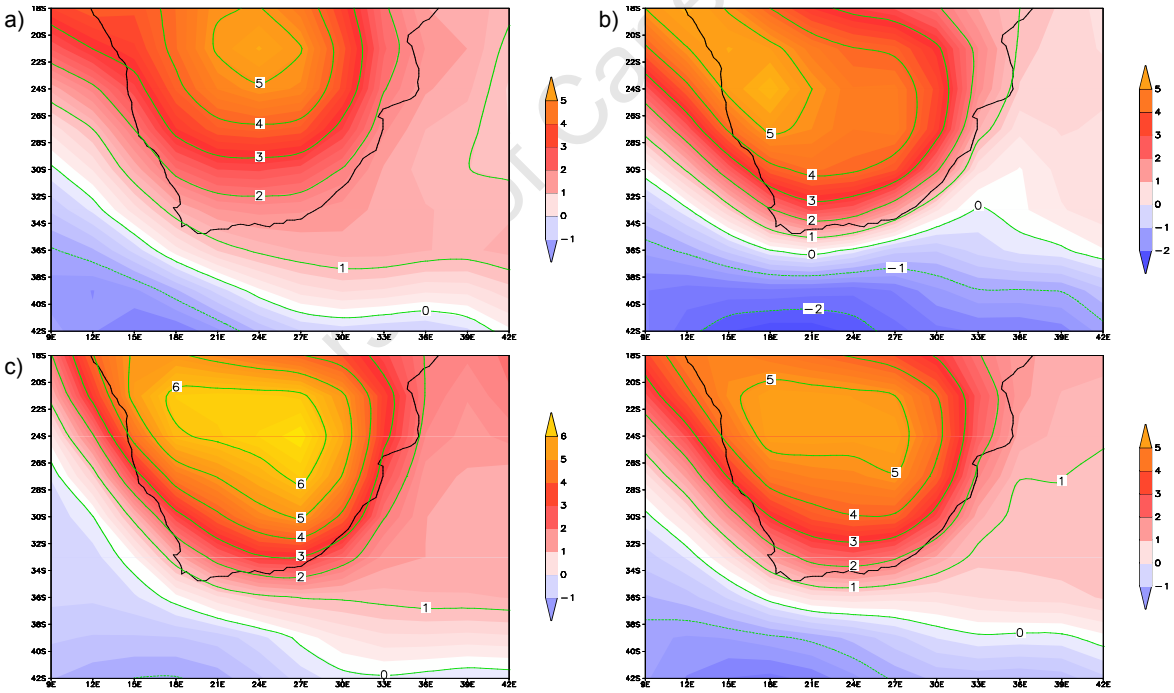


Figure 5.9: July anomalies of NCEP annual mean temperature in °C from weighted a) GFDL, b) MIROC, c) MRI and d) multi-model

October anomalies are mostly negatively biased for all models as shown in figure 5.9. The MRI model outperforms the multi-model average, reinforcing that this technique is not sufficient to improve projections on its own.

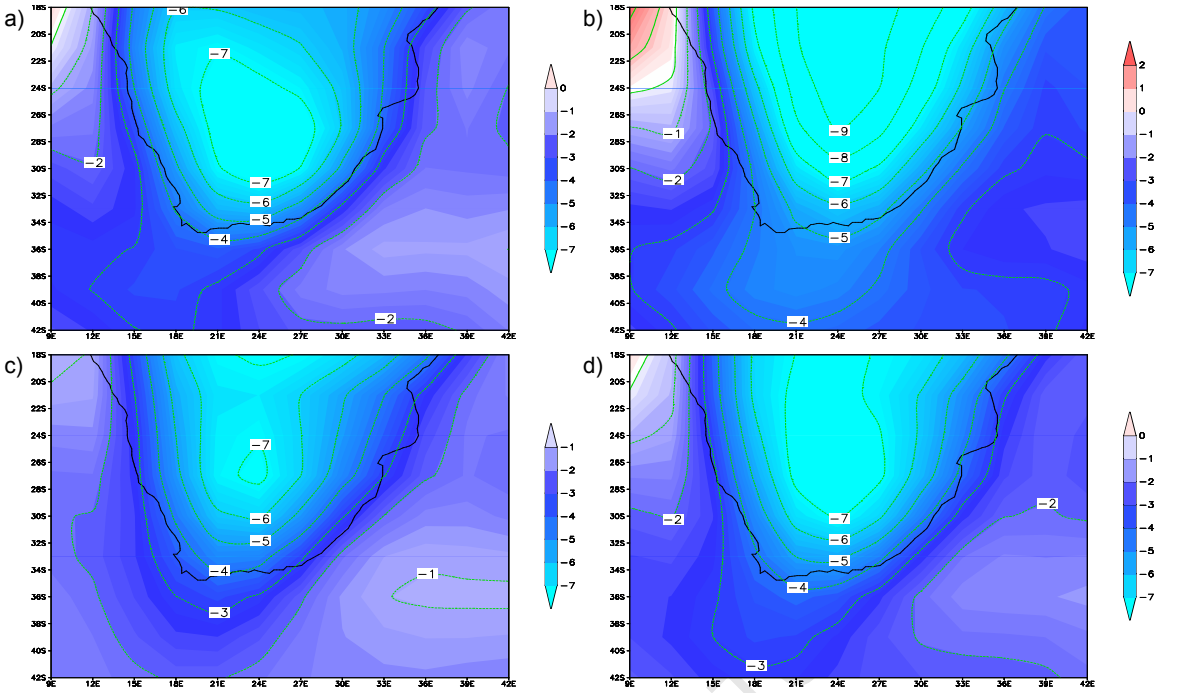


Figure 5.10: October anomalies of NCEP annual mean temperature in °C from weighted a) GFDL, b) MIROC, c) MRI and d) multi-model

5.2.2 Precipitation

In the annual data, the small weighting range produces little deviation from a multi-model average as shown in figure 5.11. All the models overestimated precipitation over the South Indian Ocean (as seen in figure 5.3). Therefore this feature is still prominent in the multi-model mean, and in fact the MRI map mean out performs the multi-model solution. The future data shows an overall reduction over the South Indian, which could indicate that this reduction will be less as this system is already over precipitating in the models.

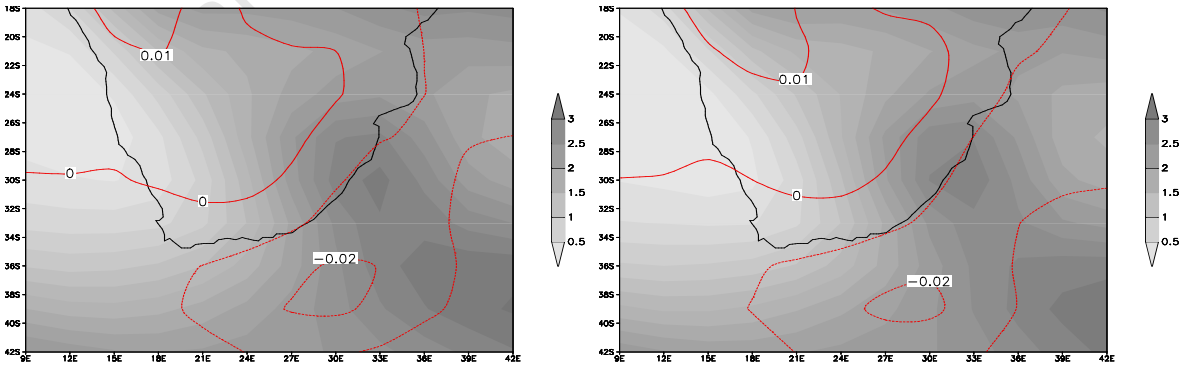


*Figure 5.11: **Multi-model annual precipitation (mm/d)**: created by combining all models weighted average control (left) and future (right) precipitation overlaid with the anomaly between weighted and average values*

In figure 5.12, which depicts the monthly means, the same limitations in the January and April data are seen, due to the similar weightings. July anomalies indicate that the multi-

model solution reduces South Indian Ocean precipitation by 0.1mm and in October increases precipitation in the same region.

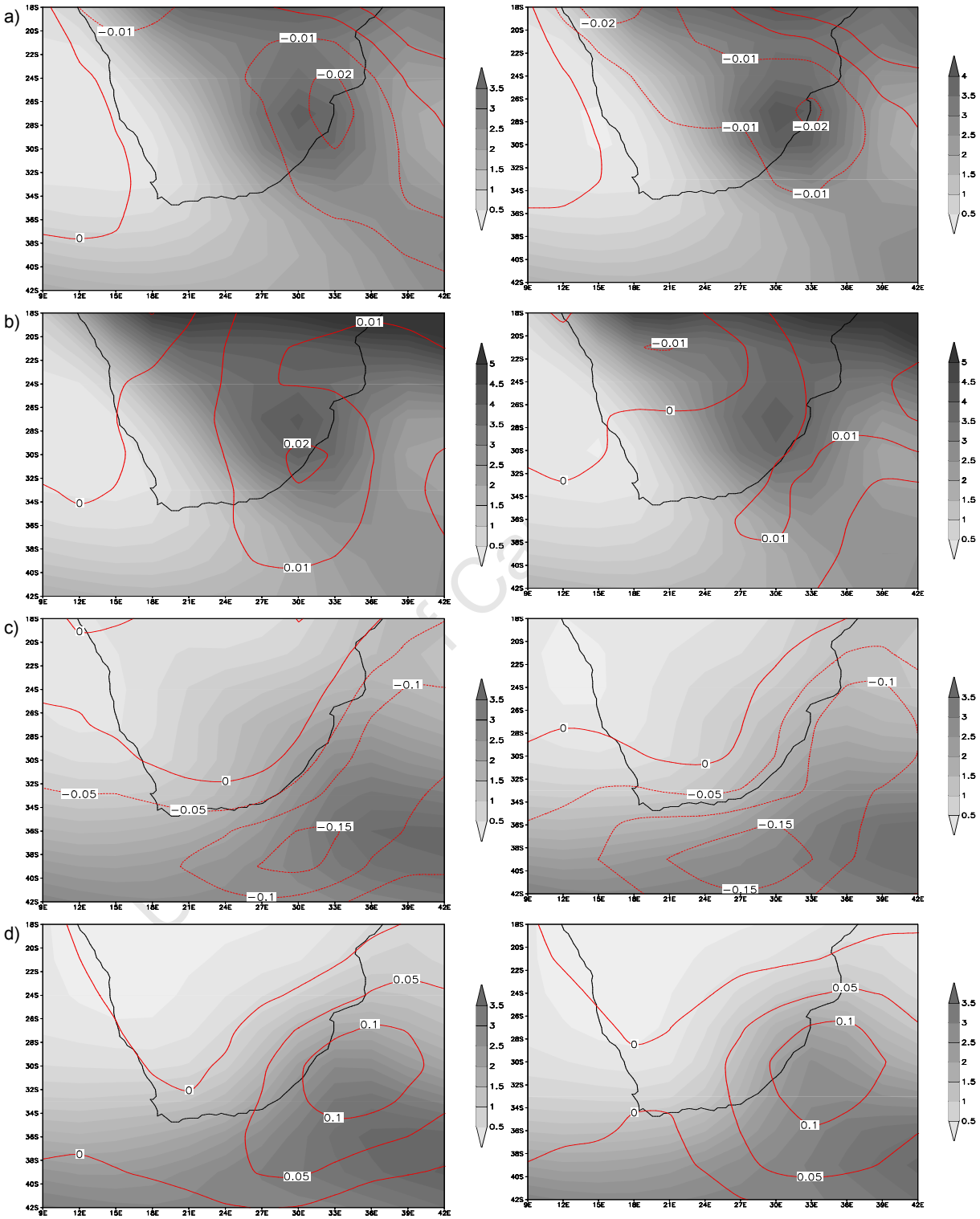


Figure 5.12: ***Multi-model precipitation (mm/d)****: created by combining all models weighted average control (left) and future (right) precipitation overlaid with the anomaly between weighted and average values for a) January, b) April, c) July and d) October*

In figure 5.13, the July anomalies between NCEP and the weighted GCM are compared to the weighted multi-model mean. These results indicate that MRI is the most accurate, outperforming the multi-model mean.

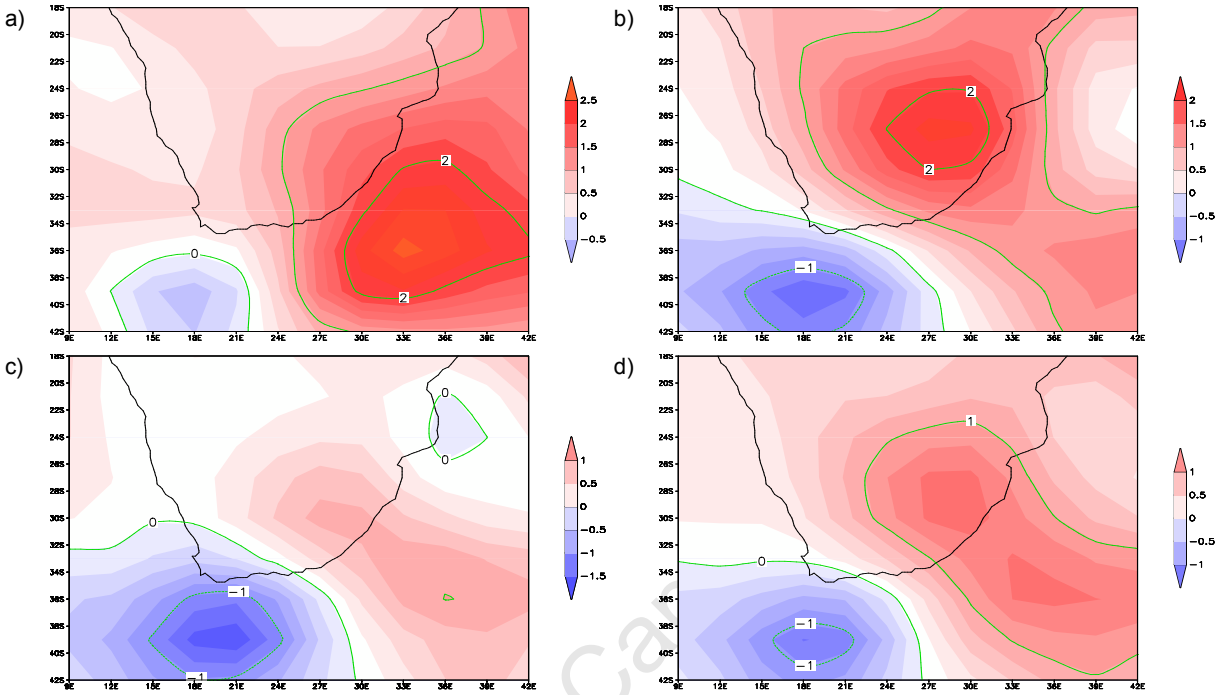


Figure 5.13 July anomalies of NCEP annual mean precipitation (mm/d) from weighted a) GFDL, b) MIROC, c) MRI and d) multi-model

The use of this technique has been shown to have a limited influence on the internal dynamics and can in fact produce results that are less accurate than a single model. For example in October (Appendix III, figure C9) there are larger anomalies in the weighted multi-model solution than in MIROC and MRI over the Southern Indian Ocean. However, as the adjustments have been for frequency alone, it is not surprising that the models systematic biases remain. If the only error in the models had been the frequency at which any given synoptic state occurs, the technique would have yielded near NCEP results. In the next section the weighting is applied to future anomaly projections and the systematic bias is corrected for by changing the baseline to fit the observed archetypes.

5.3 MEAN WEIGHTED FUTURE CLIMATE PROJECTIONS

As seen in the previous sections, the difference between the NCEP and GCMs mean node composites remain large even when frequency based weighting is applied to the GCMs. The baseline means are therefore different enough to warrant exploring the use of the NCEP baseline in conjunction with frequency weighted future change anomalies (future - control). The average of the three models weighted future change anomalies is then added

onto the NCEP baseline and compared to the average of the three models to determine how much improvement is garnered.

The systematic bias is removed from the future model average in addition to correcting for the frequency differences identified in the control. This is accomplished by taking the difference between future ($Pave$) and control ($Cave$) average projections and weighting each model anomaly with the weights (w) used in section 5.2. The weighted anomalies are then divided by the summated weights to produce a multi-model average anomaly and added to the NCEP baseline average ($Nave$) as per equation E.4.

$$\text{(E.4)} \quad Mfutr = \frac{\sum_{i=1}^{n} Wm(Pave - Cave)}{\sum_{m=1}^{3} Wm} + Nave$$

To provide some context for the climate change signals shown below, in Appendix III figure C10, all three models average future temperature are shown overlaid with the average climate change. The temperature change is positive in all models and ranges from 1-2.4°C.

The multi-model weighted future projections are compared to the multi-model average future projection in figure 5.14. On the left in the shaded plot we see the weighted anomaly added to the NCEP observed baseline. The anomaly field overlaid in red depicts the weighted climate change signal of the model. The amplitude of future change shown in red is not greatly different from those observed in the individual models, as the baseline adjustment has the larger influence on the overall result. The overall pattern of change is also not altered so that the overlay on the left is fairly similar to the difference between the average future and control projections on the right.

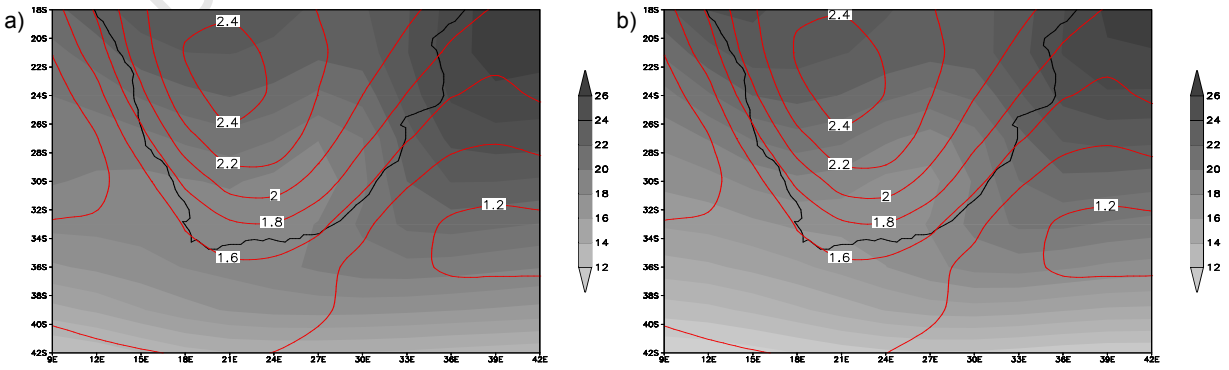


Figure 5.14: Annual future temperature projections in °C of a) weighted multi-model overlaid in red with the weighted anomaly between future and control and b) average of all models original projections overlaid in red with the anomaly between future and control

As the baseline is now from NCEP (which is warmer than the models) the resultant future temperature map (the shaded plot) is slightly different. The largest change is over the Western Cape where the models had a strong gradient of temperature increase in the multi-model average to the right whereas the gradient is shallower in the weighted projection on the left.

The changes in the future precipitation field (figure 5.15) are marked by the decrease of the systematic bias identified over the South Indian Ocean. This is almost entirely due to the adjustment of the baseline to NCEP values as there is no difference in the overlay fields, which is due to the lack of variance in the annual weights, making them behave like the average. Along the western coast the isohyets are closer together in the weighted plot, indicating a more intense moisture gradient in this region than the model average depicts.

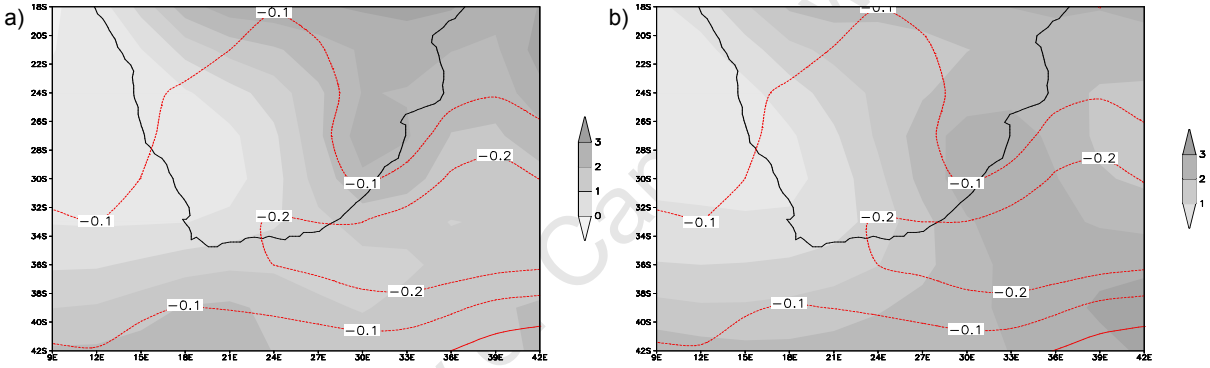


Figure 5.15: Annual future precipitation projections in °C of a) weighted multi-model overlaid in red with the weighted anomaly between future and control and b) average of all models original projections overlaid in red with the anomaly between future and control

The next step is to look at the overall difference between the weighted and unweighted future means to see what the magnitude of the adjustment is and in what direction. Blue indicates where the weighted model projection is less than the original future projection and red where the weighted model has increased temperature in comparison to the original. From the biases identified at the end of Chapter 4, it is encouraging to see that the anomalies in figure 5.16 correct for many of these biases. For example the Southern Ocean surface air temperatures, which was always too cold in the control projections, has been increased by around 2°C in the weighted temperature field. In the precipitation data, the bias for more precipitation over the South Indian Ocean is corrected for. This illustrates that the technique does make an improvement over using the average model values and corrects persistent biases very well.

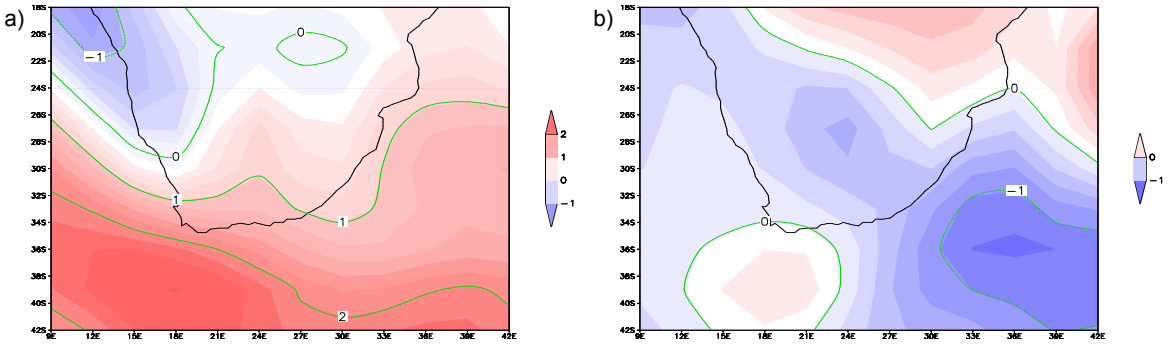


Figure 5.16: Annual weighted future projections anomaly of weighted - average of all models original projections for a) temperature (°C) and b) precipitation (mm/d)

Monthly data

From figure 5.17, we see that the climate change signal varies between the months in the amplitude and placement of the temperature increases. The January weighted climate signal is adjusted to the warmer observed summer values and the area of maximum increase has also shifted position to be centred further south, which is the result of adjustment to the NCEP baseline. In both January and April, changes are entirely a function of the baseline adjustment as the positioning of the climate signal (red contours) is not affected. In fact the two climate change signals are indistinguishable as the weights are so close that they behave as the average (as discussed in 5.2). The April baseline reduces temperature over the continent from the multi-model average.

In July and October the weighting does effect the climate change signal, with July warming increasing over Zimbabwe by about 0.2°C. The baseline adjustment cools temperatures over the interior in July and in October increases temperatures over northern Namibia, correcting for systematic biases identified in 4.4. The October pattern of change shows a slight expansion of the 3°C warming isotherm over the western interior.

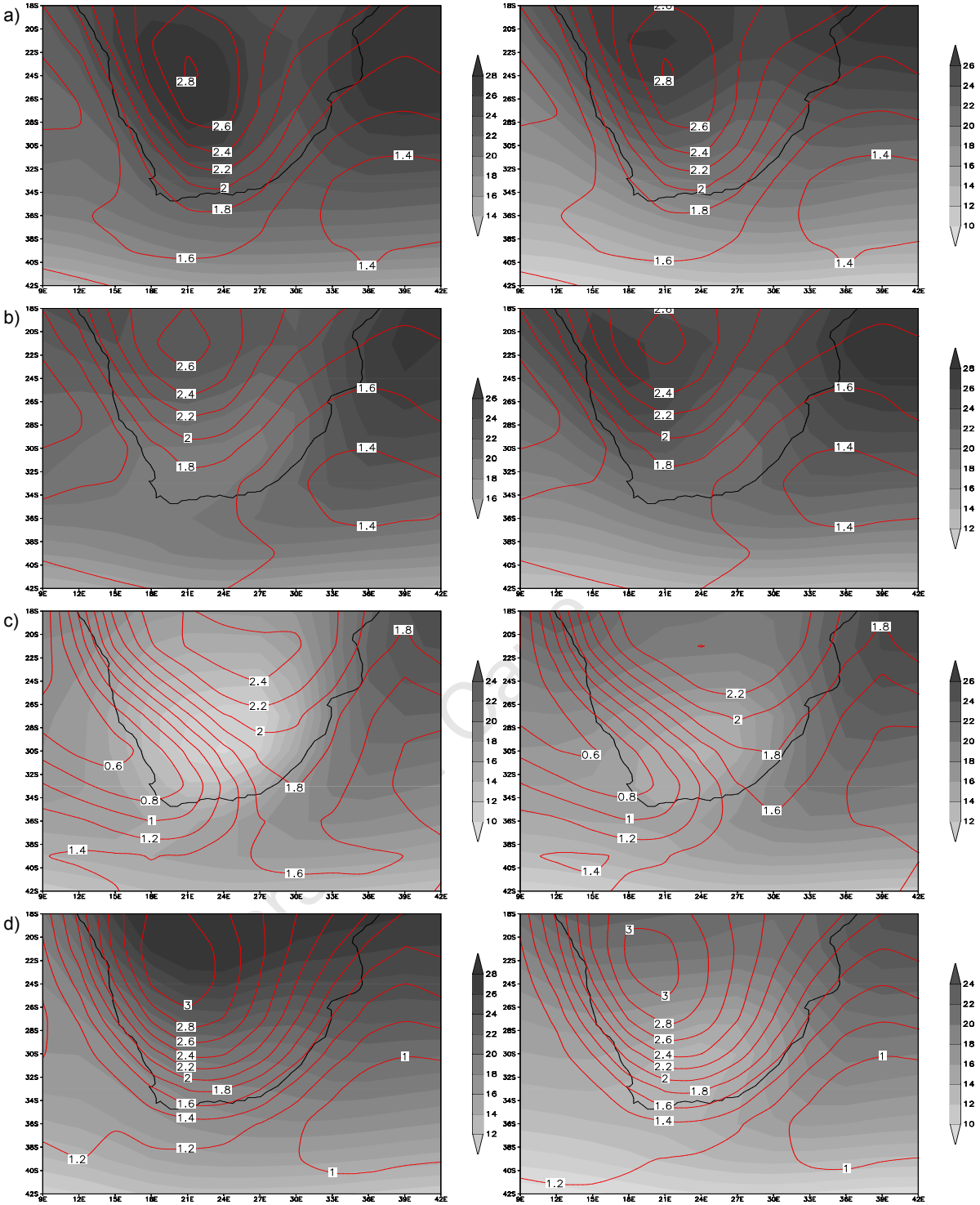


Figure 5.17: Monthly future temperature projections in °C were weighted and averaged as shown in the grey scale image and overlaid in red with the weighted anomaly between future and control (left). Average of original model projections is overlaid in red with the anomaly between future and control (right) for a) January, b) April, c) July and d) October

In the precipitation field, shown in figure 5.18, the largest change from baseline adjustment is seen in January, where the corrected field has doubled precipitation over northern Mozambique and Zimbabwe in comparison to the multi-model average. In April there is less

precipitation over the continent in the weighted map, especially on the east coast. July and October climate change signals in red contours are very similar to the multi-model average. The baseline adjustment in July reduces precipitation over the continent and dries out the north western coast by ~0.5°C. In October the precipitation decreases over the south Indian Ocean in the weighted projection.

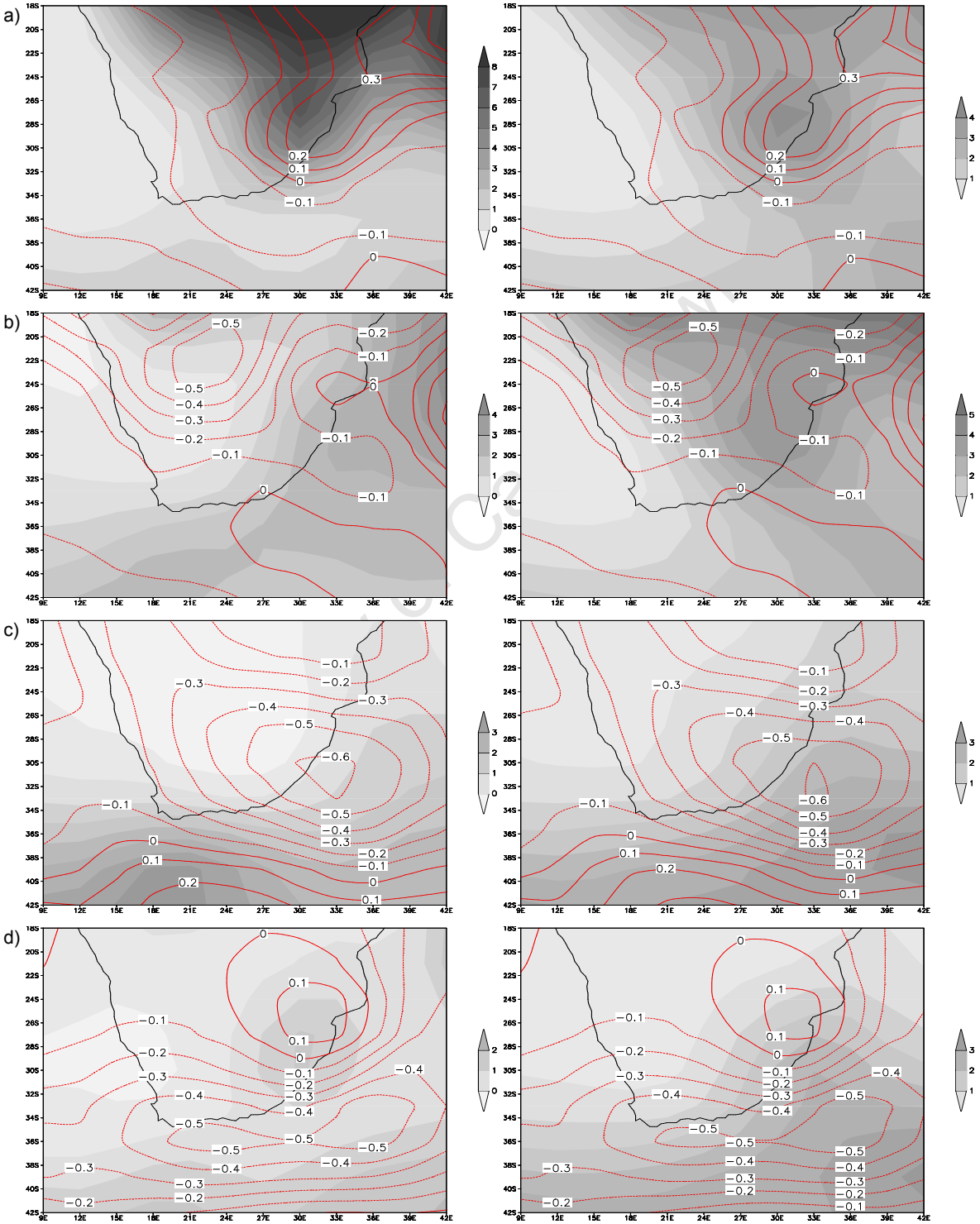


Figure 5.18: Monthly future precipitation (mm/d) projections were weighted and averaged as shown in the grey scale image and overlaid in red with the weighted anomaly between future and control (left). Average of original model projections overlaid with the anomaly between future and control (right) for a) January, b) April, c) July and d) October

The anomalies of the weighted projections from the average multi-model (figure 5.19) show that all the large systematic biases identified in Chapter 4 are being corrected for. This is entirely due to the baseline correction in January and April. In January the southern Ocean is warmed by 4°C to correct for the cold bias over this region. For precipitation there are large increases in the northeast of the domain (in the heart of the summer convective rainfall region), with a smaller decrease in south Indian Ocean precipitation. April temperature projections show a strong decrease over the west coast of 3°C and a slight warming of the Southern Ocean surface air temperatures. Precipitation is reduced, especially over the north east of the continent, which could indicate that the models over simulate convective rains in this season. Small reductions in ocean precipitation in the south Atlantic and off the Natal coast are also seen.

July corrections cool the model temperature results by 5°C over the central interior with a slight warming over the Southern Ocean. Precipitation is reduced over the east coast by 1mm and slightly increased over the South Atlantic Ocean. In October, interior temperature values are raised by as much as 8°C, indicating that the models had failed to simulate the interior's continentality effects in spring and perhaps that the cooler/winter dynamics are more persistent in the models. All surface air temperatures over the ocean are also warmer in the weighted model supporting the idea that the models do not shift seasonal boundaries at the right time. In precipitation results reductions of 1-2mm are seen over the South Indian Ocean with little change over the continent.

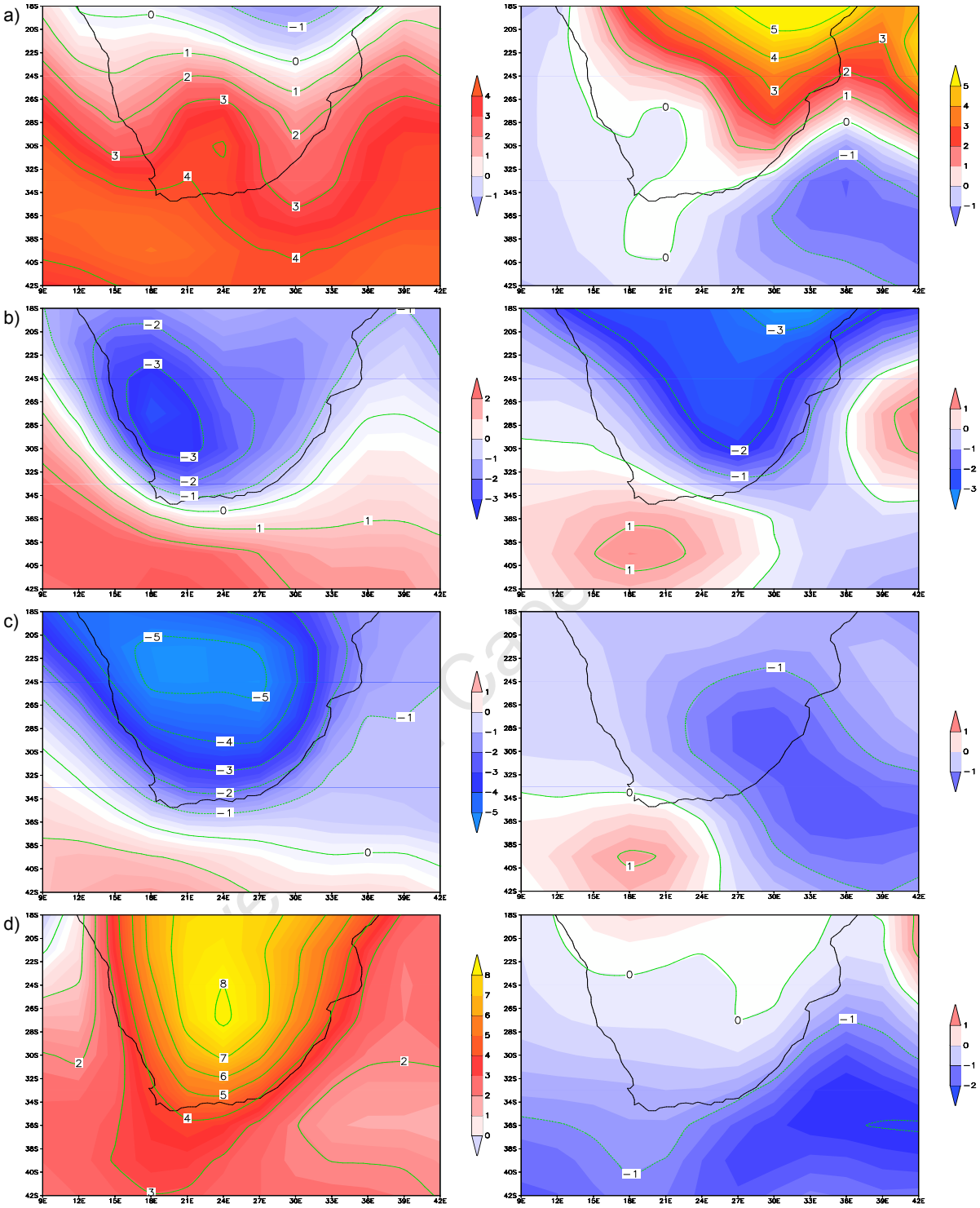


Figure 5.19: Monthly future temperature in °C (left) and precipitation in mm/day (right) anomalies of the weighted multi-model from the average of all models original projections for a) January, b) April, c) July and d) October

From this section a lot of positive results were obtained, however, they were largely due to the adjustment to an NCEP baseline. The frequency correction is hamstrung by the lack of variation in weights between models for most time slices. In the next chapter, node specific weightings will further illuminate the extent to which the correction tool can be used.

CHAPTER SIX

Synoptic specific multi-model climate projections

From the previous chapter it was seen that using a single model weight (that comprised of the average of all node weights) to adjust the mean climate state has some flaws. This is mainly due to the great affect one or two low value node weights can have on the average. Many of the nodes may have high weights, depicting greater accuracy, but the few very low weights skew the average to a low overall weight. The techniques presented here uses node specific weightings, which build on from the discussion of mean weights in the previous chapter. In this chapter each node or synoptic state will have its own weight, thereby disaggregating the biases to a synoptic level. The analysis methods are very similar to those outlined in 5.2 and 5.3. However, as the nodes are not combined into a mean weight in this method, we instead combine the models to produce weighted multi-model averages for each synoptic state.

6.1 SYNOPTIC STATE SPECIFIC WEIGHTING

This method examines how node specific weighting can improve upon the single model weight method shown in Chapter 5. The limitation of similar model weights producing a near average mean should be less problematic here as the node differences have not been averaged and are therefore more variable. An individual node weight for each model is created for each synoptic state using the fractional frequency difference between the GCM and NCEP as described by equation 1. Instead of adding up the node weights as was done in the previous chapter, they are kept separated.

(E.1) $$Wi = 1 - \frac{|Cfi - Nfi|}{100}$$ i = node number

The three models each have their own weight for each node. These are compared by putting them over a common denominator, which is the sum of all three weights for that node. Each models weight in this node specific formulation is therefore expressed as a fraction of all three models' frequency differences. This is done so that it represents the degree of similarity to NCEP in terms of frequency contributed by each model to the overall value for the node. The fractional frequency difference (Wi) is divided by the sum of the three model differences which results in a fractional value (Wi') (see Equation 5[1]).

[1] Equation numbers are drawn from Chapter 2 to maintain consistency.

(E.5) $$Wi' = \frac{Wi|}{\sum_{m=1}^{3} Wi}$$

Figure 6.1 depicts the three models node weights for the annual data. Each node has a value between 0 and 1, with those tending to 1 (which is shaded darker) representing those nodes that have the most similar frequency to the observed and the lighter nodes are those with the greatest differences. MRI and MIROC both have a large number of nodes with high weights (good accuracy). Each model performs better for some region of the SOM space, but none seems to be better overall or even better for a set of processes (e.g. the three central rows depicting typically summer states show no model has higher weights for all nine nodes). The monthly values (figures 6.9 – 6.11) confirm that no model has an even distribution of weights, with each model having some nodes that are poor and others that are good.

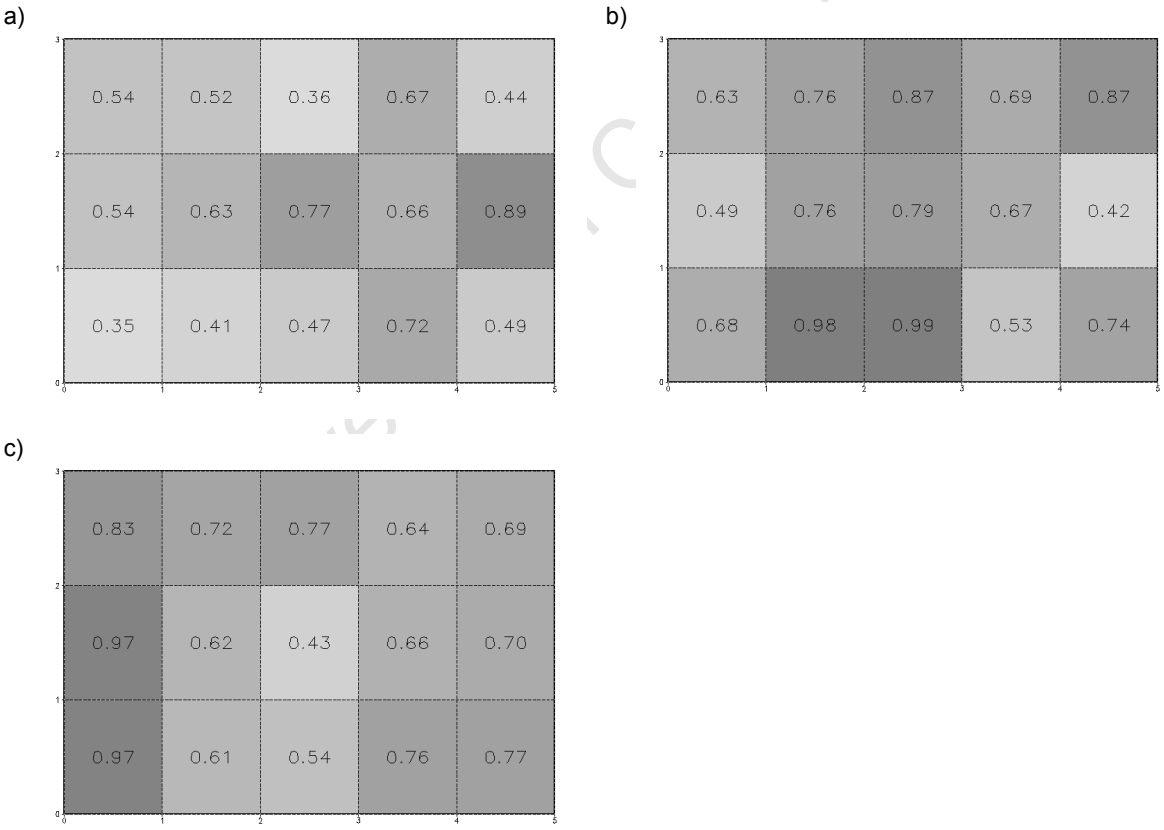


Figure 6.1: Node weights based upon each GCMs relative similarity to NCEP for annual data: a) GFDL, b) MIROC, c) MRI

The node weight (*Wi'*) is then applied to each model. The three models are then combined to make a multi-model average value for each synoptic state (see Equation 6).

$$\text{(E.6)} \quad Mfutri = \frac{\sum_{m=1}^{3} Wi'(Pi - Ci)}{\sum_{m=1}^{3} Wi'} + Ni \qquad \text{per node}$$

The benefit of combining the models at this point is that it prevents a single weak node from making a disproportionately large impact on the overall model weight, as in 5.2. To rephrase, the weight can be thought of as a proportion of all model error and allows for a model to be compared on a synoptic level. Each models weight should therefore be looked at in conjunction with the other two models to get a sense of which model has the most accurate/least accurate representation of that node.

A factor to consider in this is that each node weight can only be compared between models for that particular node. This is because the weight is a "proportion of correctness" for that node, but does not take into account the relative differences in frequency between adjacent nodes. For example, if one node has very small differences in frequency for all three models, the total of the three model weights will be small and therefore one of the models may have a very low weight compared to the other models but in fact have a frequency quiet close to NCEP's, just not as close as the other two models. The node adjacent may have a much higher weight (more accurate) but have a larger frequency difference between the model and NCEP. Therefore, just because a node has a low weight does not mean it is more "wrong" than its neighbour, but that out of the three models it was the least correct and therefore should be given a lower weight. In this way those models with the closest frequency have greater representation in the weighted results.

Monthly weights

As shown in Chapter 3, the monthly frequency anomalies are generally larger for the monthly data sets than the annual data; therefore, the variation in weights is also larger. For sake of brevity, only the April and July weights are shown in this section as there is a larger variation in weights for these months. January and October weights are to be found in Appendix IV.

Looking at April weights in figure 6.2, GFDL performs well with all weights having a value of more than 0.5 in every node. GFDL also has the most even distribution of weights in January and October. In April we see that GFDL has two nodes with the value 1 (meaning the same number of time steps were assigned to the model as to the observed resulting in the frequency of the model being identical to NCEP). This is only possible in my analysis in the April data as in all other months the number of observations for NCEP is greater than in the model, as the

models have 30 day months. However, as April has 30 days in reality as well as in the model, therefore the two can have the exact same frequency and a resultant weight of 1.

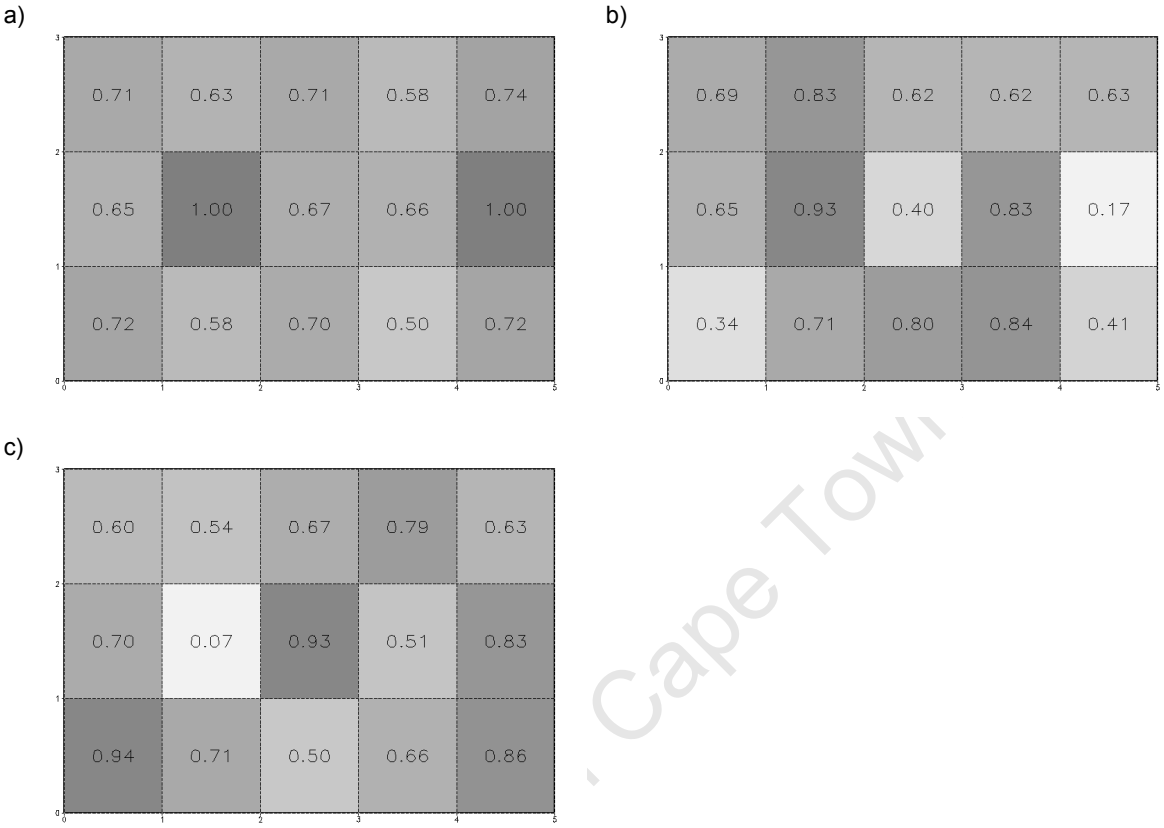


Figure 6.2: Node weights based upon each GCMs relative similarity to NCEP for April data: a) GFDL, b) MIROC, c) MRI

In July, the MIROC model has the most even weight distribution, with GFDL performing poorly in comparison to the previous months (see figure 6.3). Node 12 is a good example of how weighting can be misleading if nodes are compared. GFDL and MIROC have very high weights for this node, whereas MRI has a very low weight (0.11). When looking at the frequency difference on this node, the real difference is only 1.13%, however, the other two models have the same much smaller difference of 0.0733% (a difference of two observations from the NCEP data). In the same way, node 10 has a very low weight in GFDL as its frequency difference is much larger than the other two models. To reiterate, the weights do not take into account the original differences between nodes as they are scaled by the total difference of all three models per node.

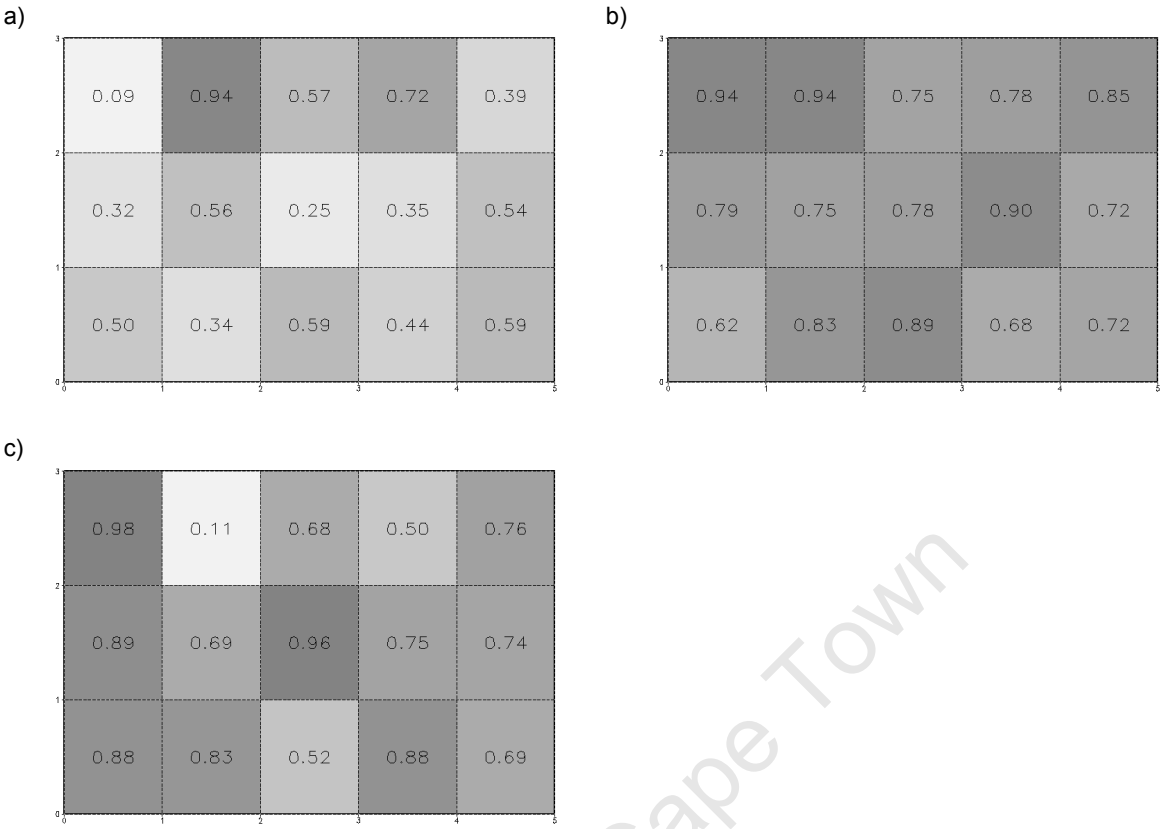


Figure 6.3: Node weights based upon each GCMs relative similarity to NCEP for July data: a) GFDL, b) MIROC, c) MRI

6.1.1 Temperature

The individual model node weights are used to weight each node's mean composite for temperature for each model. Then all models are added together to produce a weighted multi-model mean. The annual weighted temperature nodes are shown in figure 6.4. Warming is greatest in the three centre columns with cooler states in the outermost columns.

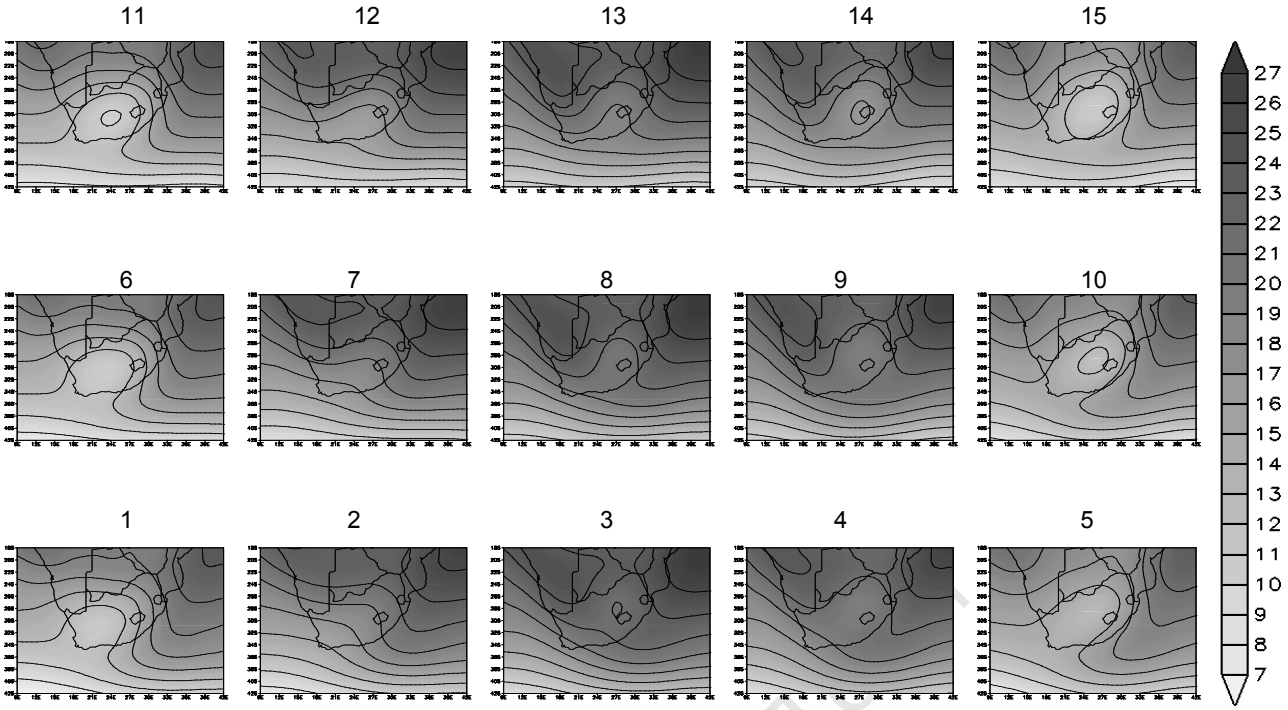


Figure 6.4: Annual node weighted multi-model temperature

Figure 6.5 shows the difference between control weighted and unweighted multi-model SOM composites. The differences are moderate, at most 0.4°C in either direction. However there is some indication that there is a seasonal component to the difference in terms of the direction of change. The outer left and right columns represent winter states and the three middle columns represent summer patterns. All the summer trough states (nodes 2, 3, 4, 7, 8 and 9) are showing a reduction in temperature. Those states with strong anticyclones (including the top row) are showing an increase in temperatures in the weighted model. This would indicate that the raw multi-model average has greater temperatures for troughs states and cooler temperature for anticyclone states.

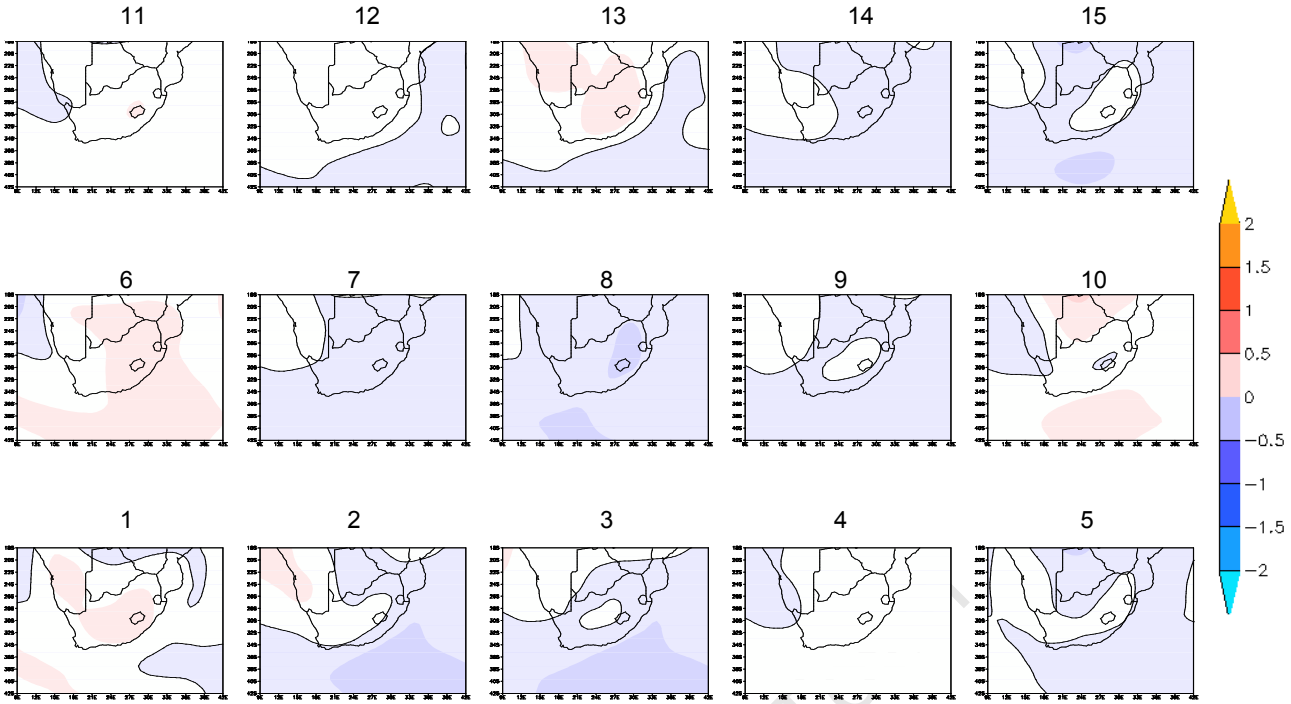


Figure 6.5: Multi-model weighted control projection – average multi-model mean for annual temperature, contour interval of 0.2°C

This implies that the frequency biases of one model allow that model's temperature bias to overly influence the multi-model average. So for example, considering the GFDL model which has a high bias over the continent for cooler temperatures in trough states and a low weight due to its erroneous frequency of these states, the weighted multi-model solution compensates for the disproportionate frequency influence of this models temperature bias.

For individual nodes, there are some exceptions to this general pattern as in node 15 which has dominant Indian Ocean anticyclone but shows a reduction in temperature in the weighted multi-model average.

The difference between the weighted multi-model variable composite nodes and NCEP observations are shown in figure 6.6. Here we see in the weighted results that there is an improvement in decreasing the magnitude of the anomalies in comparison to the single model anomalies shown in Chapter 4. The same persistent biases are seen as in the uncorrected data, for example the overestimation on the Namibian coast and underestimation over the Southern Ocean. This is because the weighting is based on the frequency differences of a node and not the internal node bias and therefore cannot correct for differences in baseline state.

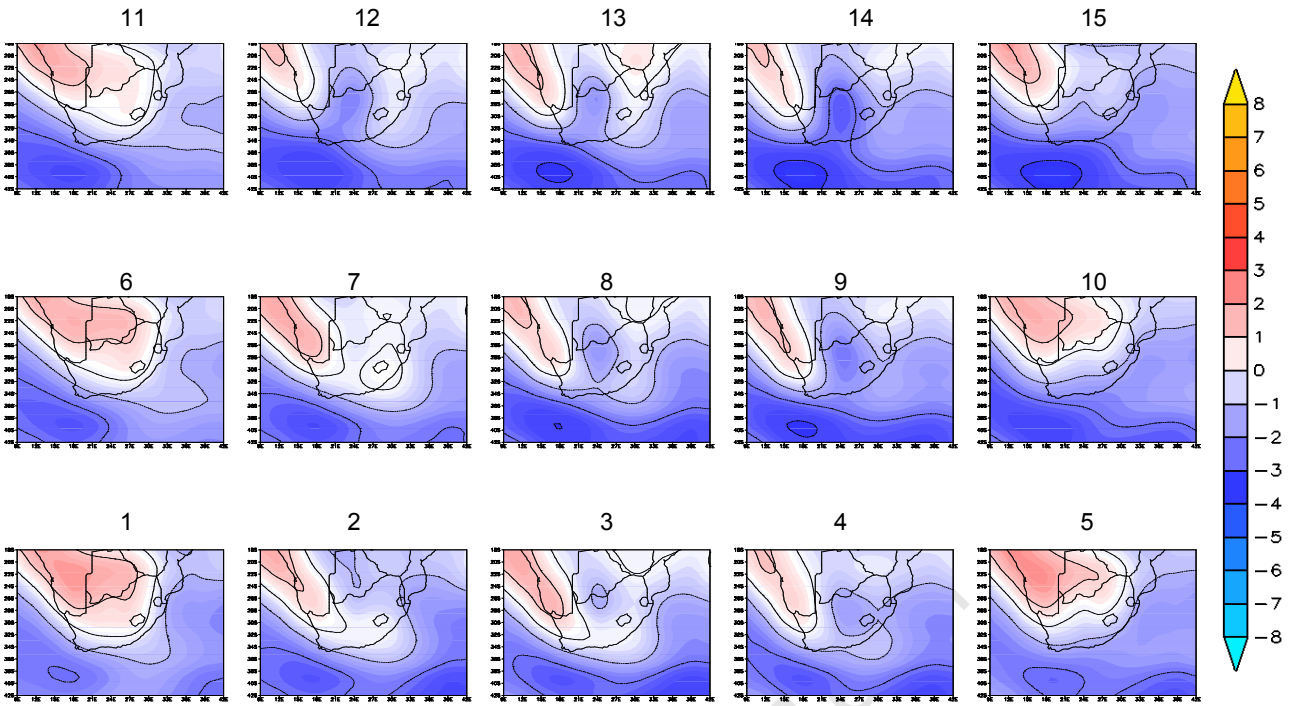


Figure 6.6: Annual node weighted multi-model – NCEP temperature, contours at 1°C interval.

If the only model error was frequency the correction would be able to match the NCEP distribution. However, as seen in Chapter 4 there are some large differences in the node maps of the models and NCEP. As the correction is based on frequency alone, which has been shown to be only part of the bias (see section 5.2) the technique cannot be expected to bring the weighted multi-model average closer to NCEP. Instead it is a way of exploring the effect of frequency on the multi-model solution and to gauge where changes in frequency generate changes in the magnitude of temperature and precipitation, which can then be linked to change in the dynamics.

Future data

The difference between the future weighted and unweighted multi-model nodes (figure 6.7) shows that the weighted results are slightly warmer in the winter states than the unweighted nodes by around 0.4°C. Strong anticyclone states show an increase in warming in the weighted projection. This is consistent with the control differences seen in figure 6.3; however, the future differences are slightly larger indicating that the frequency correction enhances the changes in the future to warmer conditions related to anticyclones and cooler conditions related to trough states.

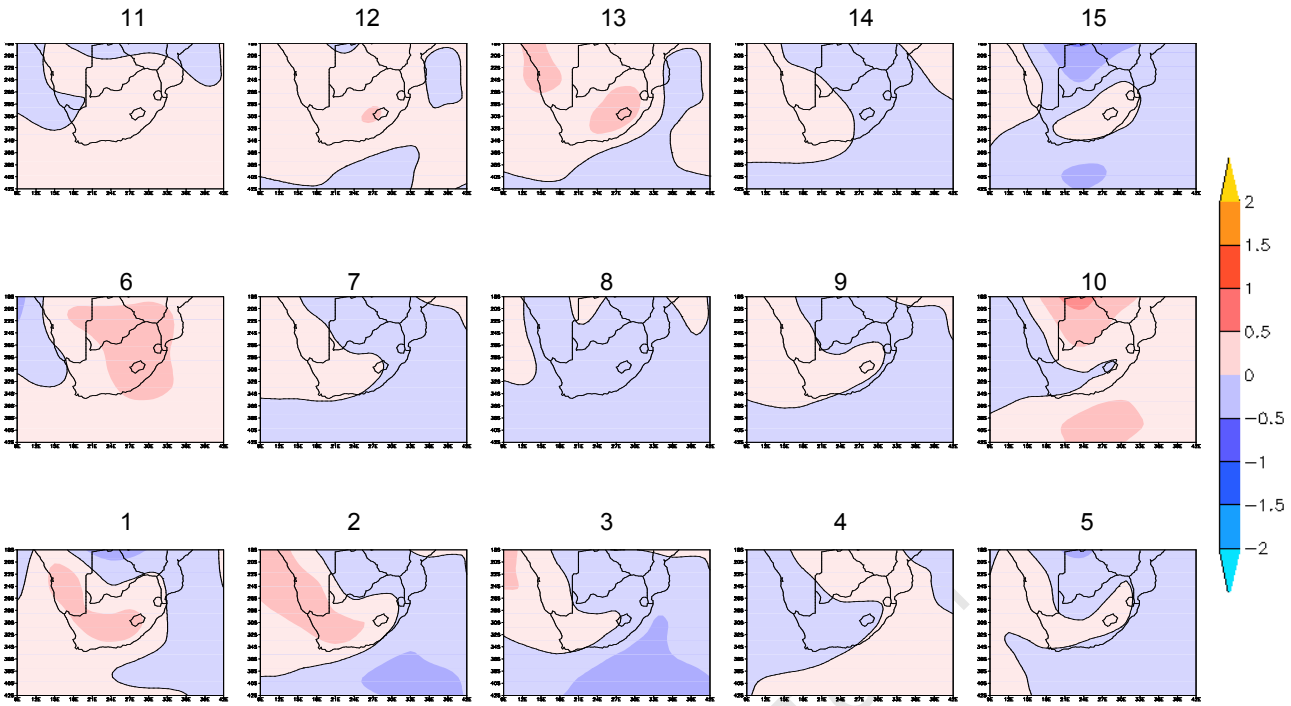


Figure 6.7: Multi-model weighted future projection – average multi-model future composite for annual precipitation, contour interval of 0.2 mm

The difference in the climate change signals shown in figure 6.8 between weighted and unweighted node composites indicates that the weighted nodes have a similar climate change anomaly to the unweighted with variation under 0.2°C in either direction. However, there is once again a state specific trend, with the weighted multi-model composite winter states showing a reduced climate change signal compared to the unweighted and greater warming, especially over the oceans, in summer states. While the differences are only small to moderate, this is only using three models and expanding to include many more models may result in more notable consequences. However, this is beyond the scope of this work.

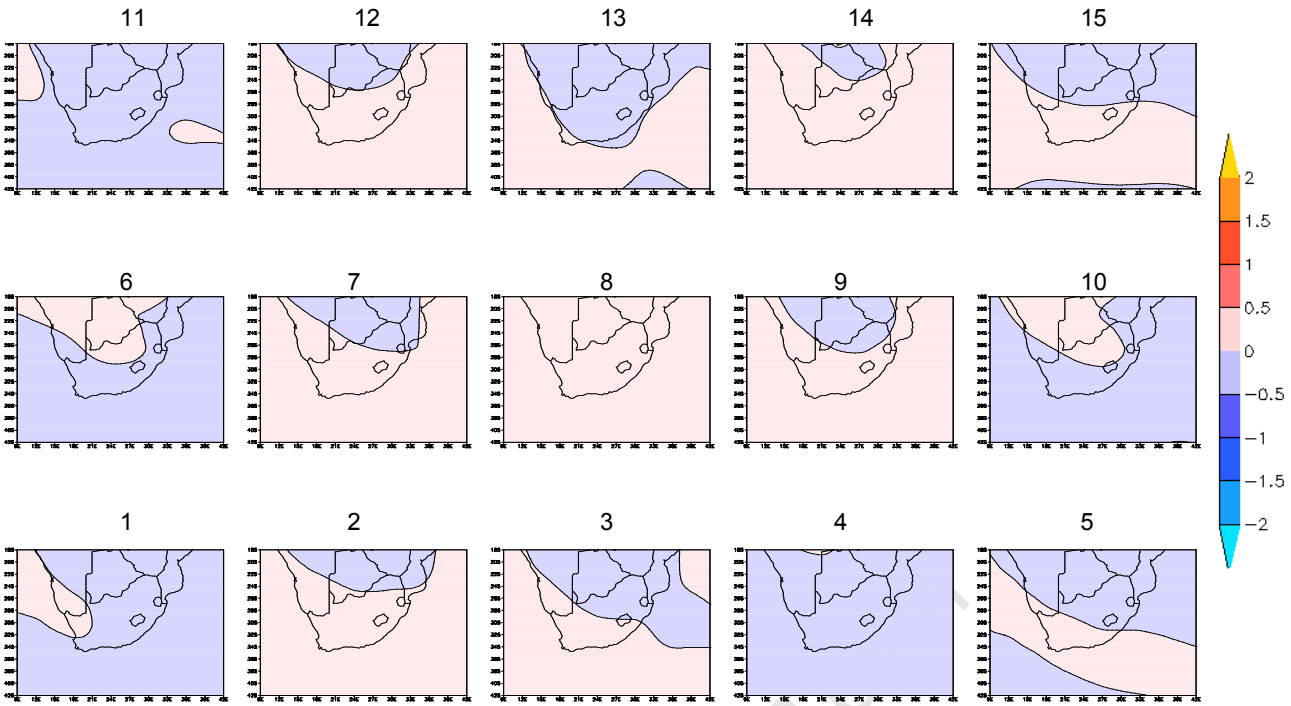


Figure 6.8: Multi-model annual weighted climate change delta (future – control) projection – average multi-model delta for annual temperature, contour interval of 0.2 °C. Red indicates where the weighted multi-model climate change signal is larger than the average multi-model composite. Blue where average is greater than weighted multi-model

Overall this suggests that the raw model data over simulates the increases in winter and under simulates summer increases in temperature.

6.1.2 Precipitation

The same technique is applied to the precipitation data and in figure 6.9 the annual weighted precipitation nodes show summer convective rainfall on the east coast in the top row (ridging anticyclones with interior heat low) and frontal rain in nodes 1 and 5 (cold front connected to westerly wave).

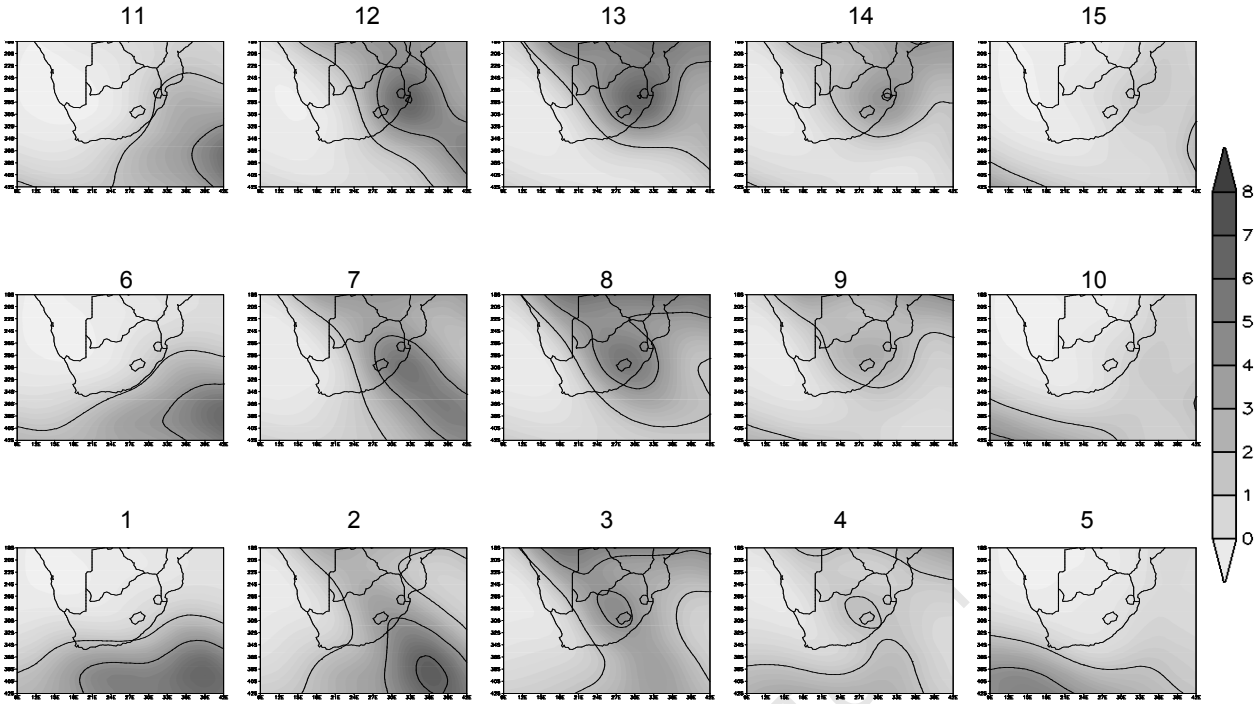


Figure 6.9: Annual node specific weighted multi-model precipitation

Figure 6.10 shows changes for most nodes with variations of around 0.2mm per day between unweighted and weighted nodes for the control climate. There is an exception in the central column of nodes, where strong continental thermal trough features are prominent, and there is slightly greater precipitation in the weighted nodes over the continent. In general, for all the states with strong anticyclones the weighted multi-model solution has lower precipitation that the multi-model average. As these states generally suppress rainfall the correction would seem to be in line with the dynamics.

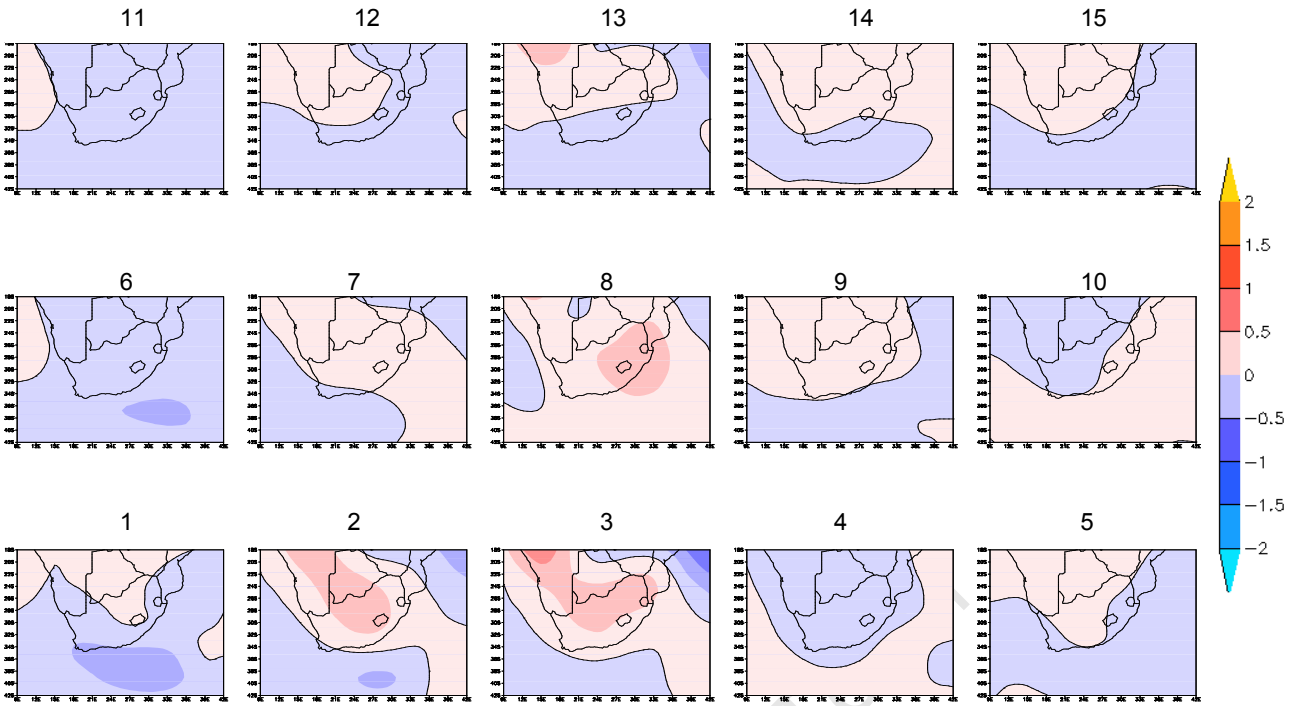


Figure 6.10: Multi-model weighted control projection – average multi-model mean for annual precipitation

As with the temperature data, the anomalies from NCEP (figure 6.11) are very similar to those of individual models. Persistent biases such as the dipole pattern in the southern Ocean in the bottom row are consistent with the patterns seen in all models in section 4.5.

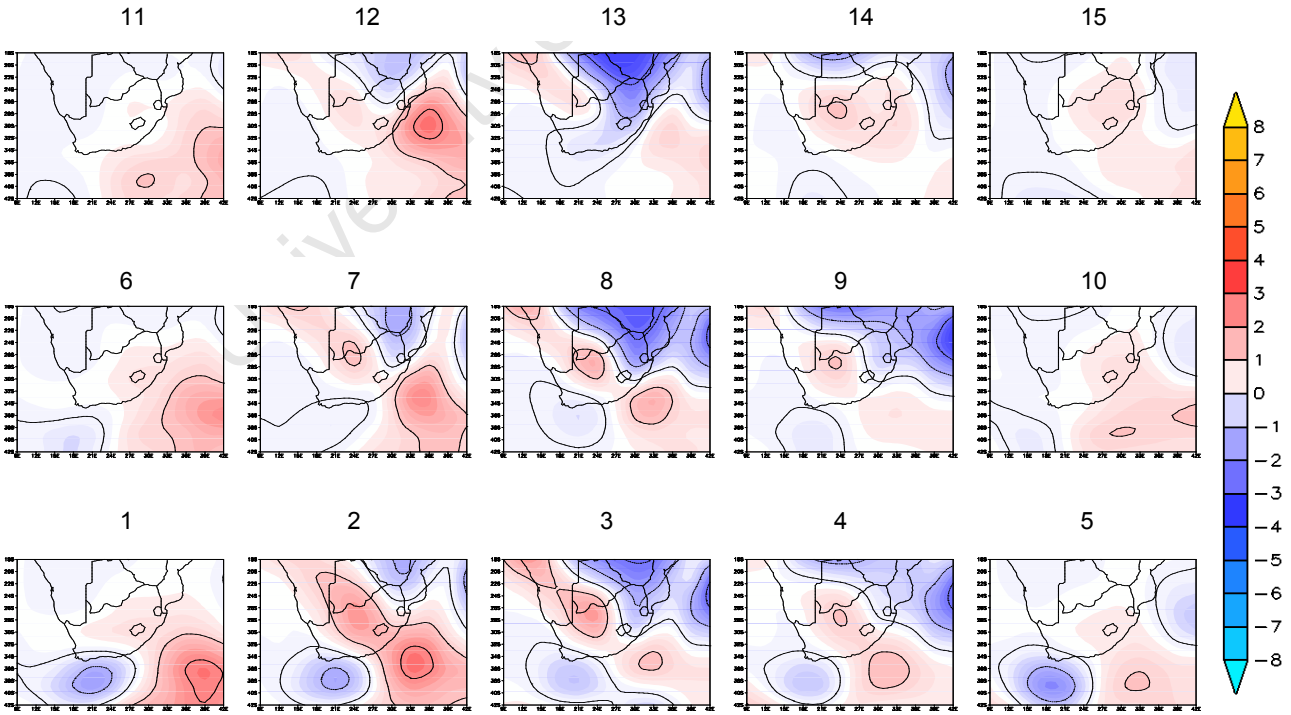


Figure 6.11: Annual node weighted multi-model – NCEP precipitation, contours at 1mm interval.

Future data

Future changes from using the node specific weighting technique (figure 6.12) show that the summer continental trough states are wetter over the continent in the weighted projections. More intense troughs are a likely explanation for the increased rainfall in these states. Decreases in Southern Ocean precipitation are seen in states with strong westerly wave features, which could indicate that the westerly wave brings less frontal rainfall. In areas with strong anticyclones the changes are small and inconsistent.

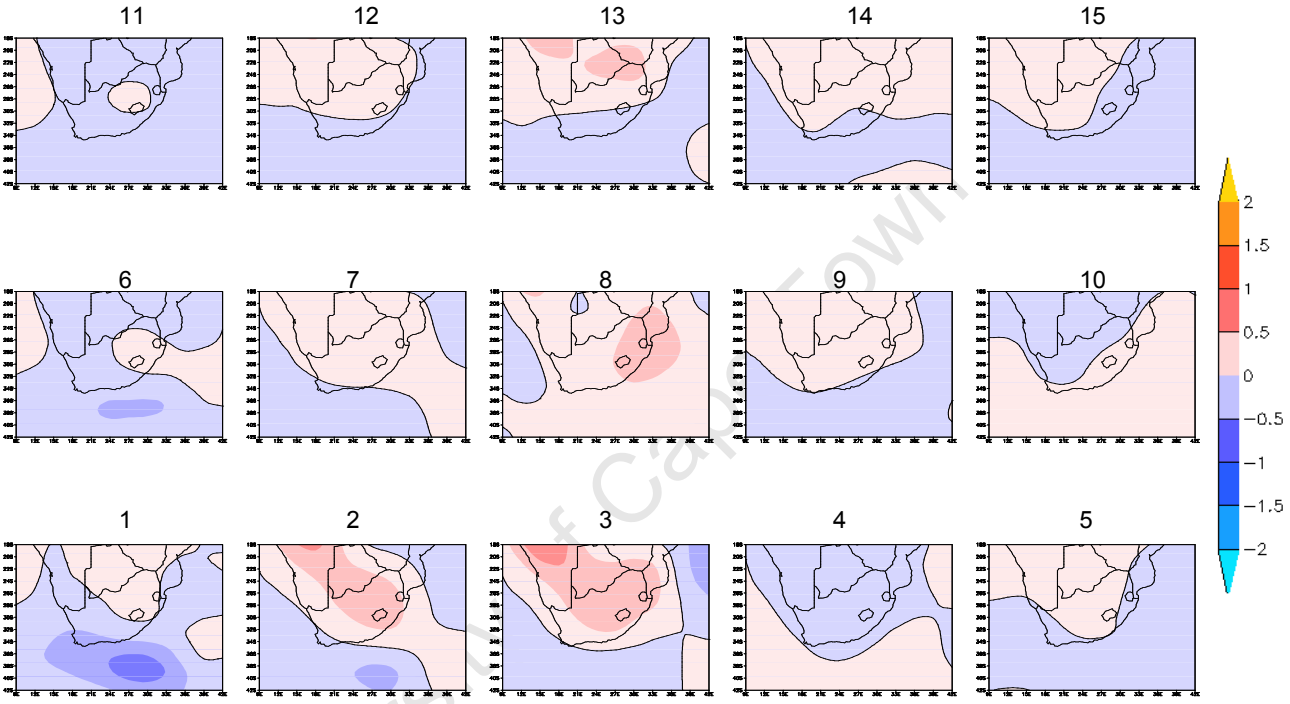


Figure 6.12: Multi-model weighted future projection – average multi-model mean for annual precipitation, contour interval of 0.2 mm

The differences between the weighted and unweighted climate change signal in figure 6.13 show small deviations both positive and negative, with the weighted solution wetter over interior than the average multi-model, by up to 0.4mm per day for the strong continental thermal trough states. The rainfall reduces in the future climate under anticyclones but this is not as consistent as in figure 6.12.

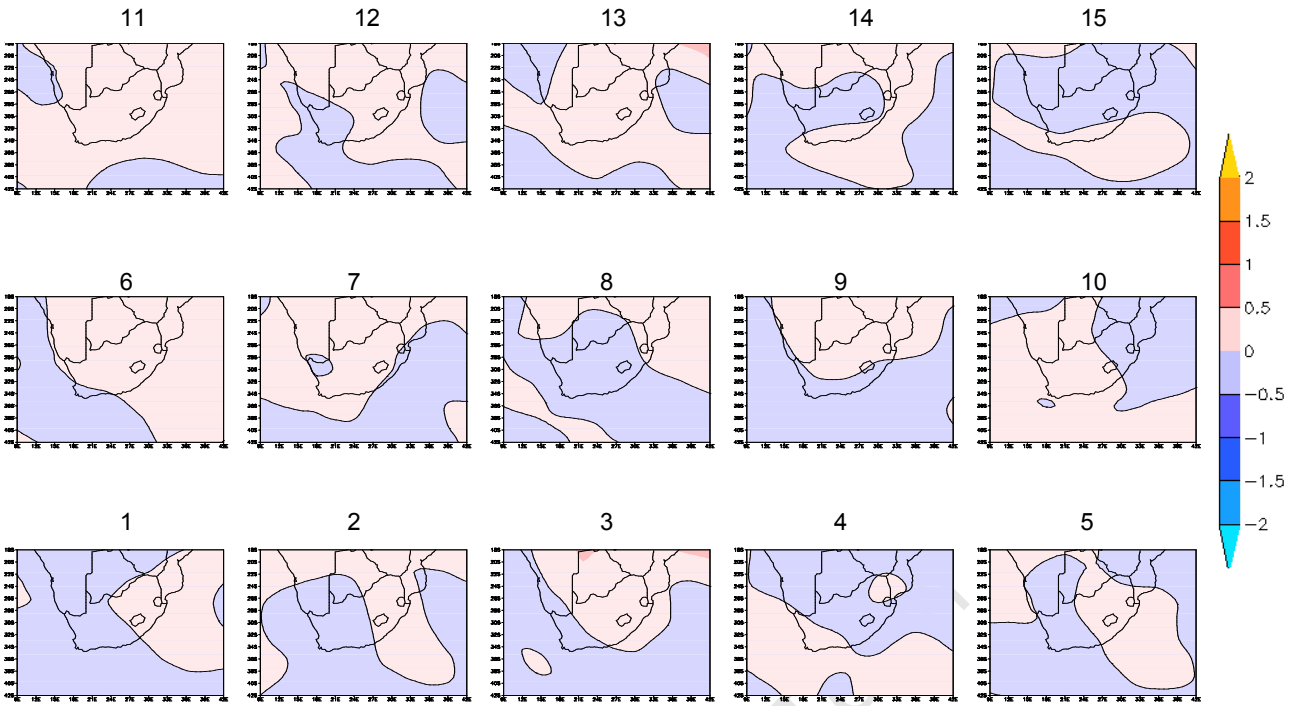


Figure 6.13: Multi-model weighted climate change delta (future – control) projection – average multi-model delta for annual precipitation, contour interval of 0.2 mm. Red indicates where the weighted multi-model climate change signal is larger than the average multi-model composite. Blue where average is greater than weighted multi-model

Monthly analyses

A quick overview of the monthly results is shown below, drawing on only a few figures of the weights and anomaly maps that indicate dynamical changes not seen in the annual SOM. The rest of the figures can be found in Appendix IV.

As an example, in the April temperature maps (Appendix IV, figure D8), node 7 shows a large improvement in skill in the weighted projection; however in the precipitation map (Appendix IV, figure D9), the same node shows a loss of skill. This is due to the fact that the technique is correcting for frequency and assumes that the model dynamics are the same as NCEP. However as discussed in 5.3, the reality is that the model has systematic biases.

Similar analyses to those shown in the previous section of the differences between weighted and unweighted projections of monthly data were undertaken. The weighted multi-model nodes were compared to the multi-model average and show similar results to the annual figures 6.5 and 6.10. Anomalies for monthly data are less consistent in terms of sign of change and cannot be related as easily to specific synoptic processes. This is likely as a result of the greater intra-seasonal variance and differences in the models seasonal boundaries, which precludes any meaningful conclusions. See Appendix IV for some examples. It is likely that some of the

inconsistencies at monthly scales are a result of differences in the timing of seasonal transition between models.

In the next section, the NCEP baseline is used in place of the model baseline. This should account for the model's systematic bias in representing the mean climate. The node weights are applied to the climate change anomaly (difference between control and future projections) to correct for frequency bias using the node specific weighting.

6.2 WEIGHTED FUTURE CLIMATE PROJECTIONS

This method builds on the positive results in section 5.3, to use an NCEP baseline for the mean conditions and to add a weighted future climate change component to this. However, in this case instead of a single mean for the whole time series as in 5.3, each node will have its own NCEP mean. The use of the NCEP baseline for each synoptic state will adjust for the systematic biases. The climate change anomaly is weighted with the frequency correction metric described in section 6.1. Weighting the future component of the projections based on each models relative frequency bias, in conjunction with the NCEP baseline, helps to create a more accurate multi model future projection. The adjusted future climate results are compared to the original unweighted SOM results to see what biases are affected by the change in baseline and frequency adjustment.

The frequency adjusted weights (taken from 6.1) for each node of each model are applied to the climate change anomaly (i.e. the difference between control and future projections) and added onto the NCEP node average (see Equation 7). This will remove the systematic bias from the model as well as correct for the frequency differences.

$$Mfutri = \frac{\sum_{m=1}^{3} Wi(Pi - Ci)}{\sum_{m=1}^{3} Wi} + Ni \qquad \text{(E.7)}$$

From a user's perspective, one may want to use the correction on the original time series source data. A possible methodology would be to use the node assignment to determine the weight (Wi), derive a climate change anomaly by subtracting the average control observation for that node (Ci) from the future observation (Pi) and using the NCEP node average as the baseline component (Ni). However, this is beyond the scope of this work.

6.2.1 Temperature

The results of the weighted multi-model adjustment for each SOM node are depicted by a greyscale shaded plot (part a). Each node's difference between the future weighted value and NCEP baseline values in colour plots (part b) which shows what component the weighted climate change anomaly contributes to the overall result. In the temperature data all contours are shades of red as the temperatures are always increasing in future projections. The lighter the colour of the contour, the less warming under future conditions.

The annual corrected temperature results are shown in figure 6.14a. In figure 6.14b, the largest increases in temperature are over the three central columns which depict moderate anticyclones and an interior heat low. The larger increases for the more stable conditions are probably related to less cloud cover when high pressure is dominant, leading to greater insolation and therefore greater warming. Increases are smaller in nodes 1 and 5 (cold fronts) which are generally associated with low pressure and more cloud cover, which also supports this idea.

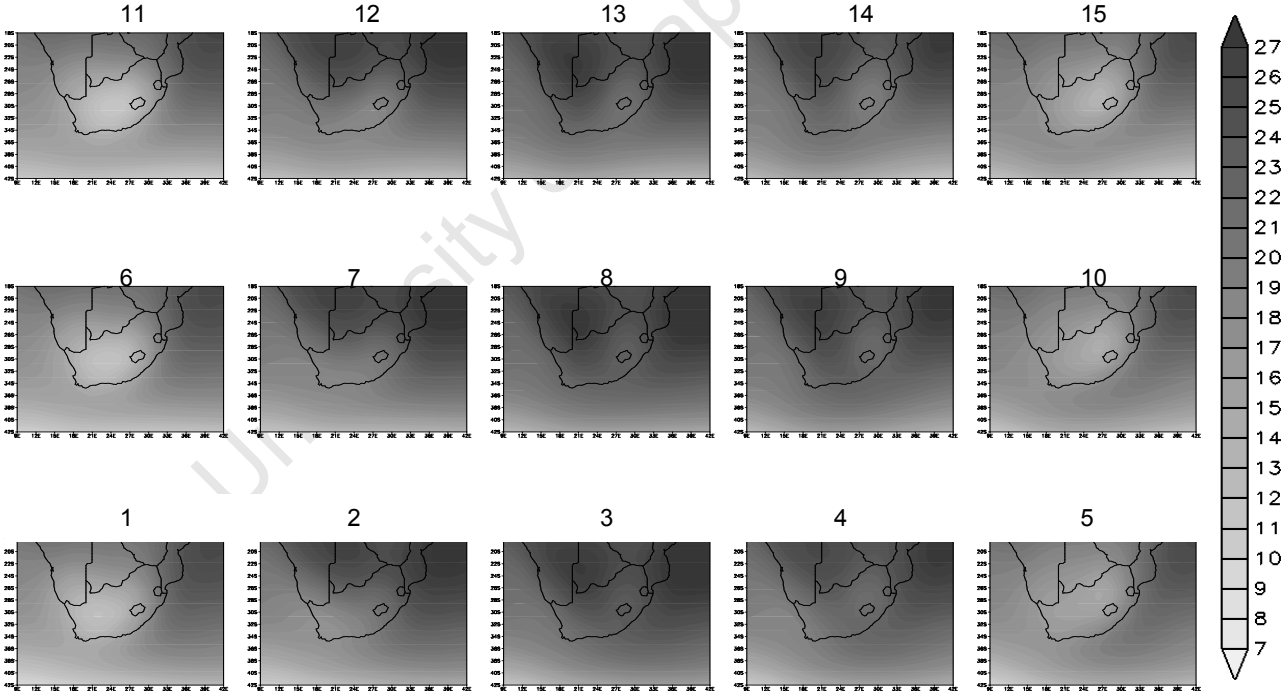


Figure 6.14a: Weighted temperature per node, contour interval 0.5°C

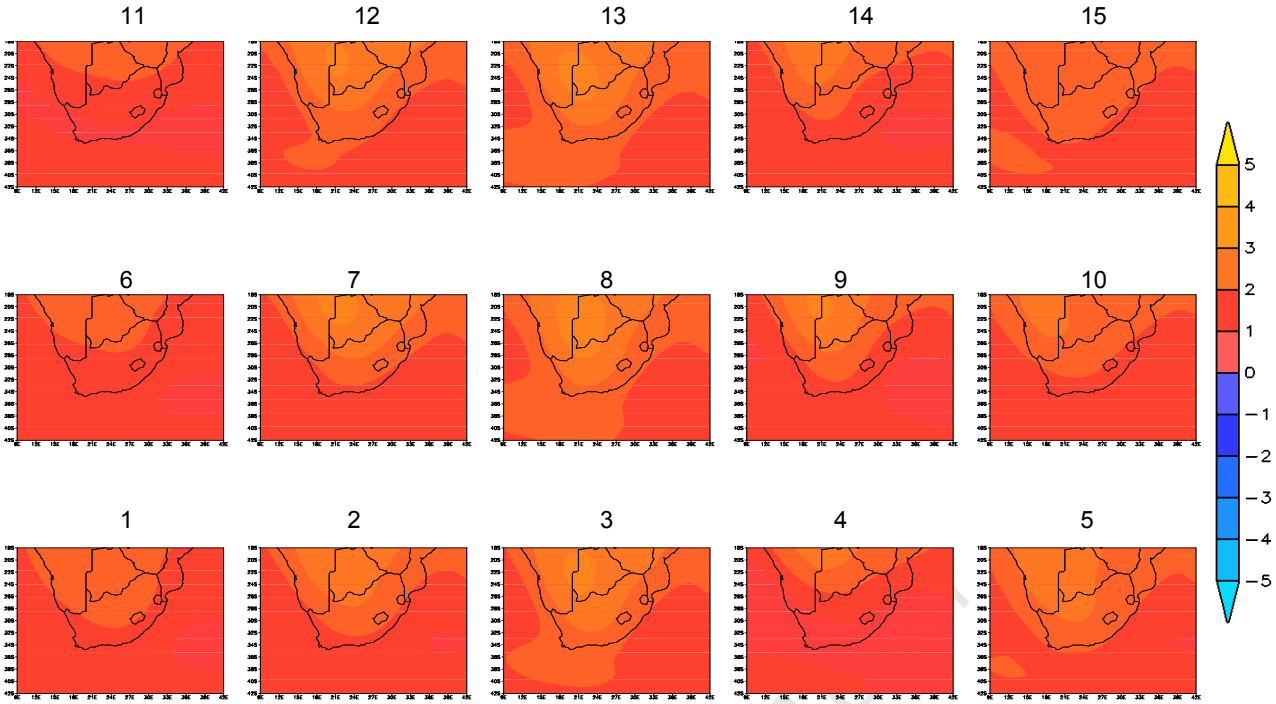


Figure 6.14b: Weighted temperature climate change signal (weighted multi-model – NCEP) per node, contour interval 0.5°C

In figure 6.15a, the original multi-model average temperature is shown. Figure 6.15b shows the difference between future and control node composites. The climate change signal is fairly consistent between the two figures which mean that on the annual scale the weighting has little effect. As the annual data has less variation between models, it is not surprising that the climate change anomaly would be similar in position and magnitude. The real magnitude is however adjusted for with the NCEP baseline, which warms the surface air temperature over the Southern Ocean in particular.

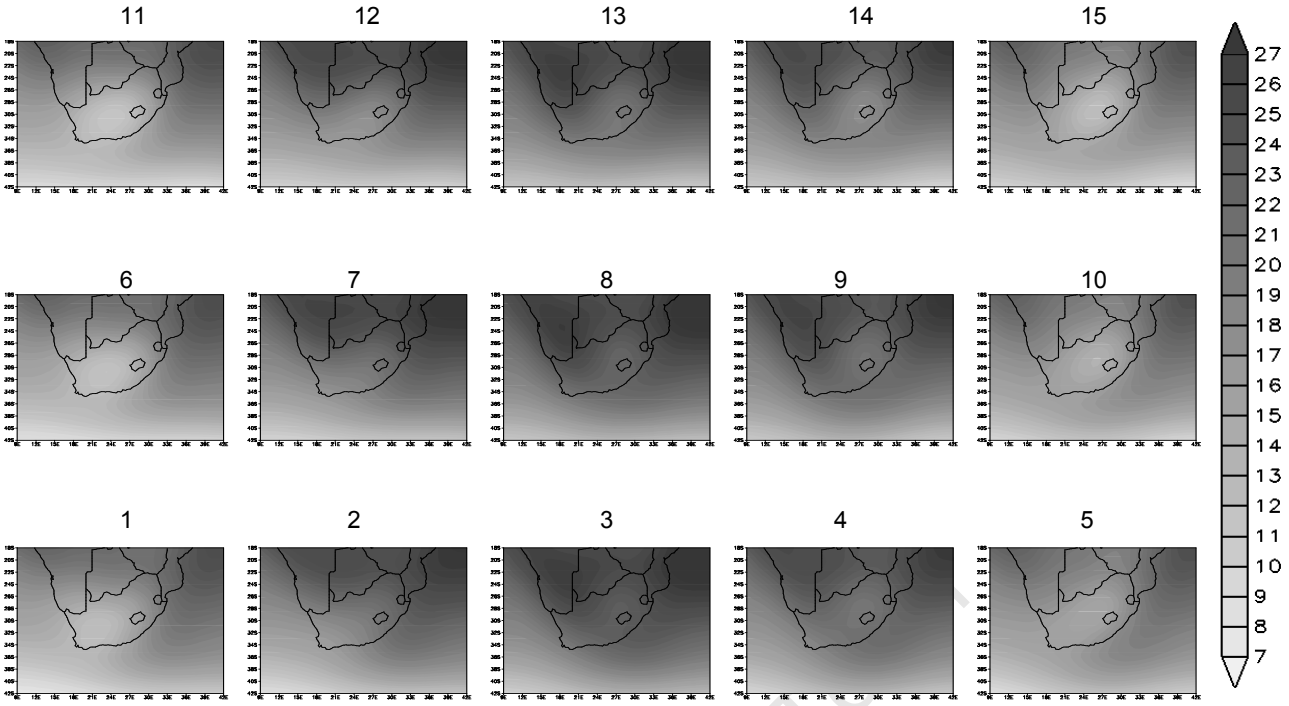


Figure 6.15a: Multi-model average annual temperature mean per node, contour interval 0.5ºC

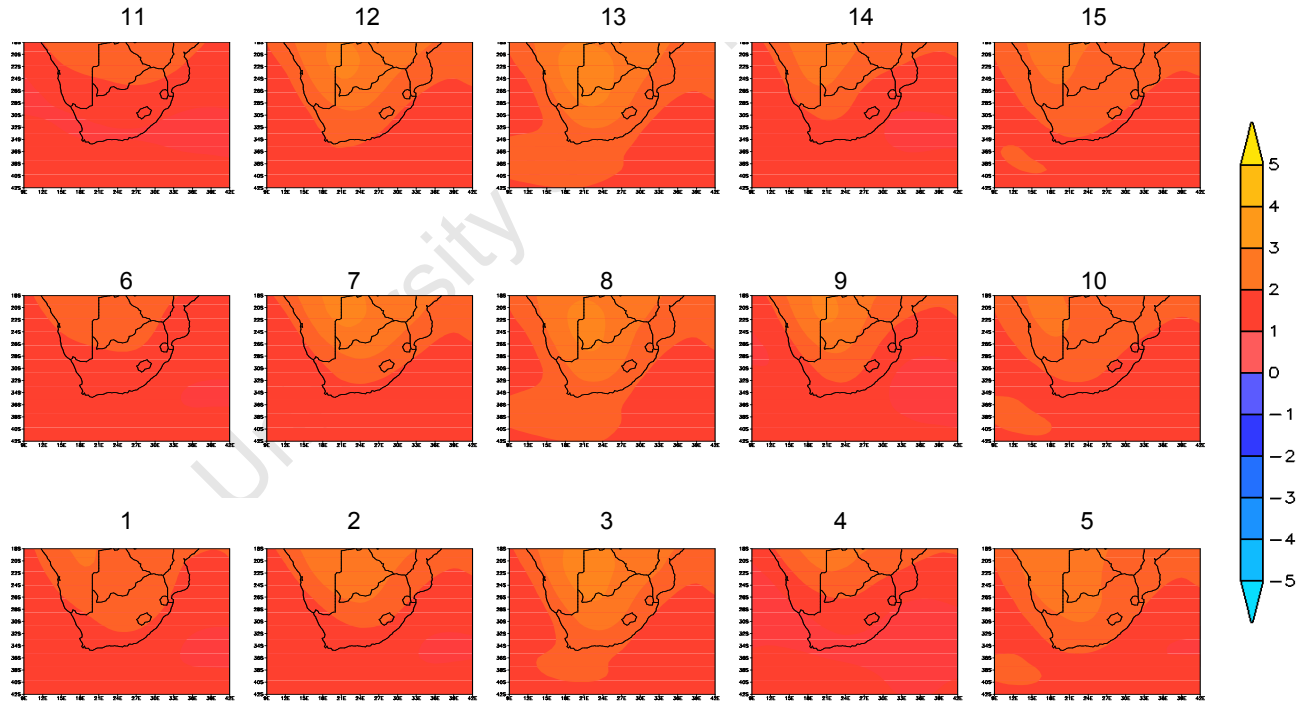


Figure 6.15b: Multi-model average annual temperature climate change signal (future – control) per node, contour interval 0.5ºC

As we know that the frequency weighting function has only made small adjustments, most of the changes are attributable to the NCEP baseline adjustment. In 6.16 we see the difference between the average multi-model and weighted multi-model results. Blue indicates where the weighted results are lower than the average and denote a cooling, whereas red indicates the

increase of temperatures. The persistent biases that were identified in Chapter 4 are corrected for as shown by the increased Southern Ocean surface air temperatures and the decrease over the northwest coast.

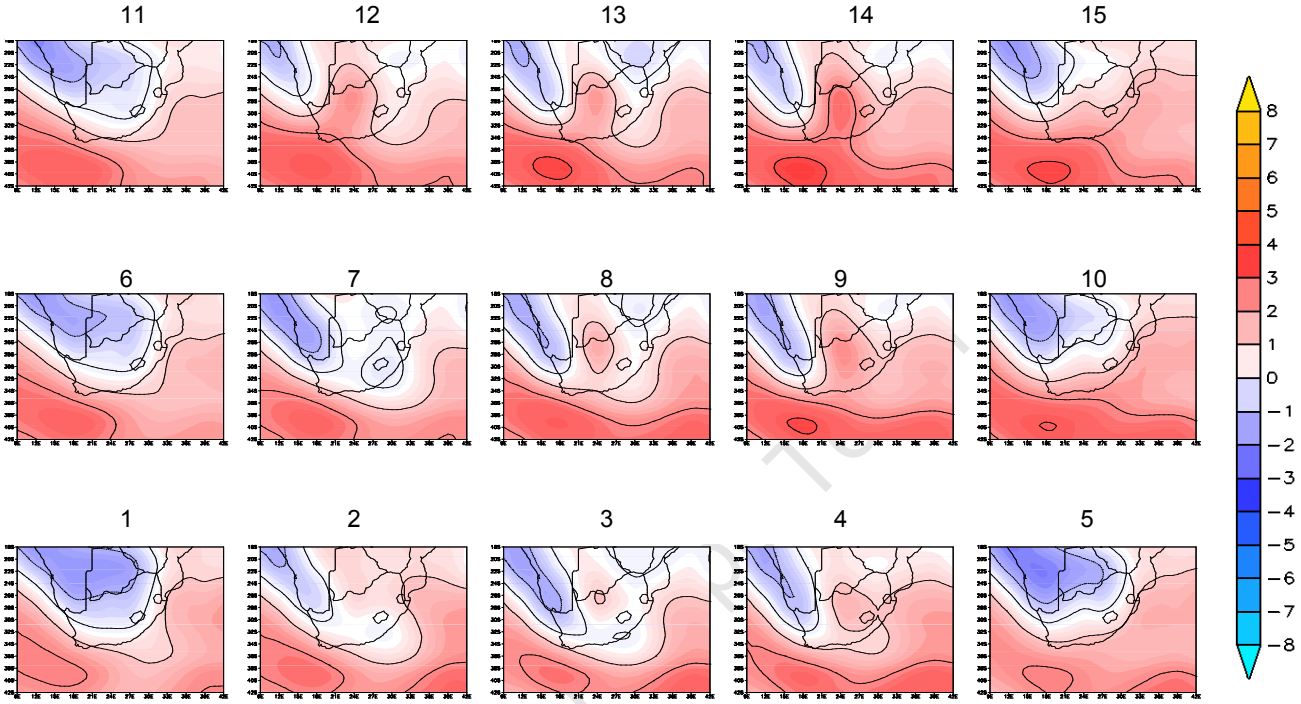


Figure 6.16: Anomaly of annual temperature weighted future projections - unweighted multi-model average, contour interval 1°C

From these results it seems that the large scale biases are much more clearly identifiable when the multi-model results are used. Persistent biases have been targeted and substantially reduced by this technique.

6.2.2 Precipitation

The same technique is applied to the precipitation data, where the internal node dynamics have had a stronger influence on the rainfall response. Therefore, less persistent biases are identifiable but rather dynamic based biases. Figure 6.17 shows the weighted precipitation field in the shaded plot (greyscale colour bar to the left), overlaid with the climate change anomaly in contours (colour bar to the right). Blue indicates areas of decreasing precipitation and red increases. In the annual data there is little change in this parameter with mostly small decreases of up to 0.5mm or an unchanged signal. Slight increases match areas of low pressure in the node which indicate that the models slightly lower rainfall associated with low pressure by about 0.2mm.

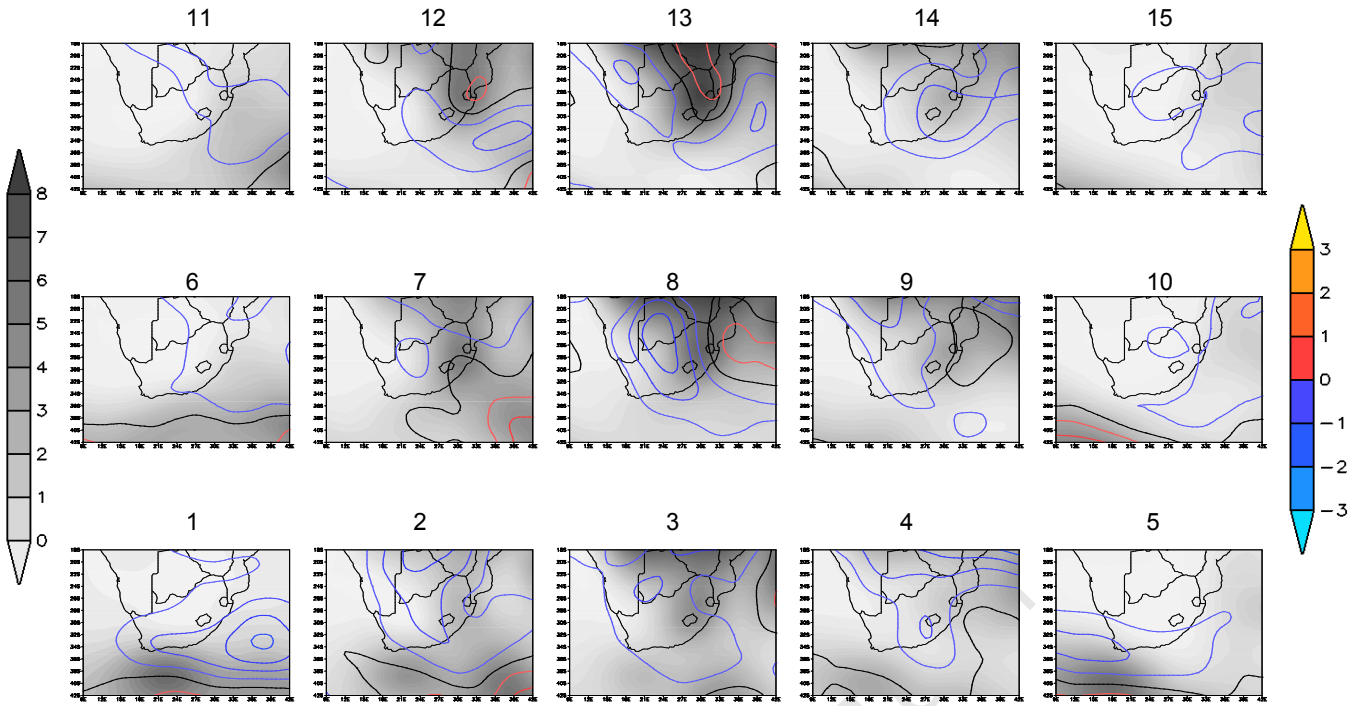


Figure 6.17: Weighted annual precipitation overlaid with climate change signal (weighted multi-model – NCEP), contour interval 0.2mm

The difference between the weighted (figure 6.17) and average (figure 6.18) multi-model climate change anomaly is small and the variance is in magnitude only, not position. This confirms that the baseline is making a greater difference to the resultant weighted values.

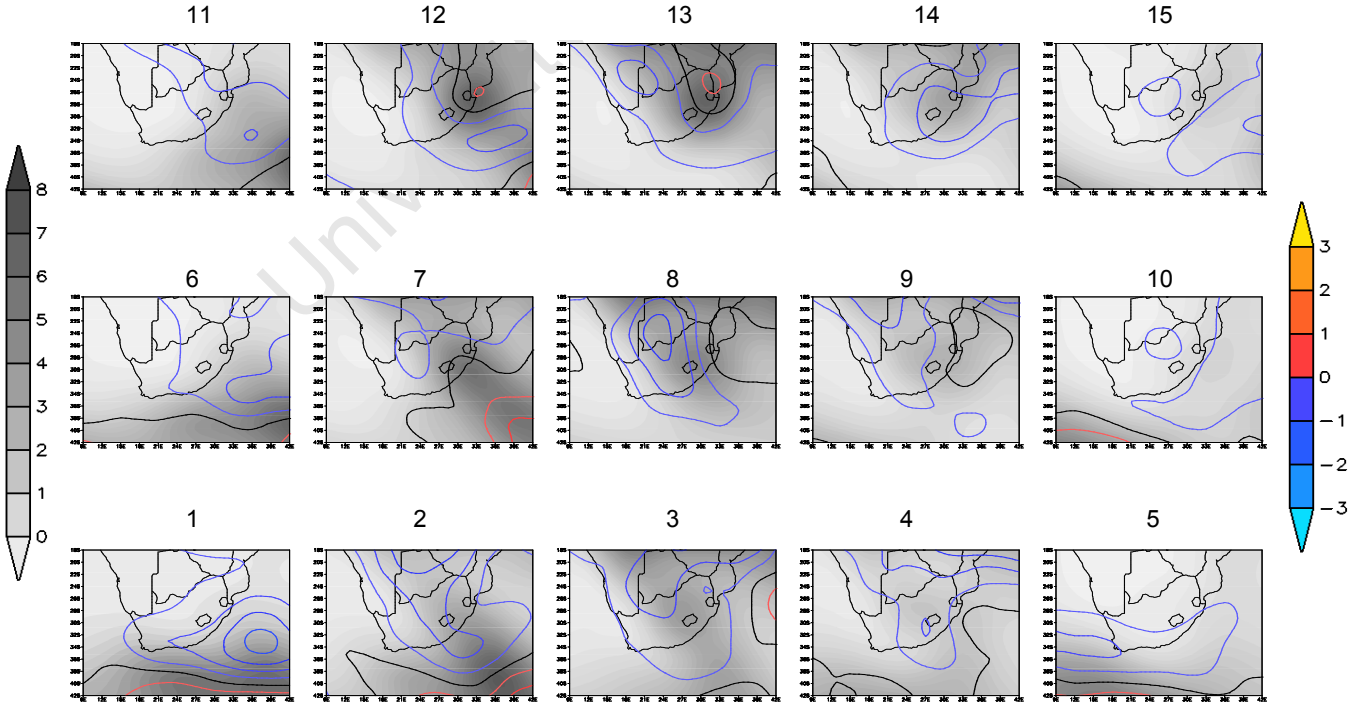


Figure 6.18: Multi-model average annual precipitation overlaid with climate change signal (future – control), contour interval 0.2mm

In figure 6.19 the difference between weighted and average results show that the centre columns of the SOM (typically summer states) have corrected for an under simulation of rains on the northeast of the continent, and decreased rains over the South Indian (one of the persistent biases in Chapter 4). The Southern Ocean dipole pattern is also adjusted to reflect the correct polarity.

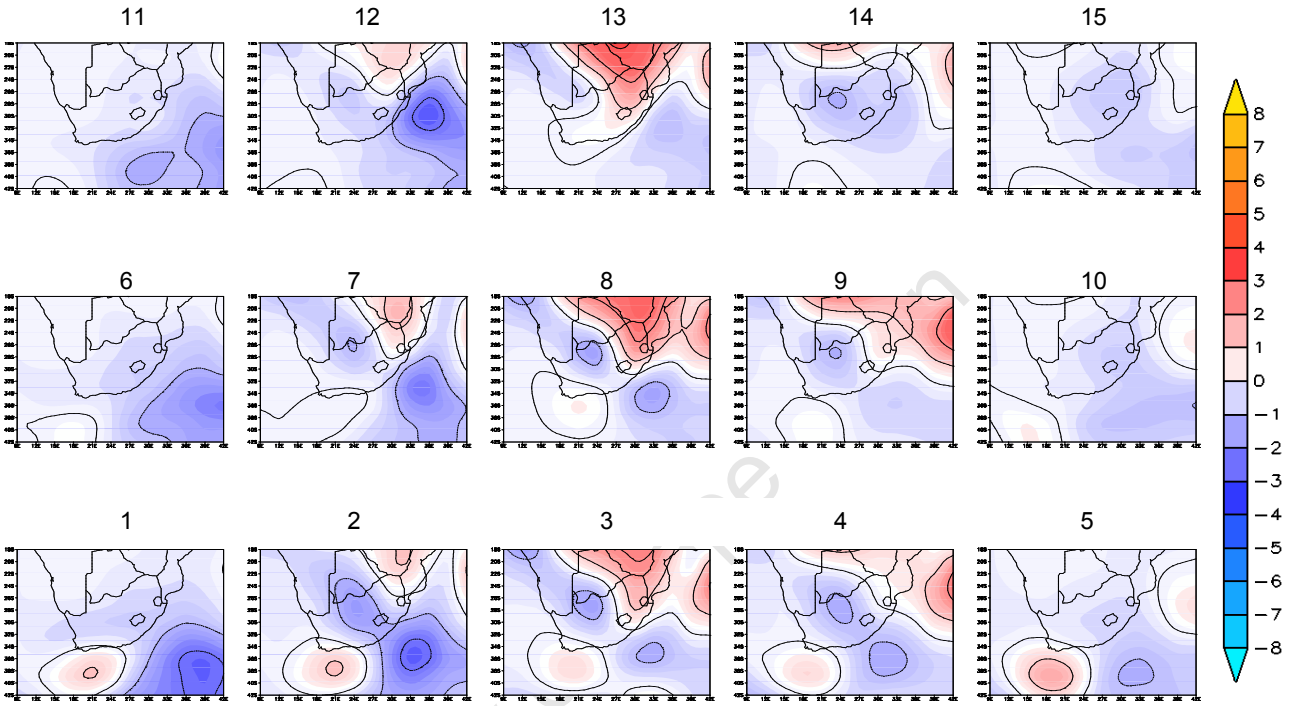


Figure 6.19: Anomaly of annual precipitation weighted future projections - unweighted multi-model average, contour interval 1mm

Because the technique adjusts the anomaly to reflect the observed baseline, the multi-model node specific weighting solution has taken out many of the model specific biases and that is why the persistent biases are more visible.

Monthly data

The same process is done for each month. Figures not shown can be found in Appendix IV. Temperature increases are generally larger over the northern continental part of the domain and lowest over the oceans due to continentality effects. For monthly precipitation, the same figures were generated and can be found in Appendix IV.

In the January temperature field (see Appendix IV, figure D21), the climate change signal is largest in the outer nodes, particularly the top row (strong anticyclones with interior low). The last column (connected tropical low) has slightly stronger warming signals than the multi-model average; however the positioning is consistent in all nodes between weighted and multi-model

average. Looking at the weighted versus multi mode average anomaly, a similar but stronger than annual warming is discernable over the Southern Ocean, with a small decrease over the northeast, centred on Zimbabwe. In January precipitation, the weighted multi-model nodes have on average higher maximum values than the multi-model average counterpart node. Almost all changes are positive especially over the east coast where increases of more than 2mm are seen for node 6 (ridging Indian high with interior low). Summer precipitation is increased by as much as 6mm due to the baseline adjustment, as shown in figure 6.20.

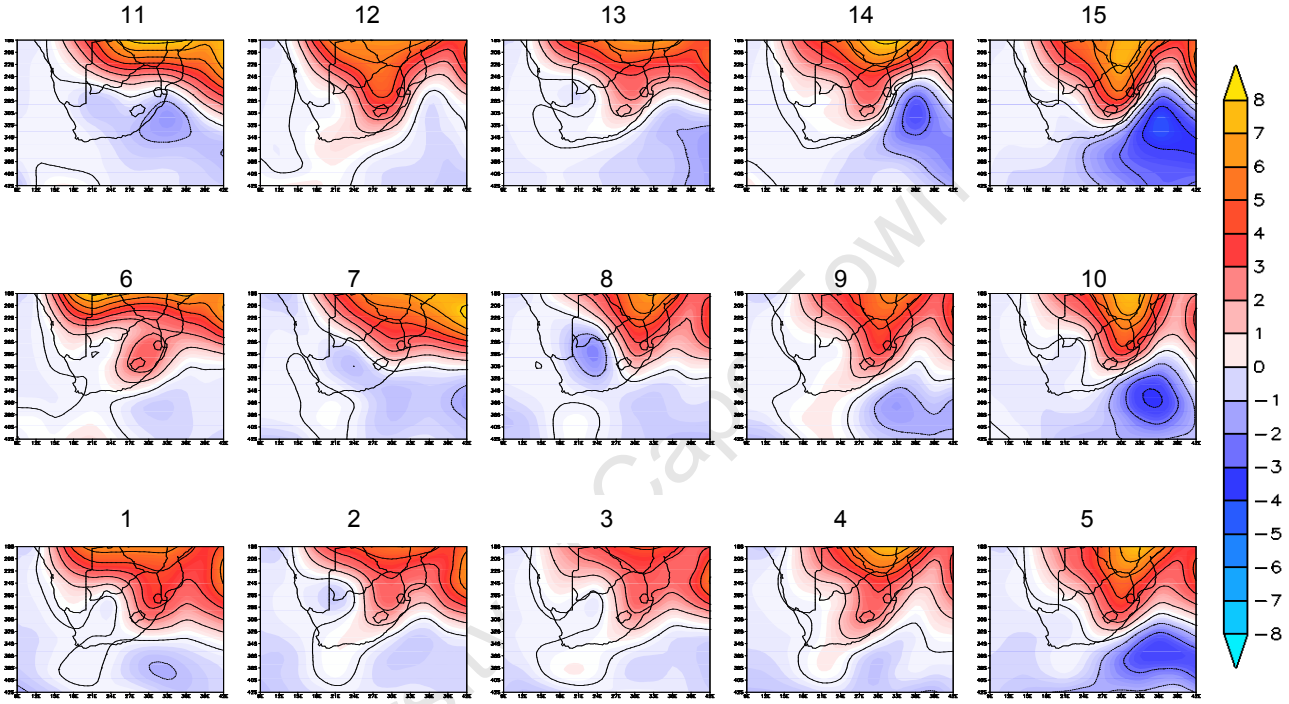


Figure 6.20: Anomaly of January precipitation weighted future projections - unweighted multi-model average, contour interval 1mm

In April the climate change warming signal is consistent between the weighted and average multi-model results. In terms of the baseline adjustment there is a small warming over the Southern Ocean, with strong decreases over the west coast (see Appendix IV D22). For precipitation, most of the nodes show little change between the weighted and average multi-model projections, except in the last column (weak troughs) where the decreases are slightly larger in the weighted projections. The decrease in the comparative heating (Appendix IV, figure D22 and D24) may account for this as the systems would have less heat driving the heat low (convective) precipitation. Increases in precipitation over the Southern Ocean are associated with troughs (figure 6.21).

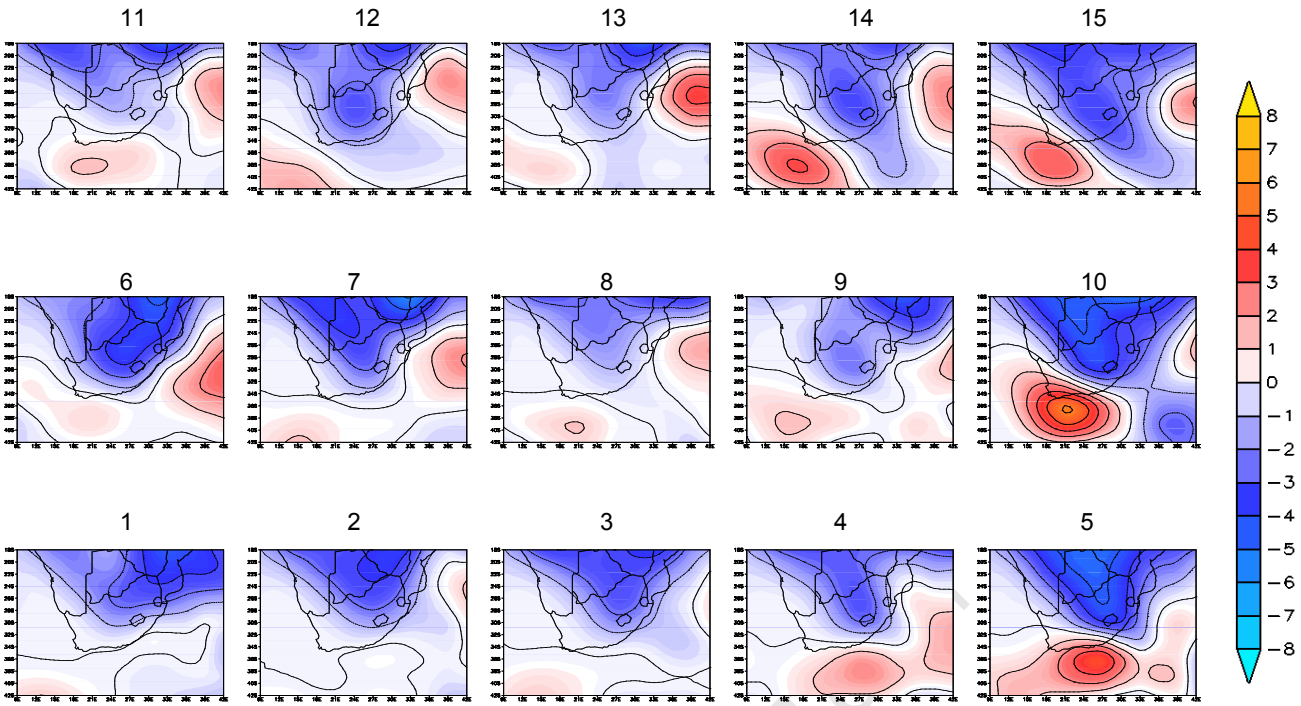


Figure 6.21: Anomaly of April precipitation weighted future projections - unweighted multi-model average, contour interval 1mm

In July the climate change temperature increase is smaller than in other months, probably due to the smaller insolation received in winter. Some nodes have a slightly different position for the climate change anomaly, especially in the left half of the top row (ridging South Atlantic anticyclones). The magnitude of change is consistent however. In figure 6.22, the warming over the Southern Ocean is greater in the top row of nodes, and shifted to the west instead of over the whole ocean. The interior shows significant decreases, the centre of which shifts dependant on which of the high pressures is dominant in the node.

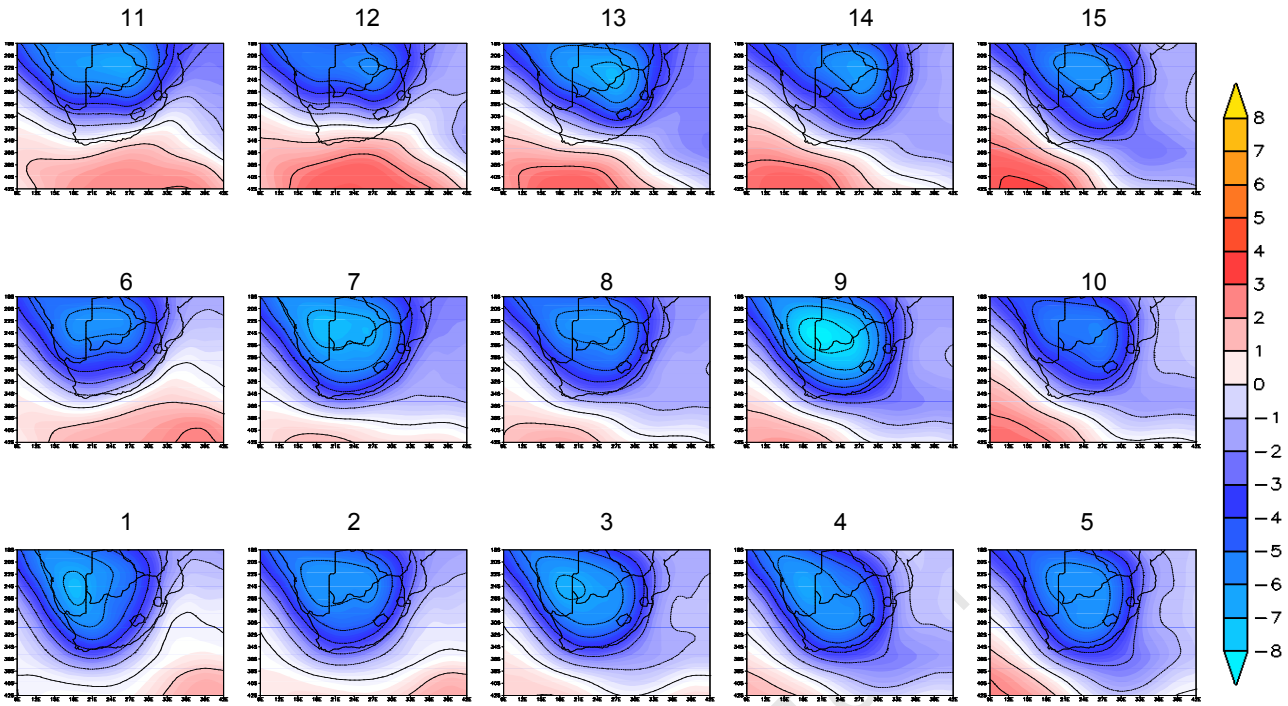


Figure 6.22: Anomaly of July temperature weighted future projections - unweighted multi-model average, contour interval 1°C

July rainfall anomalies for weighted and average multi-model projections are fairly consistent in magnitude with minor variation in position. Node 11 (strong Atlantic ridging anticyclone) has a noticeable increase over the Indian Ocean not seen in the average. Looking at the weighted anomaly field (figure 6.23), there is a strong decrease in rains over the eastern half of the domain for this node and slight increases in rains over the South Atlantic Ocean in the strong westerly wave/cold front states.

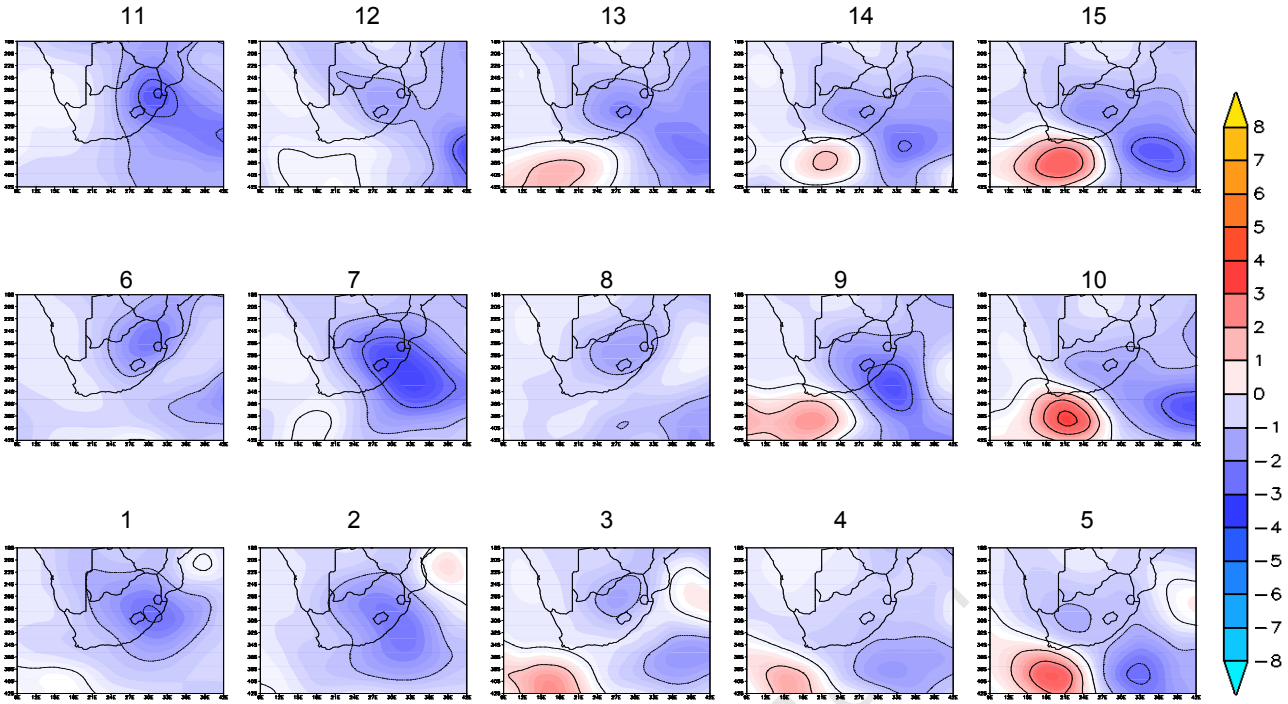


Figure 6.23: Anomaly of July precipitation weighted future projections - unweighted multi-model average, contour interval 1mm

In October the climate change signal for temperature increase is very strong, especially in the central node (stable condition with dominant highs over oceans and interior low) where warming is as great as 5.2°C. Most nodes have an increase of up to 3°C and by at least 0.6°C. The positioning of the climate change anomaly is consistent between weighted and average multi-model projections, however, the weighted model anomalies magnitude differs by up to 0.4°C in either direction. The bias identified in Chapter 4 to under estimate temperatures in October is corrected for by the baseline adjustment, resulting in increases of as much as 8°C as shown in figure 6.24.

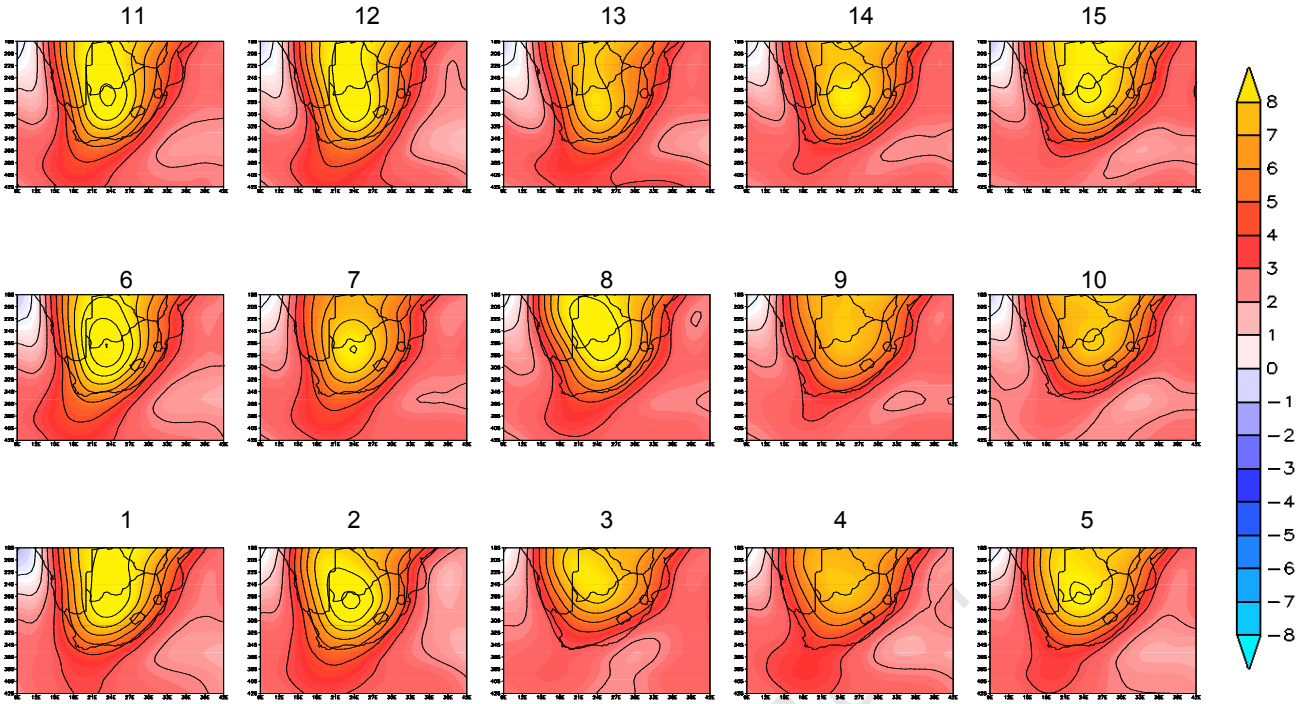


Figure 6.24: Anomaly of October temperature weighted future projections - unweighted multi-model average, contour interval 1°C

For October precipitation, the only states with deviations from the average multi-model climate change anomaly are in nodes 3 and 8 (interior heat low) where the weighted projections show slight increases of 0.5mm over the east coast in addition to the larger 1.5mm decrease over the Southern Ocean seen in both. The weighted anomaly (figure 6.25) shows that the baseline adjustment reduces the rainfall over the east coast significantly for these two nodes, therefore even though the climate change signal is positive the observed bias is so large for this node that the difference is still 9mm.

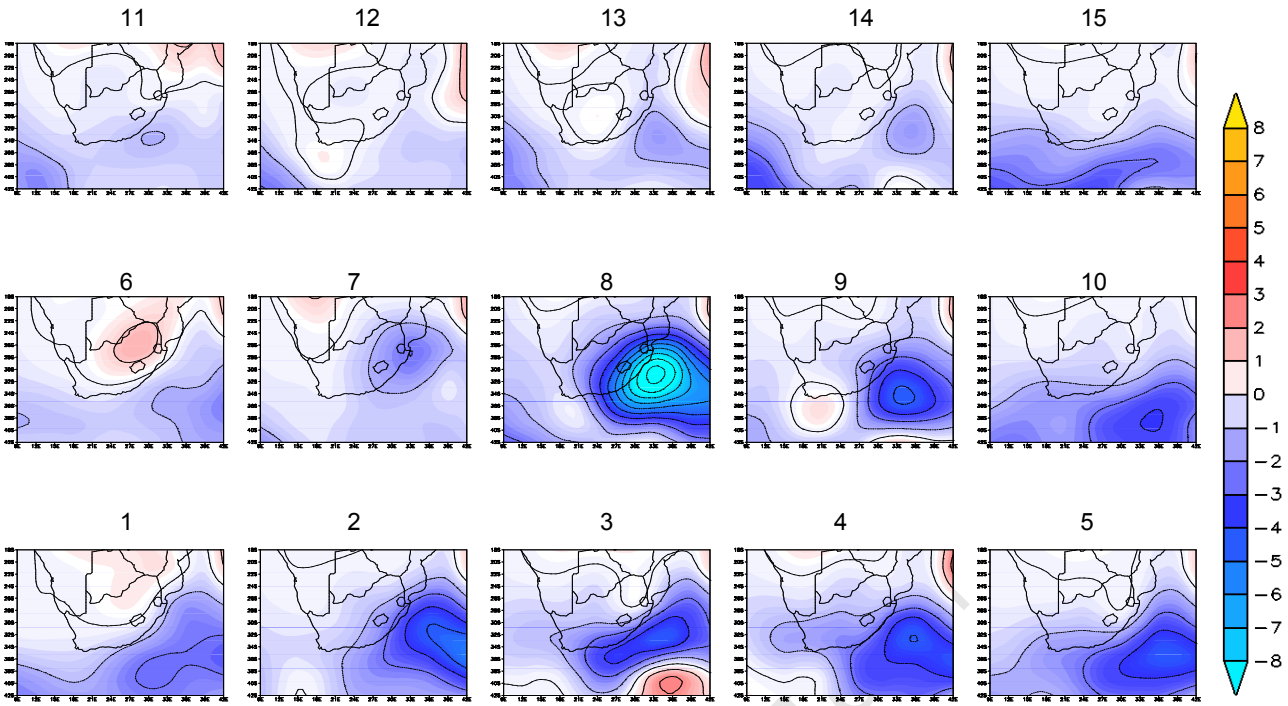


Figure 6.25: Anomaly of October precipitation weighted future projections - unweighted multi-model average, contour interval 1mm

In each month the persistent biases over the ocean and continent are modified by the position of the large scale forcings (high and low pressure systems) which accounts for the variation in anomalies while maintaining similar bias features.

6.3 CONCLUDING THOUGHTS

The adjusted results demonstrate that the original projections had many persistent biases (i.e. a bias that is seen in all states or in all models for a single state). These biases are propagated through into the future data and result in some regional processes being under or over represented.

In summary, the correction technique used in this section has been able to reduce many of the large scale, persistent biases identified in previous chapters. Examples of corrected results are:

- surface air temperature over the Southern Ocean is warming in all models for most nodes in all seasons;
- increasing the summer high temperatures;
- decreasing winter minimum temperature;
- reducing precipitation in the Indian Ocean;
- increasing summer rainfall over eastern half of continent;

These were some of the main biases identified in Chapter 4 and therefore provide credibility to using this method to obtain more accurate future predictions for a specific synoptic state. In the final technique much improvement was made to the projections through correction of the baseline to the observed value and using a weighting based on frequency bias to limit the influence of biased model results in the future multi-model projection.

When comparing the weighted versus unweighted multi-model projections, there are some implications for the interpretation of regional projections. The corrected results have substantial differences for certain synoptic states and are highly seasonal in differences of characterisation. When strong continental thermal trough features are prominent, there is a reduction in temperature and slightly greater precipitation over the continent in the weighted nodes. For states with strong anticyclones, the weighted multi-model solution has warmer temperatures and lower precipitation that the multi-model average. These tendencies are also seen in the future weighted projections. Additionally, westerly wave features are linked to decreased rain in weighted future results. Likely drivers of these results are more intense troughs and less frontal rainfall.

Comparison of the magnitude of the climate change signal showed that the weighted multi-model composite winter states showed a reduced climate change signal compared to the unweighted and greater warming and increased rainfall, especially over the oceans, in summer states. This suggests the models are over estimating winter changes in temperature and underestimating the increases in summer.

In the final chapter, the results will be discussed in terms of the overall thesis objectives with an aim to synthesise the information presented thus far.

CHAPTER SEVEN

Summary and Conclusion

This chapter presents an overview and discussion of the key findings in view of the thesis objectives, followed by a discussion of problems and caveats, as well as recommendations for future work.

The study sought to quantify aspects of uncertainty found in GCM projections of climate change, and develop an approach by which this uncertainty can be reduced. This was achieved by first examining GCM bias in terms of the frequency of daily synoptic events, and then considering the bias of surface air temperature and precipitation as a function of synoptic mode. Self-Organising Maps (SOMs) were used to identify bias in the synoptic circulation modes characterised by atmospheric circulation variables, indicative of the large scale dynamics of circulation, which is the scale at which GCMs are most skilful. This bias can be thought of as a "goodness of fit" measure of the mapping to the generalized set of synoptic states.

Existing tools to address GCM uncertainty have rarely made use of daily synoptic data. Using data at the daily timescale instead of time aggregated means (monthly or even decadal means are often used in other studies) allows for synoptic specific uncertainties to be addressed as well as uncertainty from base state. Advantages of using SOMs for synoptic mode analysis include, that the SOM is an easily visualised tool, which allows for easy spatial intercomparison between model results and addresses sub-regional scale uncertainties (as opposed to single values for whole continental or sub-continental regions as used in many existing techniques). As the SOM separates the data into synoptic archetypes, the data is much easier to manage.

Based on this analysis of the synoptic circulation, three approaches are presented to compensate for the biases in both the control and future climate simulations of temperature and precipitation. These are:

- Using the observed frequency in place of model frequency to produce an average for the SOM series
- Creating a multi-model mean by weighting each model, based on the performance of all nodes in simulating the observed climate
- Creating a multi-model mean by weighting each model per node, based upon the models' relative abilities in simulating observed climate

Biases were identified using four metrics: the average error for each synoptic state, the comparative frequency of each synoptic state in each model, the evolution of synoptic states in terms of daily transitions between states, and the similarity of the atmospheric circulation inputs, mean rainfall and surface air temperature patterns between the models and observed data according to synoptic state. Rainfall and temperature data were not direct inputs in the SOM mapping process, which was based on circulation variables. Instead average composites for each node were derived.

Of these metrics, the differences in the frequency at which models represent daily synoptic states compared to observed data proved to be the most useful in determining a weighting function for bias correction. The frequency difference was used to calculate a weight describing each model's realism in simulating a particular synoptic state, based upon the similarity between the model and the observed data's frequencies. These weights were used firstly to produce an aggregated weight for each model which depicted the average accuracy across modes in depicting the correct frequency and then for specific synoptic states. Both single model and multi-model solutions were employed, with varying degrees of success. In conjunction with this weighting technique, the future data was corrected by weighting the climate change anomaly (difference between future and control projections) according to the models relative performance in simulating the correct frequency of present climate. This weighted future anomaly was then added to the NCEP observed baseline, producing a weighted multi-model solution. The last technique proved to be the most successful in correcting bias due to differences in the frequency of occurrence of synoptic states compared with the observational data.

7.1 OVERVIEW AND SUMMARY

The aim of the study was to identify and reduce uncertainties from GCMs, with particular attention to the synoptic time scales. This section recaps the identified biases and outlines the results of the adjusted climate change projections and the implications these results have for other uncertainty studies.

7.1.1 Choosing a metric for weighting biases

In analysing the four metrics of bias, model error is measured by the average Euclidean distance of each daily entry from the SOM node. Therefore it is not a statistic of inaccuracy but rather a measure of how well the synoptic state represents the data entries assigned to it. It was seen that the error values for mapping data to a node are in general higher in the model than NCEP. This is because the data used to train the initial SOM is from NCEP, therefore the

model is not a direct input and therefore differences in the models simulation of control climate will result in higher error. Table 3.1 showed that the range of error values was smaller in the future than the control projections. Error in the control projections is highest in outer nodes, and often decreases in the future projections, which could indicate that the future dynamics are more similar to the NCEP control, as explained in section 3.2.1. Errors in the future projections do not seem to be higher for any particular synoptic state but are generally lower (at times less than NCEP) in the centre of the SOM. This is related to the middle states being more transient, less persistent states.

In terms of frequency, outer nodes also tend to have higher frequencies in the models as compared to NCEP in both annual and monthly plots. Once again, this is related to the transience of the middle states. In general, strong South Indian anticyclones were under represented by the models in all seasons. Continental thermal trough features were over represented by all models in October. Stable conditions brought on by dominant highs were favoured by the models in July, and January had increased continental lows compared to observed nodes.

The models representation of the daily transitions between synoptic states (temporal evolution) captured the same basic daily evolutionary cycle seen in the observed data, with slight variance in the average direction in SOM space and the rate of change between states, but overall maintaining the same pattern. The models behave more similarly to each other than to NCEP, although the deviations from NCEP were slight. Model similarities in terms of time-evolution are possibly due to the fact that GCM output typically has a lower variance than reanalysis data, which is constrained by observations.

The predominant differences between observed and model circulation, common to all models, include greater south-westerly flow, indicating that the high pressure cells are shifted further south than they should be and the westerly wave is stronger in the models that in observations. All models under represented easterly flow over the Mozambique coast and directly below this region they have greater easterly flow over Indian Ocean at surface level, which is strongest when Indian high is dominant over this region. All these indicate that the westerly wave is positioned further south in models which would account for the shift south of the high pressure cells and the resultant anomalies in flow. The models overestimate specific humidity in strong ridging anticyclone states over northern Namibia for the annual SOM mapping, and underestimate specific humidity on the east coast in the states with low pressure troughs over

the continent. These biases lead to larger anomalies for precipitation on the west and lower values to the east.

Temperature anomalies showed that: the Southern Ocean surface air temperatures were too cold in all models for all nodes in all seasons; the models underestimate the summer high temperatures; and overestimate winter minimum temperatures. This indicates the models have a general cooling bias. Precipitation changes showed that the models over estimated precipitation in the Indian Ocean; under estimated of summer rainfall for the eastern half of the continent; slight under estimated of rainfall associated with ridging Indian highs; and over estimated rain off the Natal coast in trough states extending south from the tropics.

7.1.2 Correcting for biases

The first technique tested the concept that the only source of model bias is the differences in the frequency at which they represent synoptic states. This was tested by using the frequency of NCEP observed data instead of the models frequency. The NCEP frequency was applied to each synoptic state's control model composite field. These were then combined into a single map mean for the control simulation for the variables of temperature and precipitation. This technique had some success in adjusting the model results towards the NCEP frequencies. However, given the fact that models are not only influenced by frequency differences and that the model means for each synoptic state are sometimes very different to the observed, especially in the monthly data, there is limited scope for improving the internal dynamics with this technique. The technique did provide a guide as to the amount of influence frequency has over mean state.

The second technique used the differences in frequency per node to determine a node weight. All the nodes weights where then combined into an aggregate model weight that represented the overall accuracy of each model. This weight was applied to each model's composite map means and then the weighted map means were averaged to produce a normalized weighted multi-model average. The potential flaw in this method was that if models have a similar weight to each other, the result may be indistinguishable from taking the multi model average. In cases where the weights were different, the technique did shift the model mean in the correct direction, however not by enough to make the multi-model solution consistently closer to the observed than any individual unweighted model mean.

The same methodology was then employed but against each synoptic state, using the node specific weighting (a measure of the difference in frequency between model and observed data)

to weight the mean composite fields of temperature and precipitation. The models node weighted mean composites were then combined for each synoptic state. As the technique is based on frequency correction alone, if the only model error was frequency the weighting would be able to match the NCEP distribution. However, there are other biases within the node that are not related to frequency but to differences in the mean node state. What can be established is the effect of frequency on the multi-model solution and from this one can gauge how changes in frequency generate changes in the magnitude of temperature and precipitation.

The third technique evolved from the two prior approaches, using node weights to correct the mean composite maps. However in this instance the weighting is only applied to the climate change anomaly (difference between future and control projections). As in the second technique the method was tested on the aggregate and node specific scale. The three models' weighted climate change anomalies were then combined and this multi model weighted anomaly is added to the NCEP baseline. In the aggregated version, the mean corrections were hampered by the similarity of weights in most of the data series (as in second technique) and therefore only the NCEP adjustment contributed to the overall improvement.

The node specific corrections combined all three models weighted climate change anomalies per node and added this to the NCEP baseline per node. The results show that the weighted anomaly does differ but that the use of the NCEP baseline accounts for most of the improvement in large scale model patterns accuracy. The technique showed that almost all the large scale biases identified were corrected for to some degree and therefore resulted in the most promising multi-model solutions.

7.1.3 Implications

Most previous uncertainty studies have examined GCM uncertainty by focusing on single grid cell differences for temperature or precipitation. As it is well known that the grid scale resolution is not especially accurate for these variables, this study attempted to address uncertainty in terms of different synoptic systems instead. This also allowed for the determination of which biases are as a result of synoptic conditions and which are present all the time.

As discussed in Chapter 1, some sources of uncertainty are due to difficulties in predicting a chaotic system, and some are introduced in the models representation of climate dynamics. One of the merits of the approach is that it separates out model bias (or error) from other sources of uncertainty, and thereby refines the range of future possibilities in a credible manner. From the results it is shown that differences in mean conditions for each synoptic state

was the largest contributing factor to the bias. Frequency bias, which had been thought to contribute much of the uncertainty, was shown to be a relatively small contributor to overall bias. This strongly suggests that baseline bias must be taken into account in uncertainty studies. Examples of the base line biases that strongly influenced the model deviations from the observations included a strong cold bias in surface air temperature, especially over the Southern Ocean. Factors that contribute to this are undoubtedly linked to the position of the large scale synoptic systems. However, the large anticyclone systems were shown to be shifted further south in the model projections, which would be more consistent with warmer not cooler temperatures. This means there are probably other factors contributing to the cold bias, most likely the radiation and convection schemes in the models.

The implication for future climate change scenarios is that a bias corrected baseline allows the model to be more useful in terms of determining realistic future impacts. Separating out different synoptic processes before applying a corrected baseline to each archetype allows for more detailed analysis of the how each state will change under future conditions (the delta function), while matching the baseline to observations (NCEP node average) imparts a more accurate representation of the current climate departure point.

7.2 CONSTRAINTS AND CAVEATS

There are a number of constraints to this study one of which is due to the reliance on NCEP data as the basis for real observations, second due to limitations imposed by the SOM methodology and lastly how the models represent the climate processes.

NCEP is a good proxy of observed data for model comparison purposes, but it should not be thought of as “reality” as it is a model itself which is forced by observed data, mostly from satellite remote sensing sources, scaled to a similar resolution of GCMs. The assumption in this work that NCEP is reality must be noted and, therefore, the corrections are to bring the data in line with NCEP, which may differ slightly from real observed regional responses.

The second set of constraints to the methodology comes from using the SOM tool. A number of alternative SOM techniques were attempted prior to identifying the methodology presented in this work. Formative methods to achieve the research objectives included using both model and NCEP inputs together to create a SOM space of all model and observed possibilities. The idea behind this was to see where the models overlapped in the SOM space with NCEP and to determine what states the models depicted that are not in the observed data and *vice versa*. Single variables and multi-variable SOM trainings were conducted, however, limited

interpretation could be made of these results as often the model and observed data occupied little or no overlapping space. When looking more closely at the results it seems this was a matter of scaling and gradients. Often the model gradients for geopotential heights and specific humidity were not as strong as the observed gradients, which is why the SOM assigned them to adjacent areas in the SOM space. For some variables such as specific humidity there was almost no overlap between the models and the observed due to the differences in the gradients, with the models unable to represent the very low and very high values. Geopotential heights had more overlap but due to the limitations with other variables the technique was modified to the observed data training only.

A second attempt to understand the variability within the models and to justify the method used was to train same size SOMs with non-standardised data from each data source separately and to measure the SOM space they occupied. This showed that the size of the SOM space was relatively consistent and that once again internal gradients played a bigger role. Both these issues were considered in the study, resulting in the small number of nodes used to force the model to choose a similar state even if the gradients are not entirely synchronous. The models must still be able to choose states that are similar enough to their daily patterns otherwise the error value would be much higher if the states were too dissimilar. The 15 node size was settled upon to account for both these concerns.

Other difficulties relate to the GCMs representation of climate. The accuracy of the analysis is constrained by the quality of the climate projections used. As GCMs have a very coarse resolution there are often difficulties in capturing the correct global modes of variability. This can impact the SOM mapping of the baseline conditions for example. This also extends to problems of stationarity, as explained in Chapter 1, where future climate dynamics may not operate in same way as the observed dynamics, resulting in statistical relationships altering. It was decided to proceed with the assumption that the statistical relationships between variables seen in observations will continue in the future, as at this time it is unclear how these relations might change. Lastly, the seasonal boundaries in the models are often slightly different to the observed conditions.

Some recommendations to expand this work include using more models and scenarios and running the SOM for different future time periods up to 2100. The data inputs could be expanded to different variables, such as geopotential heights and other large scale dynamical variables.

7.3 CONCLUSION

The thesis objectives are restated below, followed after each with a short synopsis of how each objective was achieved and how they relate to the results.

1. How can synoptic scale circulation, as an integrator of processes, be used to identify synoptic specific and persistent model bias?

The models had trouble simulating the correct position of the two dominant highs. The analysis seemed to indicate that these cells are shifted further south than they should be and that the westerly wave is stronger in the models that in observations. A persistent anomaly over the Mozambique coast that under represents the easterly flow could be linked to the shift south of the high pressures, thus reducing onshore easterlies associated with the anticyclonic flow around a high. This would also explain the surface level increases in easterly flow just below this region which is strongest when Indian high is dominant over this region. Temperature anomalies generally indicate that the models have a cold bias, which is most evident in the Southern Ocean, which has some of the largest biases, and is too cold in all seasons.

The models overestimate specific humidity along the west coast and underestimate on the east coast leading to larger anomalies for precipitation on the west and lower values to the east. The displaced Indian high leads to changes in the position of the moisture that is brought on shore on the northerly limb of the anticyclone and explains the reduced precipitation over this area in almost all nodes in the annual SOM. Precipitation anomalies indicated an over simulation of precipitation in the Indian Ocean (as a result of the shifted position of the anticyclone) and under estimation of summer rainfall over the eastern half of continent. This last is probably related to difficulties in resolving convective systems in this region.

2. What synoptic features do the models fail to capture and what features do they simulate accurately?

The models had trouble simulating the correct position of the two dominant highs, which would indicate that the westerly wave is displaced slightly. Temperature anomalies in the Southern Ocean are too cold in all seasons and generally the models have a cold bias. Precipitation anomalies indicated an over simulation of precipitation in the Indian Ocean and under estimation of summer rainfall over eastern half of continent. This is probably related to difficulties in resolving convective systems in this region.

The models tend to be more accurate in depicting the areas of high pressure, even though the exact positioning and magnitude may not be correct in comparison to the more dynamically

driven states. However, it was also seen that models often do not share the same strengths, which is why the weighted multi-model solution can be very useful as it allows the models to contribute more where they are skilful and less where they do not match the observed dynamics well.

3. How does the projected regional climate change when using a metric based on the realism of the synoptic circulation, and does this reduce the uncertainty envelope of future regional projections?

The use of synoptic specific weighting was shown to be a superior tool to using a single model weight, given the lack of variability in most aggregated model weights. However, the weighting function on its own was shown to have a limited effect in correcting the projections. The importance of the weighted projections proved to be in determining the influence of frequency on the overall results.

When combining the weighting with an adjusted baseline the technique produced good results. By adjusting the climate change anomaly only, the technique can make adjustments to the signal, based on the frequency bias of the control. By using the observed baseline the dynamical errors in positioning are accounted for. Differences between the results for weighted and unweighted projections showed that strong continental thermal trough features were associated with a reduction in temperature and slightly greater precipitation over the continent. For states with strong anticyclones, the weighted multi-model solution increased temperatures and lower precipitation that the multi-model average. Additionally, westerly wave features had decreased rain in weighted future results, probably due to more intense troughs and less frontal rainfall.

Comparison of the magnitude of the climate change signal showed that the weighted multi-model composite winter states had a reduced climate change signal compared to the unweighted and in summer states greater warming and increased rainfall, especially over the oceans. This suggests the models are over estimating winter changes in temperature and underestimating the increases in summer.

Synoptic events are the primary driver of climate change impacts. Therefore errors in synoptic state will have a notable influence on the climate change projections and need to be fully considered in any climate change impact study. Initial changes in the global climate system will inevitably come about due to changes in frequency intensity and persistence of synoptic

systems. Therefore a synoptic based approach to multi-model analysis explicitly addresses the first order response of the climate system.

The study met all of its objectives and provides a novel method for reducing uncertainty, both in terms of identifying model biases and in creating meaningful multi-model solutions. The benefits of using the SOM technique are reflected in the ease of comparison and that the resolution used of synoptic scale patterns, which are more skilful than single grid cell variables commonly presented in climate change results from GCMs.

Appendix I

Figures supporting Chapter 3:

Circulation fields, frequency, error and trajectories plots

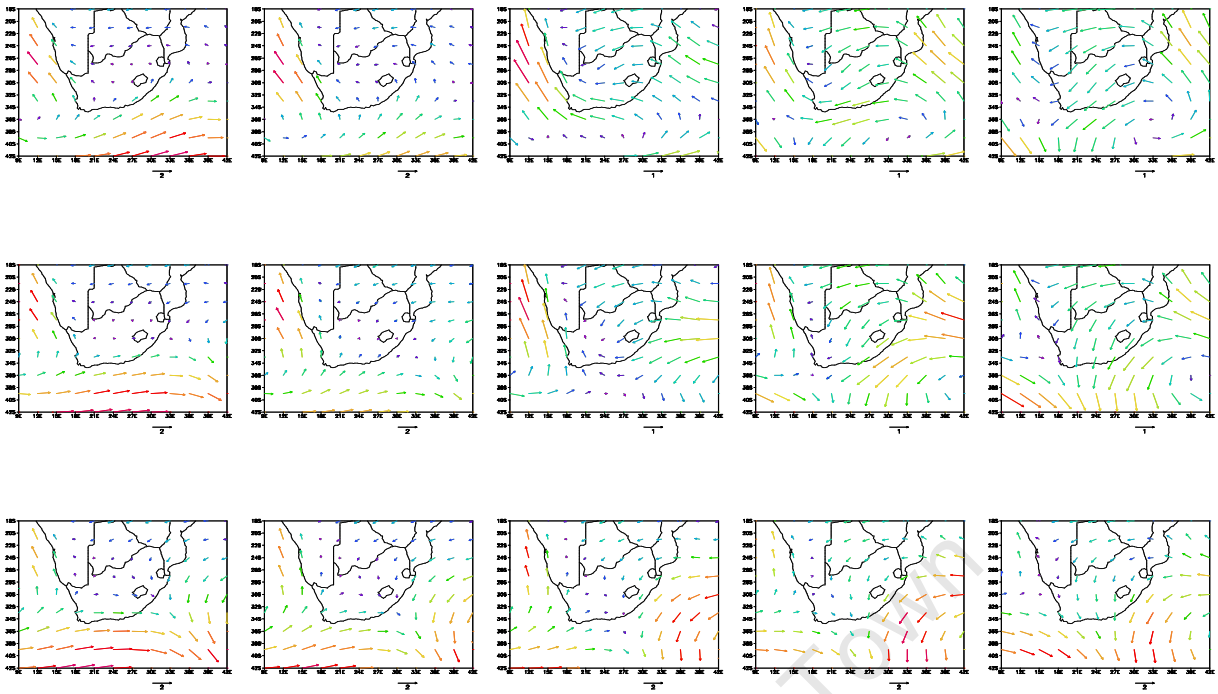


Figure A1: NCEP input data for annual SOM for surface level, winds depicted by vector direction and arrow length, colour represents speed.

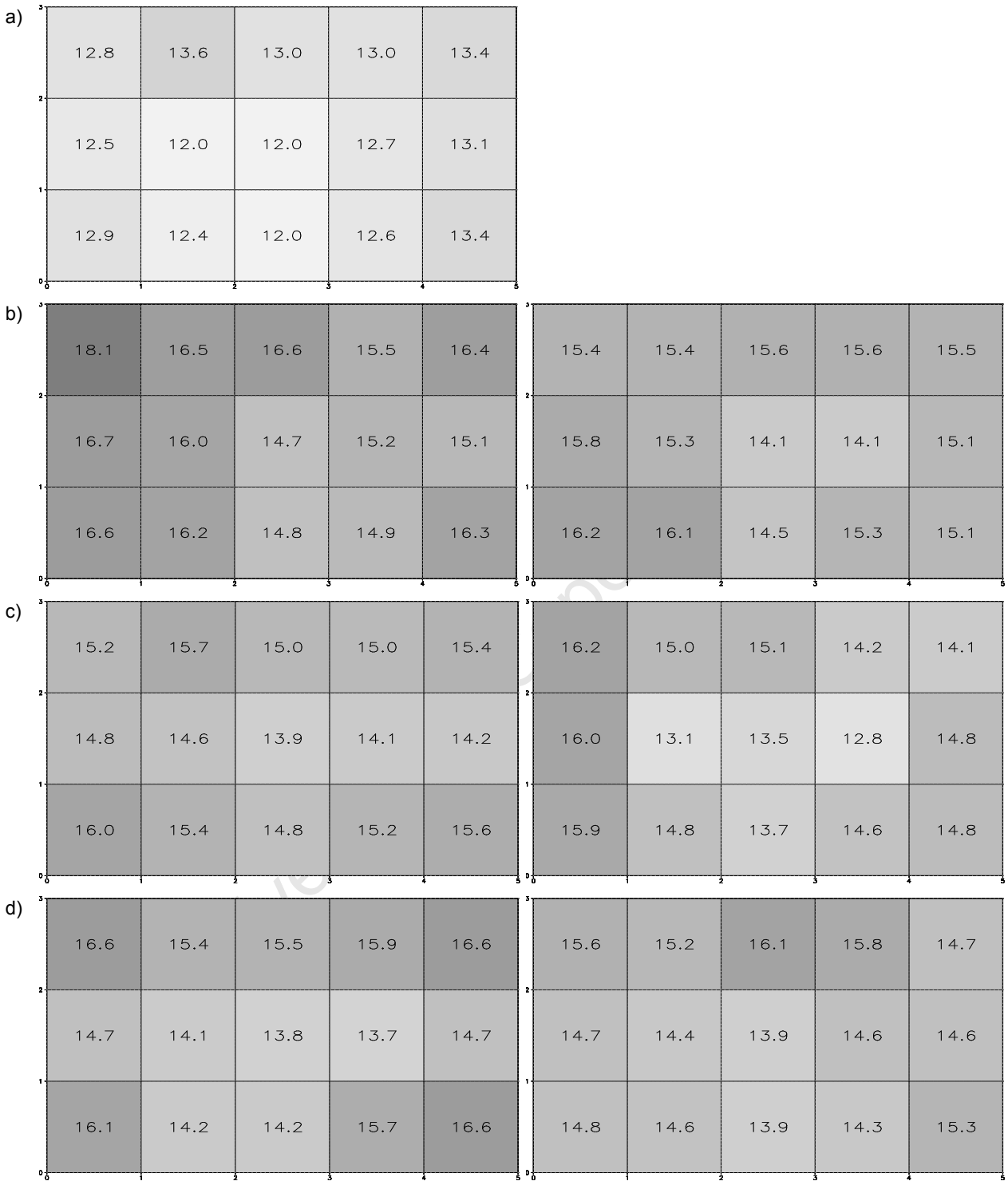


Figure A2: January error map for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

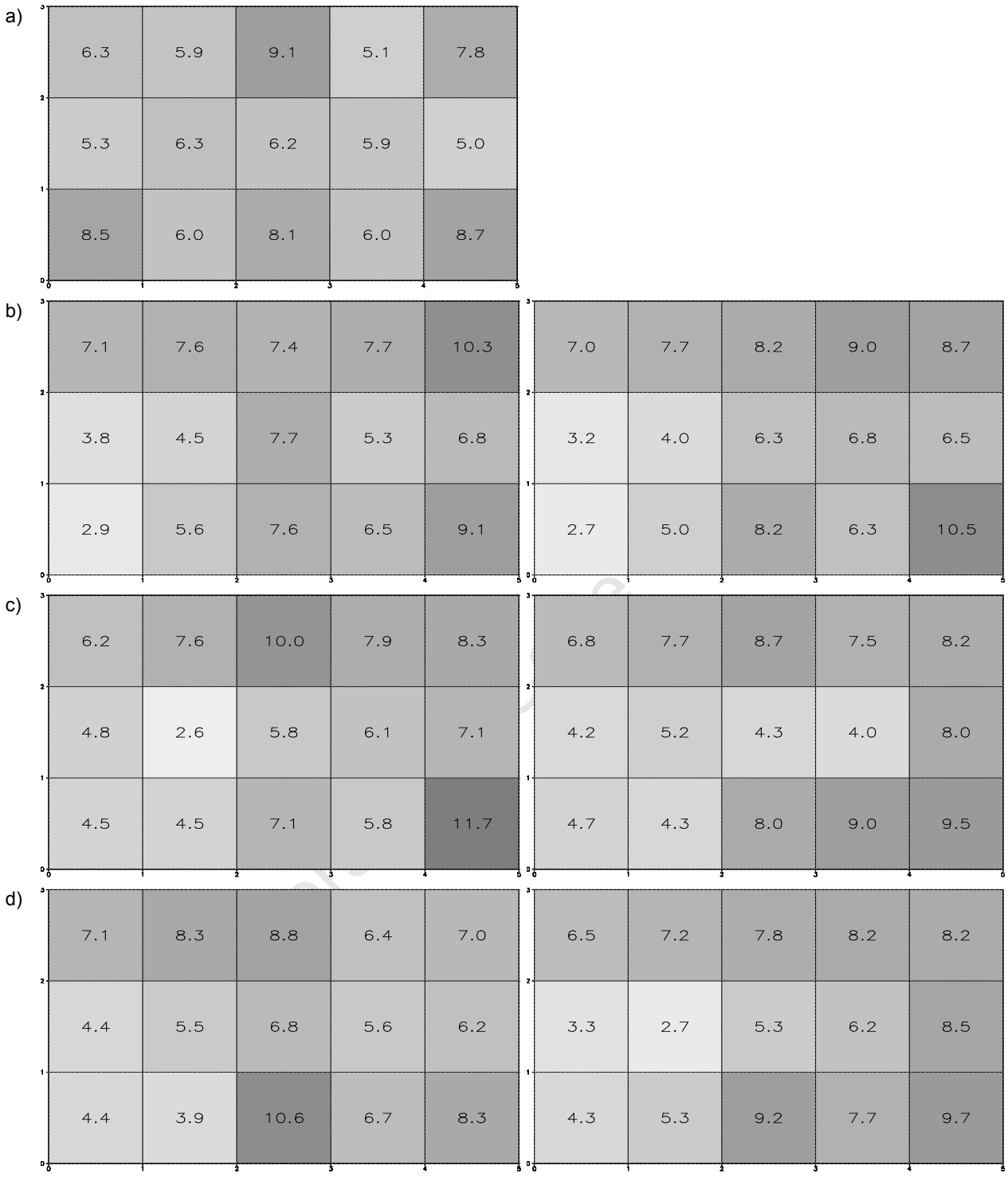


Figure A3: January frequency map for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

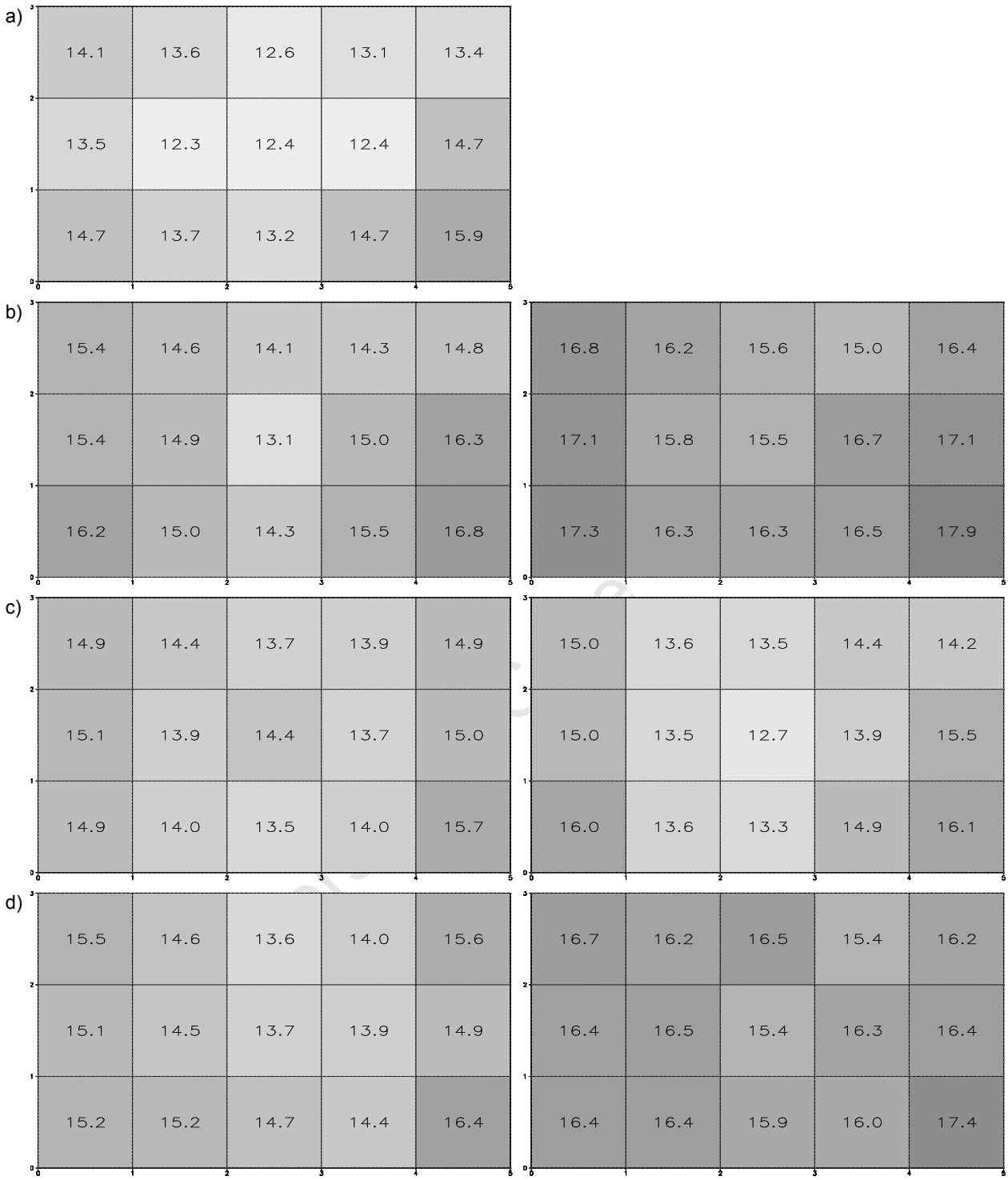


Figure A4: April error maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

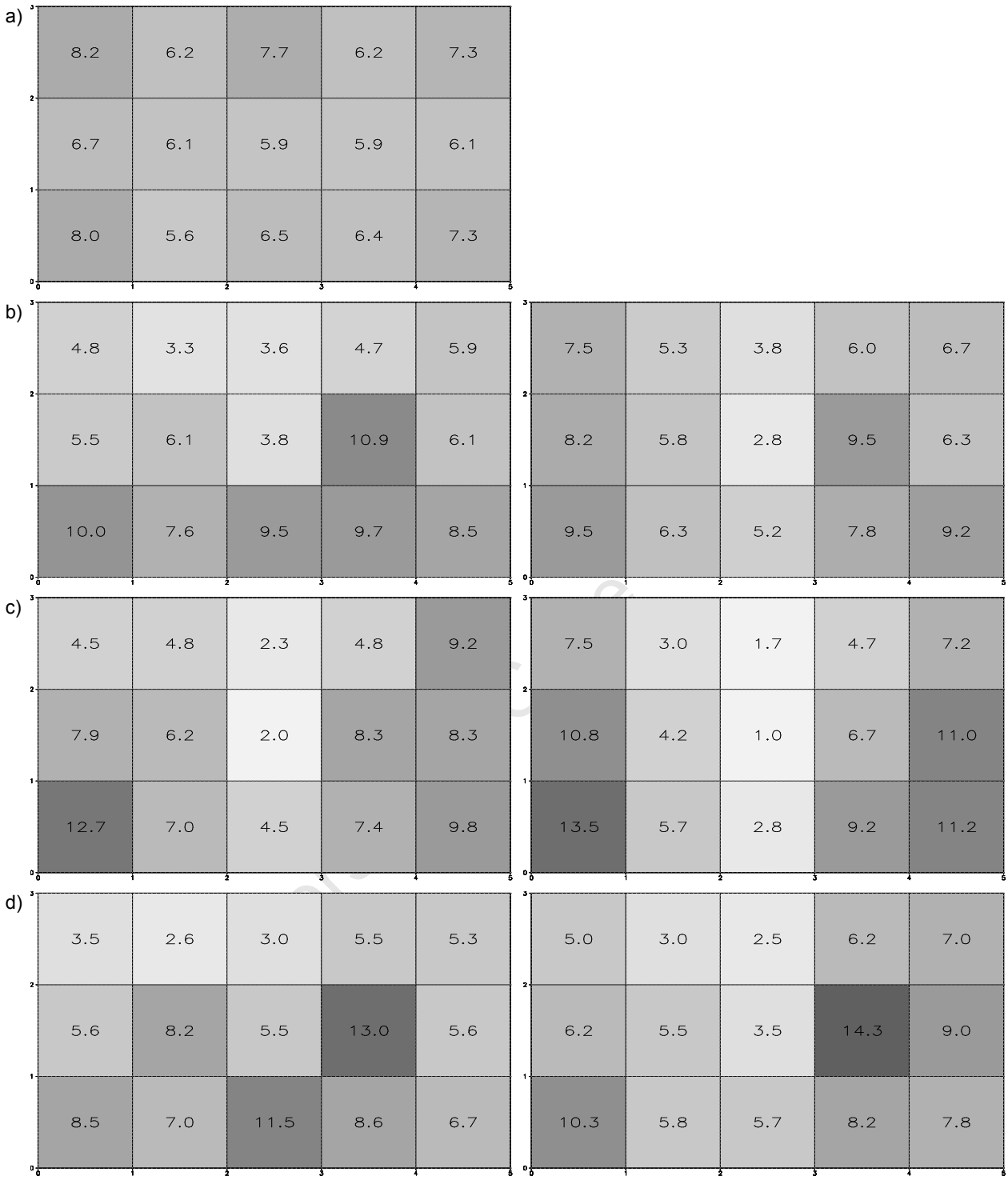


Figure A5: April frequency maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

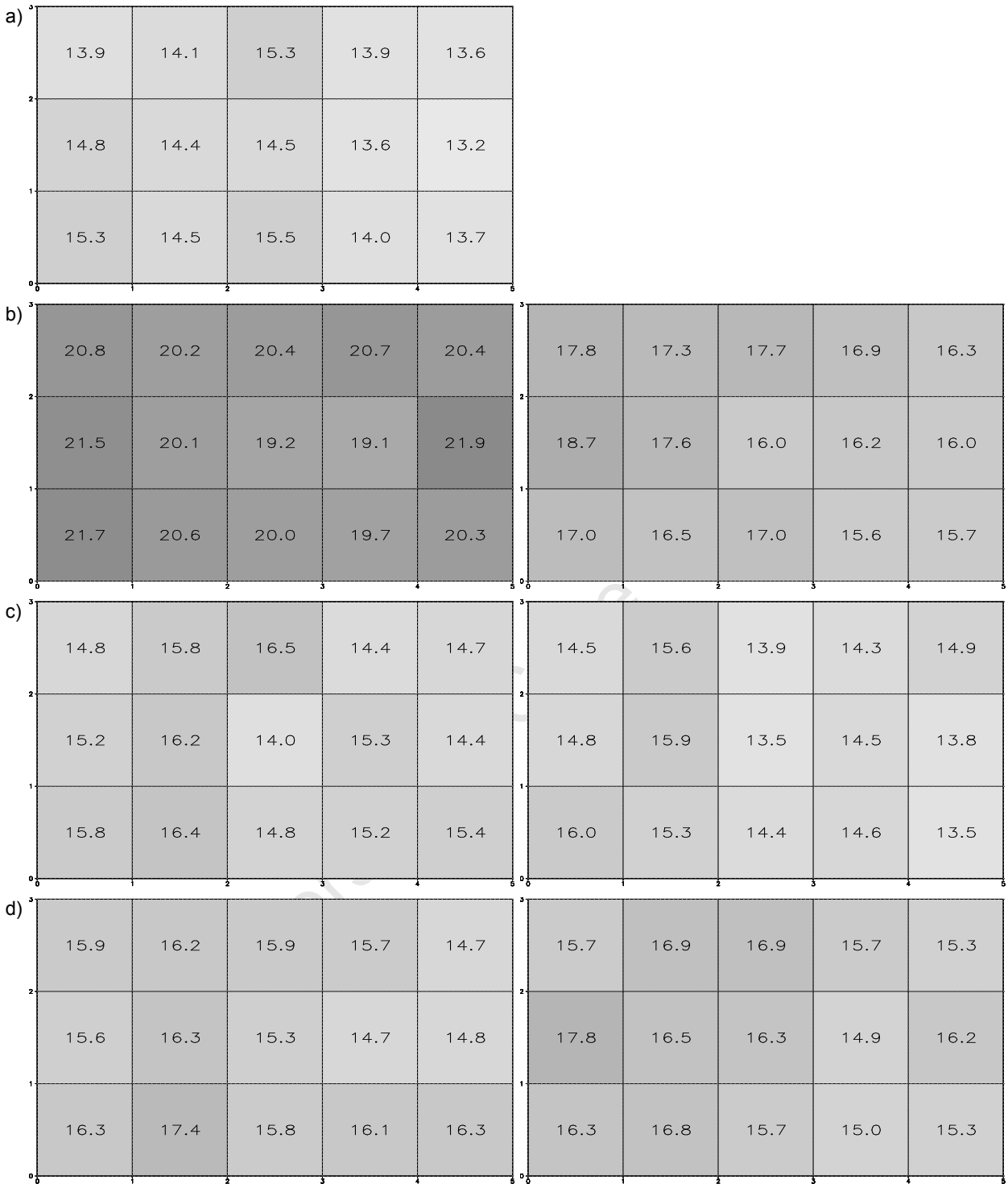


Figure A6: July error maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

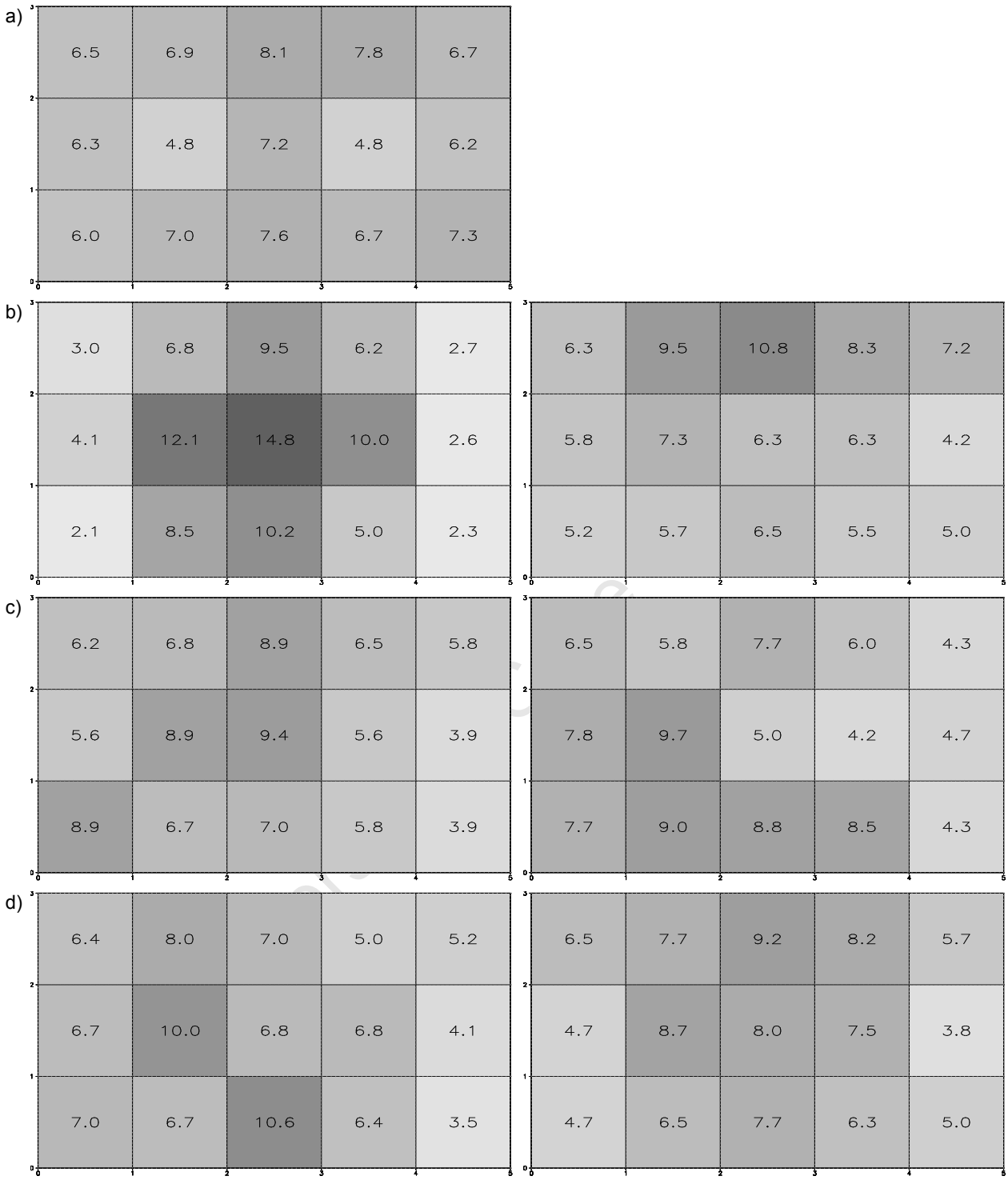


Figure A7: July frequency maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

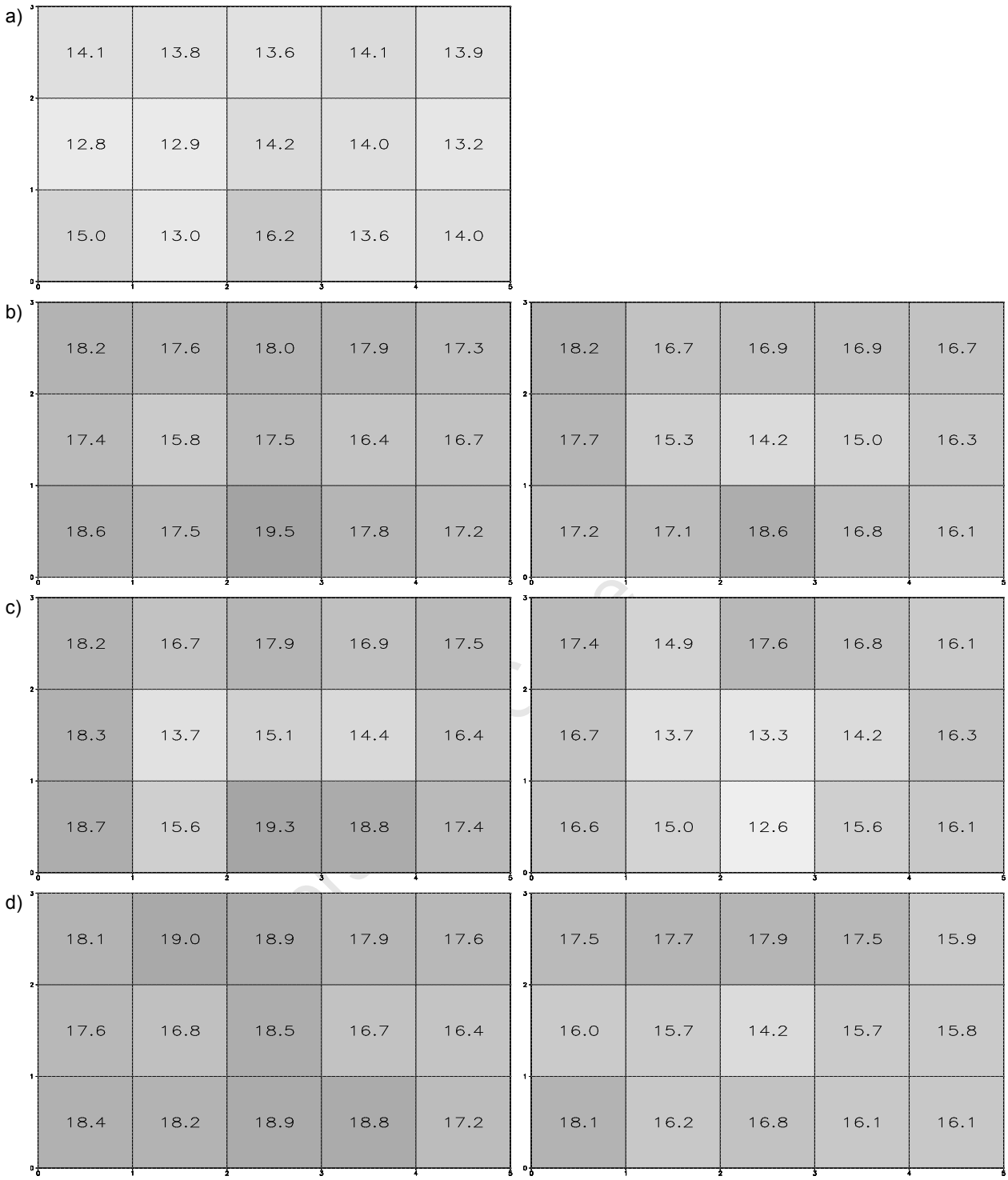


Figure A8: October error maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

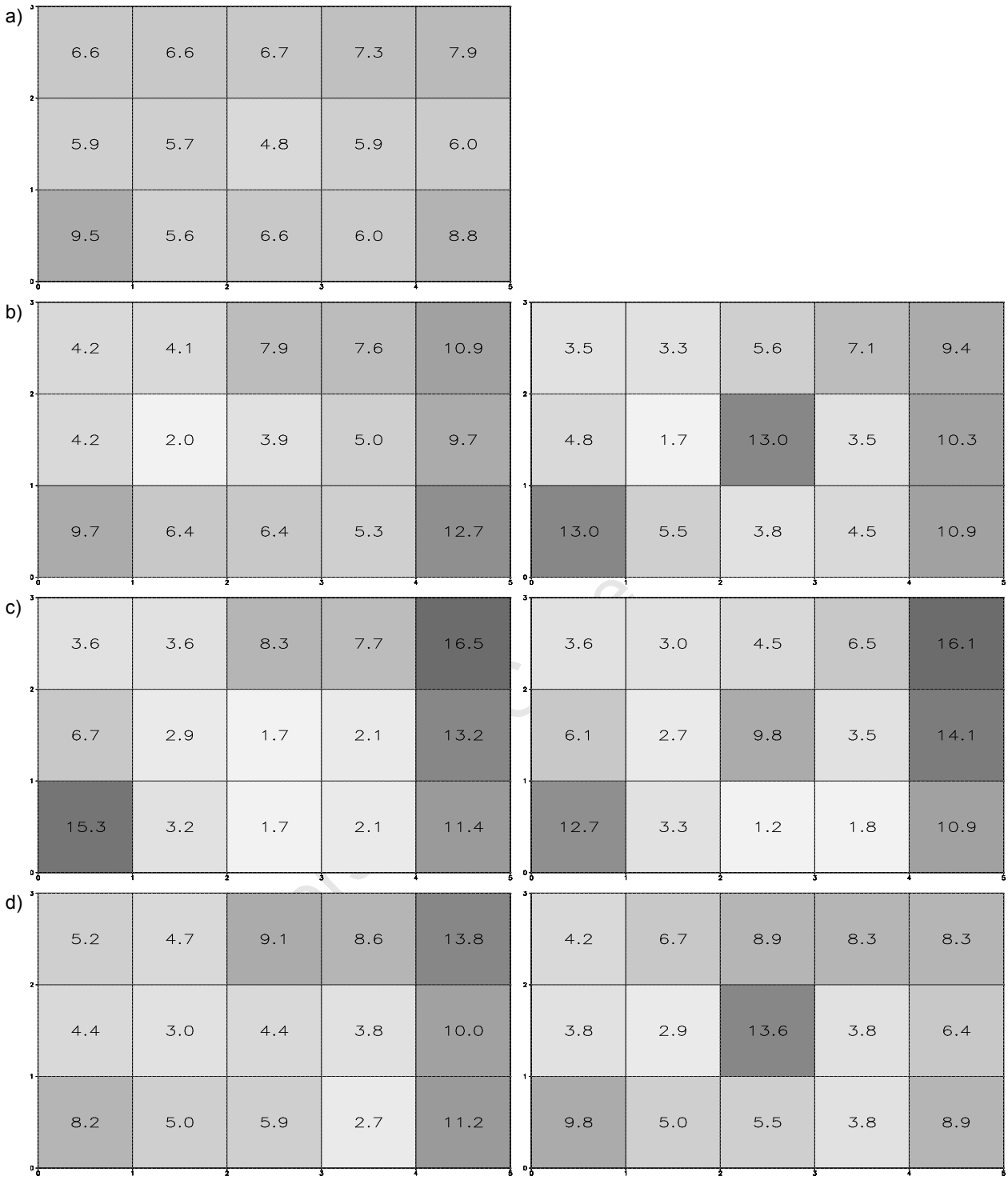
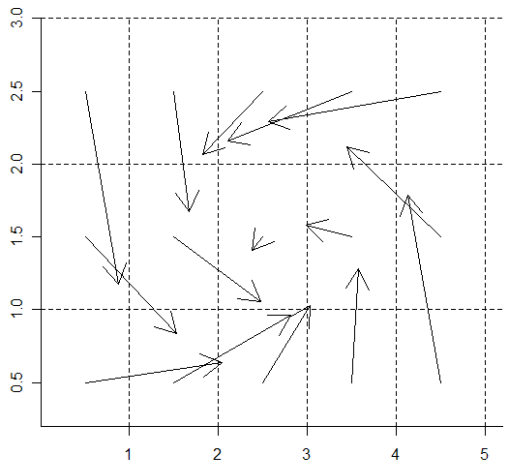
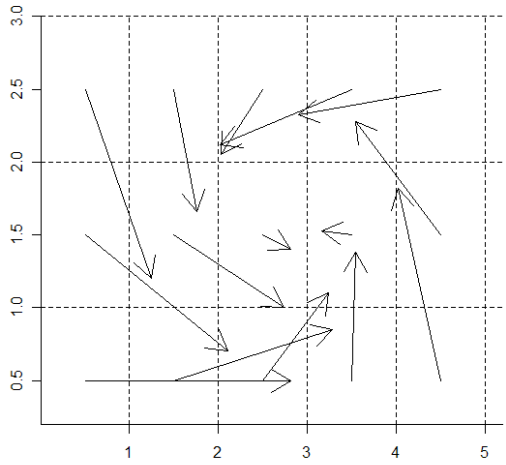


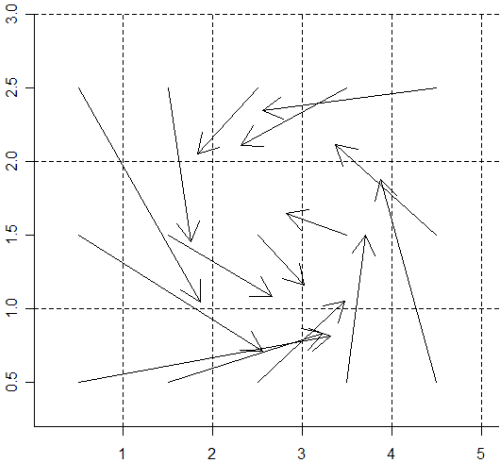
Figure A9: October frequency maps for NCEP observed and GCM control (left) and future (right), expressed as percentages for a) NCEP, b) GFDL, c) MIROC and d) MRI

a)

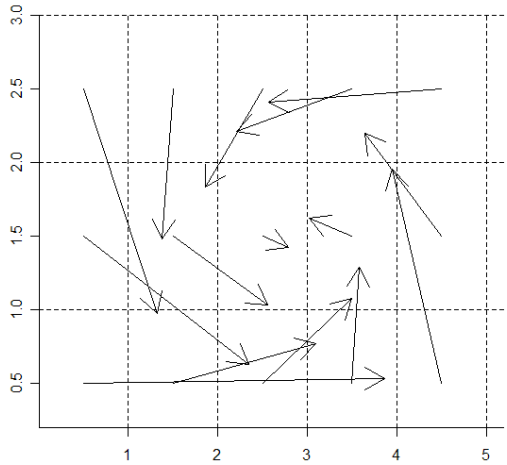

3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

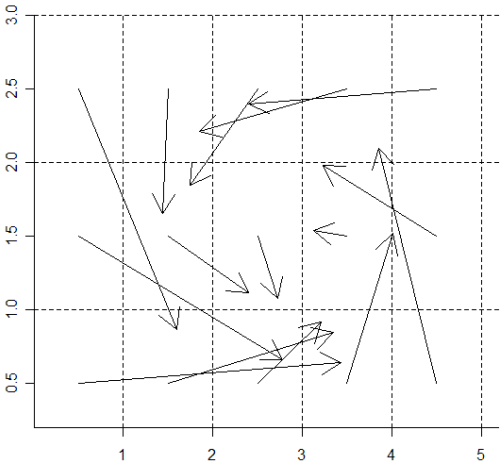
b)


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

c)


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

d)

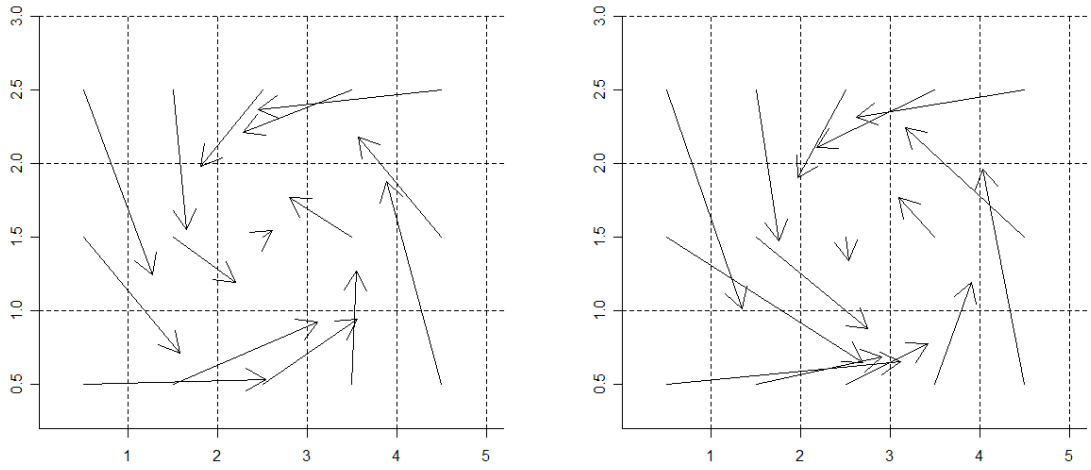
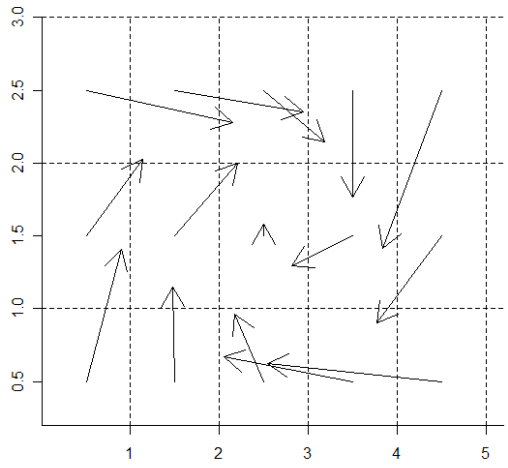
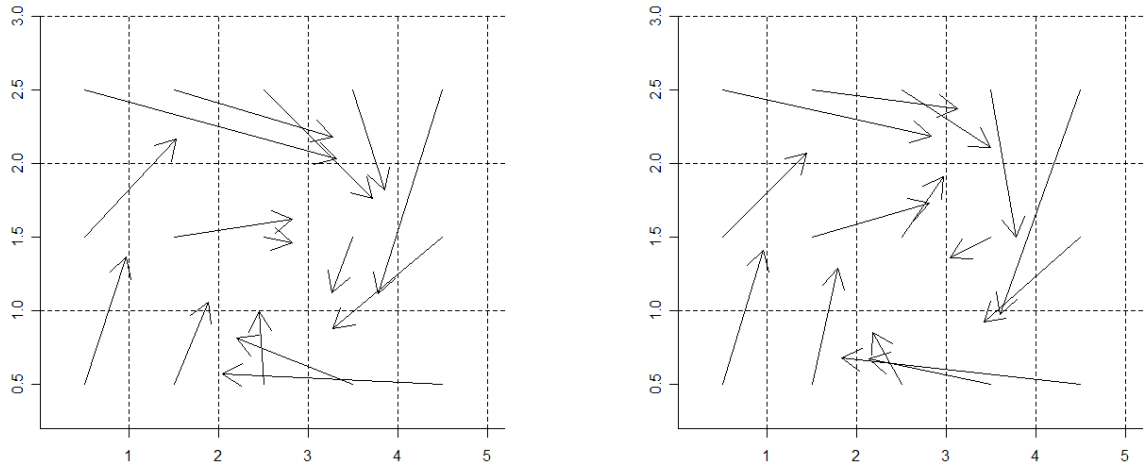


Figure A10: January trajectory maps for NCEP observed and GCM control (left) and future (right), arrows direction and length indicate transition direction and speed for a) NCEP, b) GFDL, c) MIROC and d) MRI

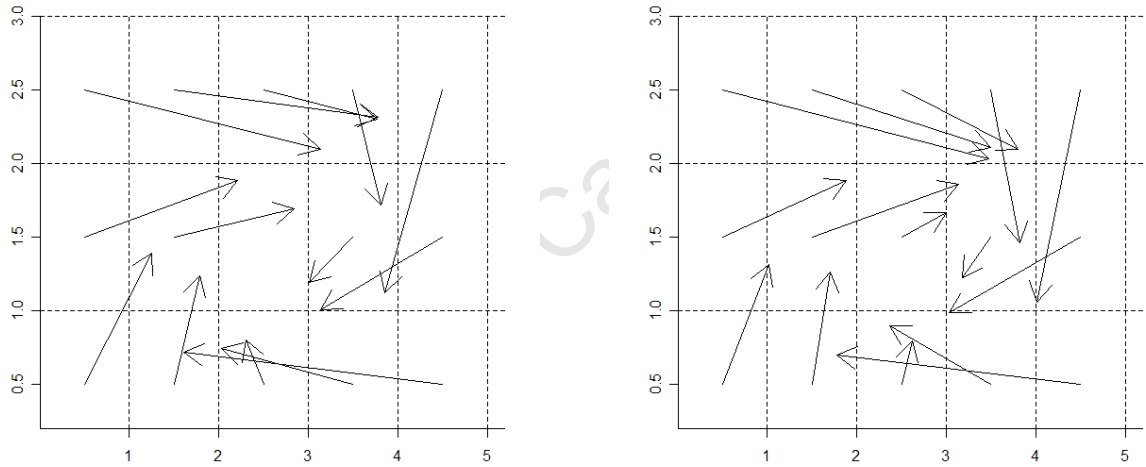
a)



b)



c)



d)

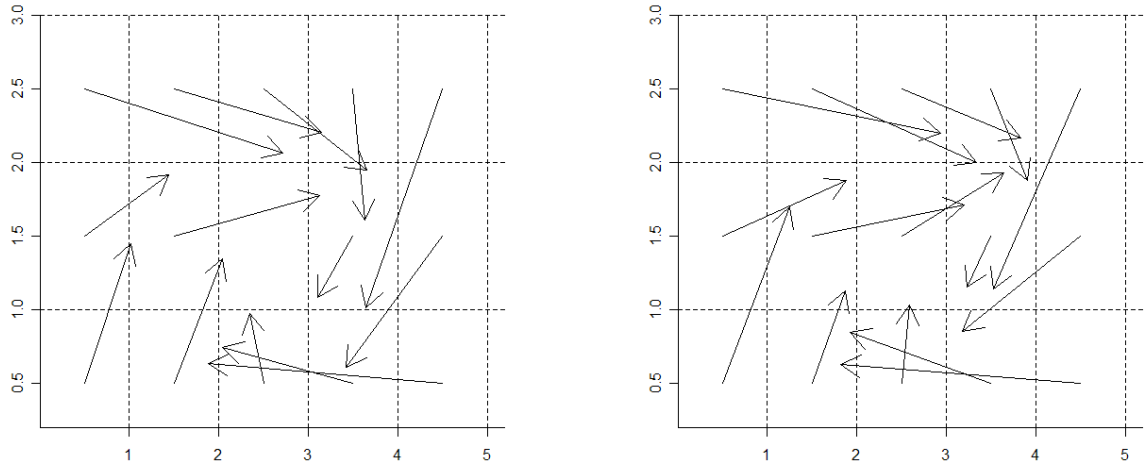
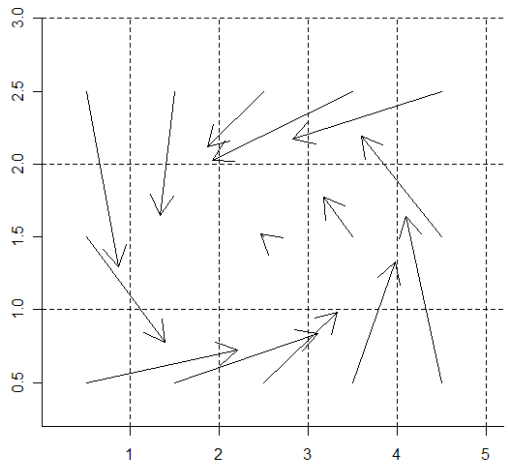
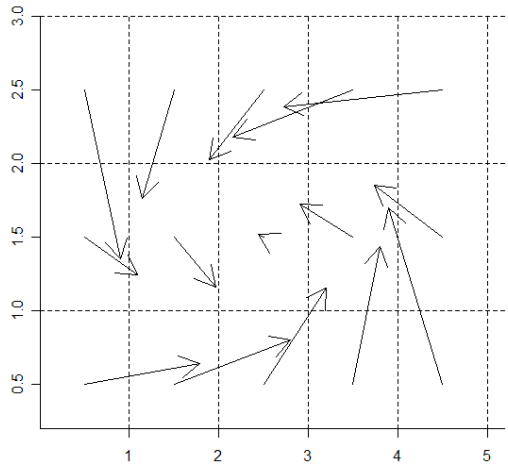


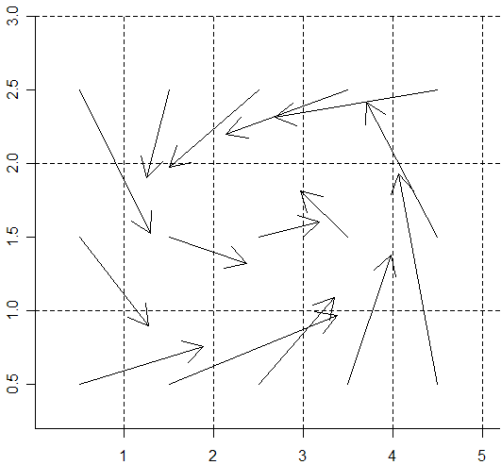
Figure A11: April trajectory maps for NCEP observed and GCM control (left) and future (right), arrows direction and length indicate transition direction and speed for a) NCEP, b) GFDL, c) MIROC and d) MRI

a)

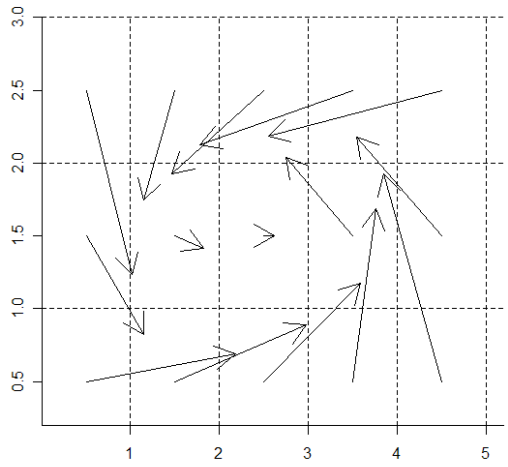

3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

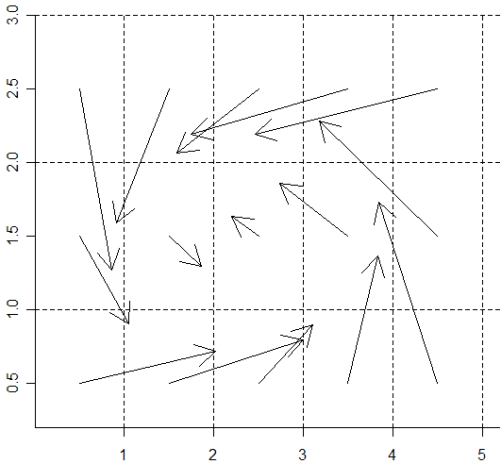
b)


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

c)


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5


3.0
2.5
2.0
1.5
1.0
0.5
1
2
3
4
5

d)

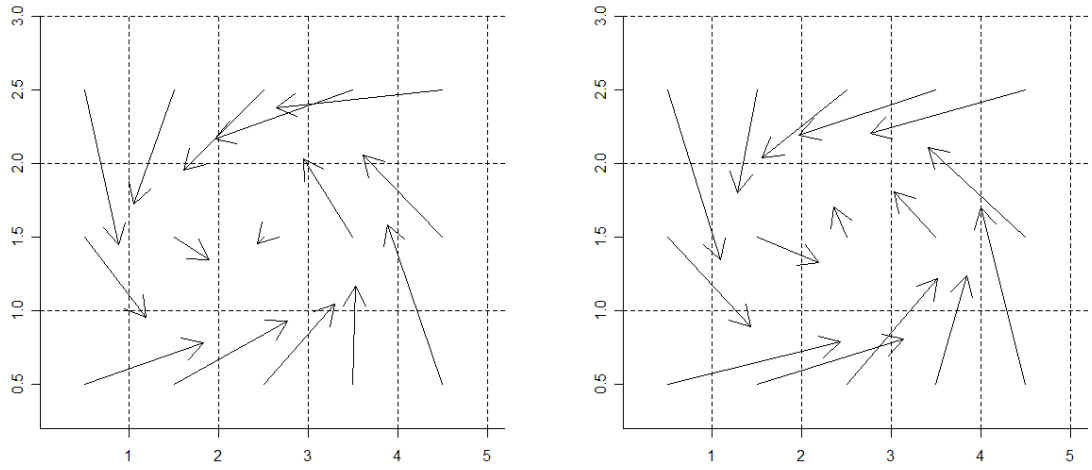
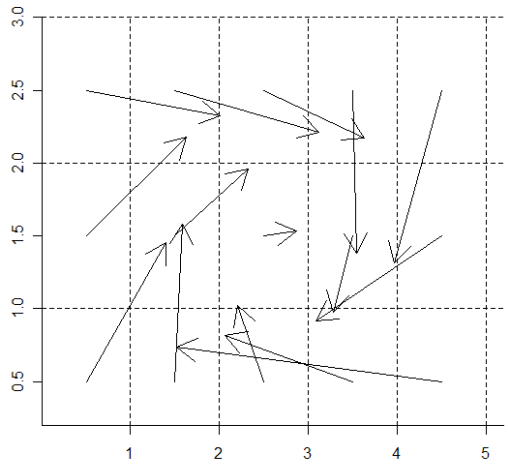
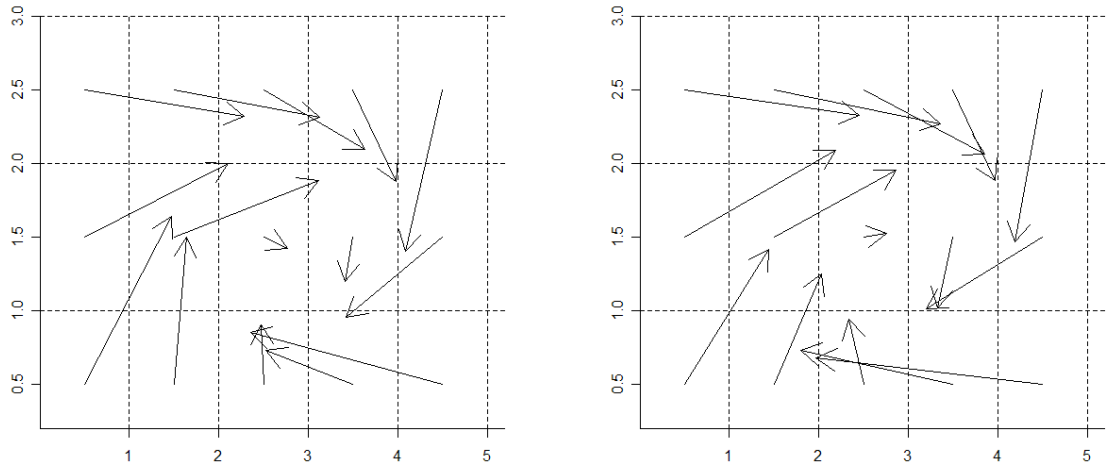


Figure A12: July trajectory maps for NCEP observed and GCM control (left) and future (right), arrows direction and length indicate transition direction and speed for a) NCEP, b) GFDL, c) MIROC and d) MRI

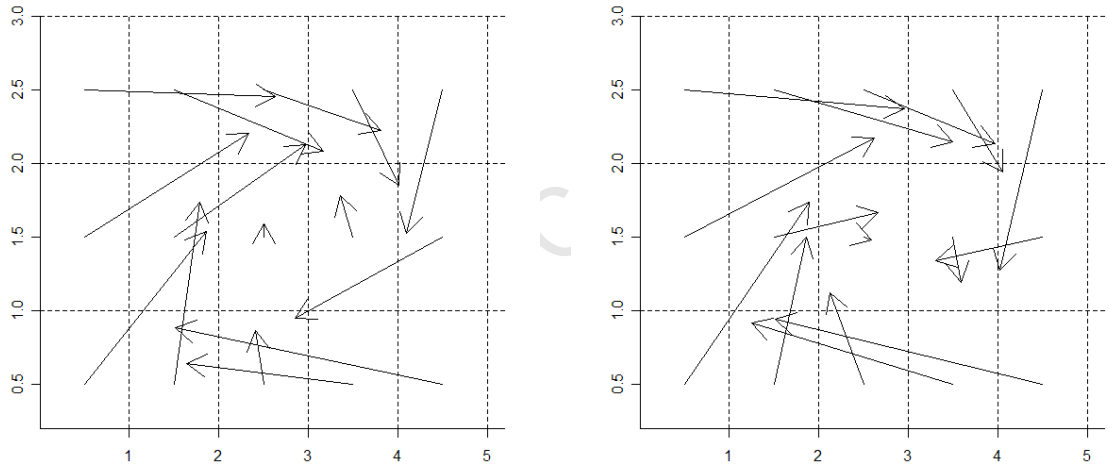
a)



b)



c)



d)

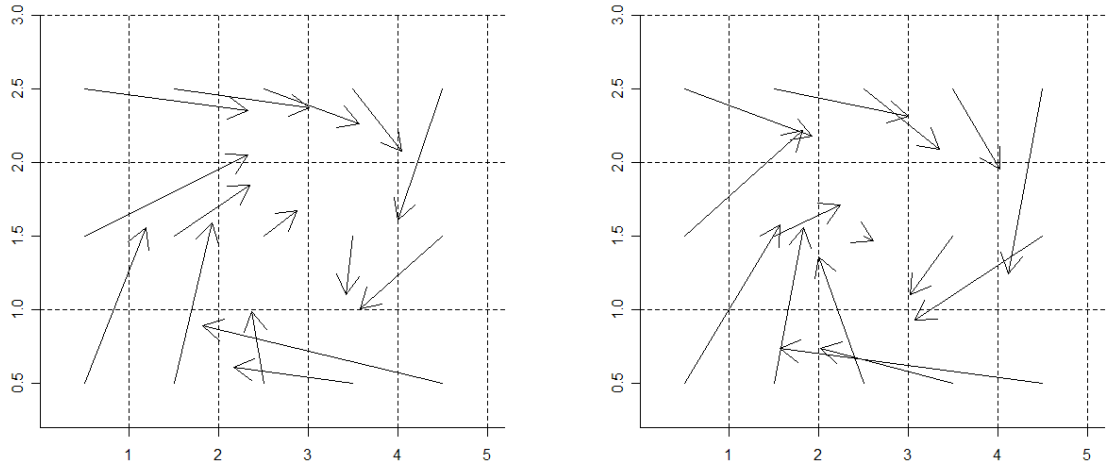


Figure A13: October trajectory maps for NCEP observed and GCM control (left) and future (right), arrows direction and length indicate transition direction and speed for a) NCEP, b) GFDL, c) MIROC and d) MRI

REFERENCES

Allen, M. R. and D. A. Stainforth (2002). "Towards Objective Probabilistic Climate Forecasting." *Nature* **419**: 228.

Berliner, L. M., R. A. Levine, et al. (2000). "Bayesian Climate Change Assessment." *Journal of Climate* **13**: 3805-3820.

Brekke, L. D., M. D. Dettinger, et al. (2008). "Significance of Model Credibility in Eastimating Climate Projection Distributions for Regional Hydroclimatological Risk Assessments." *Climatic Change* **89**: 371-394.

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**: 1506-1523.

Cavazos, T. (2000). "Using Self-Organizing Maps to Investigate Extreme Climate Events: An Application to Wintertime Precipitation in the Balkans." *Journal of Climate* **13**: 1718-1732.

Cavazos, T., A. C. Comrie, et al. (2002). "Intraseasonal Variability Associated with Wet Monsoons in Southeast Arizona." *Journal of Climate* **15**: 2477-2490.

Christensen, J. H., B. C. Hewitson, et al. (2007). Regional Climate Projections. Climate Change 2007: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller. Cambridge and New York, Cambridge University Press**:** 848-940.

Cracknell, A. (2001). *Remote Sensing and Climate Change: The Role of Earth Observation*. Chichester, Praxis Publishing Ltd.

Crane, R. G. (1978). "Seasonal Variations of Sea Ice Extent in the Davis Strait-Labrador Sea Area and Its Relationship with Synoptic-Scale Atmospheric Circulation." *Arctic* **31**: 437-447.

Crane, R. G. and B. C. Hewitson (2003). "Clustering and Upscaling of Station Precipitation Records to Regional Patterns Using Self-Organizing Maps (SOMs)." *Climate Research* **25**: 95-107.

Delworth, T., A. J. Broccoli, et al. (2006). "GFDL's CM2 Global Coupled Climate Models – Part 1: Formulation and Simulation Characteristics." *Journal of Climate* **19**: 643-674.

Denman, K. L., A. Brasseur, et al. (2007). Couplings between Changes in the Climate System and Biogeochemistry. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning,

Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller. Cambridge and New York, Cambridge University Press**:** 499-587.

Giorgi, F., B. C. Hewitson, et al. (2001). Regional Climate Information - Evaluation and Projections. Climate Change 2001: The Scientific Basis. J. T. Houghton, Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, X. Dai, K. Maskell, and C.A. Johnson (eds.). Cambridge, UK, Cambridge University Press**:** 881.

Giorgi, P. and L. Mearns (2002). "Calculation of Average, Uncertainty Range, and Reliability of Regional Climate Changes from AOGCM Simulations Via the "Reliability Ensemble Averaging" (Rea) Method." *Journal of Climate* **15**: 1411-1158.

Glecker, P. J., K. E. Taylor, et al. (2008). "Performance Metrics for Climate Models." *Journal of Geophysical Research* **113**.

Greene, A. M., L. Goddard, et al. (2005). "Probabilistic Multimodel Regional Temperature Change Projections." *Journal of Climate* **19**(17): 4326-4343.

Gnanadesikan, A. et al (2004). "GFDL's CM2 Global Coupled Climate Models–Part 2: The Baseline Ocean Simulation." *Journal of Climate* **19**: 675-697.

Gutiérrez, J. M., R. Cano, et al. (2005). "Analysis and Downscaling Multi-Model Seasonal Forecasters in Peru Using Self-Organizing Maps."

Hayhoe, H., P.Frumhoff, et al. (2006). Regional Assessment of Climate Impacts on California under Alternative Emission Pathways - Key Findings and Implications for Stabilisation. Avoiding Dangerous Climate Change. Cambridge, UK, Cambridge University Press.

Hewitson, B. C. and R. G. Crane (2002). "Self-Organizing Maps: Applications to Synoptic Climatology." *Climate Research* **22**: 13-26.

Hewitson, B. C. and R. G. Crane (2005). Gridded Area-Averaged Daily Precipitation Via Conditional Interpolation. AIACC Working Paper.

Hewitson, B. C. and R. G. Crane (2006). "Consensus between GCM Climate Change Projections with Empirical Downscalling: Precipitation Downscaling over South Africa." *International Journal of Climatology* **26**(10): 1315-1337.

Hope, P. K. (2006). "Projected Future Changes in Synoptic Systems Influencing Southwest Western Australia." *Climate Dynamics* **26**: 765-780.

Hope, P. K., W. Drosdowsky, et al. (2006). "Shifts in the Synoptic Systems Influencing Western Australia." *Climate Dynamics* **26**: 751-764.

IPCC (2001). *Climate Change 2001: The Scientific Basis, Contribution of Working Group 1 to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press.

IPCC (2007). *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, Ii and Iii to the Fourth Assessment*. Geneva, Switzerland, IPCC.

K-1 model Developers (2004). K-1 Coupled Model (MIROC) Description. K-1 Technical Report 1. [H. Hasumi, and S. Emori. (eds)] Center for Climate System Research, University of Tokyo, Tokyo, Japan 34pp.

Kalnay, E., M. Kanamitsu, et al. (1996). "The NCEP/NCAR 40-Year Reanalysis Project." *Bulletin of the American Meteorological Society* **77**(3): 437-471.

Kohonen, T. (1995). *Self-Organizing Maps*, Springer-Verlag.

Kohonen, T., J. Hynninen, et al. (1996). SOMpak: The Self-Organizing Map Package, V. 3.2. Helsinki Finland, Laboratory of Computer and Information Science, University of Helsinki.

Krishnamurti, T. N., C. M. Kishtawal, et al. (2000). "Multimodel Ensemble Forecasts or Weather and Seasonal Climate." *Journal of Climate* **13**: 4196-4216.

Main, J. (1997). Seasonality of Circulation in Southern Africa Using Kohonen Self-Organising Map. MSc Thesis, University of Cape Town, Cape Town.

Malgrem, B. A. and A. Winter (1999). "Climate Zonation in Puerto Rico Based on Principal Components Analysis and Artificial Neural Network." *Journal of Climate* **12**: 977-985.

McGuffie, K. and A. Henderson-Sellers (1997). *A Climate Modelling Primer*. England, John Wiley and Sons Ltd.

Meehl, G. A., T. F. Stocker, et al. (2007). Global Climate Projections. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller. Cambridge and New York, Cambridge University Press**:** 747-845.

Midgley, G., R. Chapman, et al. (2005). Western Cape Report: A Status Quo, Vulnerability and Adaptation Assessment of the Physical and Socio-Economic Effects of Climate. Cape Town.

Nakic´enovic´, N., J. Alcamo, et al. (2000). *IPCC Special Report on Emissions Scenarios*. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press.

New, M. and M. Hulme (2000). "Representing Uncertainty in Climate Change Scenarios: A Monte-Carlo Approach." *Integrated Assessment* **1**: 203-213.

Peixoto, J. P. and A. H. Oort (1992). *Physics of Climate*. New York, USA, American Institute of Physics.

Randall, D. A., R. A. Wood, et al. (2007). Climate Models and Their Evaluation. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to

the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller. Cambridge, United Kingdom and New York, NY, USA, Cambridge University Press**:** 589-662.

Stainforth, D. A., T. Aina, et al. (2005). "Uncertainty in Predictions of the Climate Response to Rising Levels of Greenhouse Gases." *Nature* **433**: 403-406.

Stainforth, D. A., M. R. Allen, et al. (2007). "Confidence, Uncertainty and Decision-Support Relevance in Climate Predictions." *Philosophical Transactions of The Royal Society A* **365**: 2145-2161.

Stern, N. (2006). Stern Review: The Economics of Climate Change, Cambridge University Press.

Stott, P. A., J. A. Kettleborough, et al. (2006). "Uncertainty in Continental-Scale Temperature Predictions." *Geophysical Research Letters* **33**(L02708).

Tadross, M. A., W. J. Gutowski, et al. (2006). "Mm5 Simulations of Interannual Change and Diurnal Cycle of Southern African Regional Climate." *Theoretical and Applied Climatology* **86**: 63-80.

Tebaldi, C., R. L. Smith, et al. (2005). "Quantifying Uncertainty in Projections of Regional Climate Change: A Bayesian Approach to the Analysis of Multimodel Ensembles." *Journal of Climate* **18**: 1524-1540.

Tennant, W. J. and B. C. Hewitson (2002). "Intra-Seasonal Rainfall Characteristics and Their Importance to the Seasonal Prediction Problem." *International Journal of Climatology* **22**: 1033-1048.

Tennant, W. J. and C. J. Reason (2005). "Associations between the Global Energy Cycle and Regional Rainfall in South Africa and Southwest Australia." *Journal of Climate* **18**: 3032-3047.

Thomas, D. S. G., C. Twyman, et al. (2007). "Adaptation to Climate Change and Variability: Farmer Responses to Intra-Seasonal Prediction Trends in South Africa." *Climatic Change* **83**: 301-322.

Trenberth, K. E., P. D. Jones, et al. (2007). Observations: Surface and Atmospheric Climate Change. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Forth Assessment Report of the Intergovernmental Panel on Climate Change. S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M.Tignor and H.L. Miller. Cambridge and New York, Cambridge University Press**:** 235-335.

Tyson, P. D. and R. A. Preston-Whyte (2000). *The Weather and Climate of Southern Africa*. Cape Town, Oxford University Press Southern Africa.

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." *International Journal of Climatology* **7**: 231-246.

Yarnal, B. (1984). "Relationships between Synoptic-Scale Atmospheric Circulation and Glacier Mass Balance in Southwestern Canada During the International Hydrological Decade." *Journal of Glaciology* **30**: 188-198.

Yukimoto, S. et al (2001). "The New Meteorological Research Institute Global Ocean-Atmosphere Coupled GCM (MRI-CGCM2)--Model Climate and Variability." *Papers in Meteorology and Geophysics* **51**: 47-88.

Appendix II

Figures supporting Chapter 4:

NCEP and GCM bias comparisons

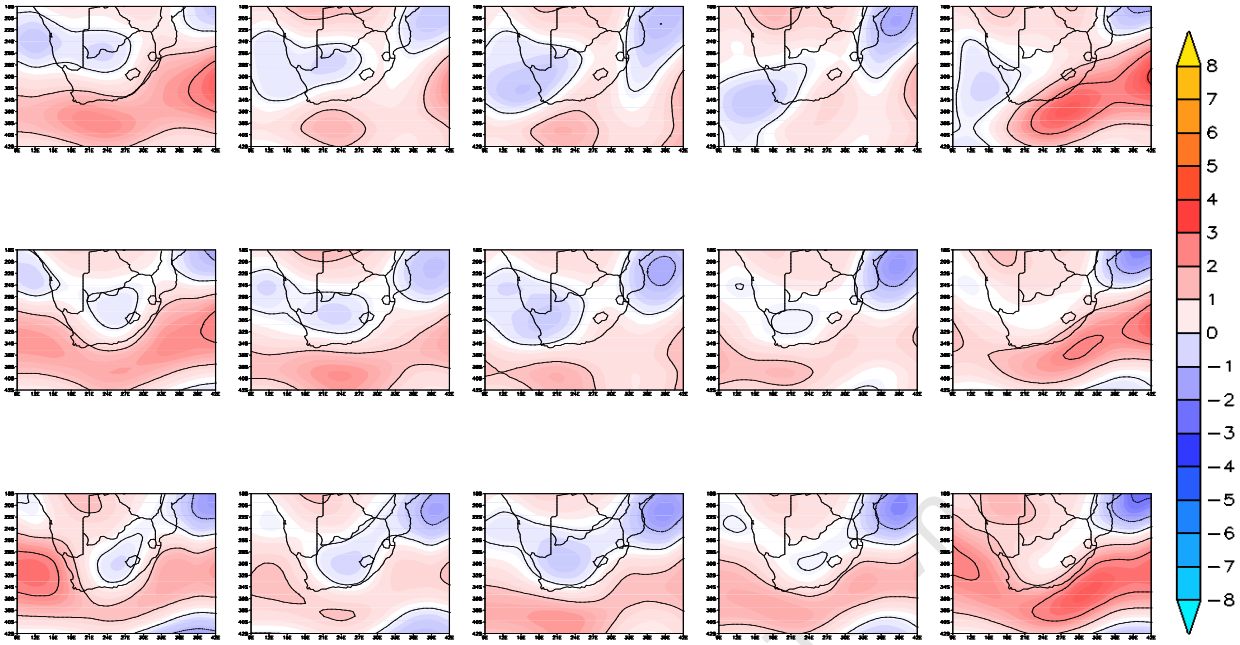


Figure B1: Anomaly of GFDL-NCEP for each SOM node for annual average u winds at the surface

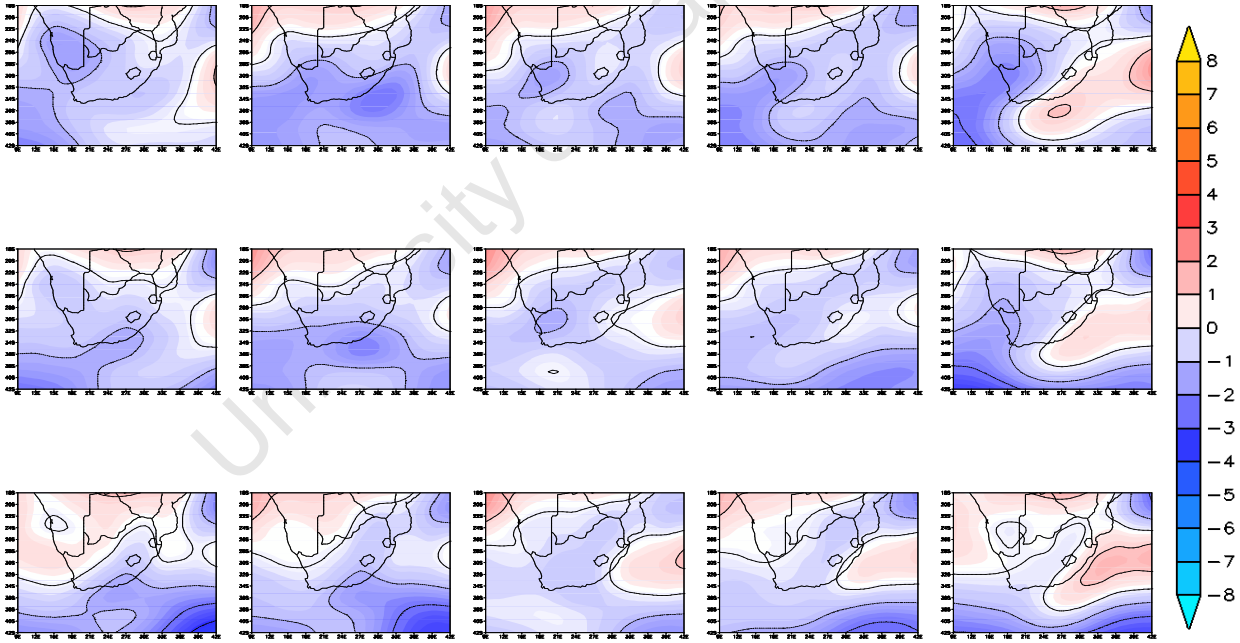


Figure B2: Anomaly of MRI -NCEP for each SOM node for annual average u winds at the surface

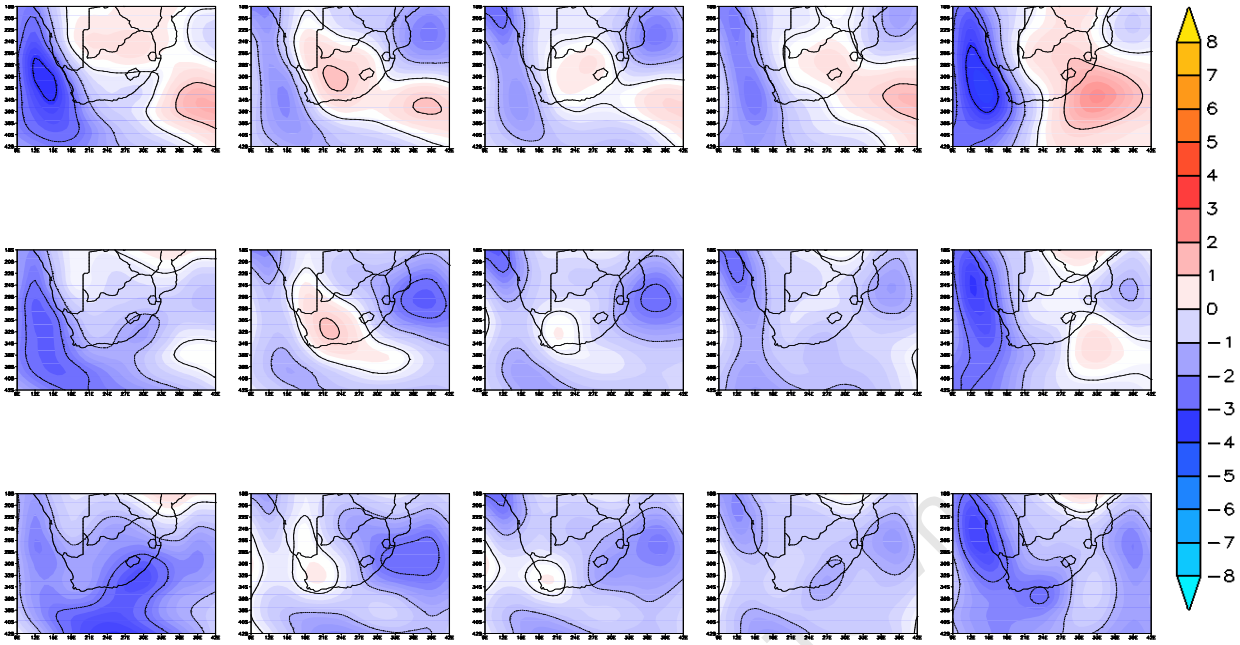


Figure B3: Anomaly of MIROC -NCEP for each SOM node for annual average v winds at the surface

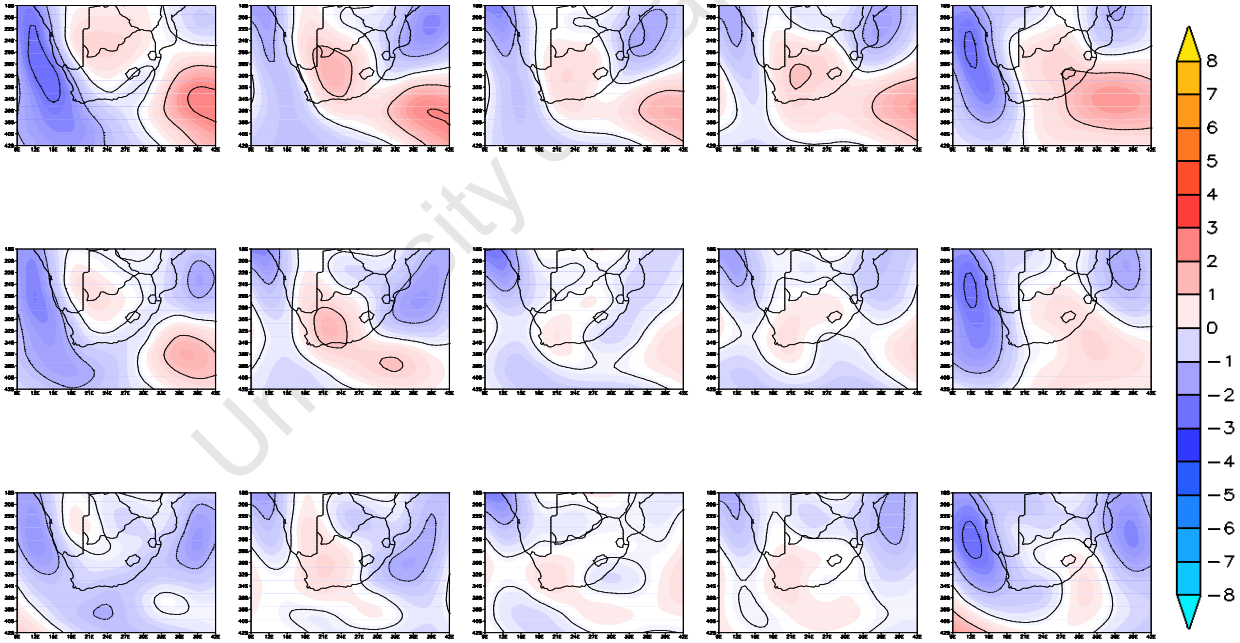


Figure B4: Anomaly of MRI-NCEP for each SOM node for annual average v winds at the surface

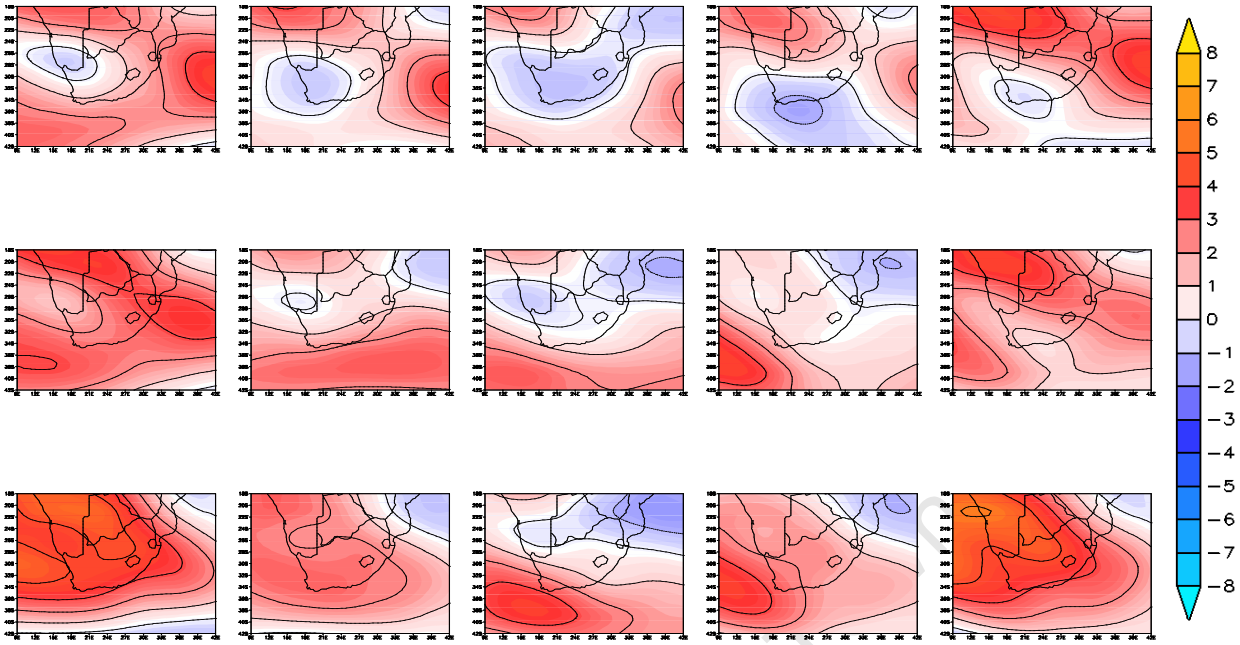


Figure B5: Anomaly of GFDL-NCEP for each SOM node for annual average u winds at 700 hPa

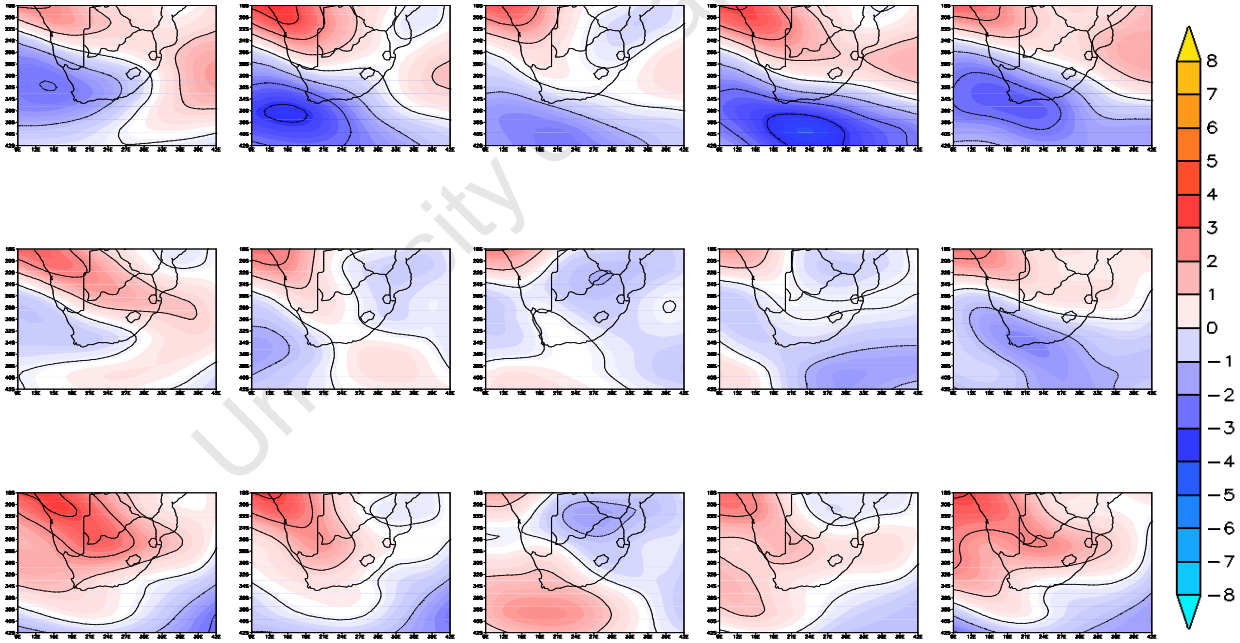


Figure B6: Anomaly of MRI-NCEP for each SOM node for annual average u winds at 700 hPa

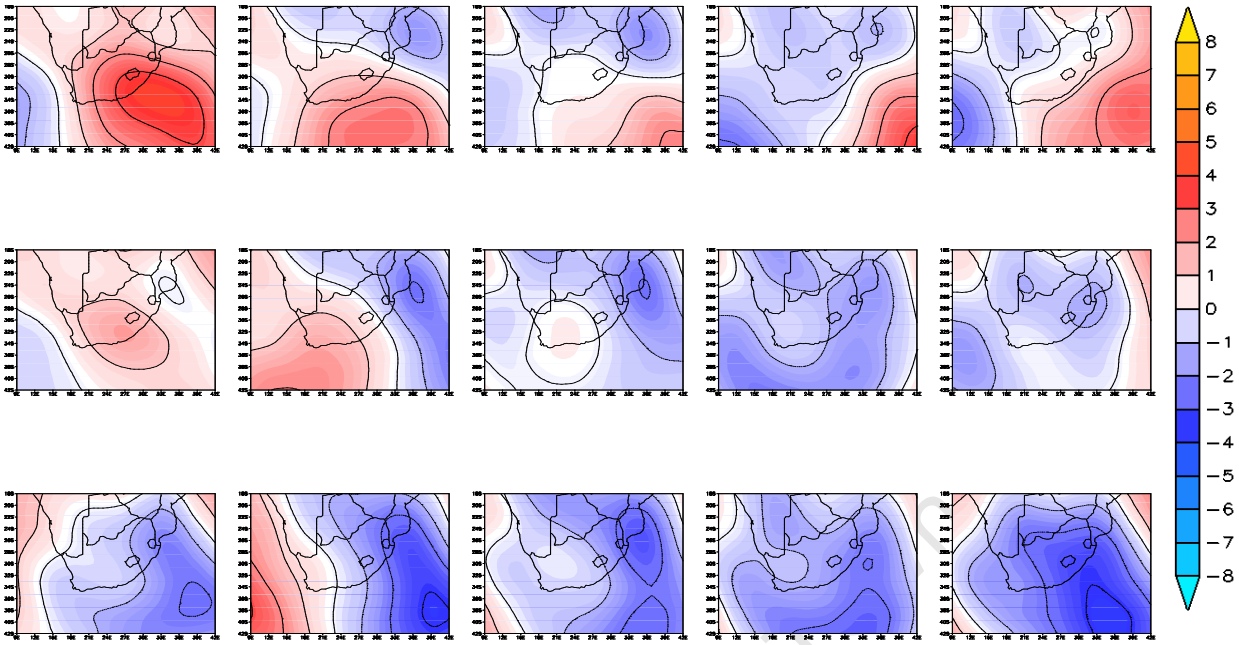


Figure B7: Anomaly of MIROC -NCEP for each SOM node for annual average v winds at 700 hPa

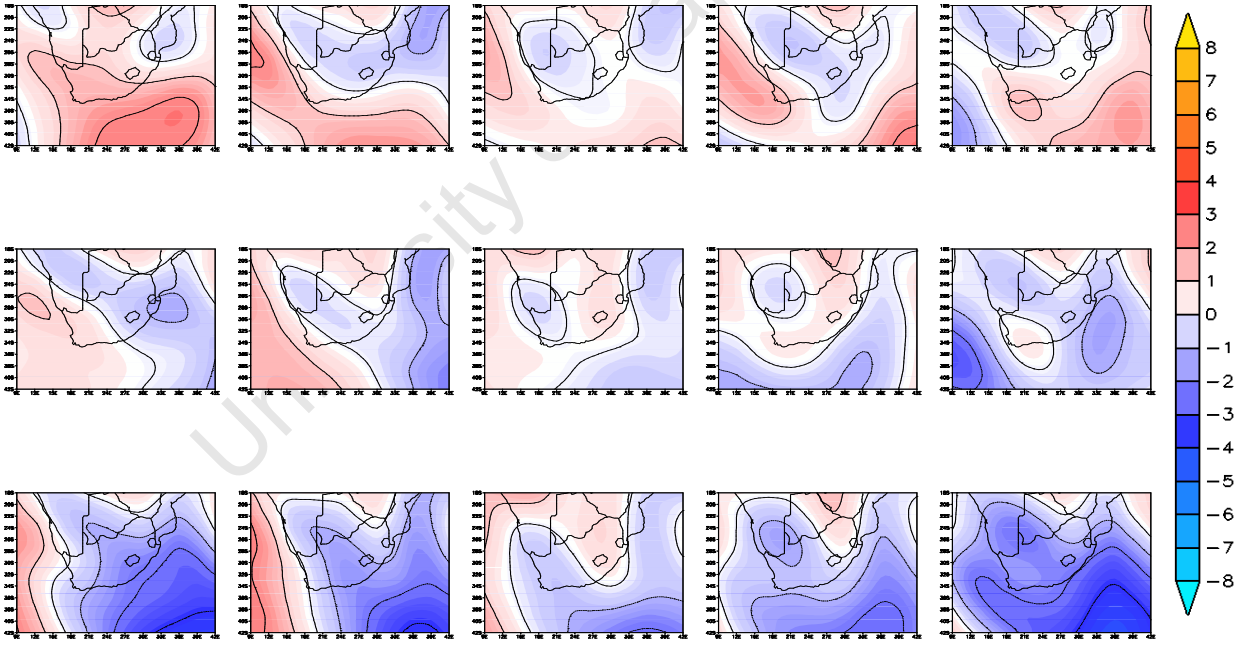


Figure B8: Anomaly of MRI-NCEP for each SOM node for annual average v wind at 700 hPa

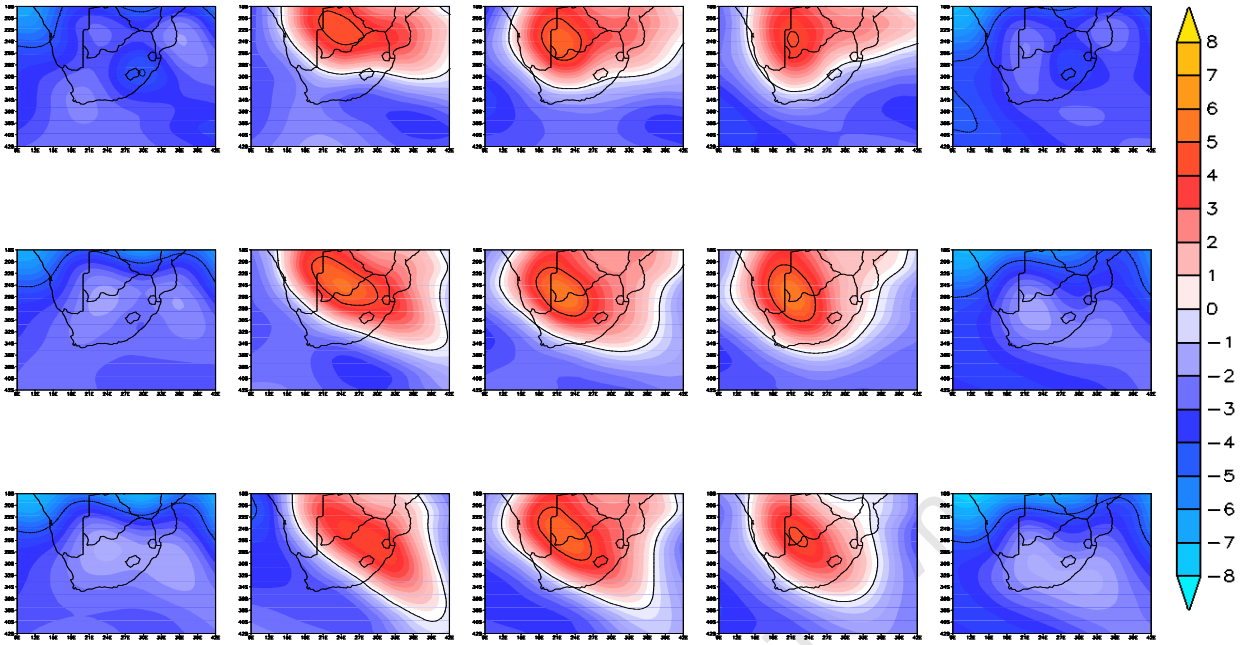


Figure B9: Anomaly of MIROC-NCEP for each SOM node for annual average specific humidity at 700 hPa

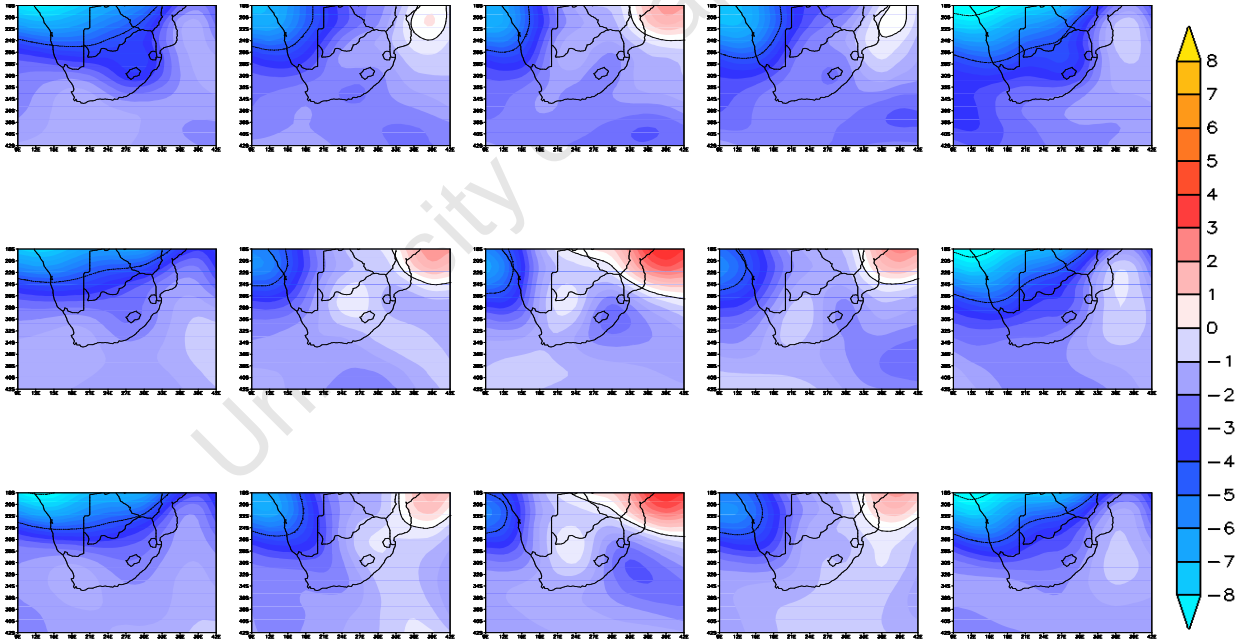


Figure B10: Anomaly of MRI-NCEP for each SOM node for annual average specific humidity at 700 hPa

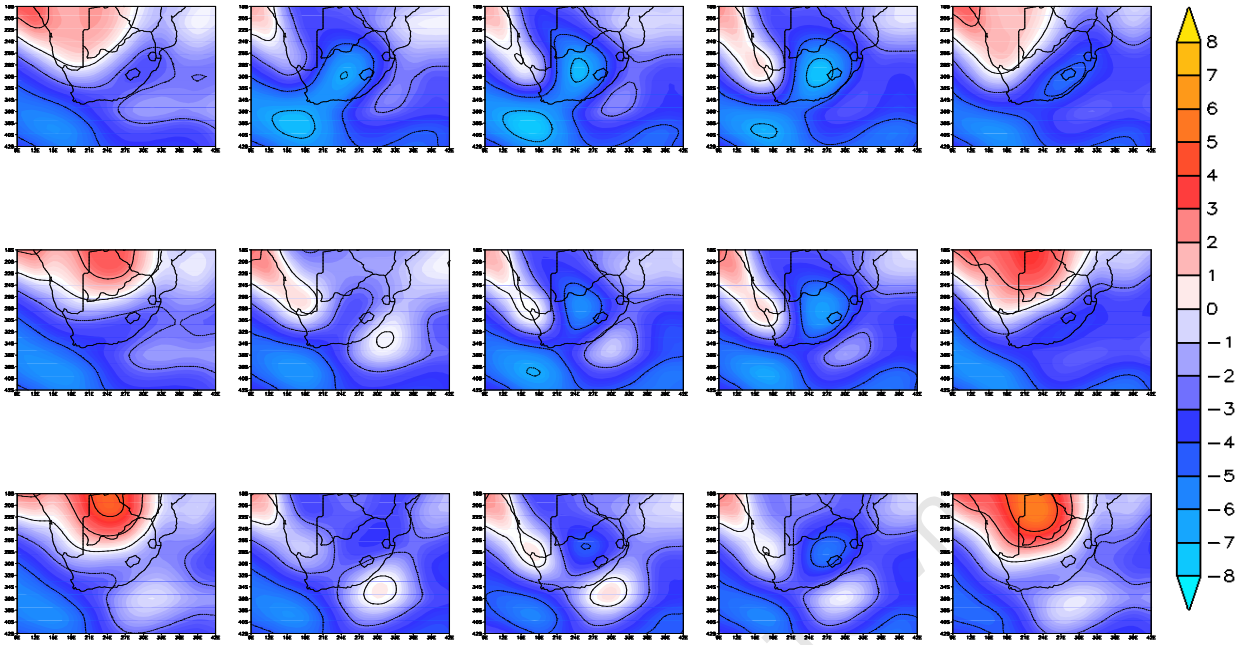


Figure B11: Anomaly of GFDL-NCEP for each SOM node for annual average surface air temperature, contour interval 1°C

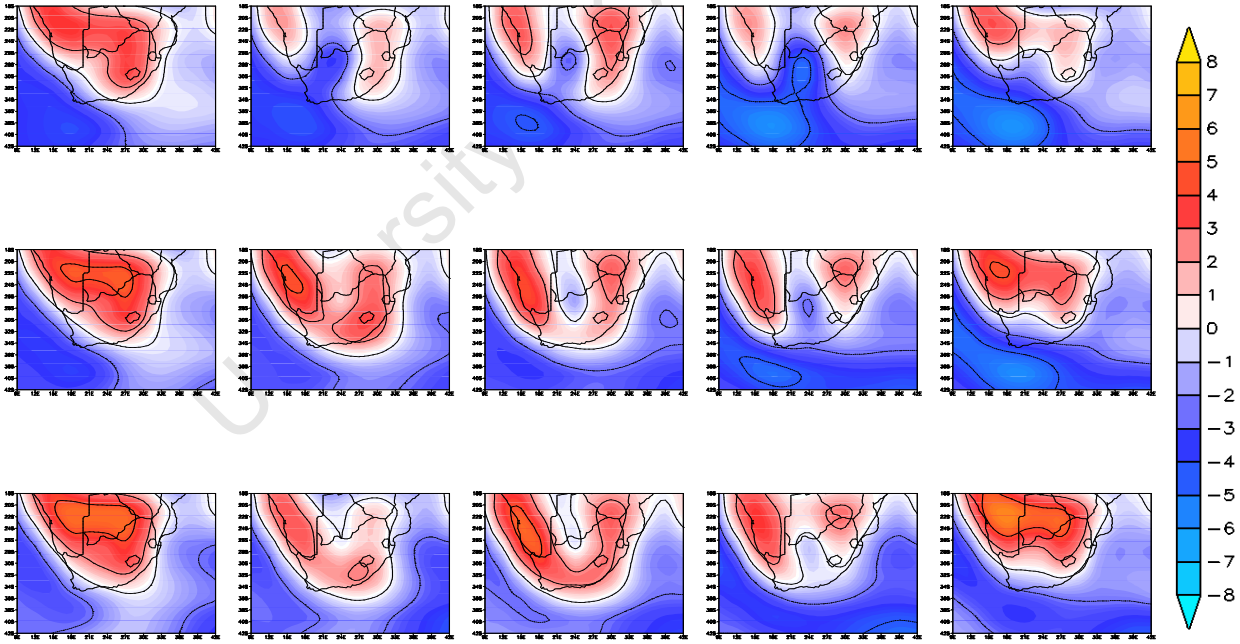


Figure B12: Anomaly of MRI -NCEP for each SOM node for annual average surface air temperature, contour interval 1°C

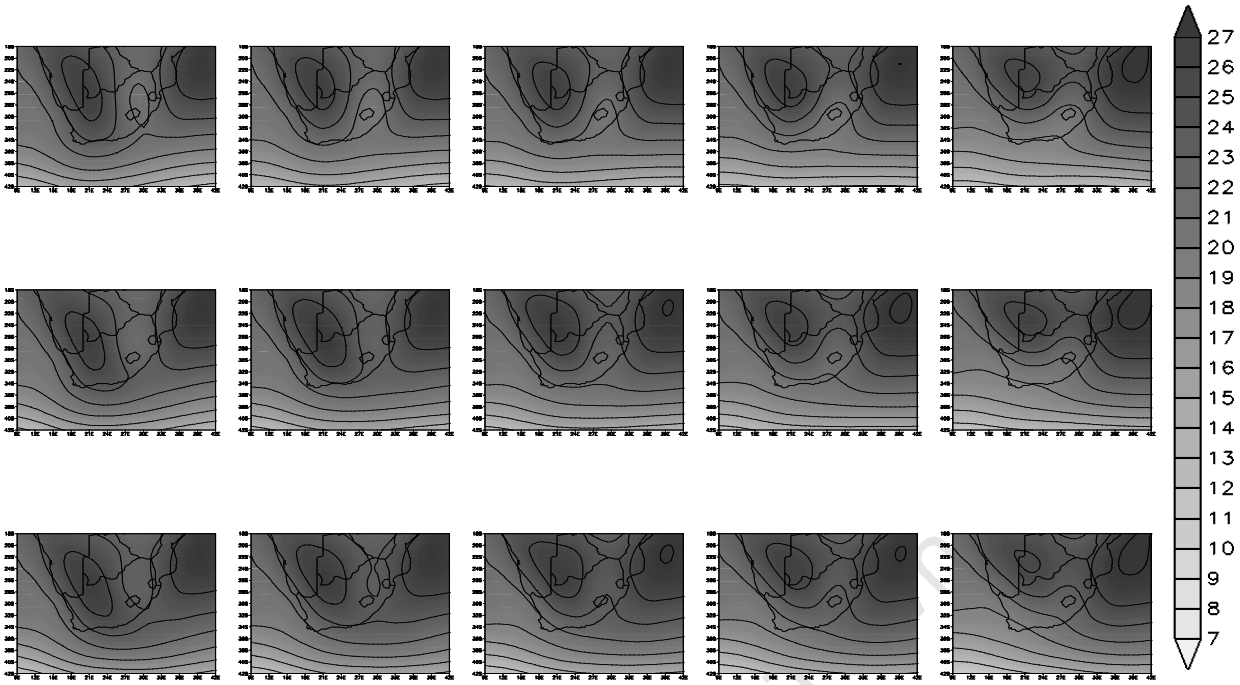


Figure B13: NCEP master SOM showing composite map of January average surface air temperature per node, contour interval of 2°C

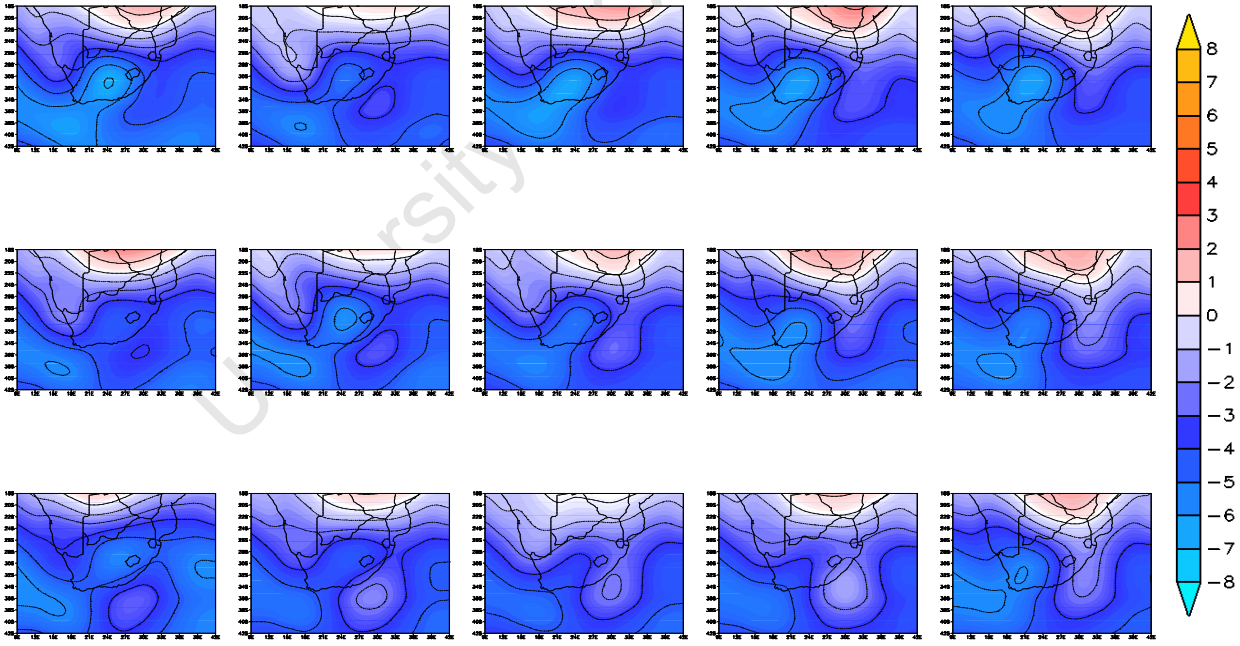


Figure B14: Anomaly of GFDL -NCEP for each SOM node for January average surface air temperature, contour interval 1°C

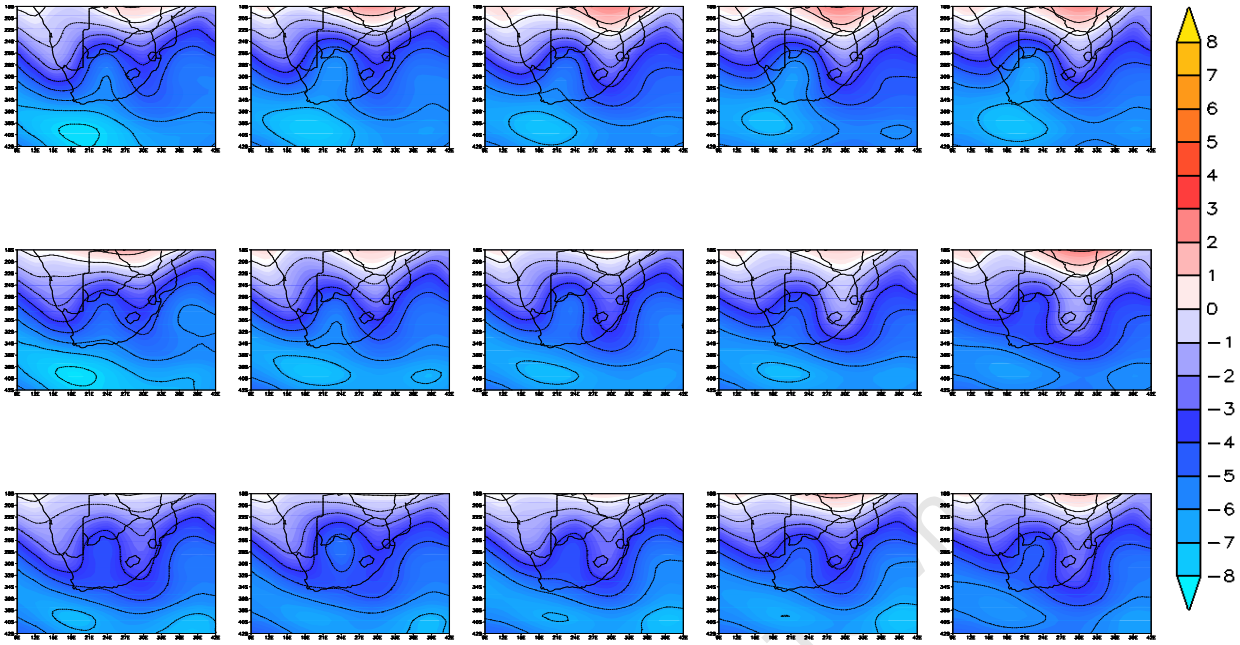


Figure B15: Anomaly of MIROC-NCEP for each SOM node for January average surface air temperature, contour interval 1°C

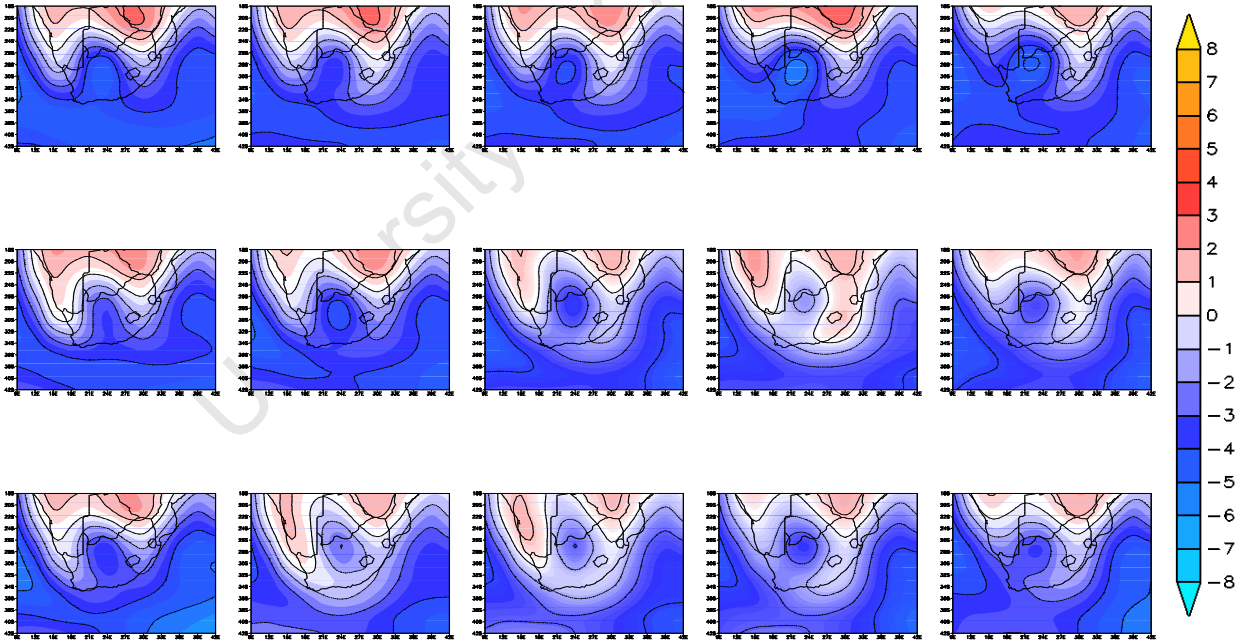


Figure B16: Anomaly of MRI-NCEP for each SOM node for January average surface air temperature, contour interval 1°C

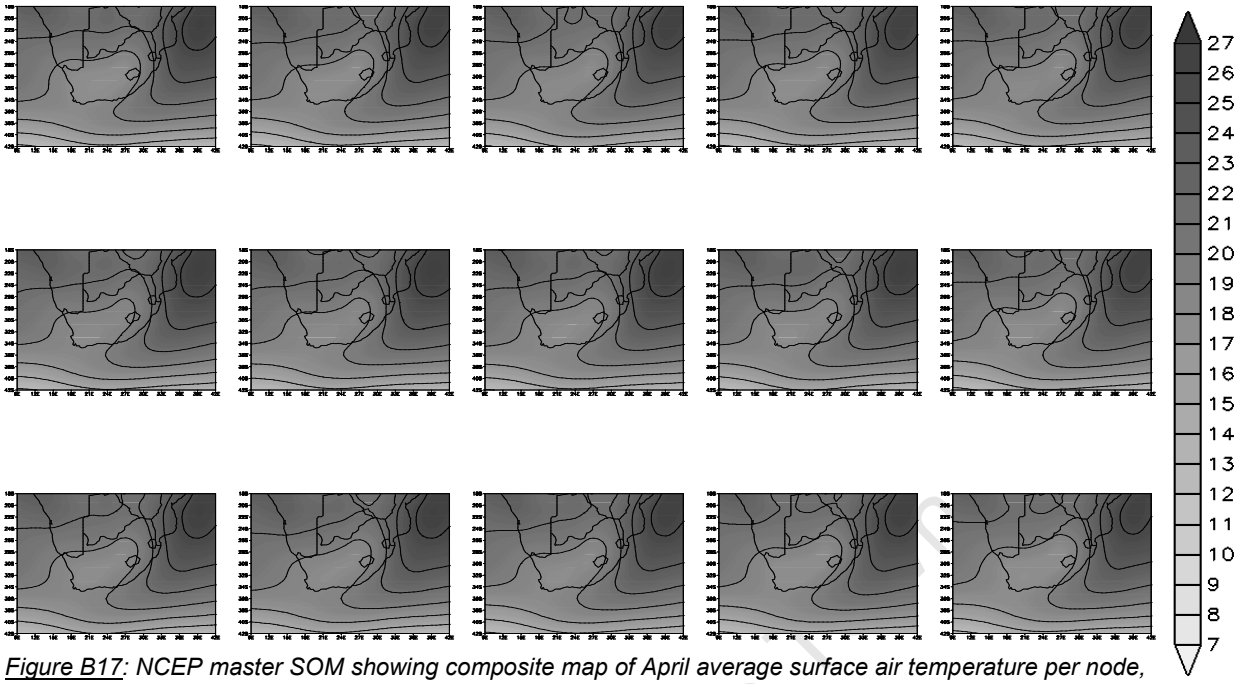


Figure B17: NCEP master SOM showing composite map of April average surface air temperature per node, contour interval of 2°C

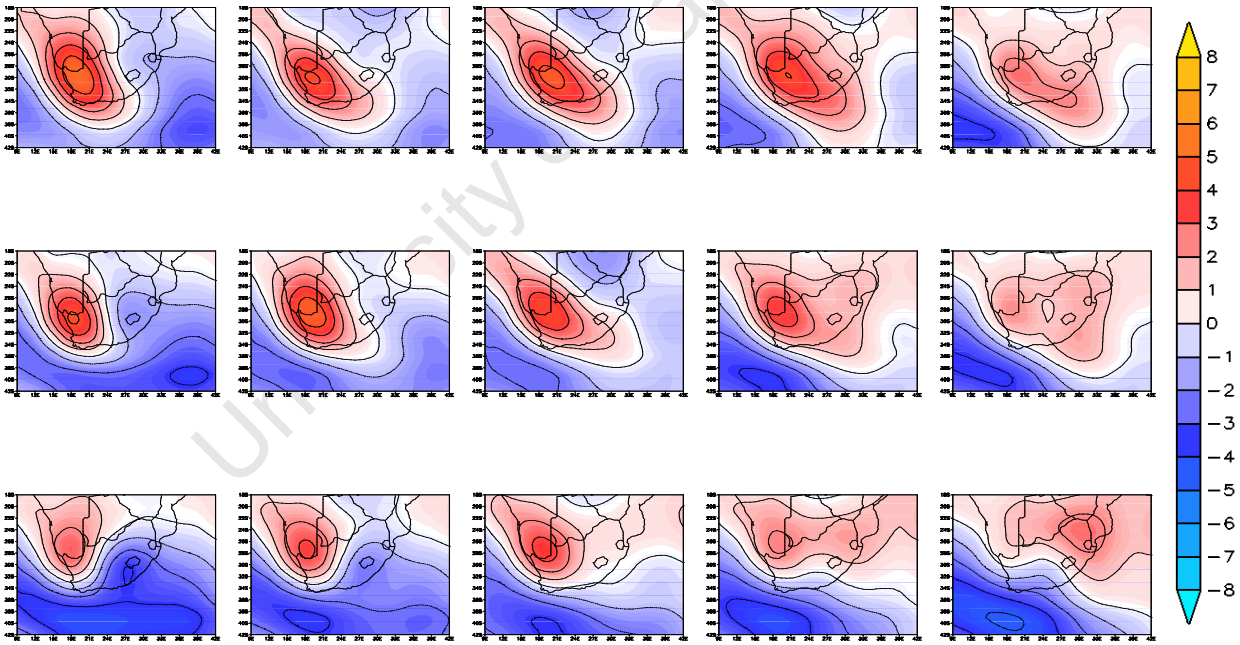


Figure B18: Anomaly of GFDL-NCEP for each SOM node for April average surface air temperature, contour interval 1°C

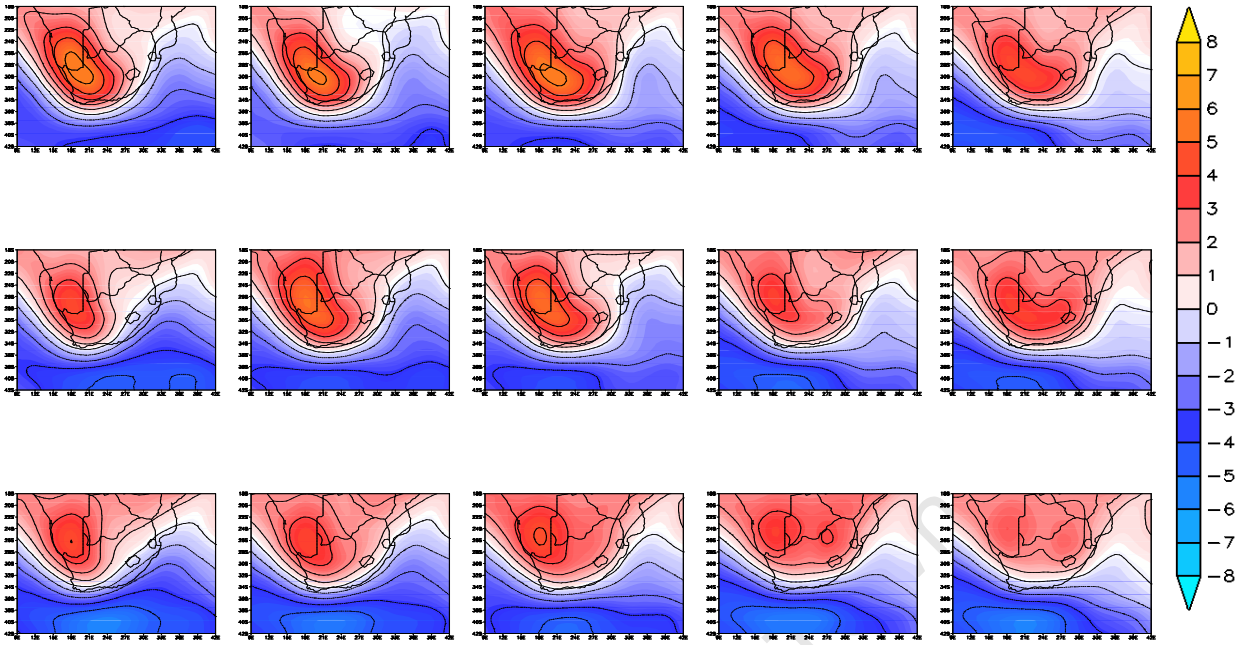


Figure B19: Anomaly of MIROC-NCEP for each SOM node for April average surface air temperature, contour interval 1°C

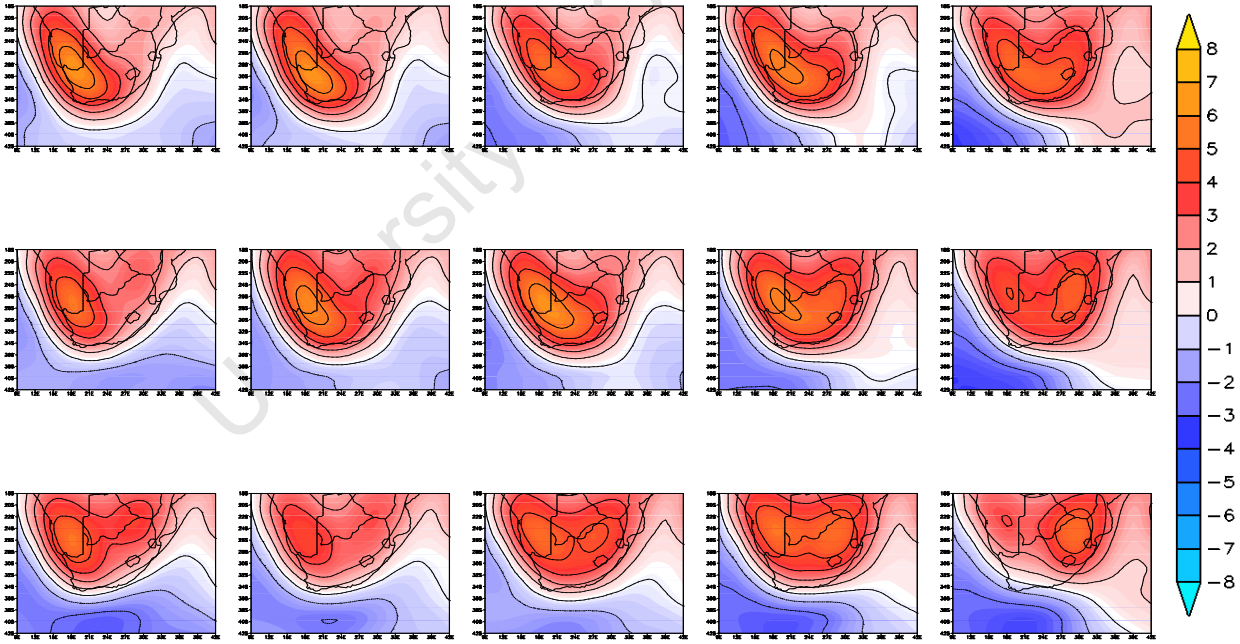


Figure B20: Anomaly of MRI-NCEP for each SOM node for April average surface air temperature, contour interval 1°C

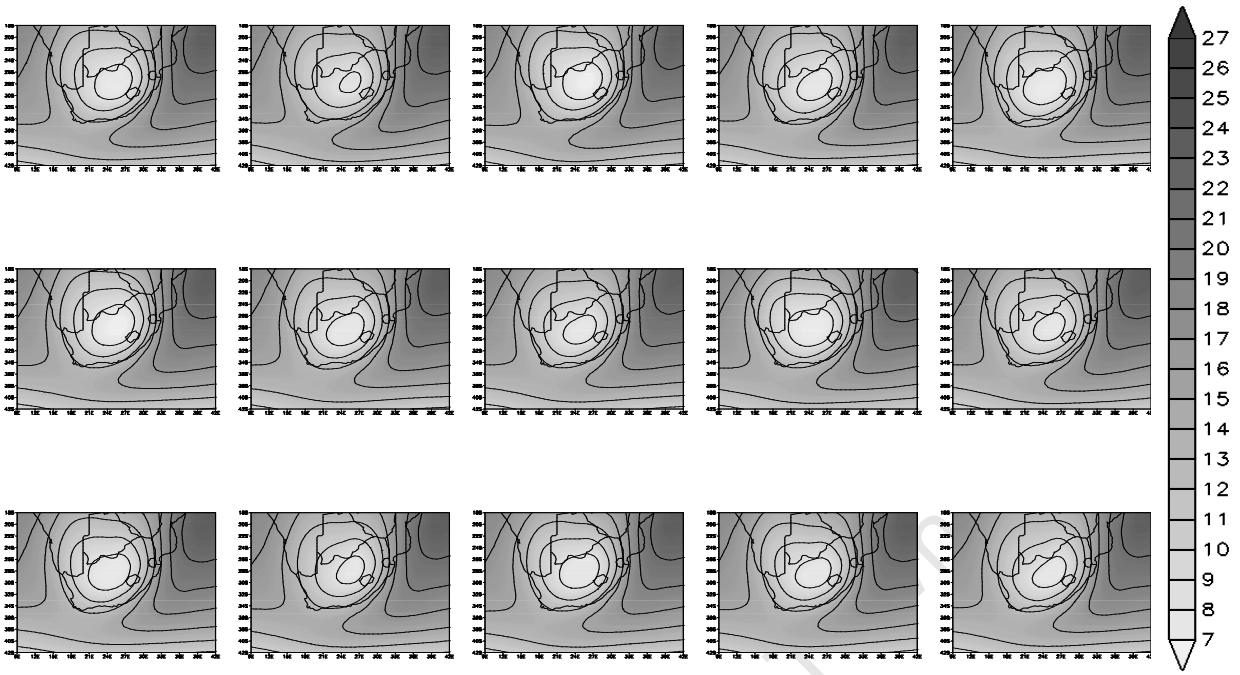


Figure B21: NCEP master SOM showing composite map of July average surface air temperature per node, contour interval of 2°C

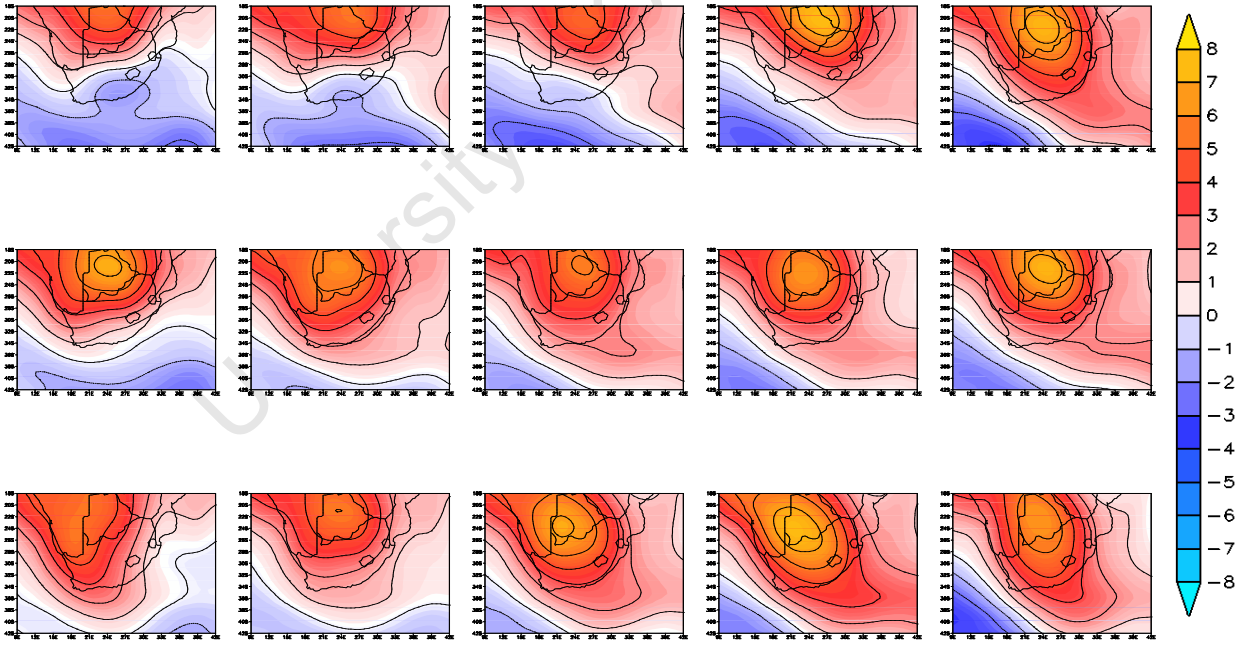


Figure B22: Anomaly of GFDL-NCEP for each SOM node for July average surface air temperature, contour interval 1°C

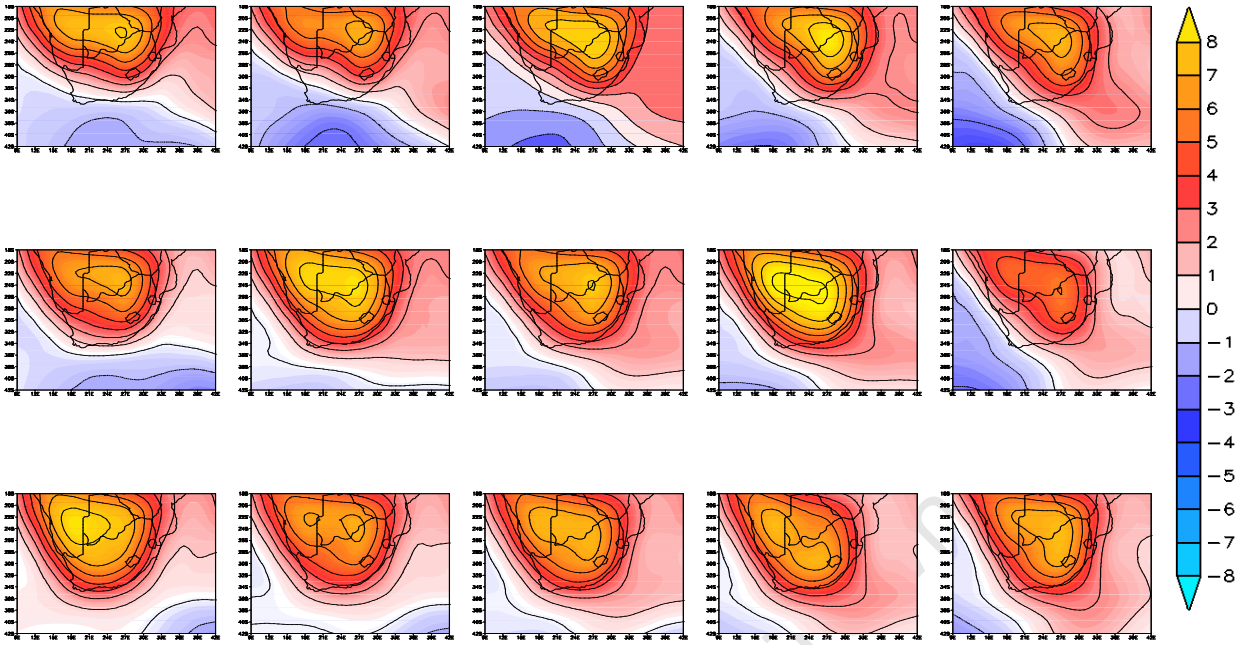


Figure B23: Anomaly of MRI-NCEP for each SOM node for July average surface air temperature, contour interval 1°C

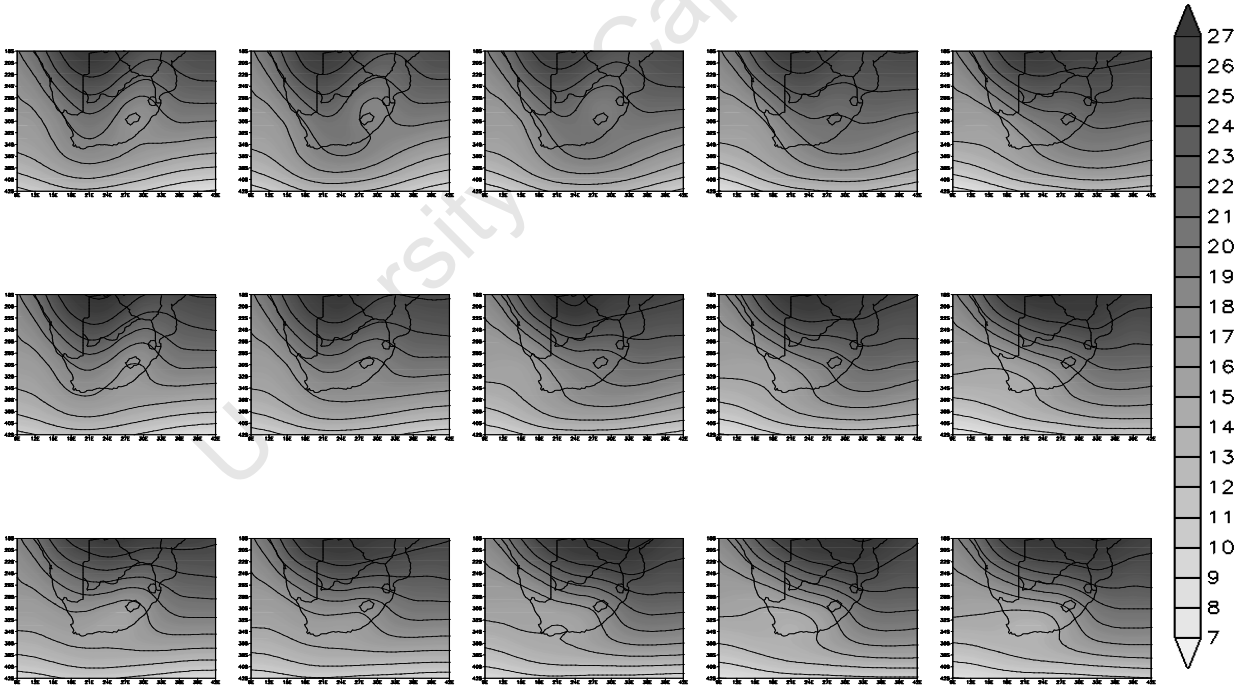


Figure B24: NCEP master SOM showing composite map of October average surface air temperature per node, contour interval of 2°C

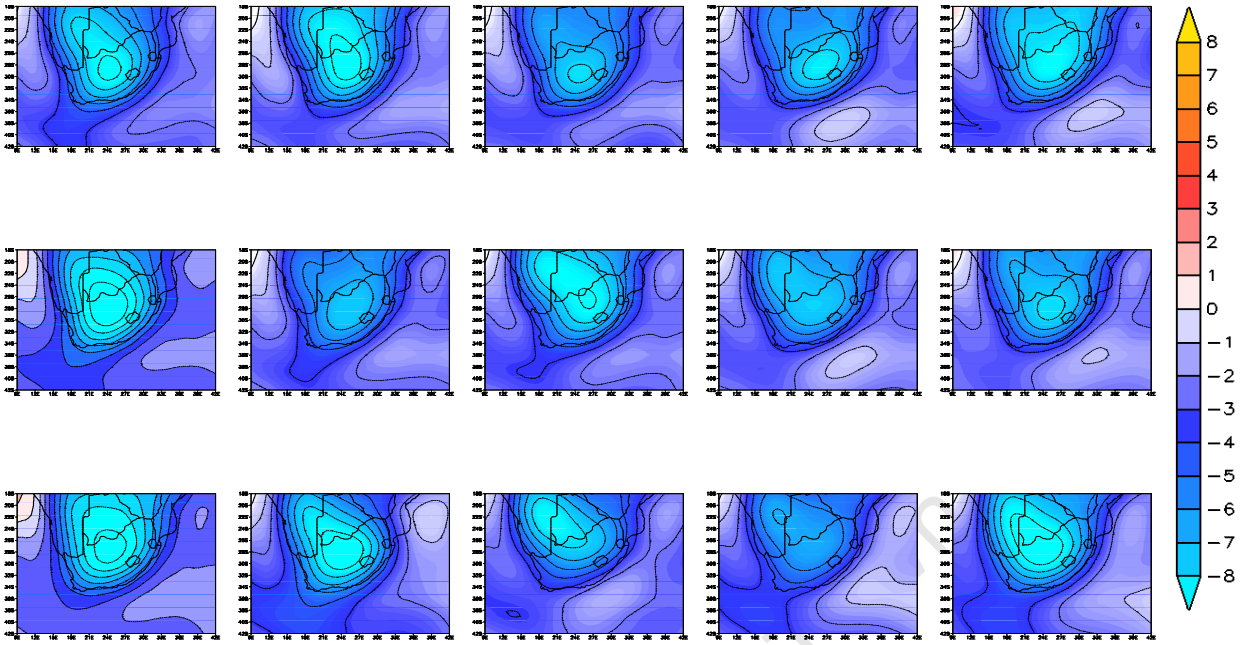


Figure B25: Anomaly of GFDL-NCEP for each SOM node for October average surface air temperature, contour interval 1°C

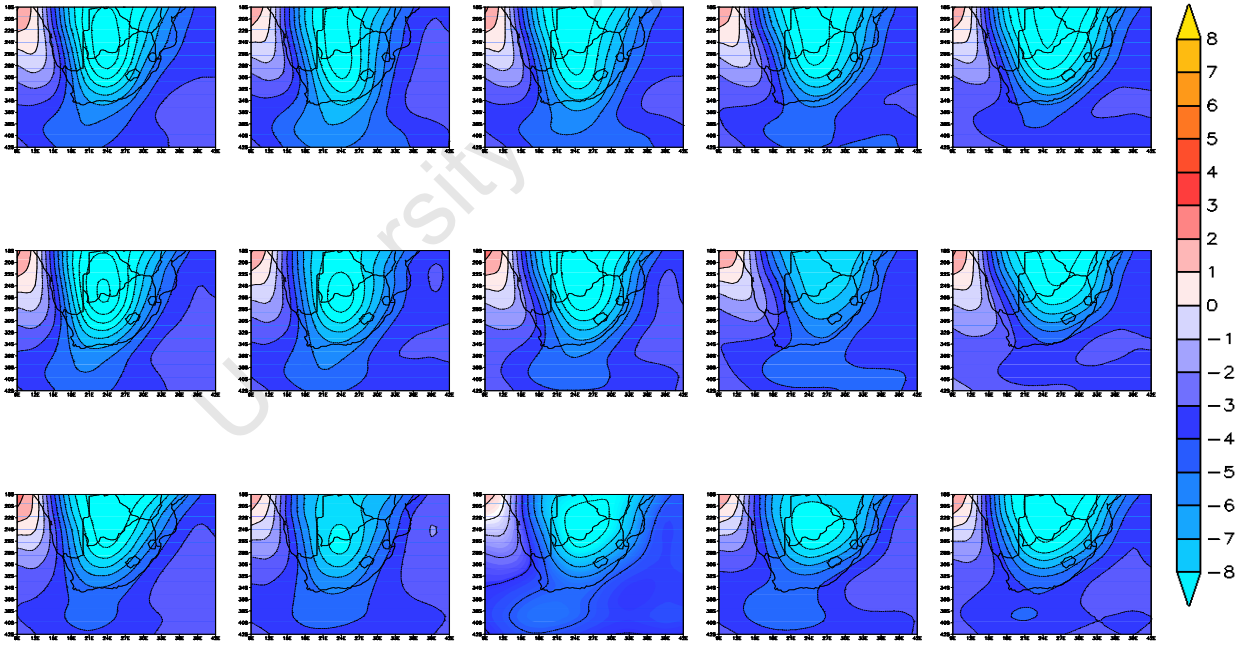


Figure B26: Anomaly of MIROC-NCEP for each SOM node for October average surface air temperature, contour interval 1°C

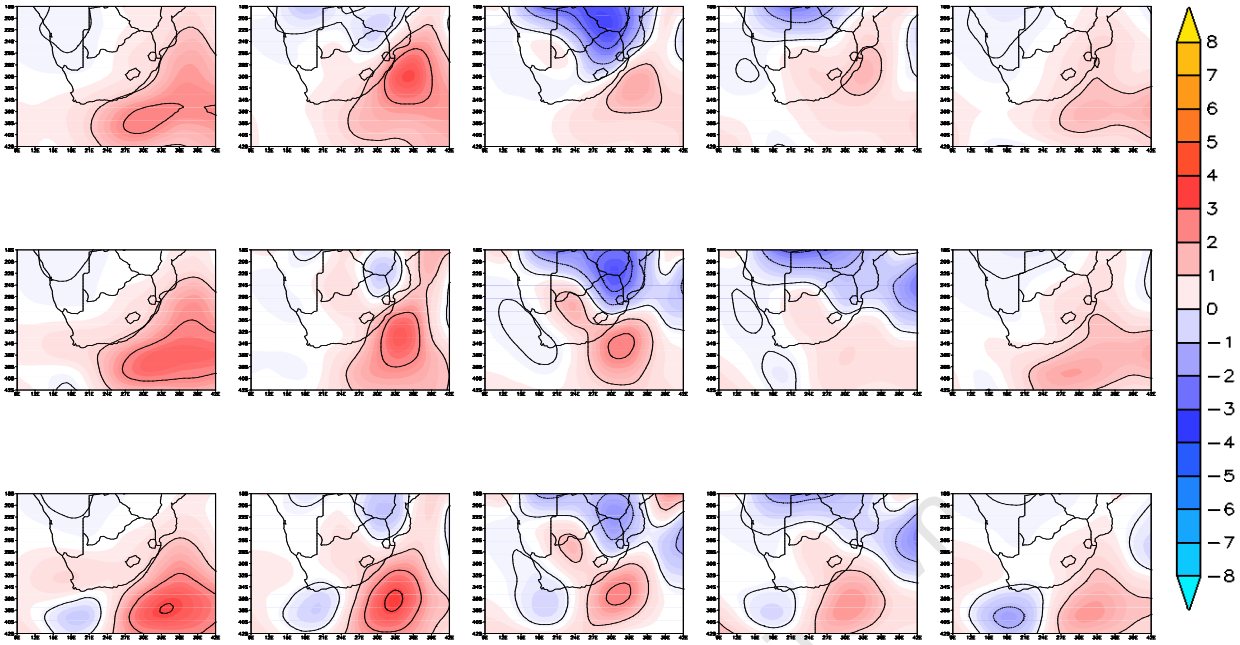


Figure B27: Anomaly of GFDL-NCEP for each SOM node for annual average precipitation, contour interval 1mm

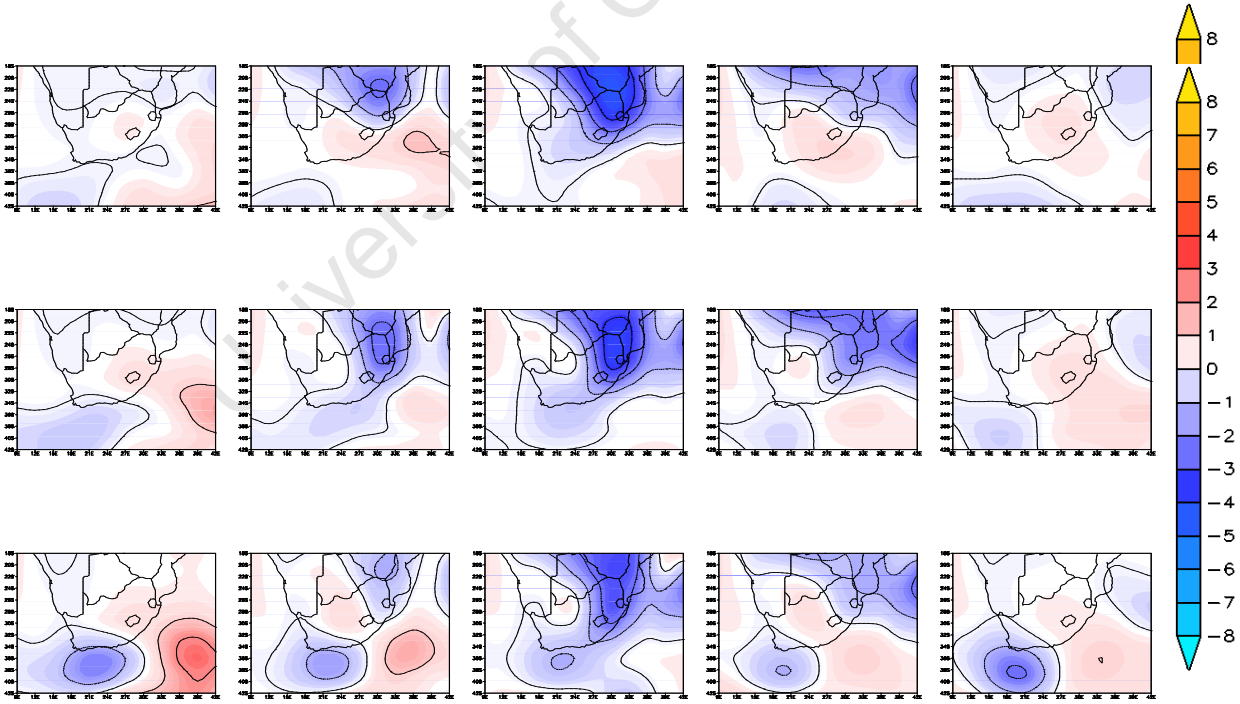


Figure B28: Anomaly of MRI-NCEP for each SOM node for annual average precipitation, contour interval 1mm

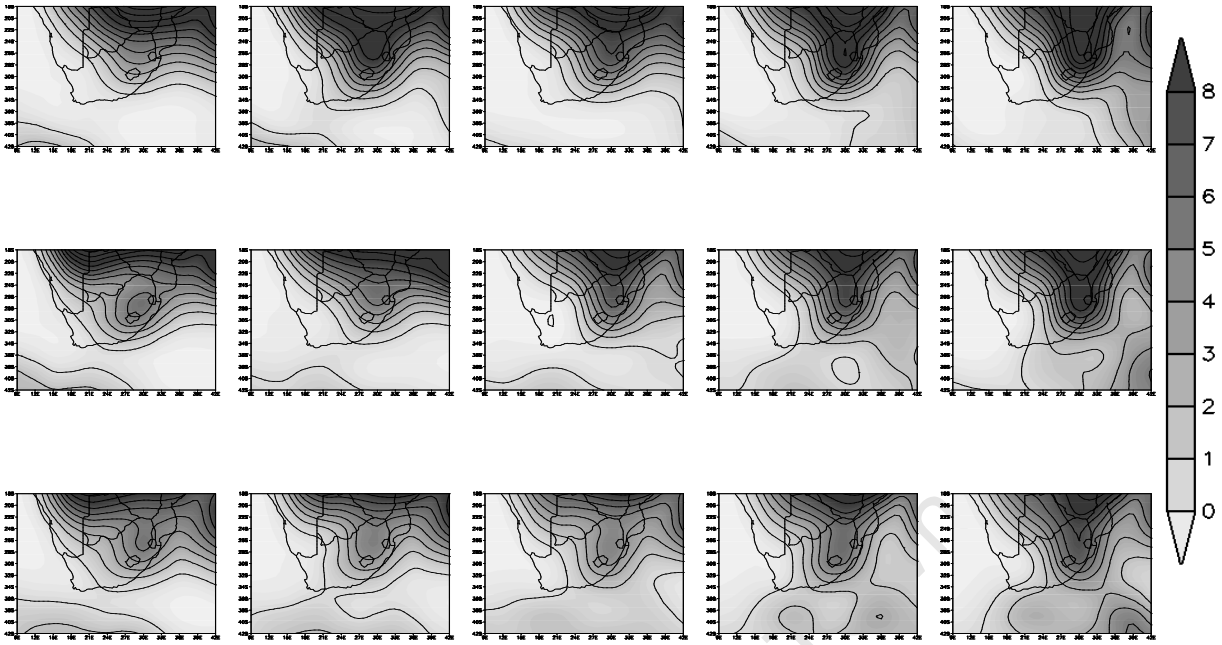


Figure B29: NCEP master SOM showing composite map of January average precipitation per node, contour interval 1mm

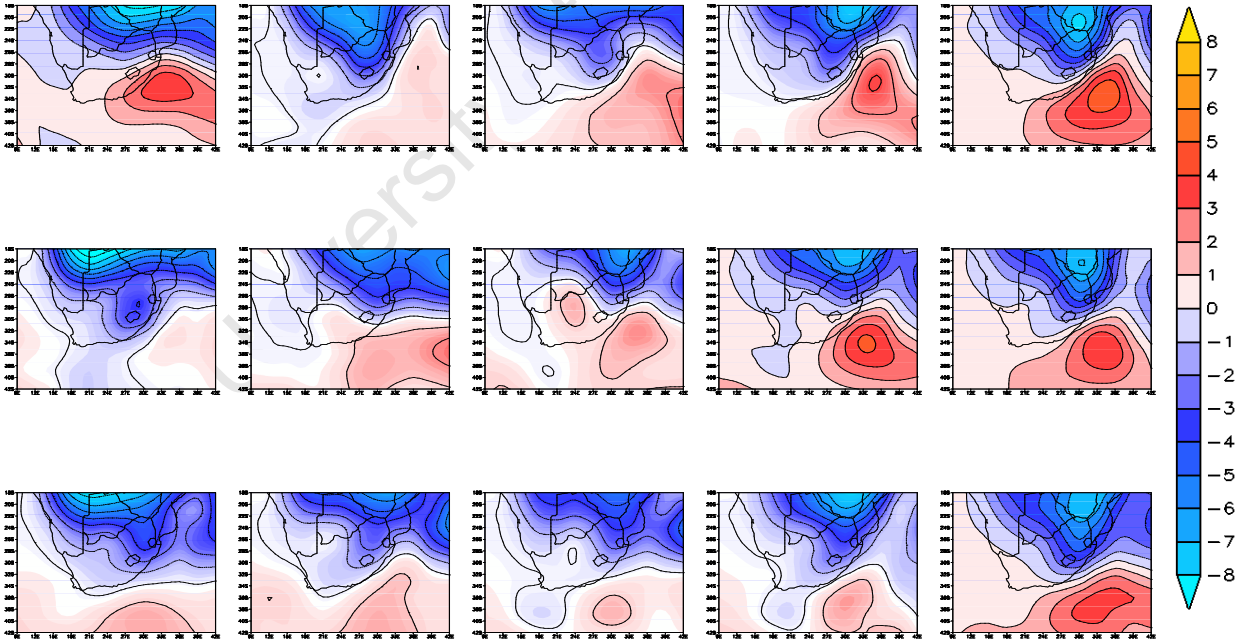


Figure B30: Anomaly of GFDL-NCEP for each SOM node for January average precipitation, contour interval 1mm

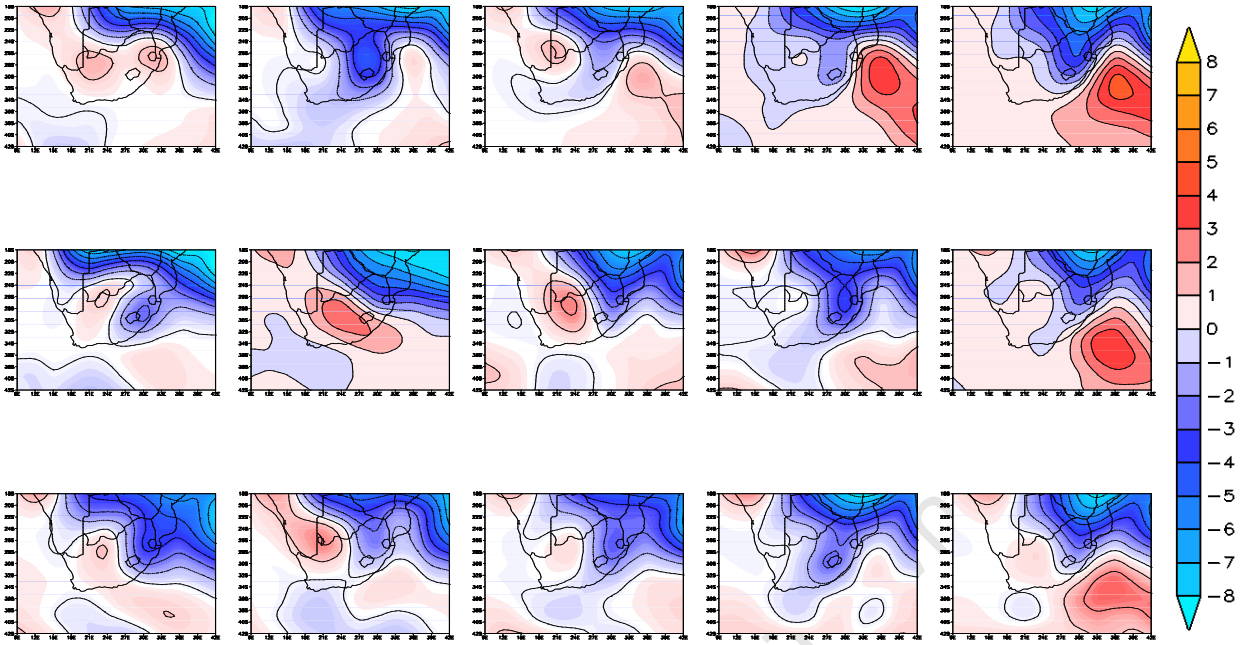


Figure B31: Anomaly of MIROC-NCEP for each SOM node for January average precipitation, contour interval 1mm

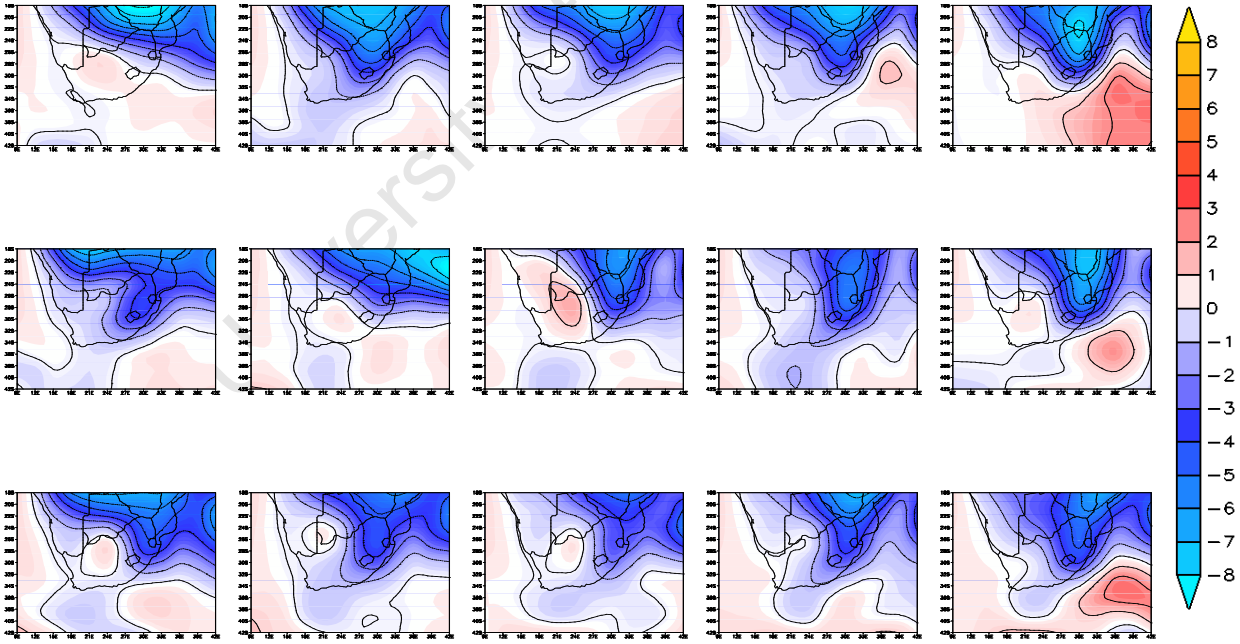


Figure B32: Anomaly of MRI-NCEP for each SOM node for January average precipitation, contour interval 1mm

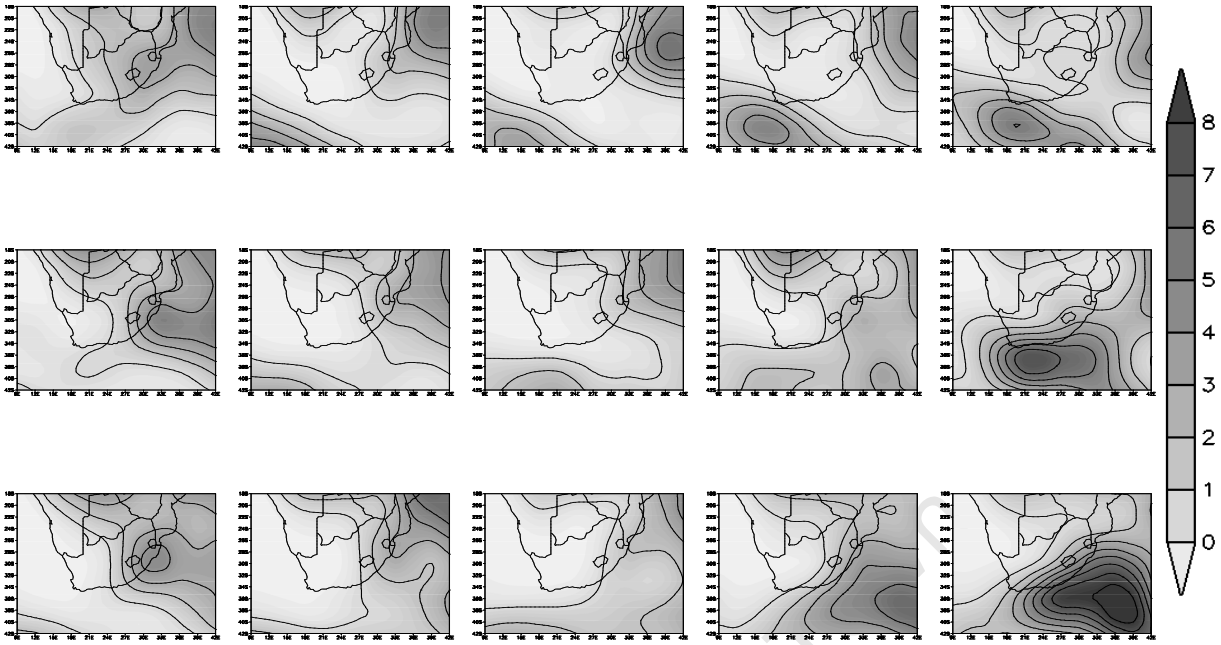


Figure B33: NCEP master SOM showing composite map of April average precipitation per node, contour interval 1mm

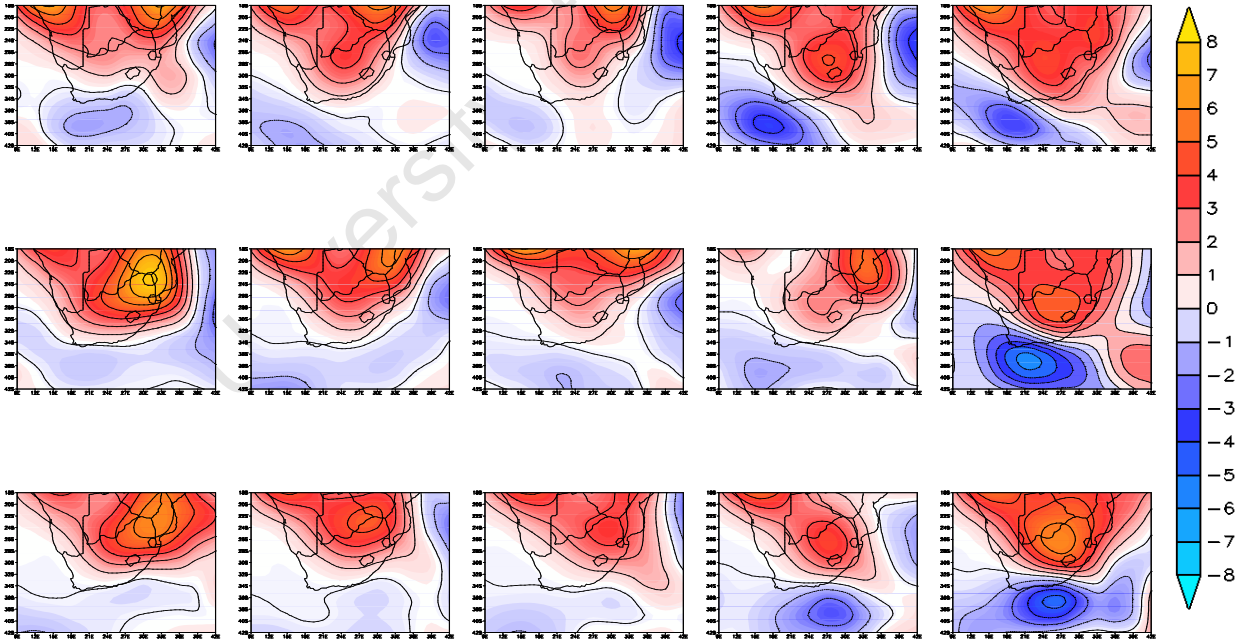


Figure B34: Anomaly of MIROC-NCEP for each SOM node for April average precipitation, contour interval 1mm

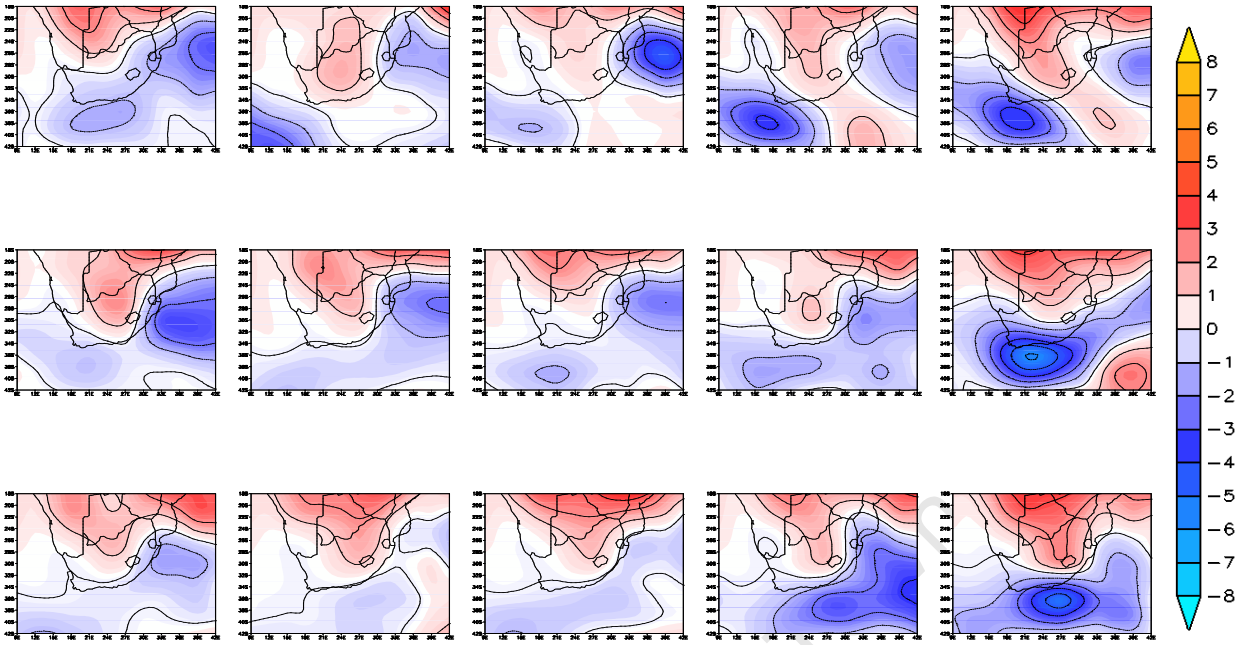


Figure B35: Anomaly of MRI-NCEP for each SOM node for April average precipitation, contour interval 1mm

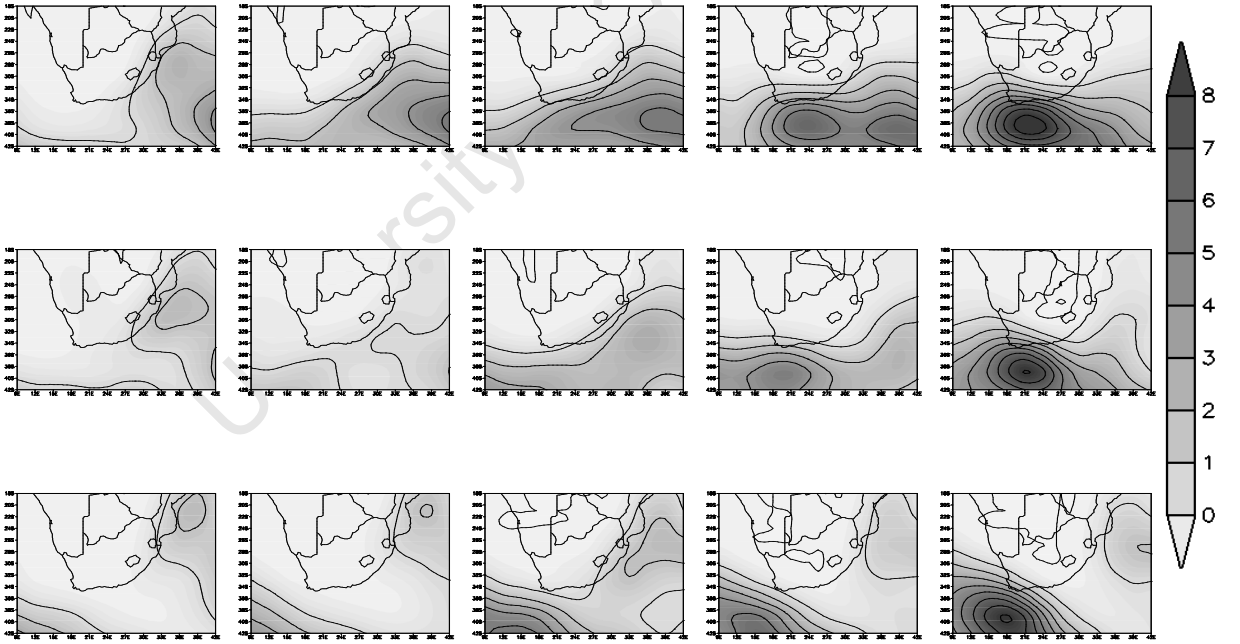


Figure B36: NCEP master SOM showing composite map of July average precipitation per node, contour interval 1mm

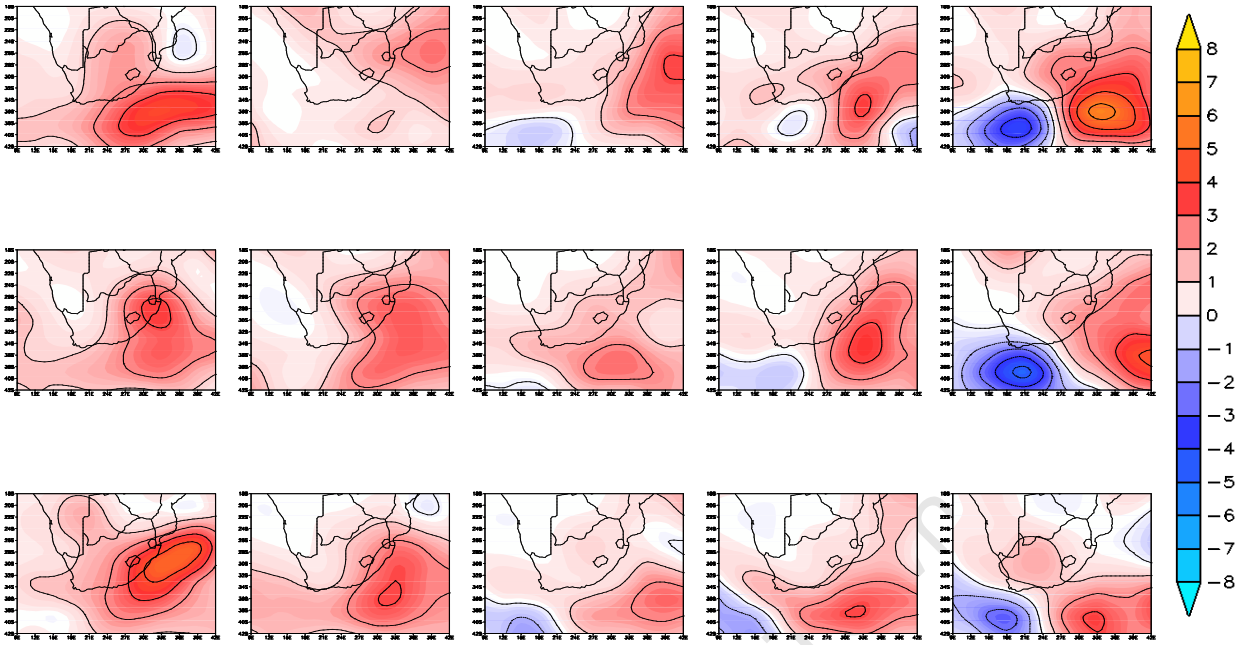


Figure B37: Anomaly of GFDL-NCEP for each SOM node for July average precipitation, contour interval 1mm

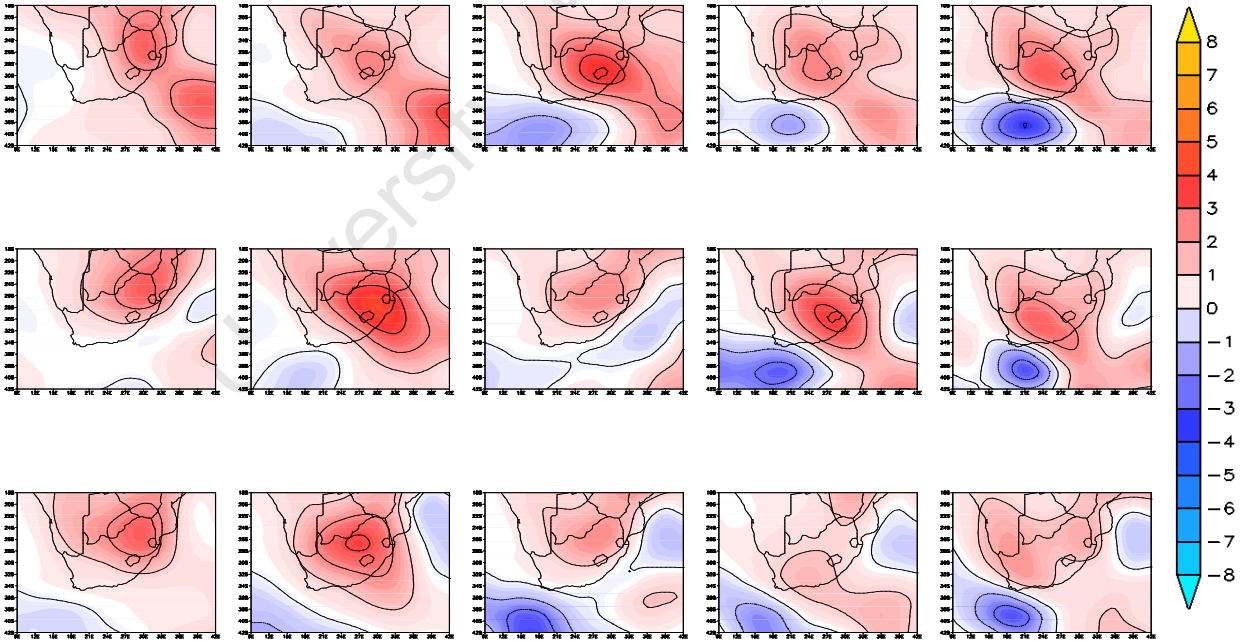


Figure B38: Anomaly of MIROC-NCEP for each SOM node for July average precipitation, contour interval 1mm

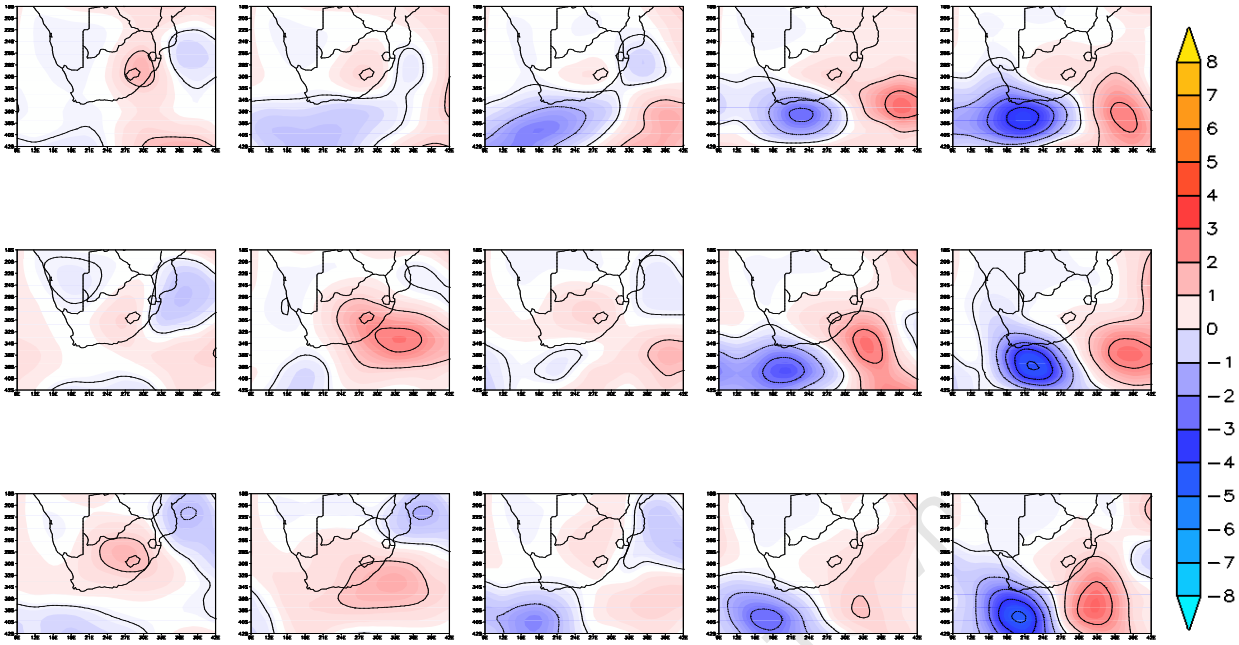


Figure B39: Anomaly of MRI-NCEP for each SOM node for July average precipitation, contour interval 1mm

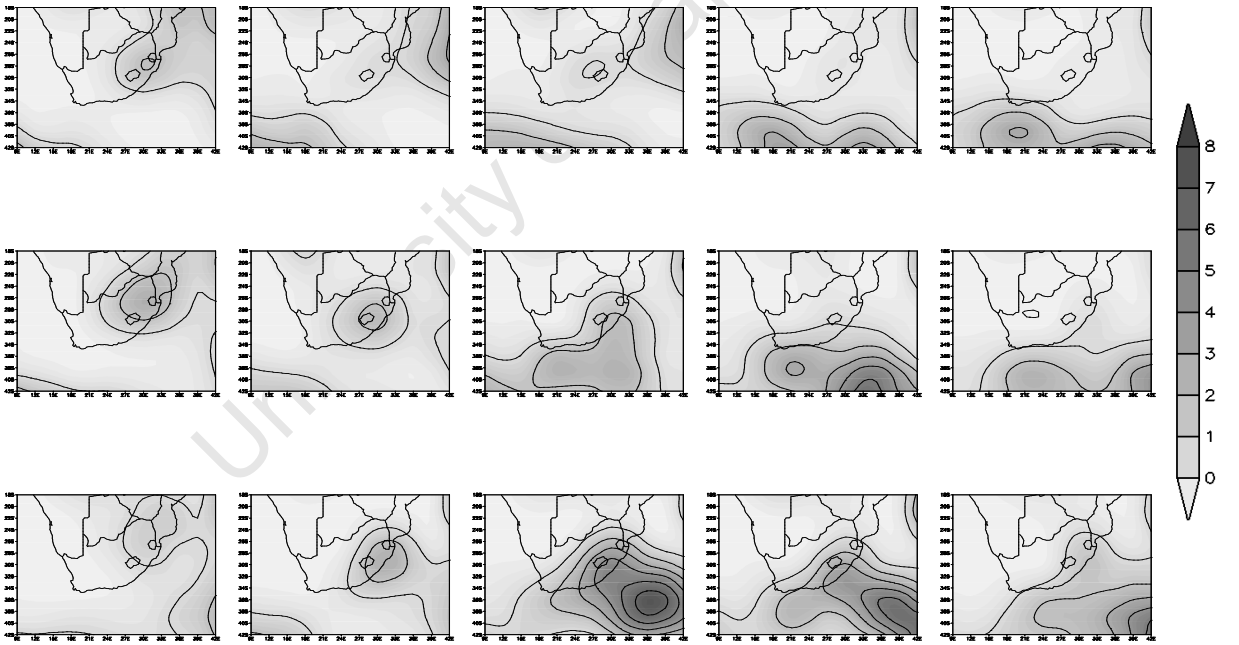


Figure B40: NCEP master SOM showing composite map of October average precipitation per node, contour interval 1mm

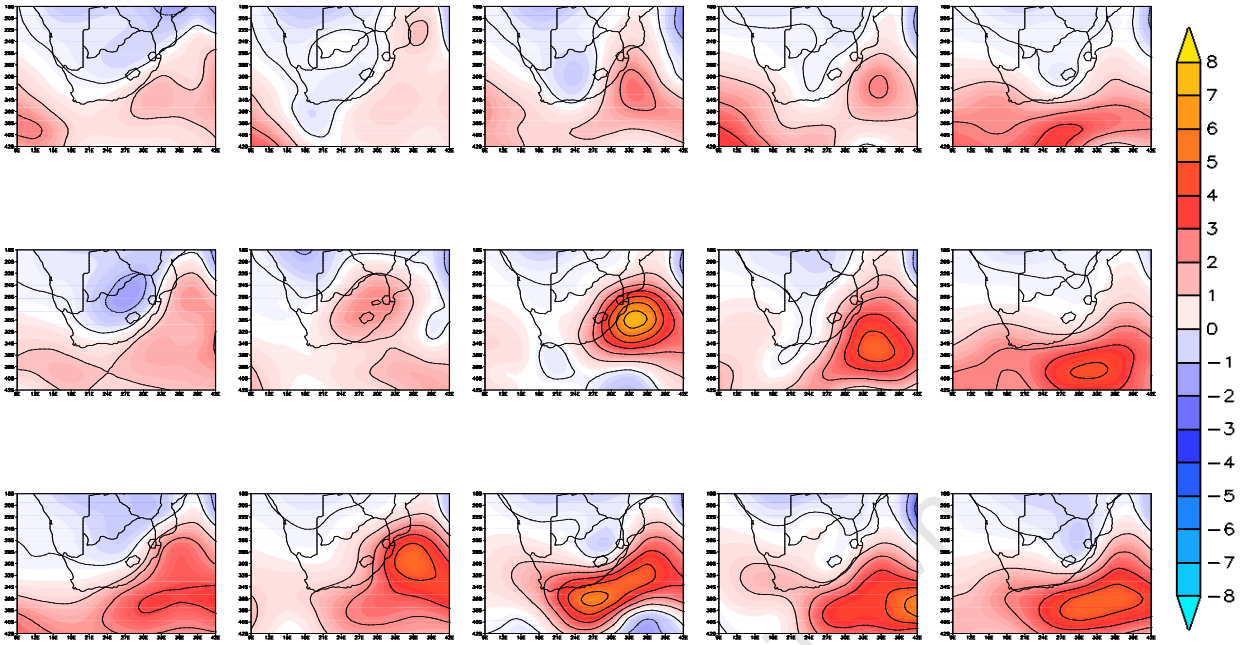


Figure B41: Anomaly of GFDL-NCEP for each SOM node for October average precipitation, contour interval 1mm

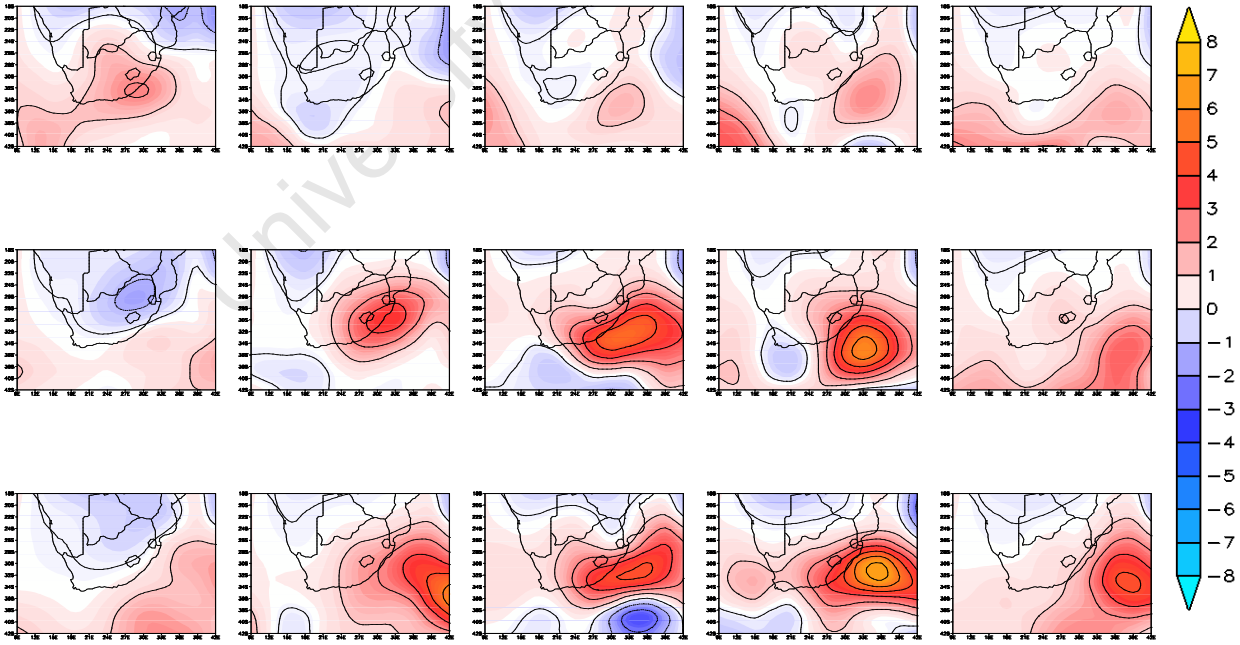


Figure B42: Anomaly of MRI-NCEP for each SOM node for October average precipitation, contour interval 1mm

Appendix III

Figures supporting Chapter 5:

Weights and map means

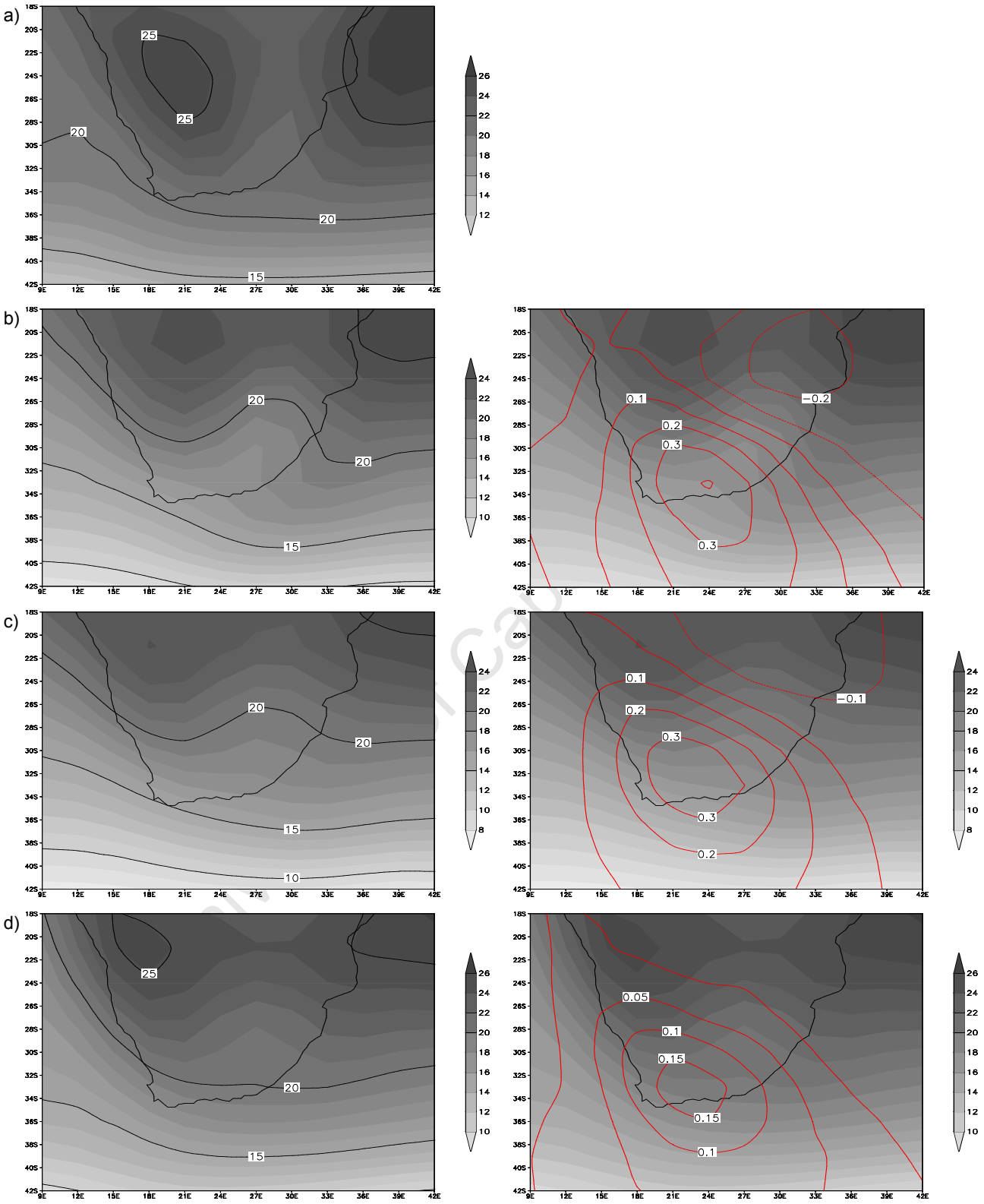


Figure C1: ***Frequency corrected temperature (°C):*** *a) NCEP January mean temperature compared to annual control (1979–2000) frequency corrected temperature (left) and original control (1979–2000) mean temperature in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

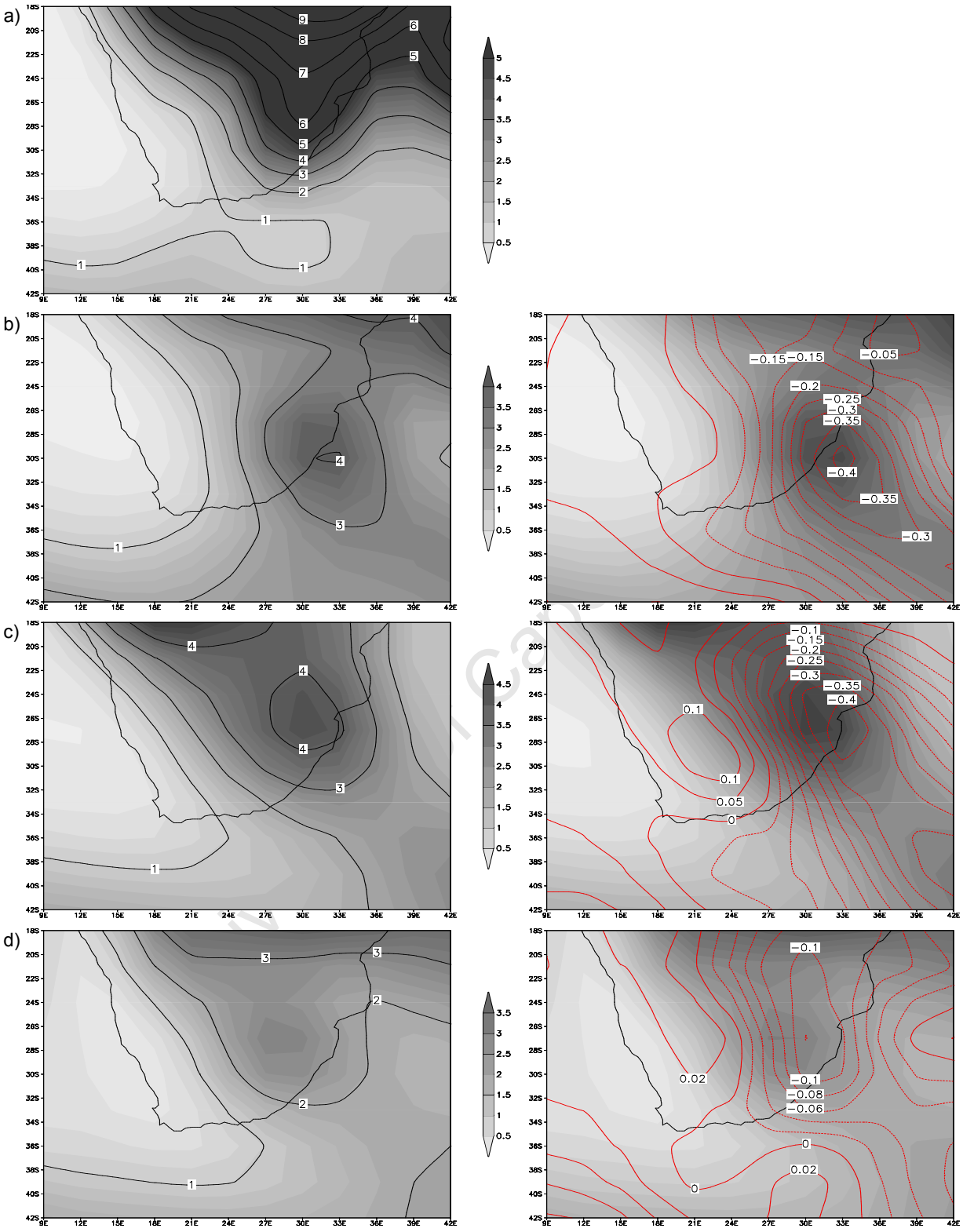


Figure C2: ***Frequency corrected precipitation (mm/d):*** *a) NCEP January mean precipitation compared to annual control (1979–2000) frequency corrected precipitation (left) and original control (1979–2000) mean precipitation in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

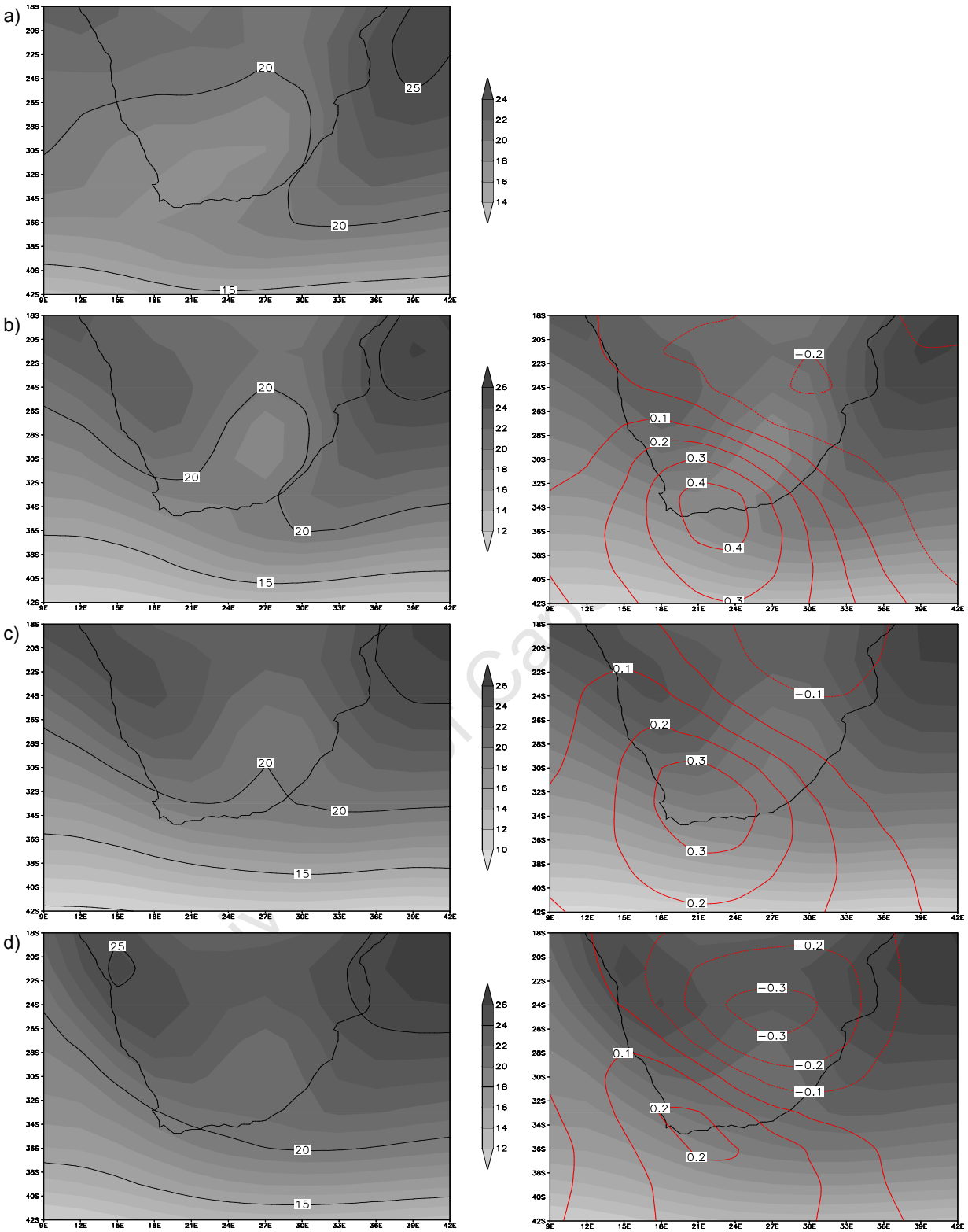


Figure C3: ***Frequency corrected temperature (°C):*** *a) NCEP April mean temperature compared to annual control (1979–2000) frequency corrected temperature (left) and original control (1979–2000) mean temperature in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

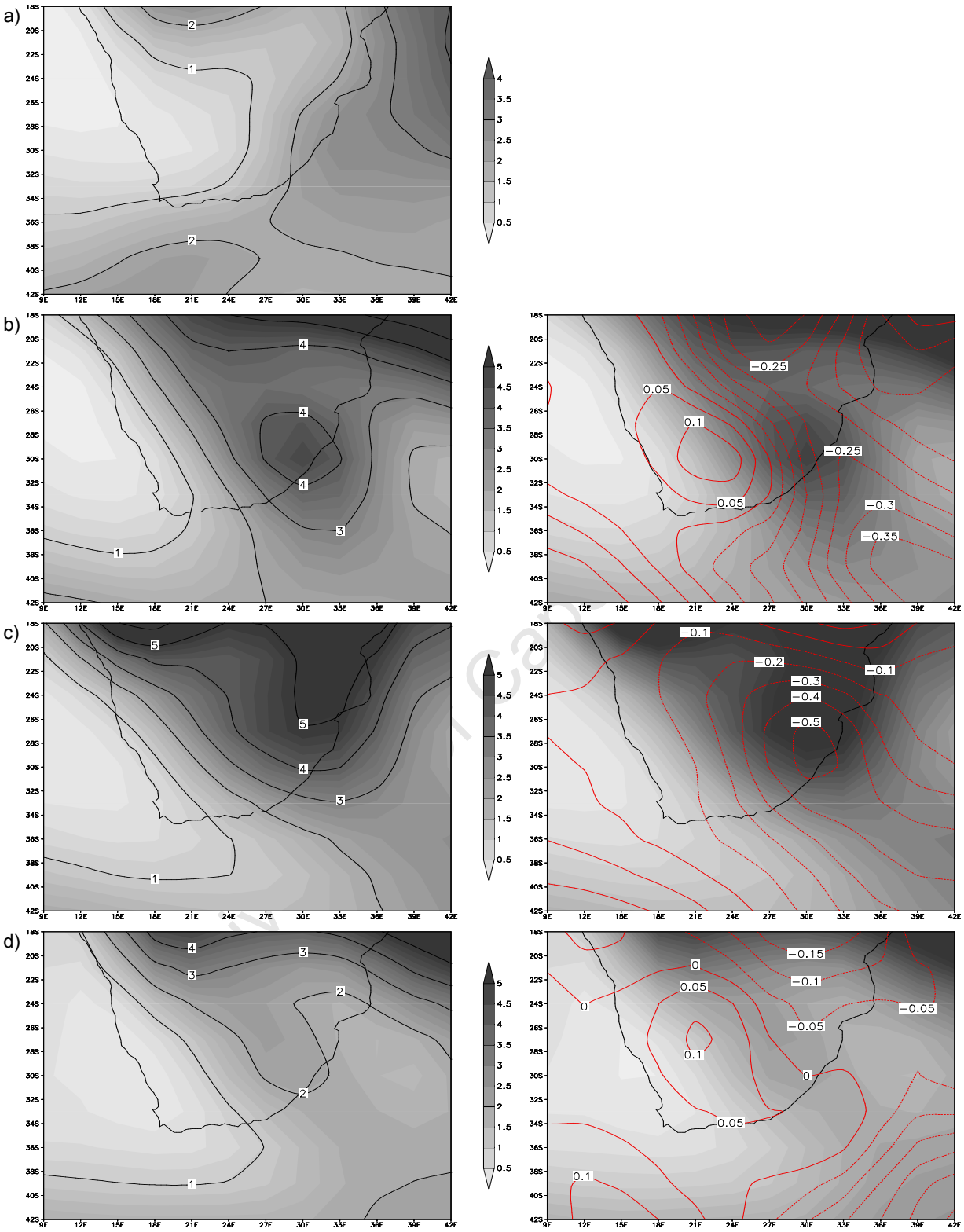


Figure C4: ***Frequency corrected precipitation (mm/d):*** *a) NCEP April mean precipitation compared to annual control (1979–2000) frequency corrected precipitation (left) and original control (1979–2000) mean precipitation in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

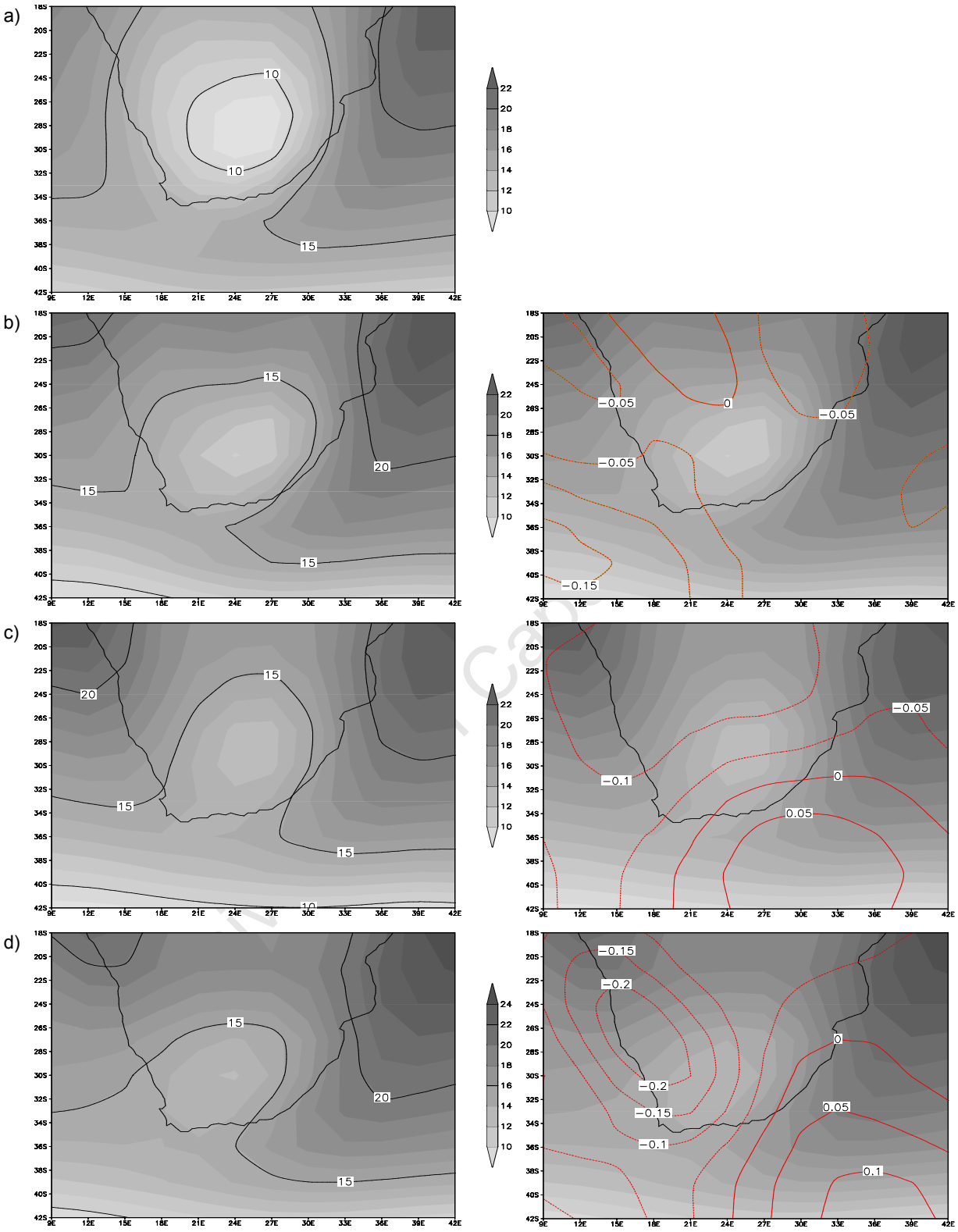


Figure C5: ***Frequency corrected temperature (°C):*** *a) NCEP July mean temperature compared to annual control (1979–2000) frequency corrected temperature (left) and original control (1979–2000) mean temperature in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

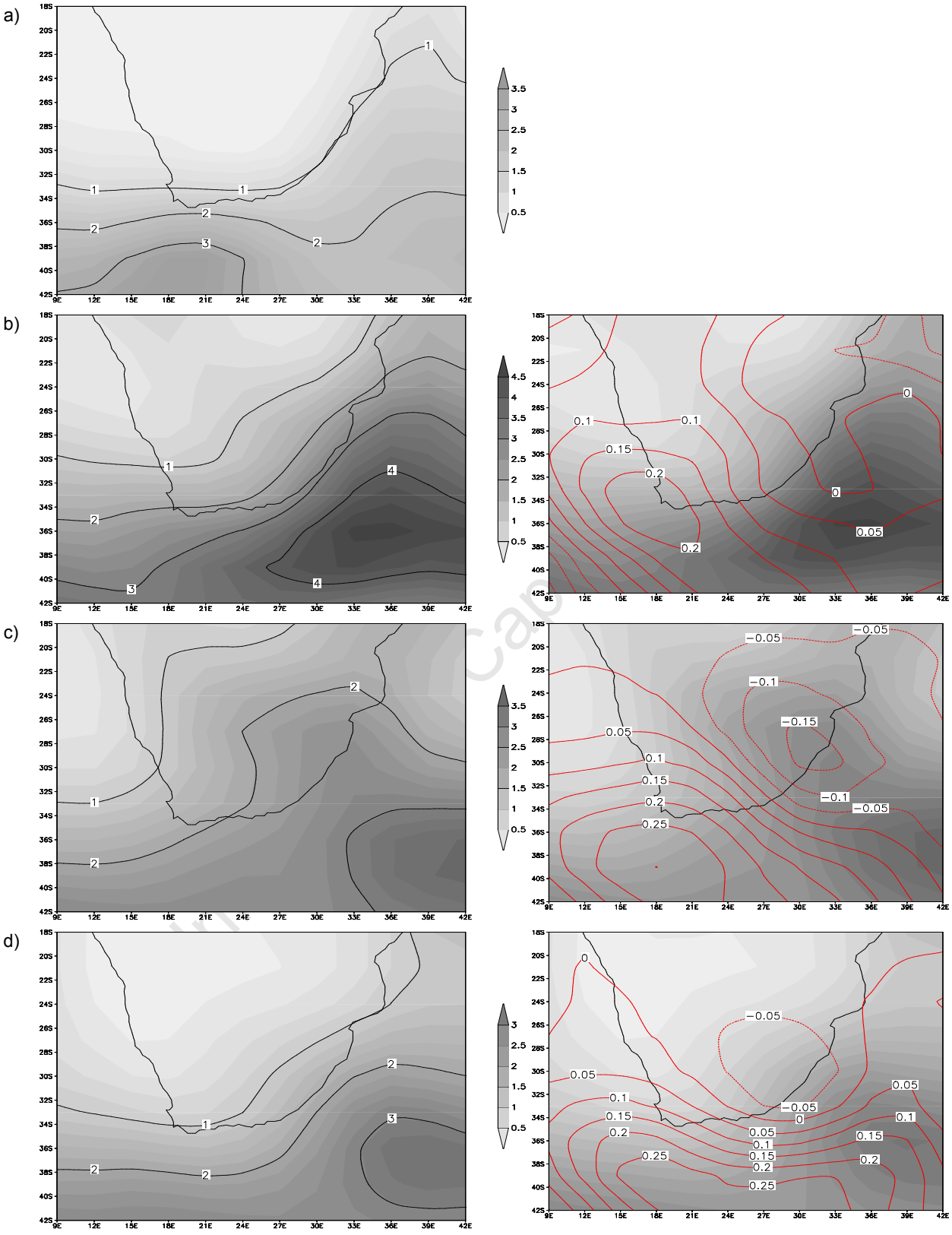


Figure C6: ***Frequency corrected precipitation (mm/d):*** *a) NCEP July mean precipitation compared to annual control (1979–2000) frequency corrected precipitation (left) and original control (1979–2000) mean precipitation in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

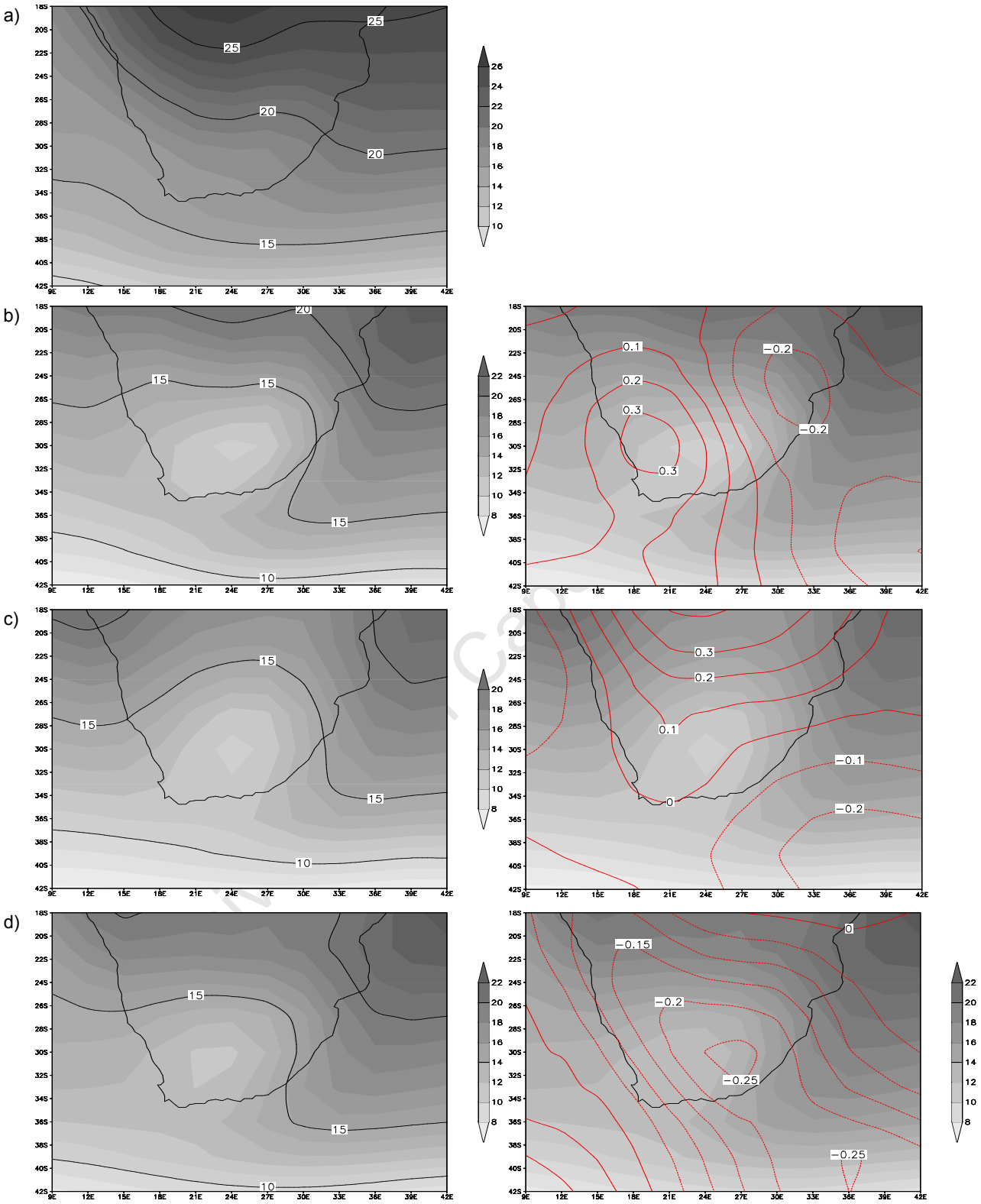


Figure C7: ***Frequency corrected temperature (°C):*** *a) NCEP October mean temperature compared to annual control (1979–2000) frequency corrected temperature (left) and original control (1979–2000) mean temperature in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

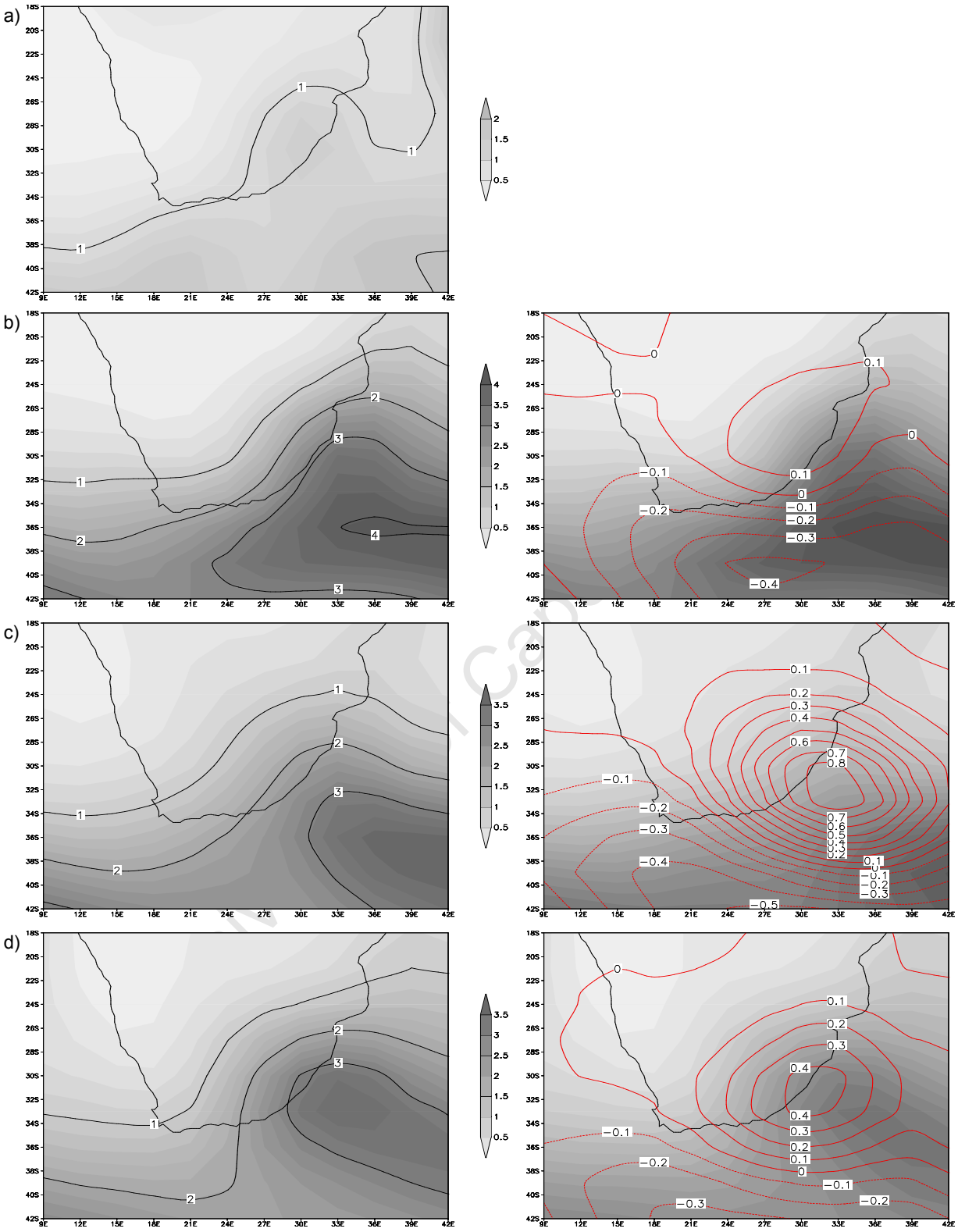


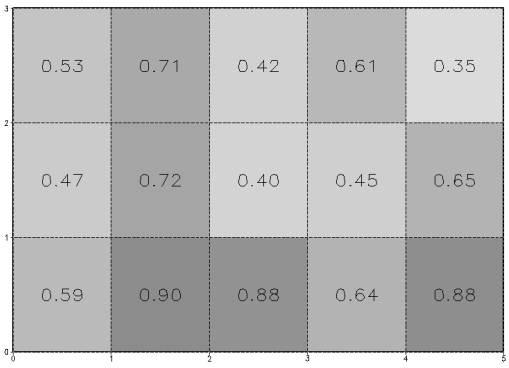
Figure C8: ***Frequency corrected precipitation (mm/d):*** *a) NCEP October mean precipitation compared to annual control (1979–2000) frequency corrected precipitation (left) and original control (1979–2000) mean precipitation in shading with difference between corrected and mean in overlaid contours (right) for b) GFDL, c) MIROC, d) MRI*

Appendix IV

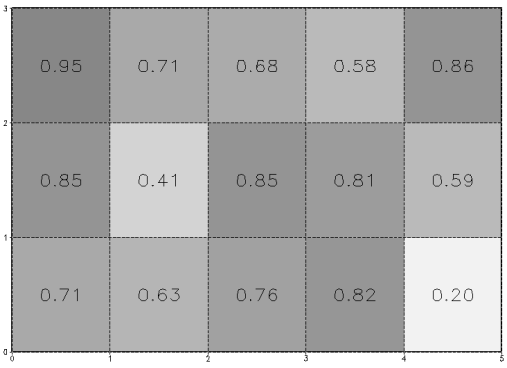
Figures supporting Chapter 6:

Synoptic specific multi-model plots

a)



b)



c)

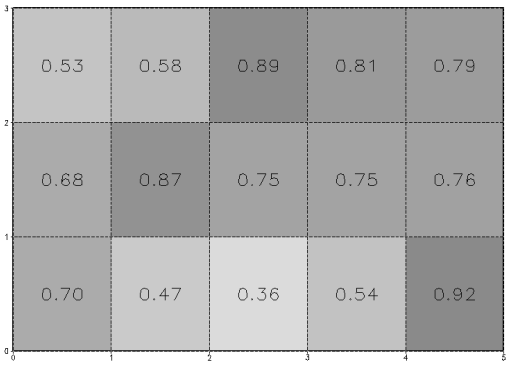


Figure D1: Node weights for January data: a) GFDL, b) MIROC, c) MRI

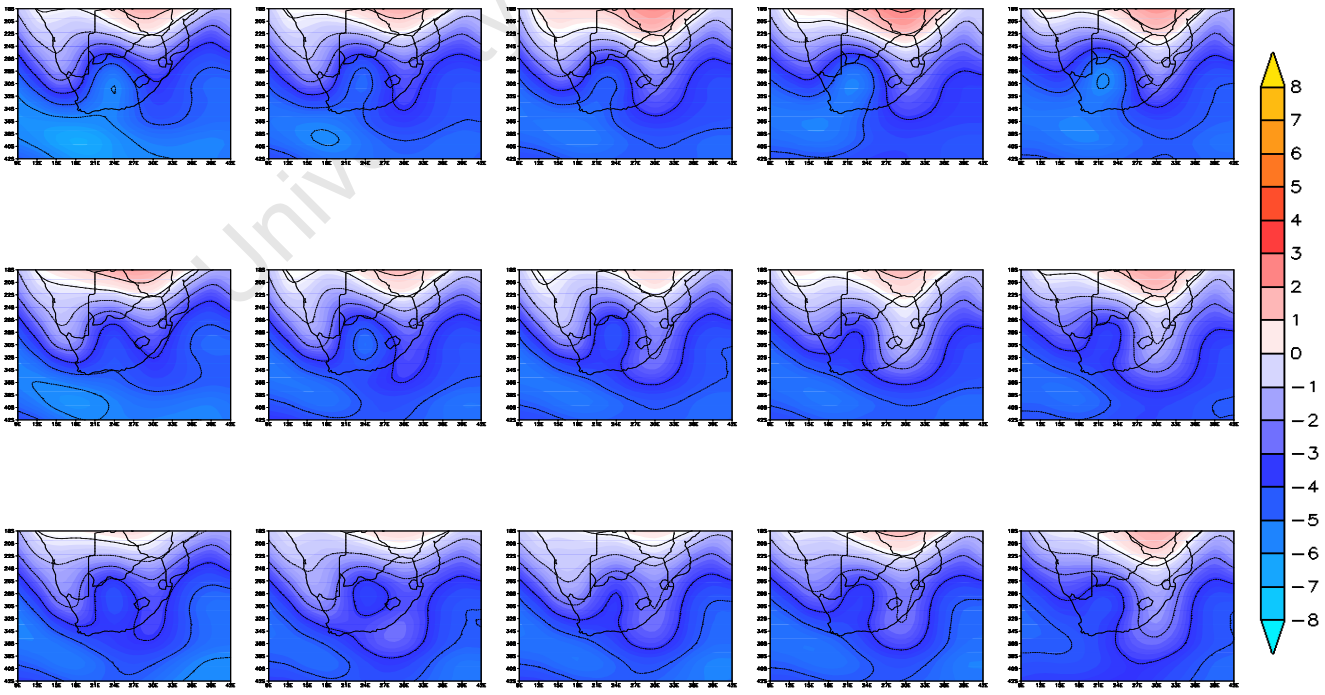


Figure D2: January node weighted multi model – NCEP temperature contours at 1°C interval

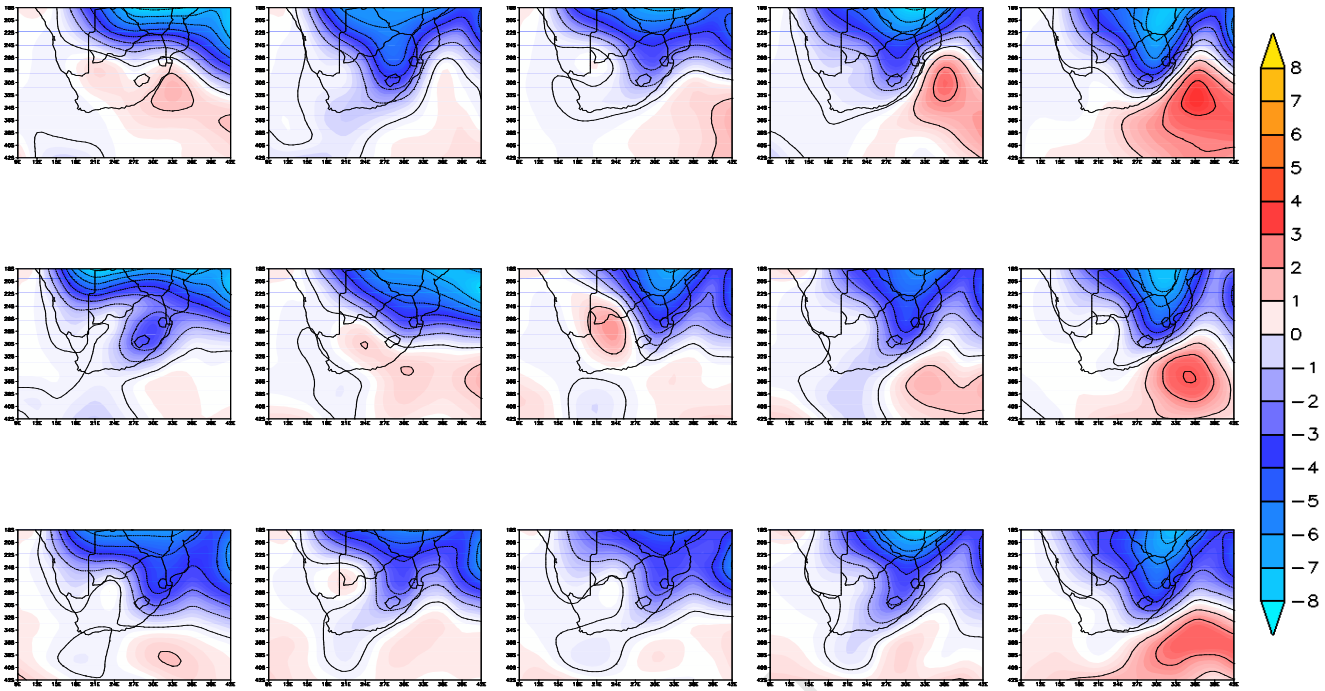


Figure D3: January node weighted multi model – NCEP precipitation contours at 1mm interval

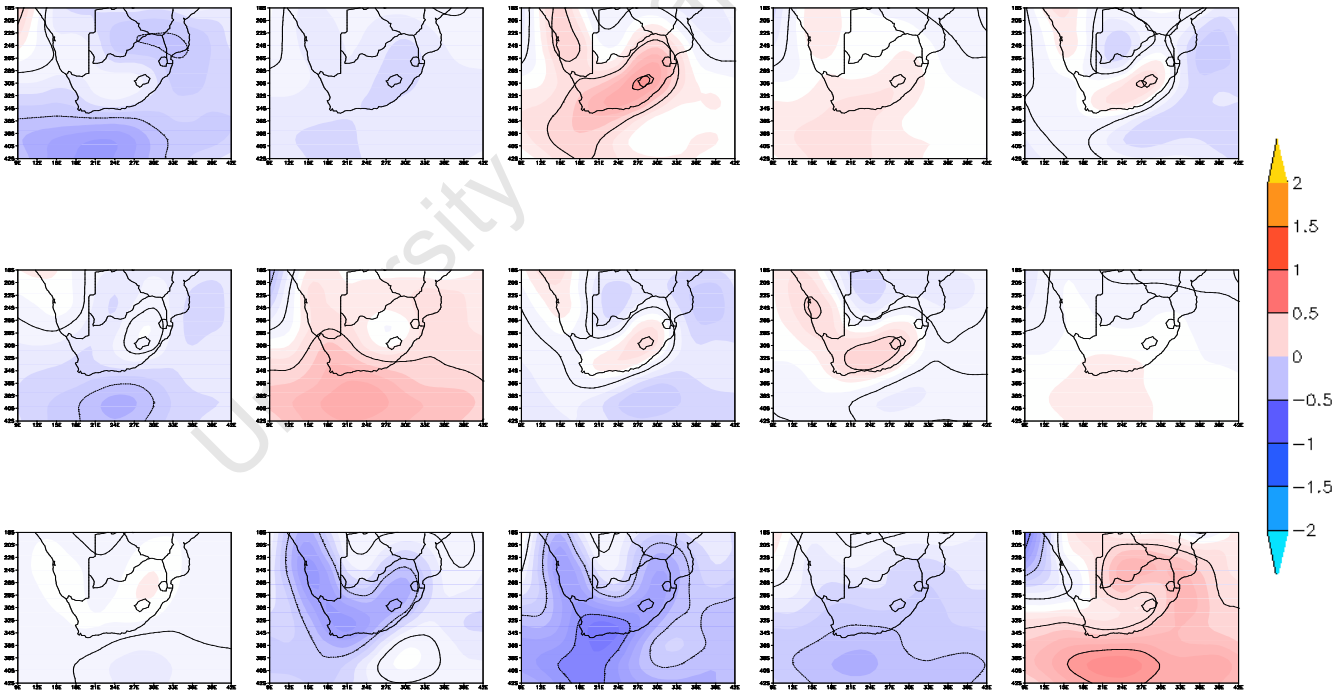


Figure D4: January node weighted multi model – multi model average temperature contours at 0.2C interval

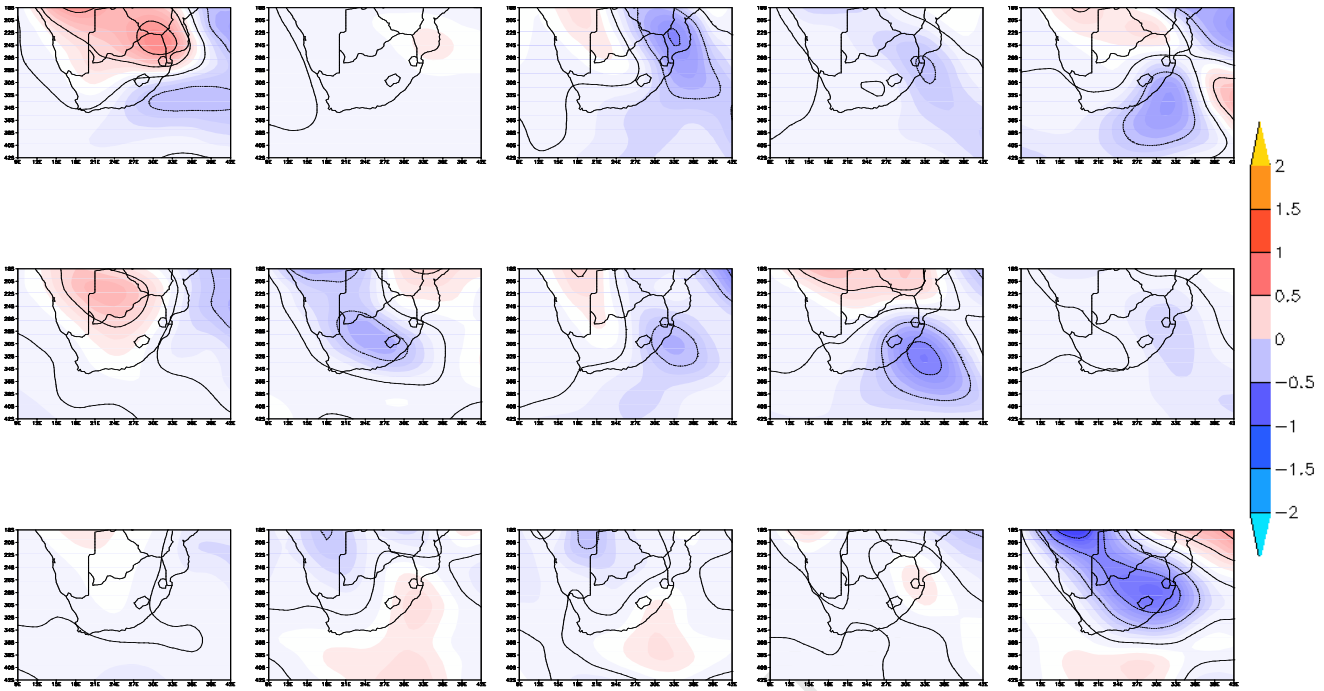


Figure D5: Anomaly from NCEP January node weighted multi model – multi model average precipitation contours at 0.2mm interval

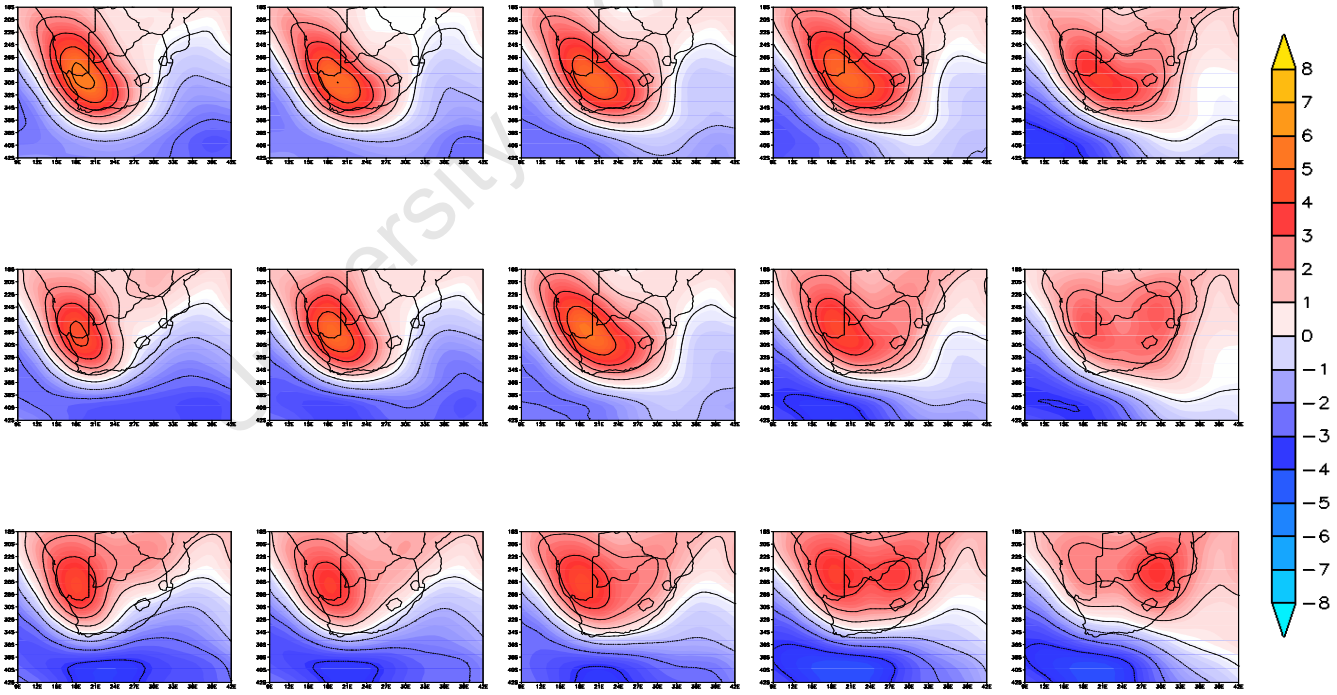


Figure D6: April node weighted multi model – NCEP temperature contours at 1°C interval

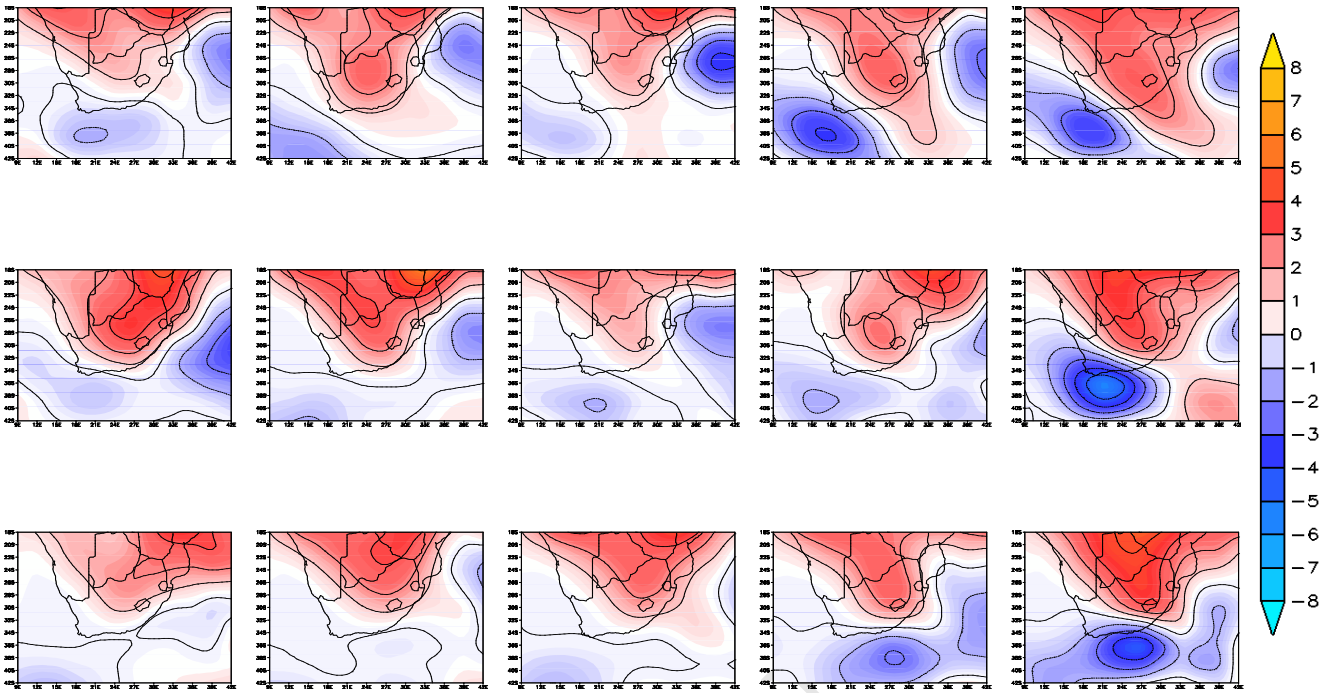


Figure D7: April node weighted multi model – NCEP precipitation contours at 1mm interval

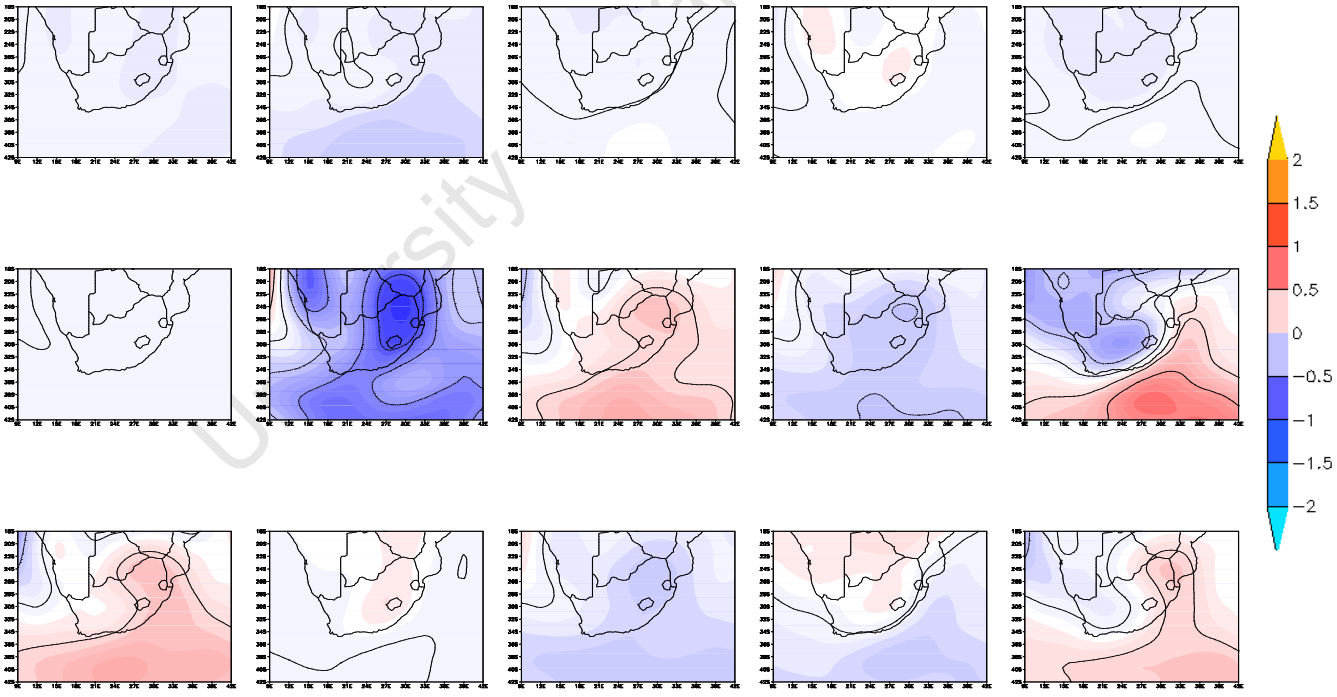


Figure D8: April node weighted multi model – multi model average temperature contours at 0.2°C interval

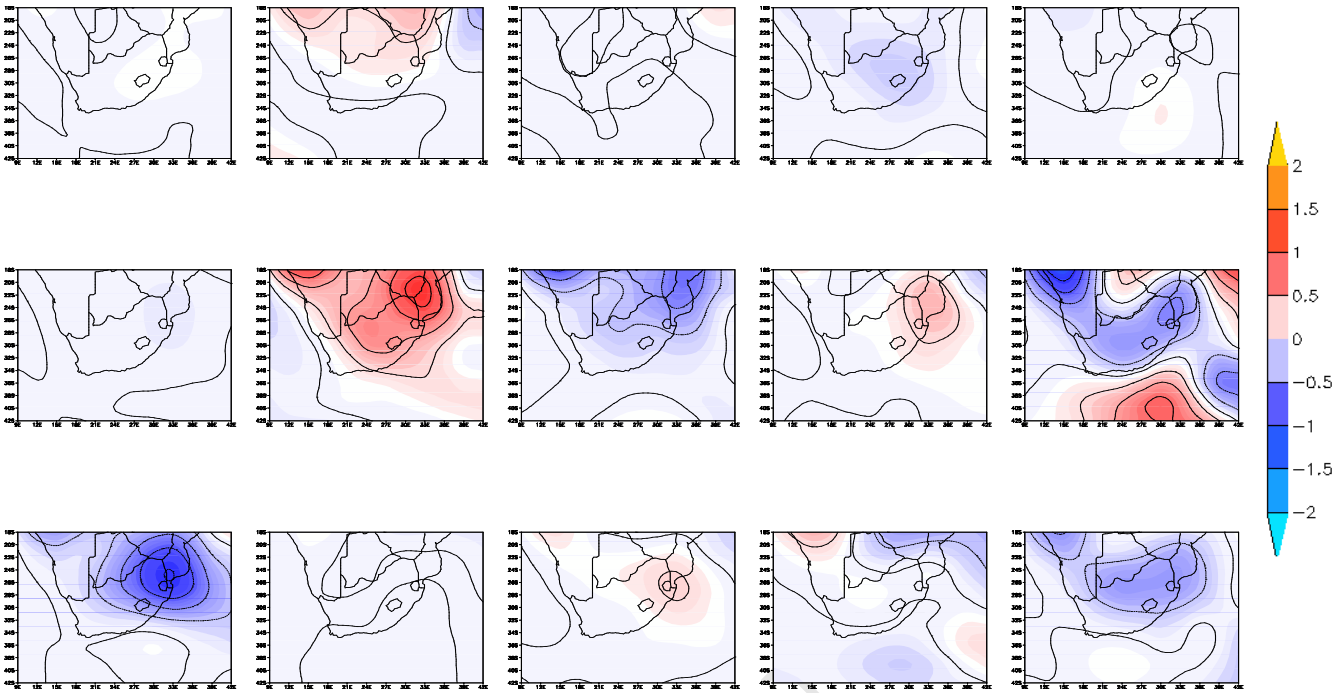


Figure D9: April node weighted multi model – multi model average precipitation contours at 0.2mm interval

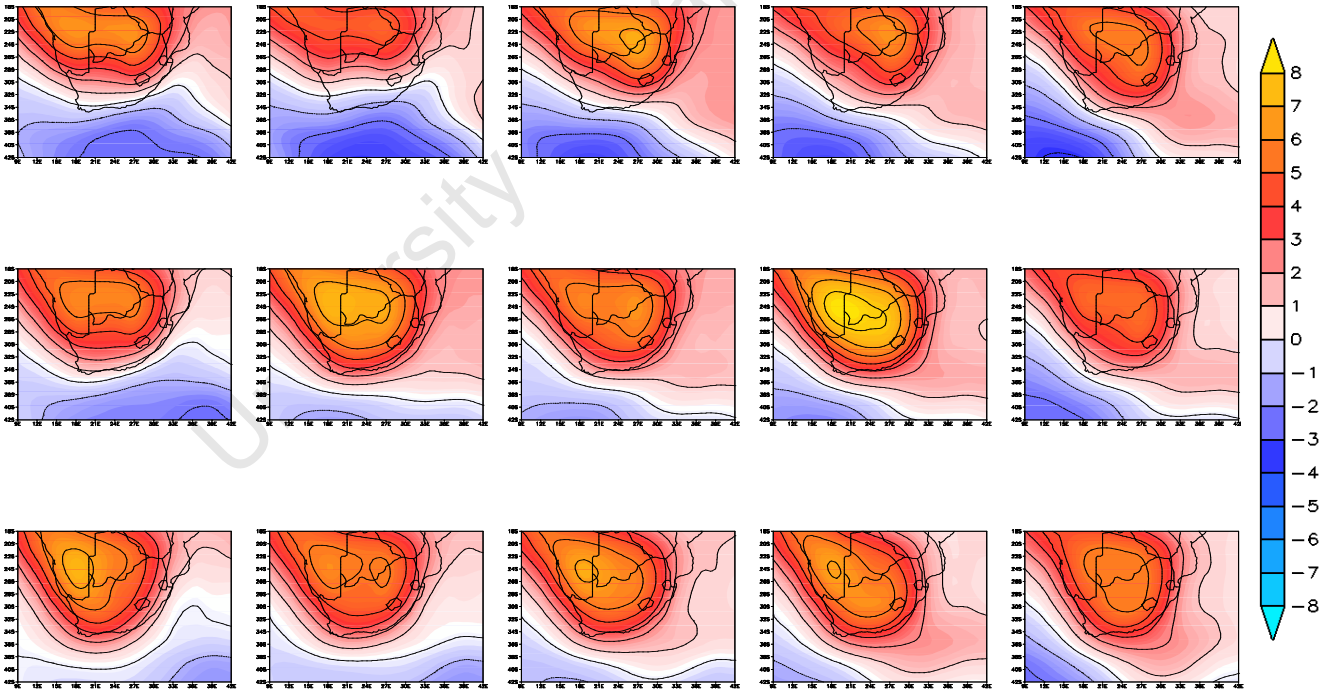


Figure D10: July node weighted multi model – NCEP temperature contours at 1°C interval

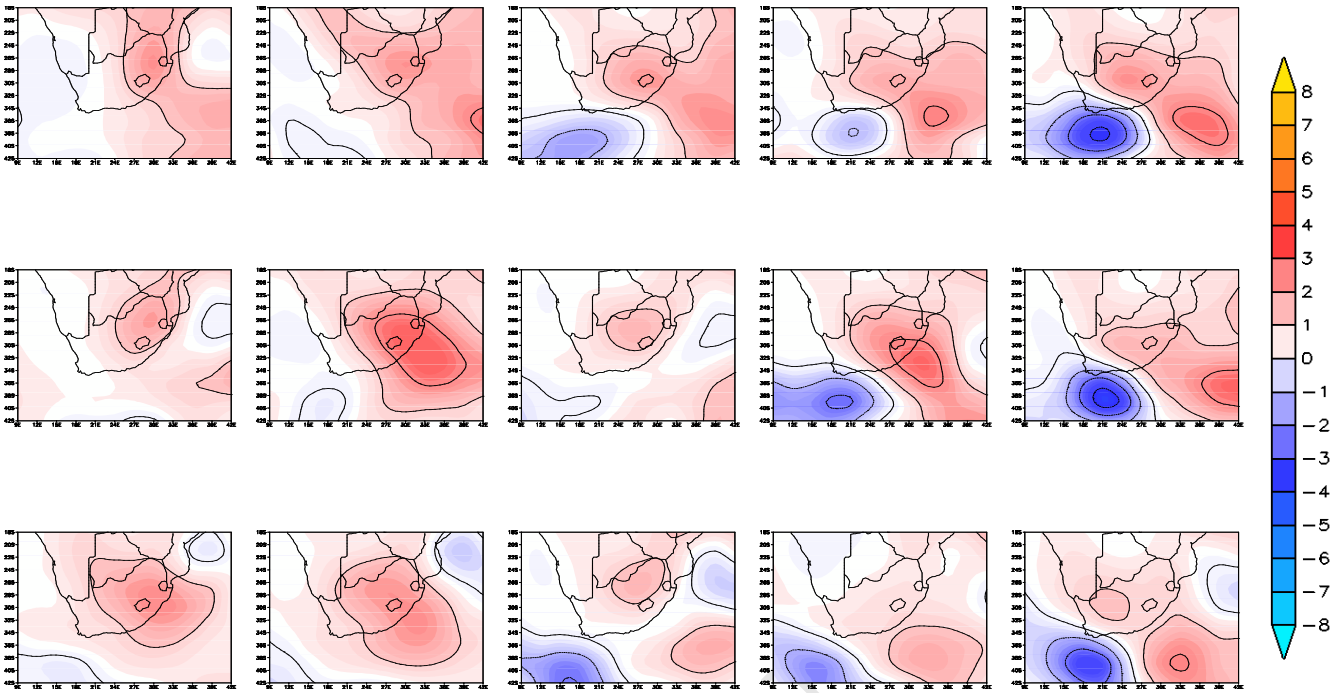


Figure D11: July node weighted multi model – NCEP precipitation contours at 1mm interval

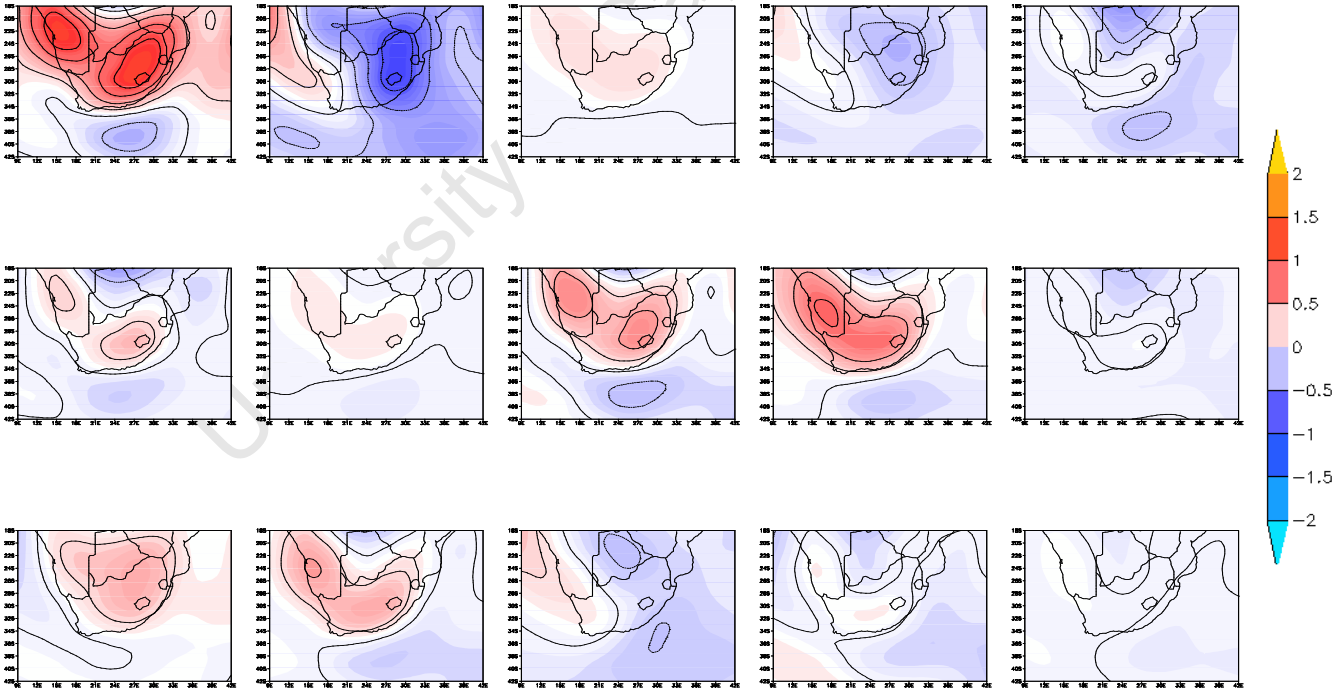


Figure D12: July node weighted multi model – multi model average temperature contours at 0.2°C interval

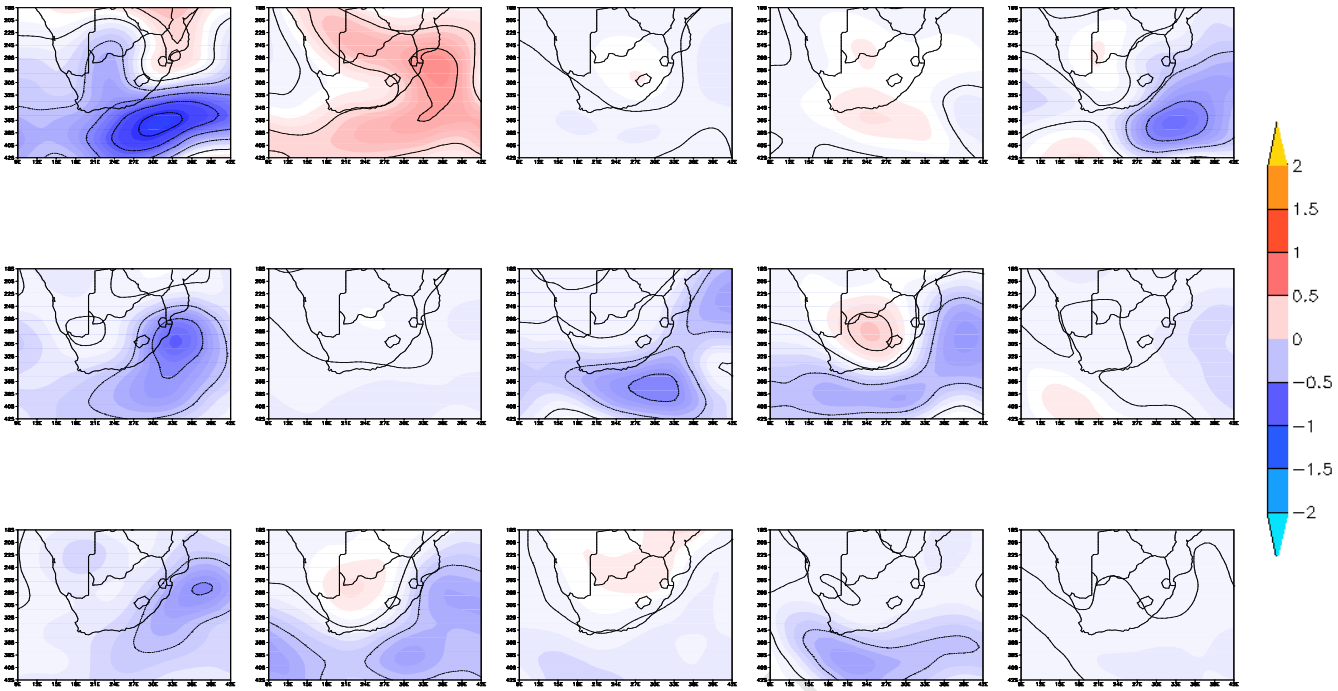


Figure D13: July node weighted multi model – multi model average precipitation contours at 0.2mm interval

a)

0.65	0.66	0.78	0.87	0.83
0.58	0.60	0.80	0.87	0.75
0.98	0.79	0.96	0.91	0.56

b)

0.56	0.60	0.69	0.80	0.51
0.79	0.69	0.30	0.44	0.52
0.21	0.36	0.16	0.51	0.71

c)

0.79	0.74	0.54	0.33	0.66
0.62	0.71	0.90	0.69	0.73
0.81	0.85	0.88	0.58	0.73

Figure D14: Node weights for October data: a) GFDL, b) MIROC, c) MRI

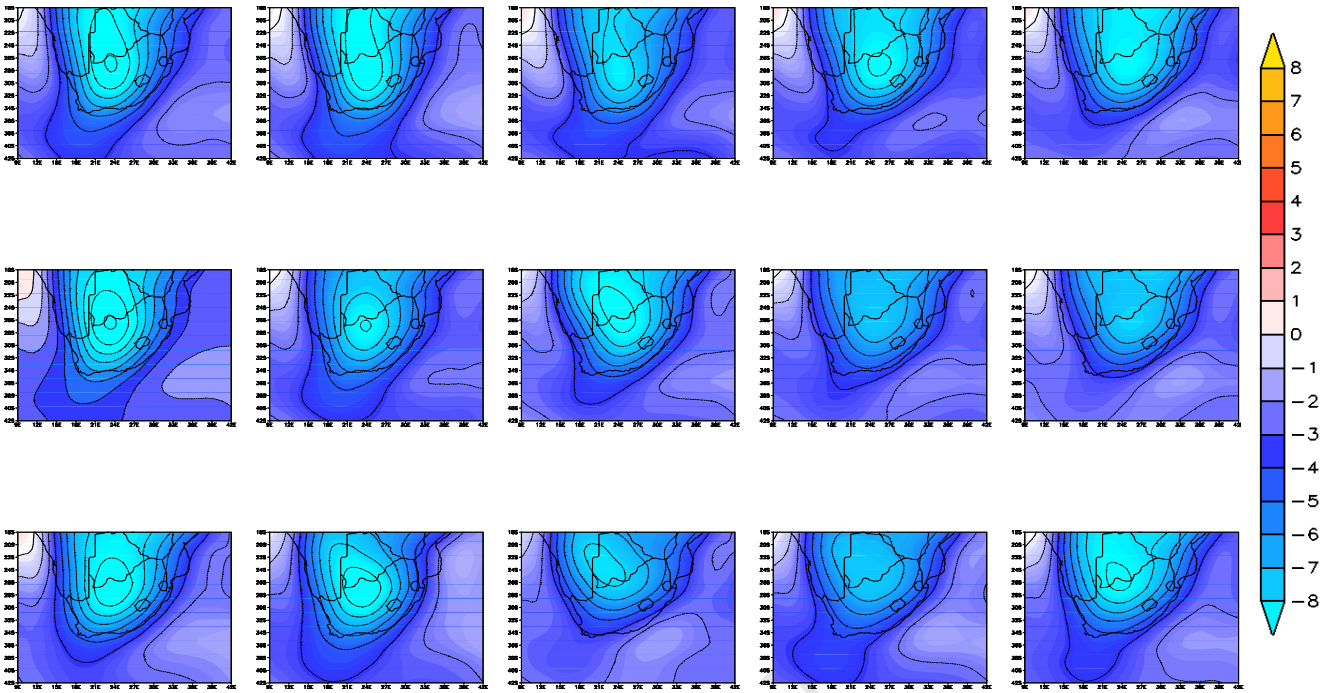


Figure D15: October node weighted multi model – NCEP temperature contours at 1°C interval

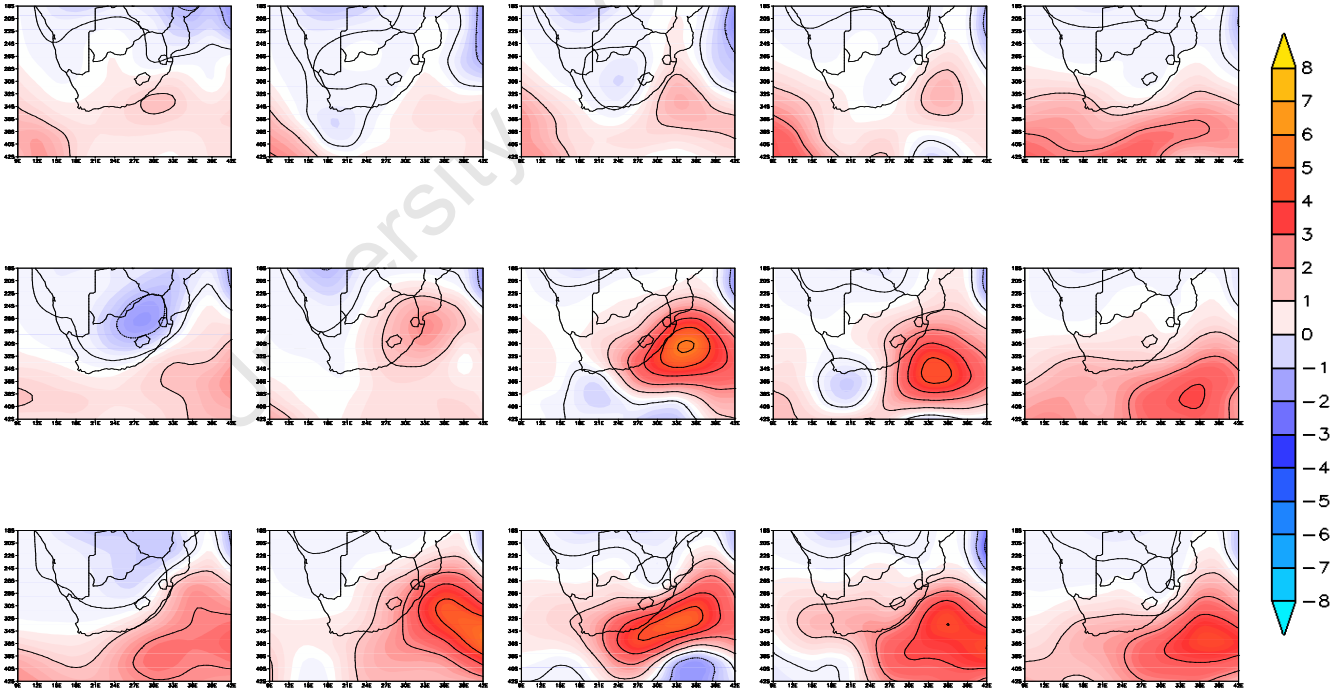


Figure D16: October node weighted multi model – NCEP precipitation contours at 1mm interval

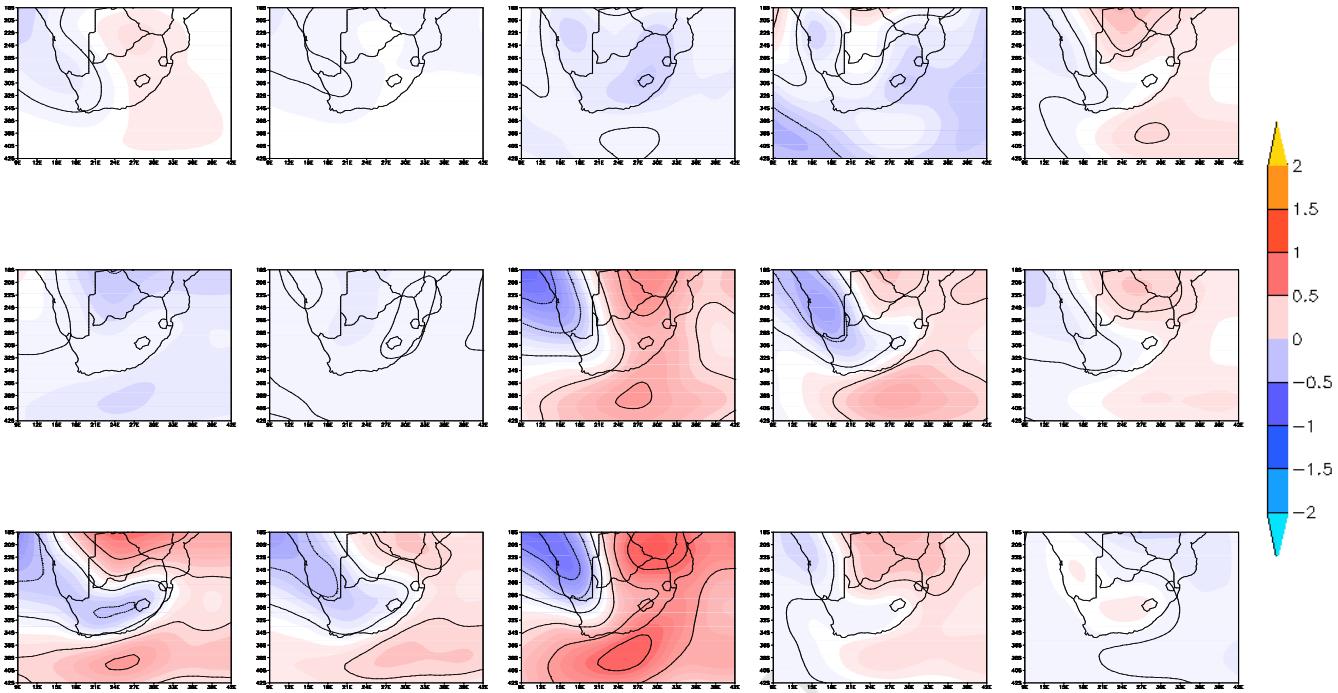


Figure D17: October node weighted multi model – multi model average temperature contours at 0.2°C

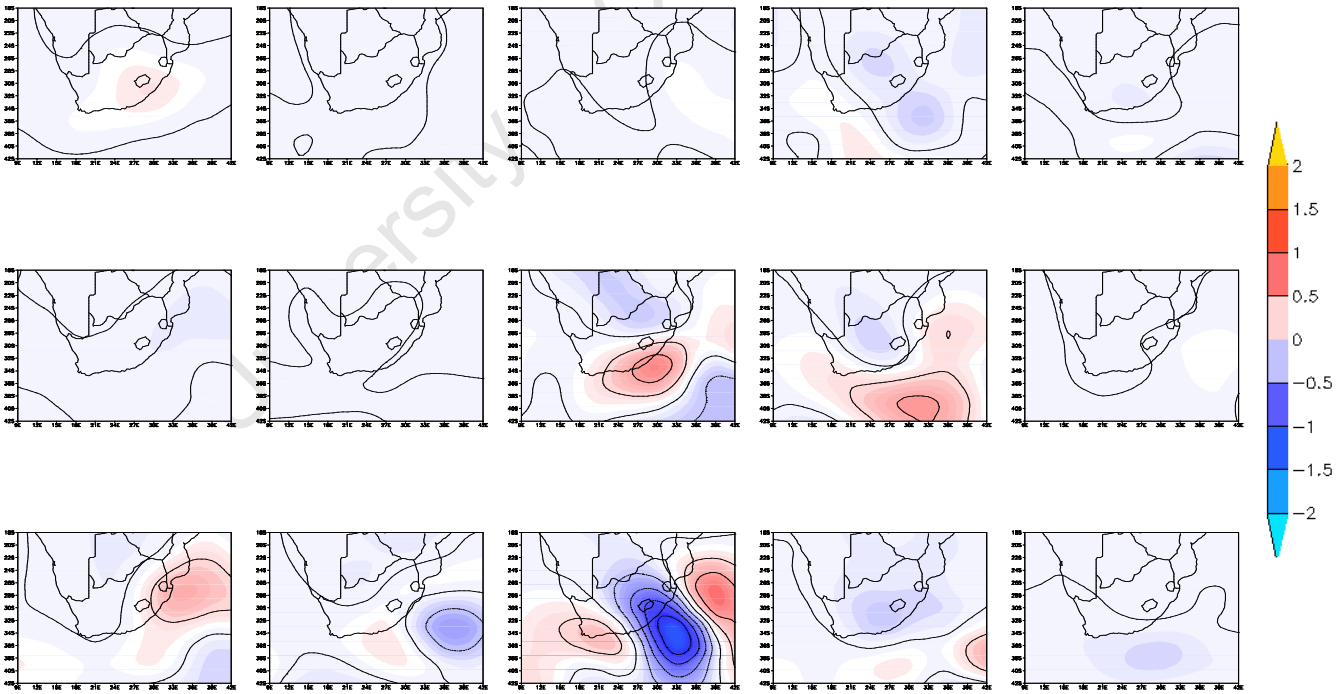


Figure D18: October node weighted multi model – multi model average precipitation contours at 0.2mm interval

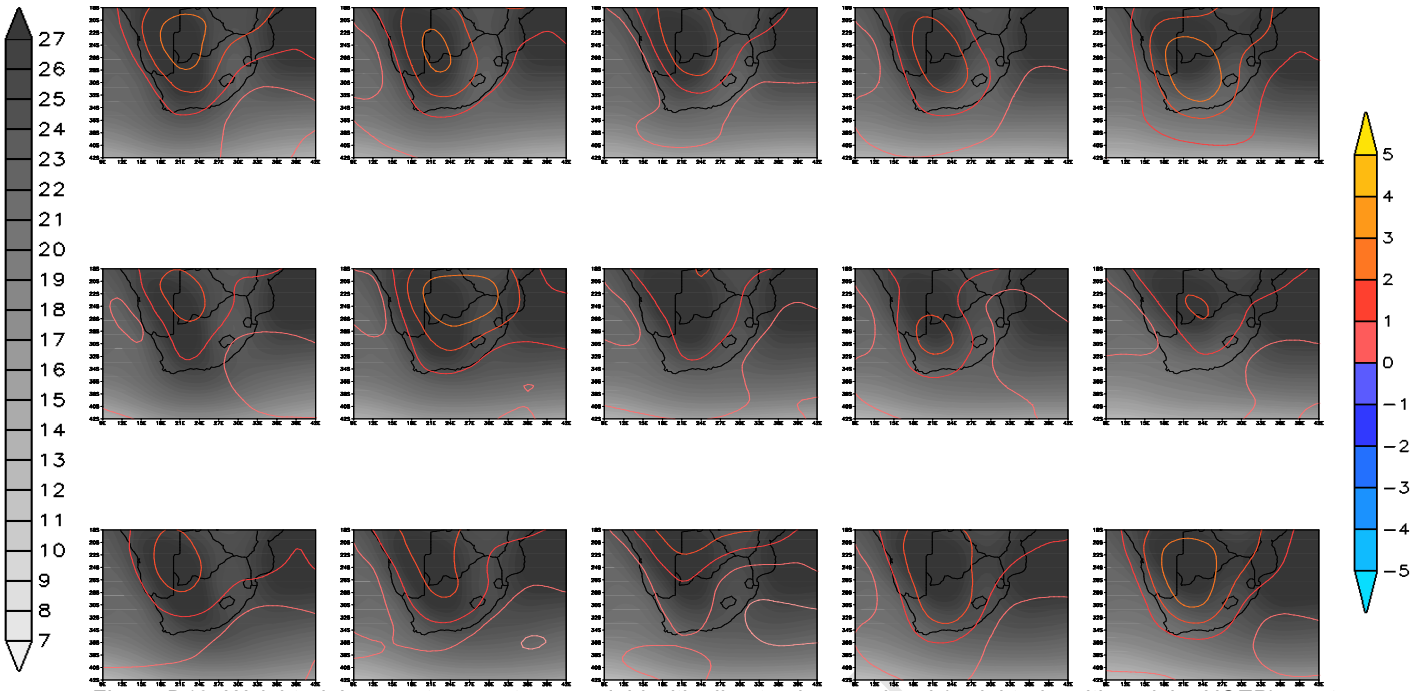


Figure D19: Weighted January temperature overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5°C

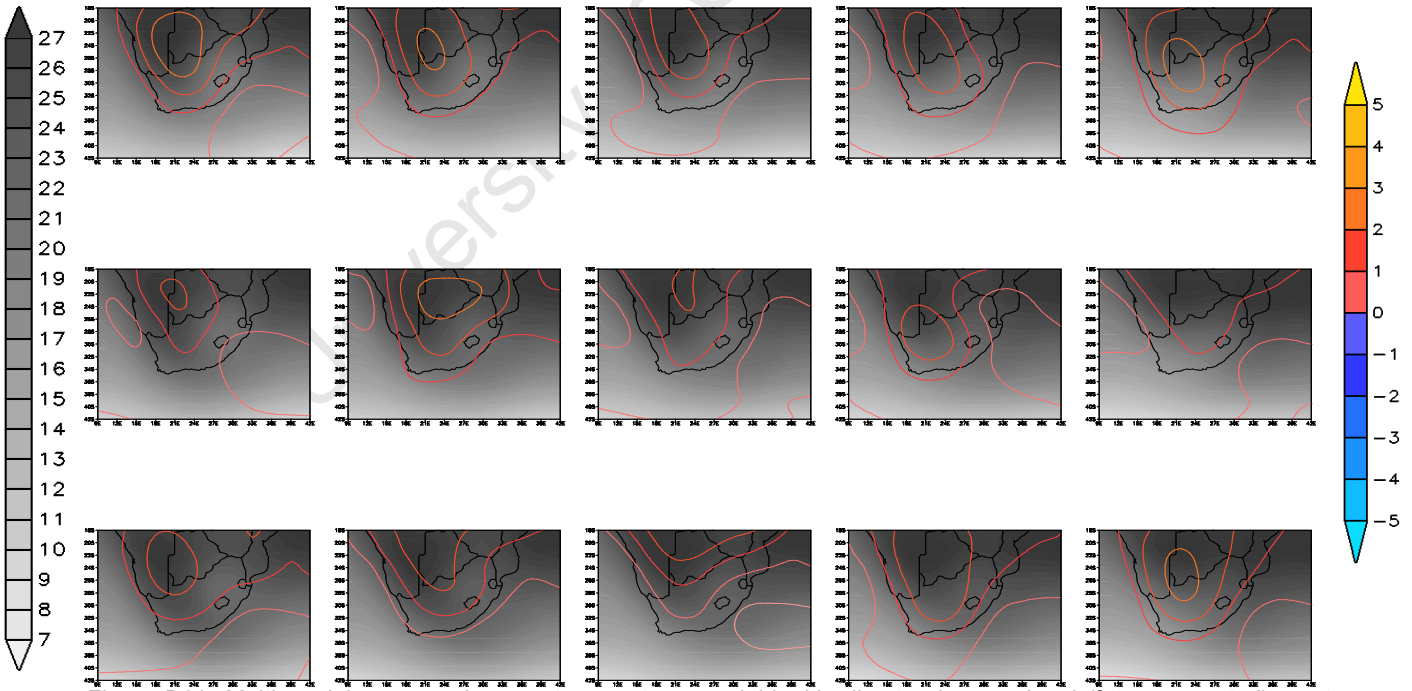


Figure D20: Multi model average January temperature overlaid with climate change signal (future – control), contour interval 0.5°C

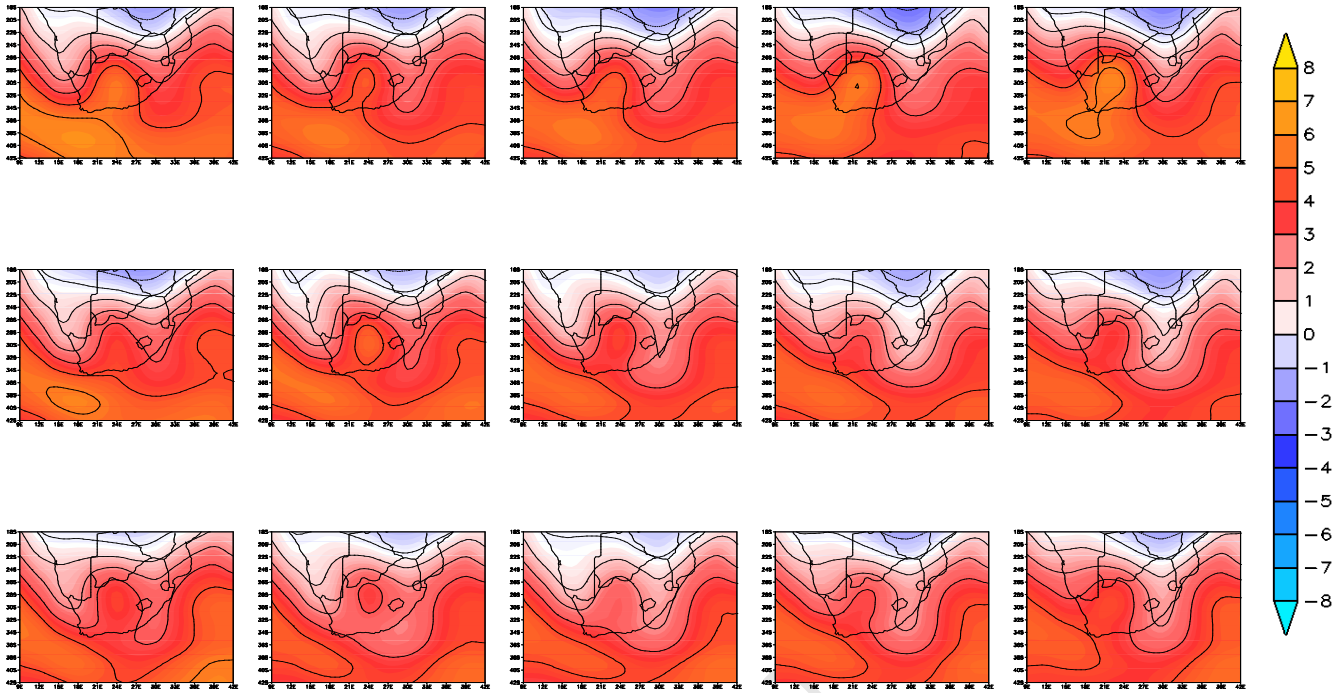


Figure D21: January temperature weighted future projections average of original all model projections overlaid with the anomaly between future and control, contour interval 1°C

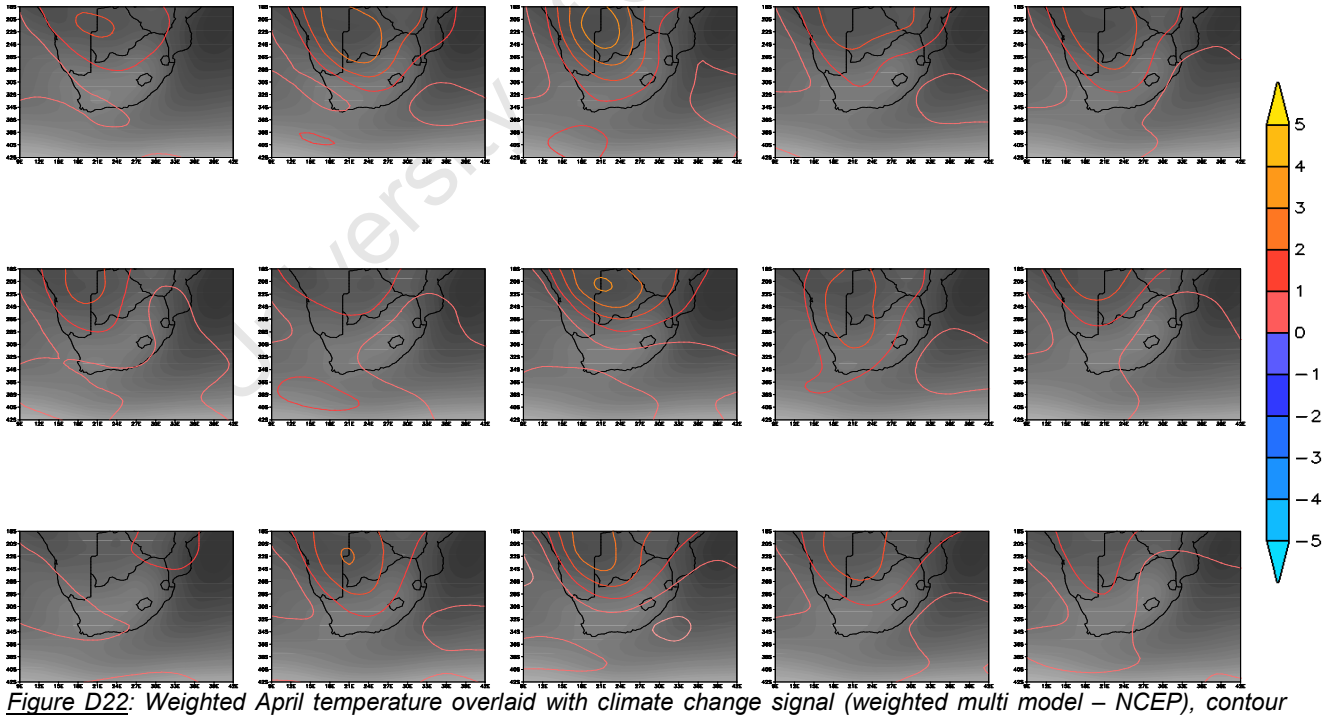


Figure D22: Weighted April temperature overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5°C

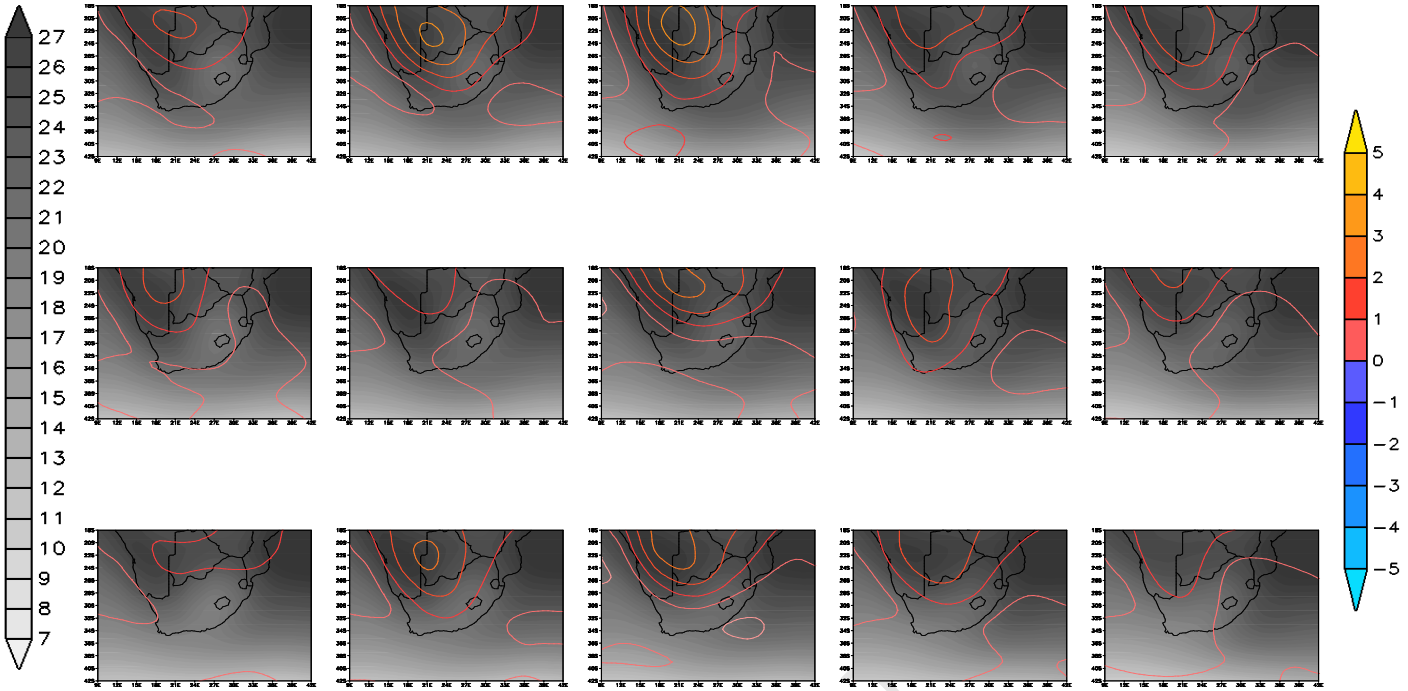


Figure D23: Multi model average April temperature overlaid with climate change signal (future – control), contour interval 0.5°C

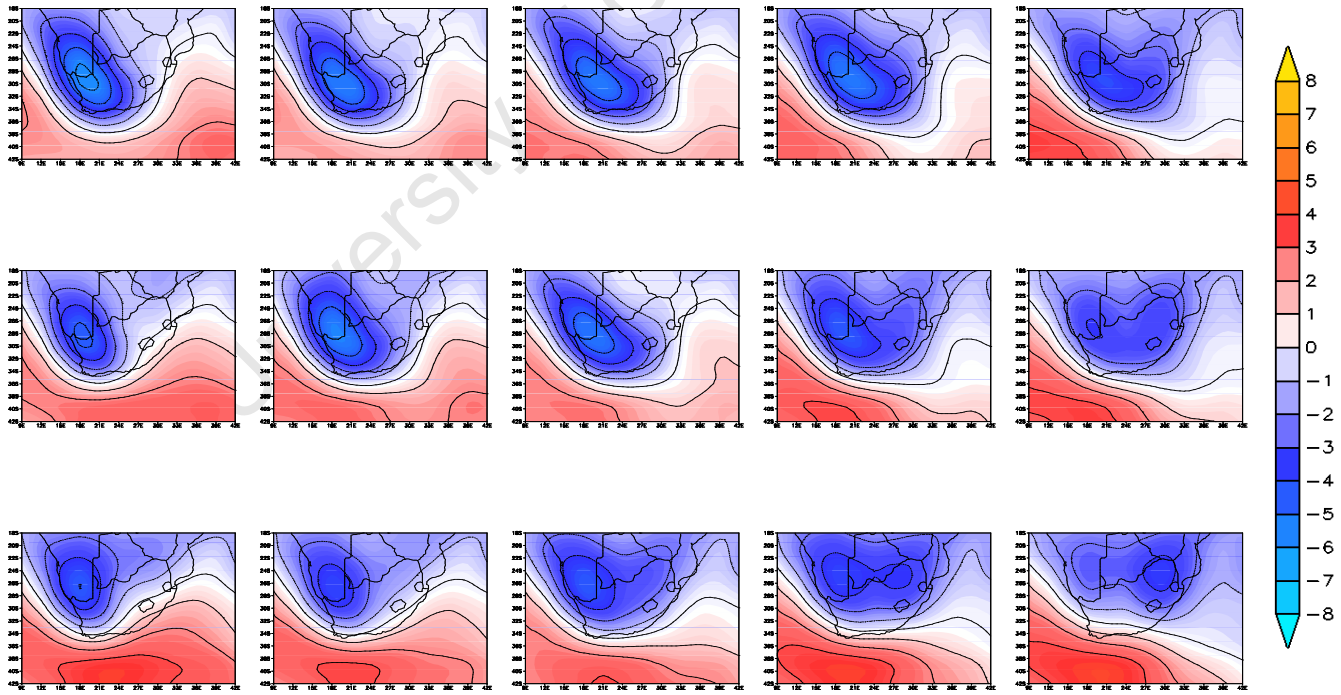


Figure D24: April temperature weighted future projections average of original all model projections overlaid with the anomaly between future and control, contour interval 1°C

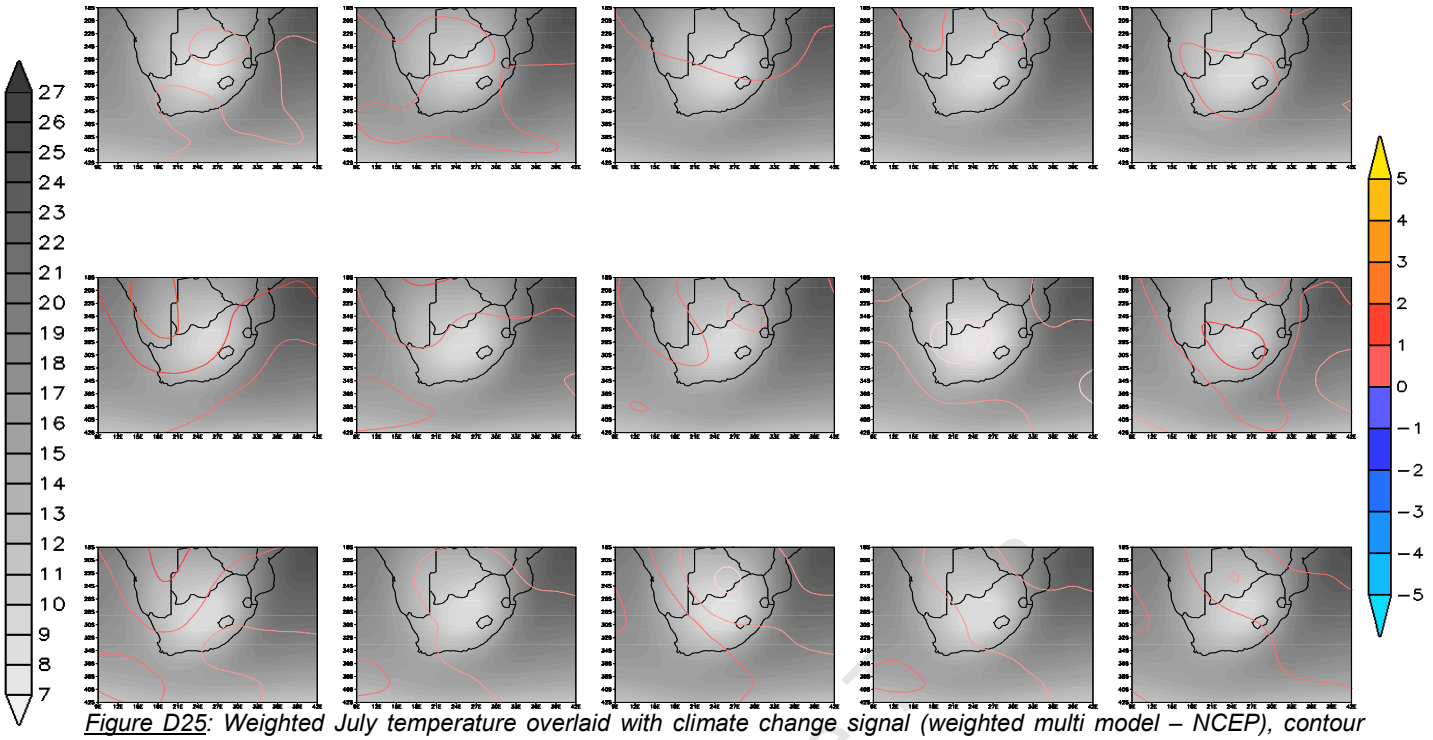


Figure D25: Weighted July temperature overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5°C

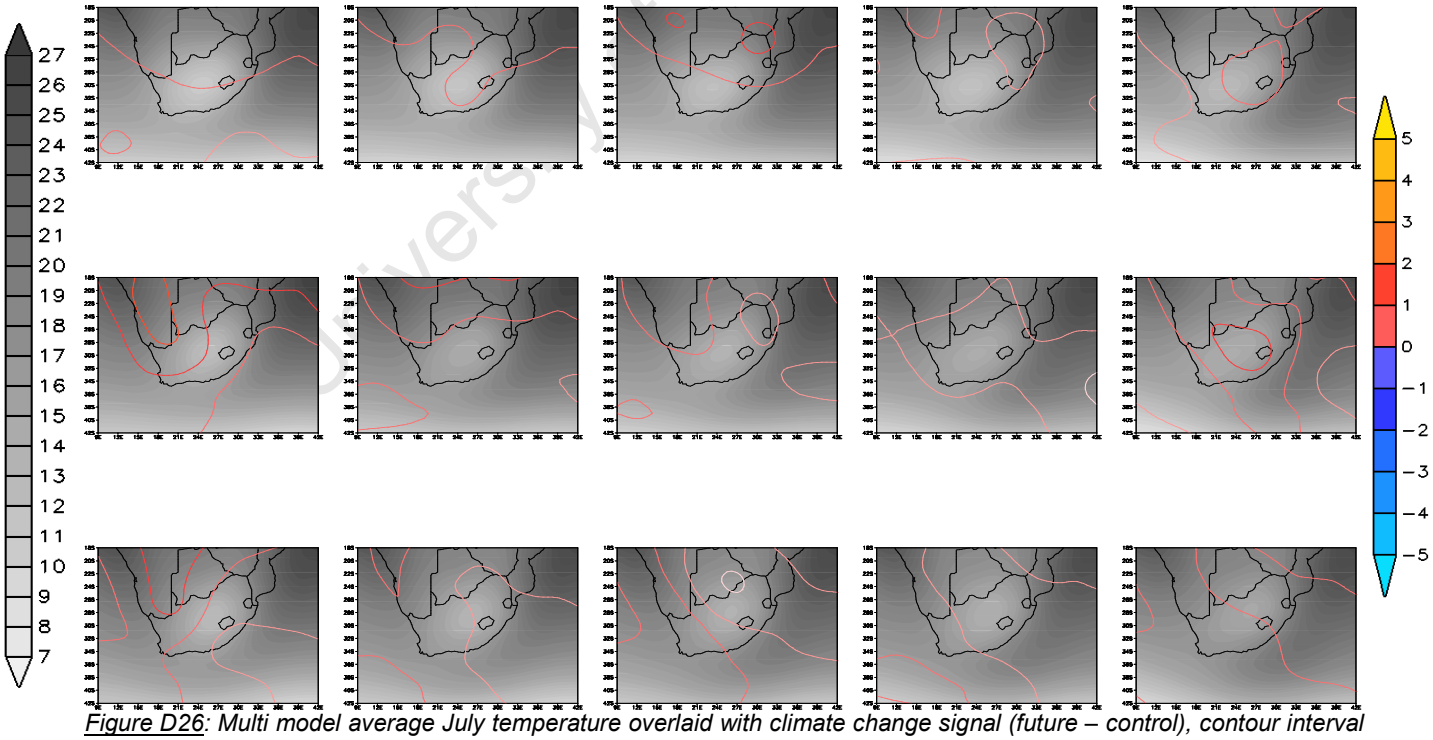


Figure D26: Multi model average July temperature overlaid with climate change signal (future – control), contour interval 0.5°C

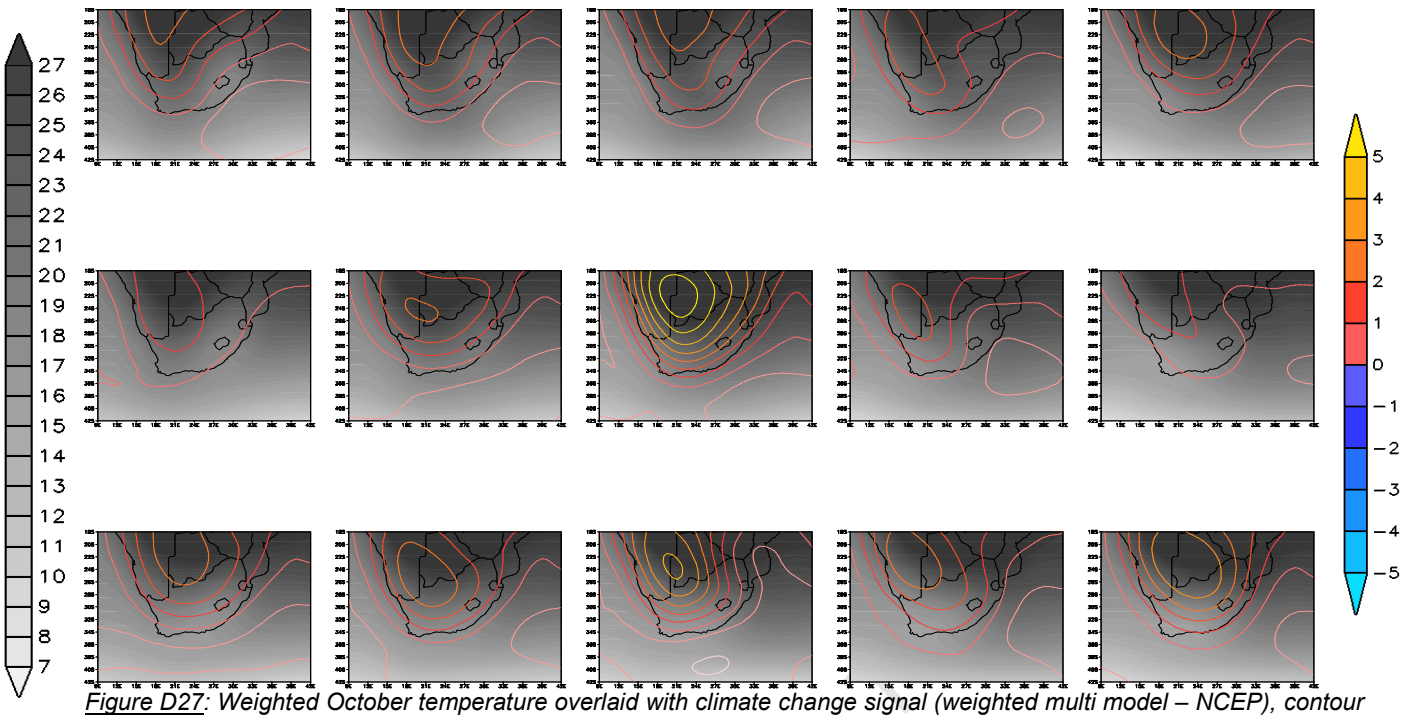


Figure D27: Weighted October temperature overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5°C

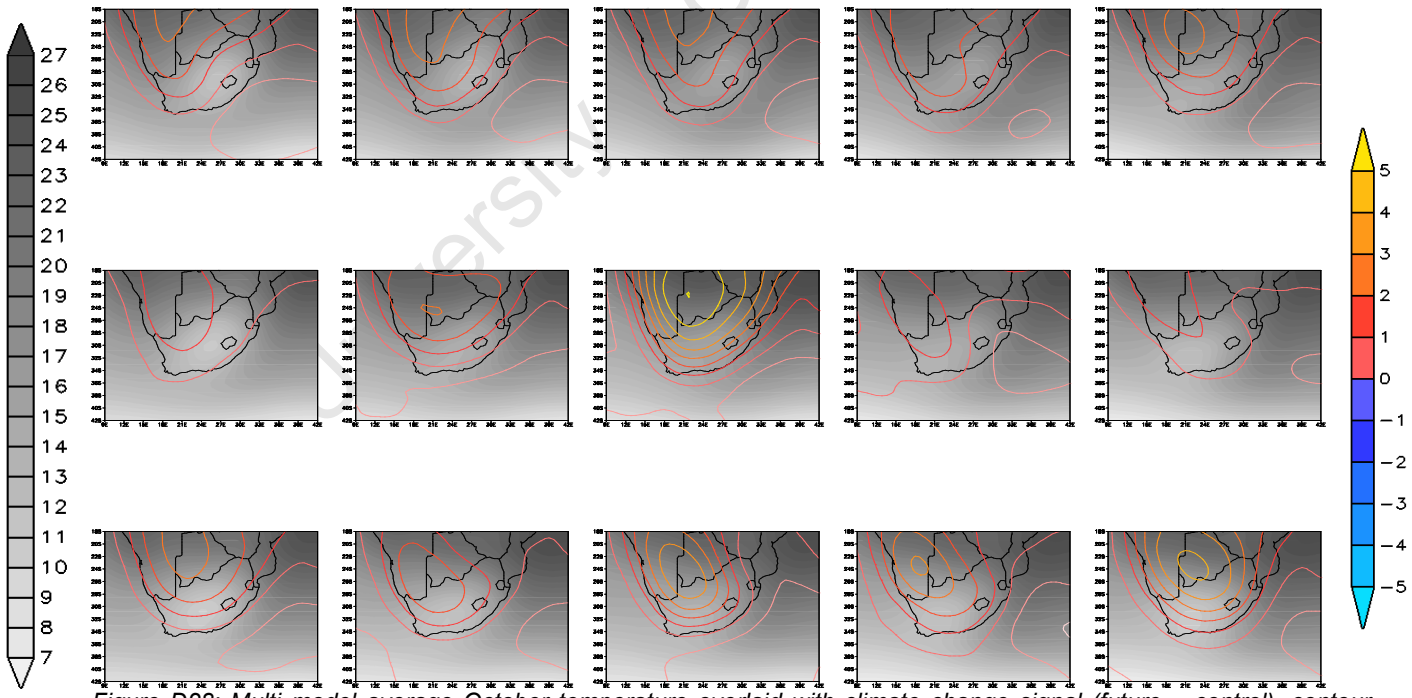


Figure D28: Multi model average October temperature overlaid with climate change signal (future – control), contour interval 0.5°C

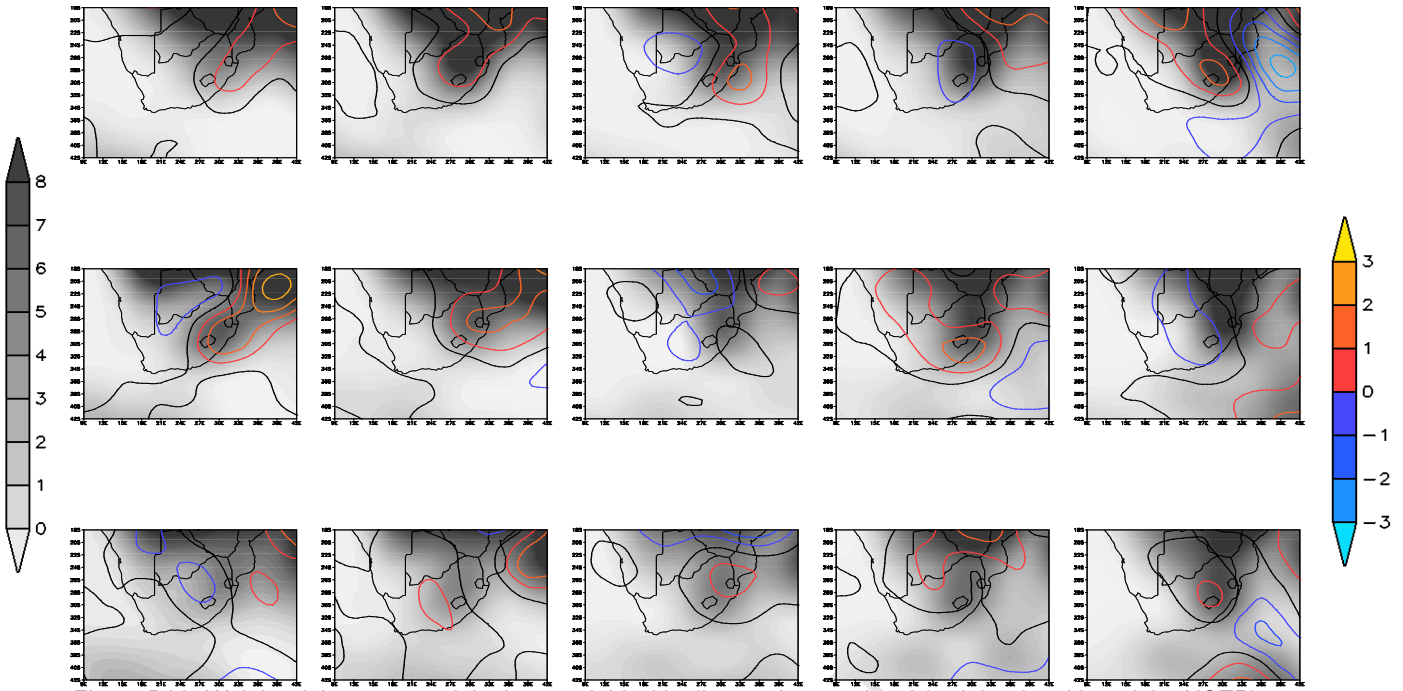


Figure D29: Weighted January precipitation overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5mm

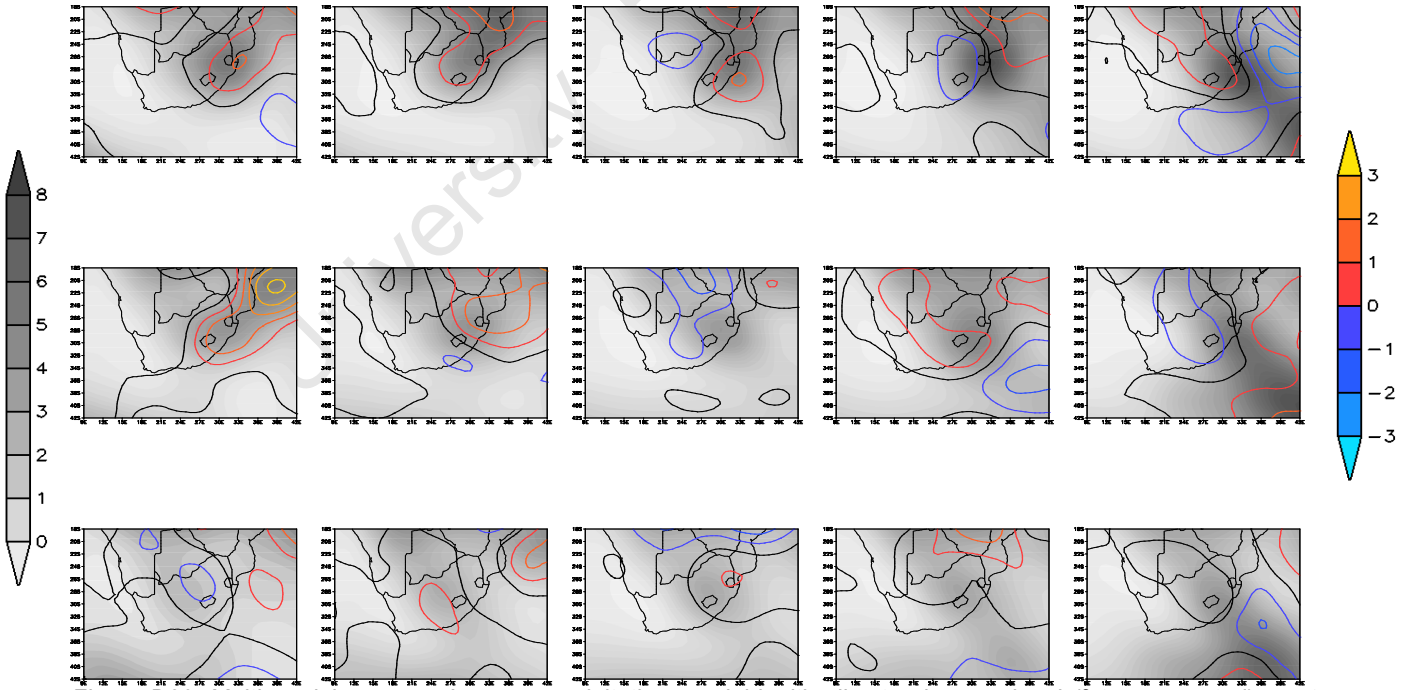


Figure D30: Multi model average January precipitation overlaid with climate change signal (future – control), contour interval 0.5mm

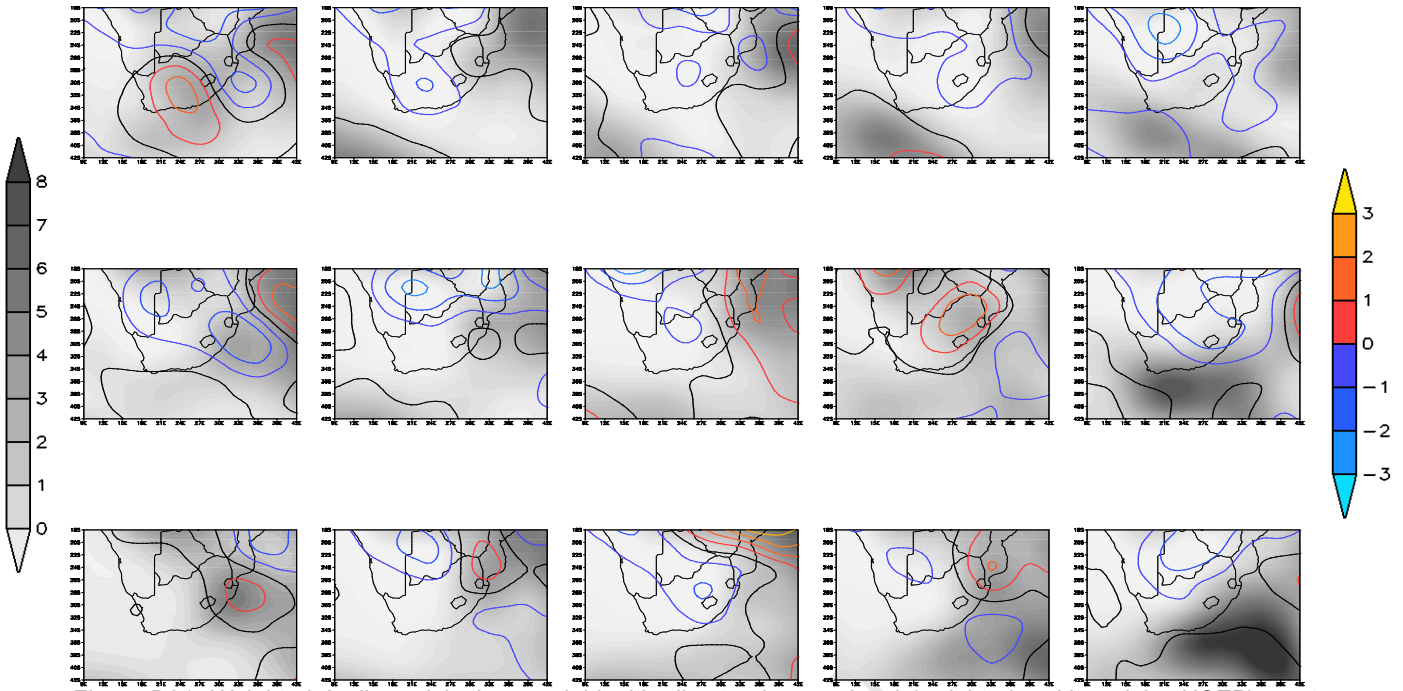


Figure D31: Weighted April precipitation overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5mm

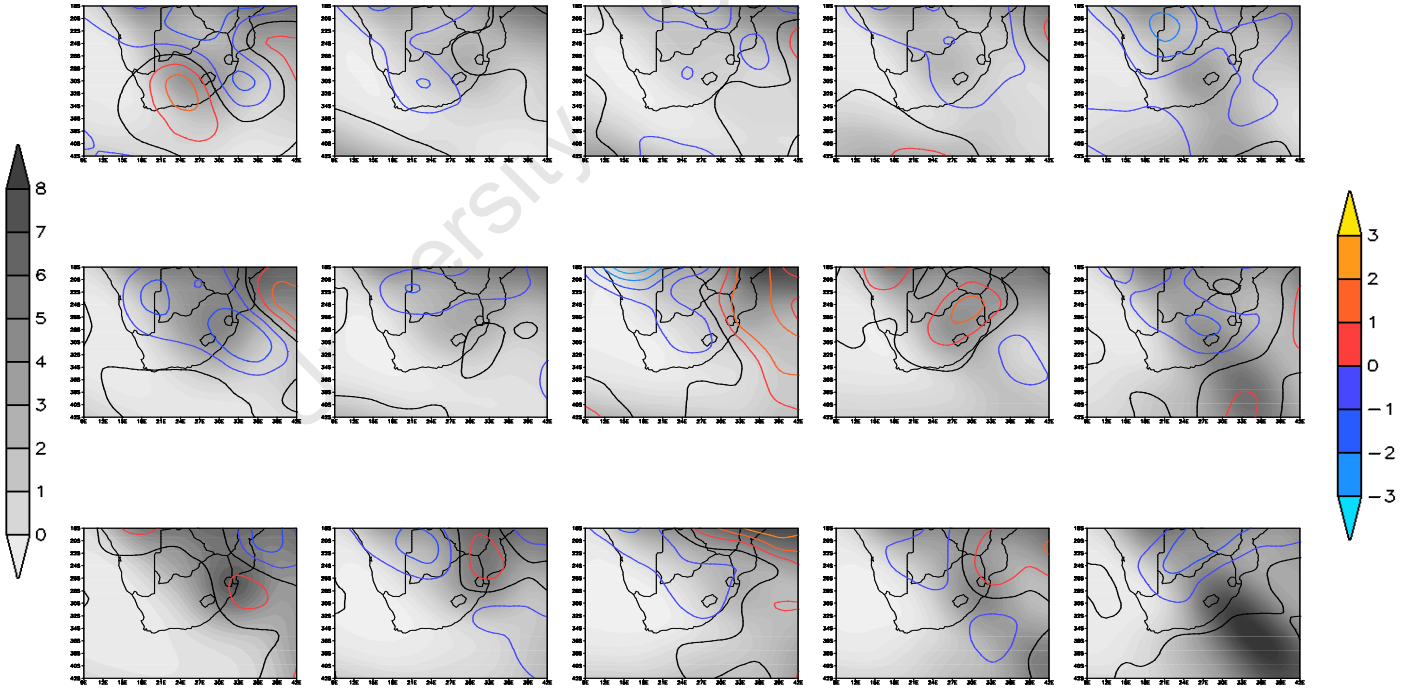


Figure D32: Multi model average April precipitation overlaid with climate change signal (future – control), contour interval 0.5mm

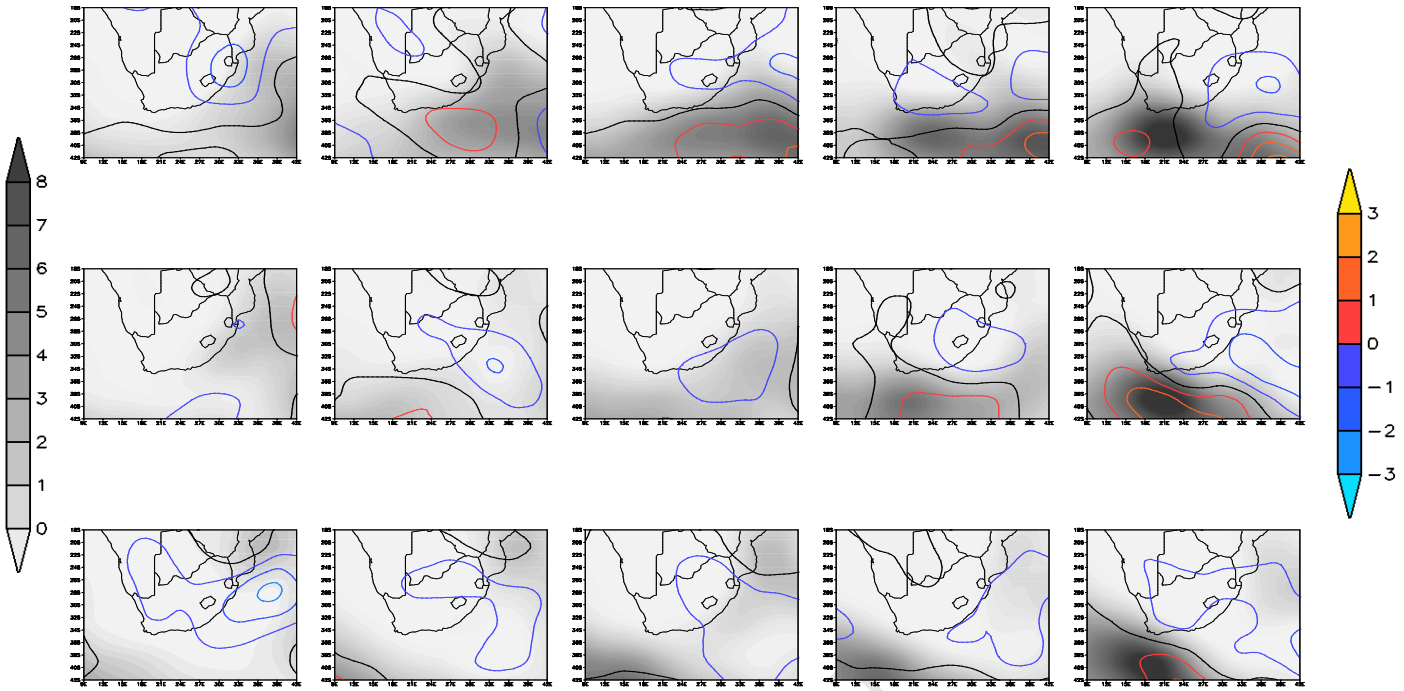


Figure D33: Weighted July precipitation overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5mm

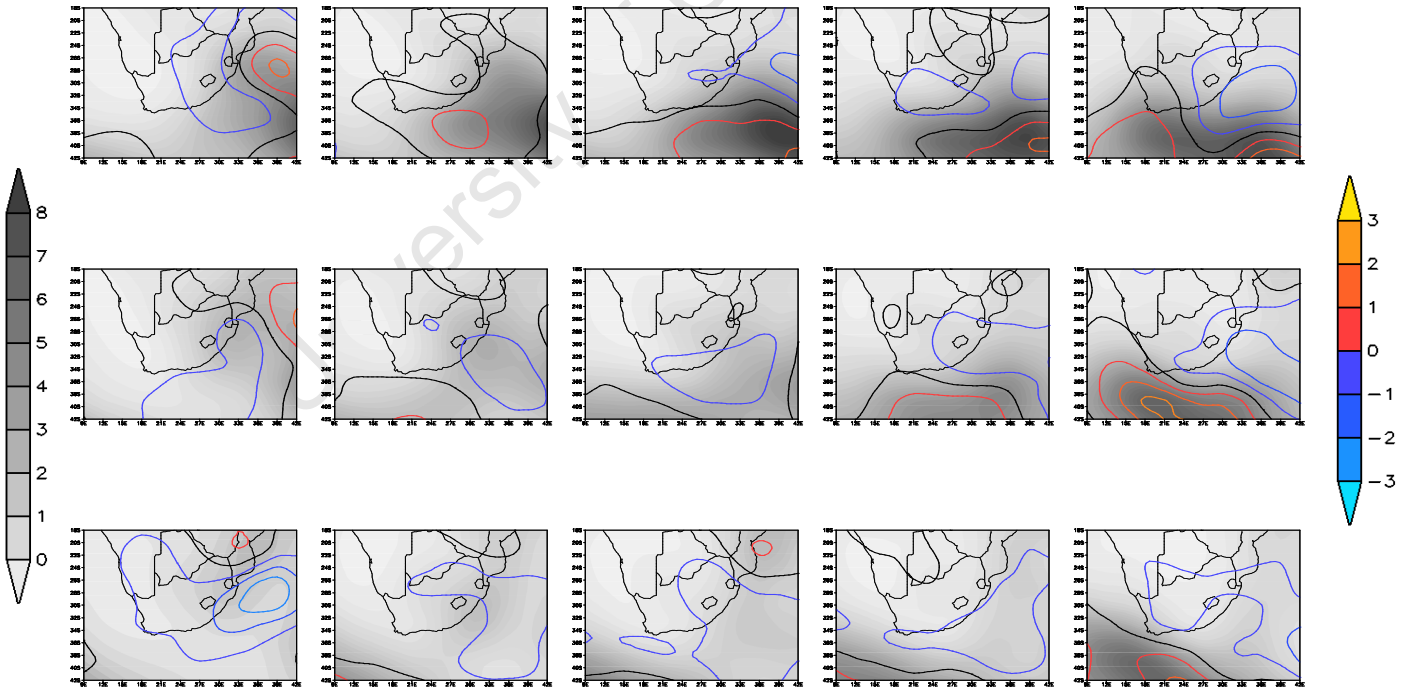


Figure D34: Multi model average July precipitation overlaid with climate change signal (future – control), contour interval 0.5mm

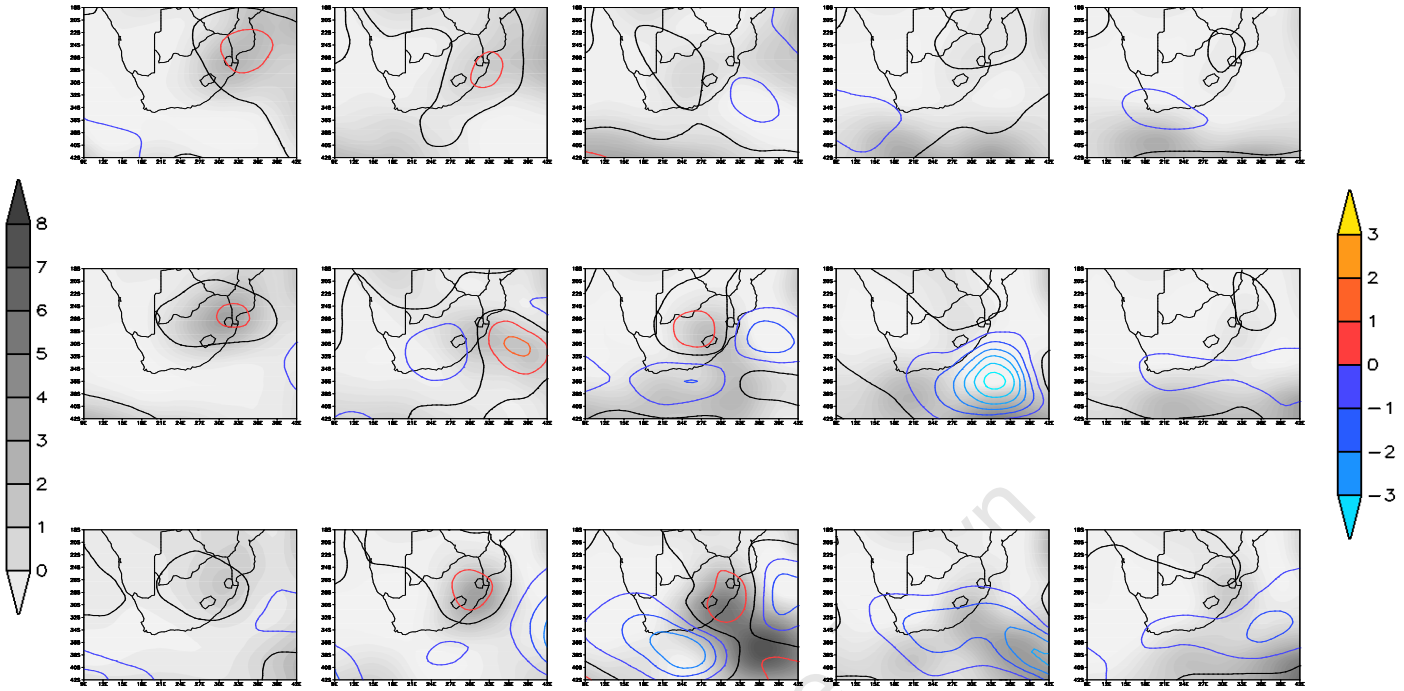


Figure D35: Weighted October precipitation overlaid with climate change signal (weighted multi model – NCEP), contour interval 0.5mm

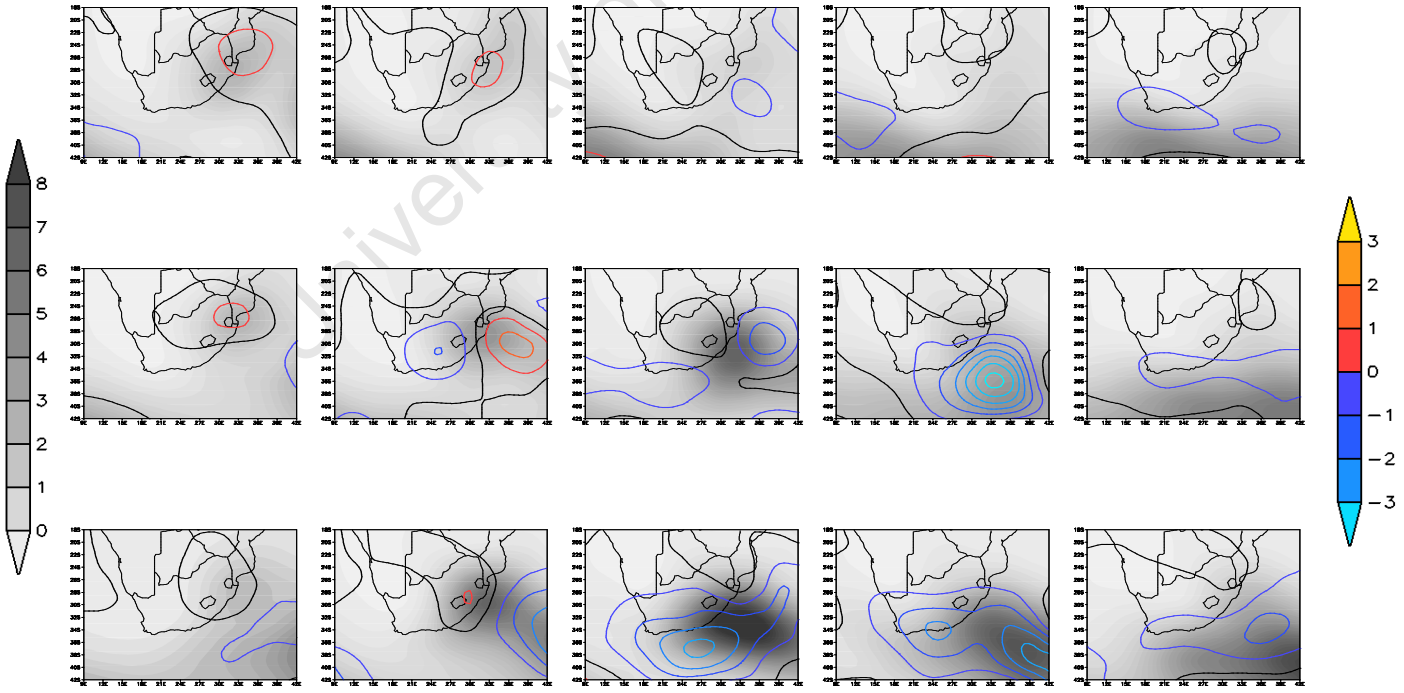


Figure D36: Multi model average October precipitation overlaid with climate change signal (future – control), contour interval 0.5mm