Investigating Projected Changes in the Nature of Extreme Rainfall Over South Africa During the 21st Century Through an Assessment of Their Associated Synoptic Environments

Report to the
Water Research Commission

by

CJ Lennard, I Pinto & M Nyak
Climate System Analysis Group
University of Cape Town
South Africa

WRC Report No. 2240/1/15
ISBN 978-1-4312-0757-2

March 2016

Obtainable from
Water research Commission
Private Bag X03
Gezina, 0031

orders@wrc.org.za or download from www.wrc.org.za

Executive Summary

Extreme rainfall events are often associated with significant societal and infrastructural impacts. An Intergovernmental Panel on Climate Change special report states 'It is likely that the frequency of heavy precipitation or the proportion of total rainfall from heavy falls will increase in the 21st century over many areas of the globe.' Climate models, that simulate the global and regional climate, provide data to assess historical and projected changes in rainfall and extreme rainfall and their driving synoptic environments. However, rainfall is a difficult variable to simulate correctly. It is a diagnostic variable that is dependent on a number of other prognostic variables (such as pressure, temperature and humidity) as well as model parameterization schemes – small errors in the prognostic variables have a large effect on simulated rainfall. In this study, we quantify the prognostic synoptic environments associated with extreme rainfall and their projected changes as a result of greenhouse gas warming. We suggest this is a more defensible methodology to assess and understand potential changes in the characteristics of extreme rainfall.

Extreme rainfall in South Africa is associated with particular synoptic environments and these have been characterized here. We find cut-off lows associated with extreme rainfall are deeper and westerly-wave driven in austral winter whereas in summer they are shallower and warmer suggesting a greater convective influence. Similarly, mid-latitude cyclones are deeper in the core winter months (June-July-August) and shallower in spring and autumn months. Extreme rainfall associated with tropical temperate troughs display clear intra-seasonal characteristics during summer where a much higher number occur in late summer (January-February-March) than in the early summer (October-November-December). Seven homogenous extreme rainfall regions have been identified over South Africa that include extreme rainfall associated with the three above synoptic environments as well as those influenced by tropical cyclones. This is the first such regionalization of extreme rainfall over South Africa and the results are complimentary to regionalizations of general rainfall over South Africa.

In observed data, trends in the frequency of occurrence of extreme rainfall events show a general increase across South Africa and both statistically and dynamically downscaled results replicate these trends. Dynamically downscaled future projections indicate synoptic environments associated with extreme precipitation over southern Africa are likely to increase in the tropical and sub-tropical summer rainfall region with the magnitude of the change higher in the tropics. Over South Africa, increases in frequency and intensity of extreme rainfall are projected over the eastern parts of South

Africa, especially over the Drakensberg while decreases are projected over the western and southern parts of the country.

Dynamically and statistically downscaled projections indicate decreases in the frequency and intensity of extreme rainfall over the western and southern parts of South Africa as a consequence of fewer circulation states associated with extreme rainfall. However, over the eastern Drakensberg the results diverge – statistical downscaling suggests a decreased frequency of extreme rainfall events and the dynamical downscaling an increase in both intensity and frequency of extreme rainfall. A likely explanation of this discrepancy is that the statistical downscaling cannot capture the enhanced Clausius-Clapeyron in a warmer world. This is a thermodynamic relationship that describes how a warmer atmosphere has a higher water vapour content, which increases the probability of extreme rainfall. The statistical downscaling method is trained using the observed thermodynamic relationships and cannot accommodate changes in this relationship in a warmer world, this problem of stationarity is a known issue in many statistical downscaling methods. As the dynamical downscaling is able to capture the enhanced Clausius-Clapeyron relationship, and we have understood the changes in the synoptic environments associated with extreme rainfall produced by these models, we suggest the message from the dynamical downscaling is more defensible in this region. We do however suggest that further study into the reasons for the differences between each methodology is needed.

The projected increase in the intensity and frequency of extreme rainfall over the eastern mountains of South Africa is an important message as this region supplies much of the Gauteng area with water. Similarly, the projected decreases of these events in the southern and western parts of the country have implications for water supply in this region. This study has demonstrated that changes in diagnostic variables like rainfall, when understood in the context of the prognostic synoptic environment, provides understandable and robust messages of future changes in the nature of extreme rainfall.

Acknowledgements

We are grateful to the following groups and persons for their contributions in various ways to the project:

1. The Water Research Commission for their funding and support.

2. Elsa de Jager and Colleen de Villiers South African Weather Service for the provision of station data.

3. I am particularly grateful to my WRC Reference Group for their valuable input into the project.

- Mr W Nomquphu (Chairman, WRC))
- Dr W Landman (CSIR)
- Dr L Dyson (University of Pretoria)
- Dr C Engelbrecht (ARC)
- Prof G Pegram (University of KwaZulu-Natal)
- Dr F Engelbrecht (CSIR)
- Dr G Green (Private)
- Dr A Kruger (SAWS)

This page was left blank intentionally

Table of contents

Executive Summary **iii**
Acknowledgements **v**
Table of Contents **vii**
List of Figures **ix**
List of Tables **x**
List of Abbreviations **xii**

Chapter One: Introduction **1**
1.1 The changing nature of extreme rainfall as a result of anthropogenic greenhouse gas warming 1
1.2 Synoptic environments associated with extreme precipitation in South Africa 2
1.3 Aims and objectives 2

Chapter Two: Data and downscaling methods **4**
2.1 Introduction 4
2.2 Observational data 4
2.2.1 Station data 4
2.2.2 Gridded rainfall data 5
2.2.3 Reanalysis data 5
2.3 Statistical downscaling 6
2.4 Dynamical downscaling – CORDEX 7

Chapter Three: Regionalization of extreme rainfall over South Africa **9**
3.1 Introduction 9
3.2 Methodology and data 9
3.3 Results 13
3.4 Discussion and conclusions 13

Chapter Four: Characterizing known synoptic environments associated with extreme rainfall **15**
4.1 Introduction 15
4.2 Data and Methodology 15
4.2.1 Identification of synoptic states associated with extreme rainfall 15
4.2.2 Self Organizing maps 17
4.2.3 Training Data 17
4.2.4 The SOM methodology 18
4.2.5 Using quantization error as a threshold for identifying characteristic synoptic types 19
4.3 Results 20
4.3.1 Cut-off lows 20
4.3.2 Mid-latitude cyclones 25
4.3.3 Tropical Temperate Toughs 28
4.3.4 Quantization error method for identifying extreme rainfall synoptic environments 30
4.4 Discussion and conclusions 31

Chapter Five: Extreme rainfall characteristics over the observational period **32**
5.1 Introduction 32
5.2 Methods 32

5.2.1 Station data trend analysis 32
5.2.2 Evaluation of regional model rainfall data using ETCCDI indices 33
5.2.3 Evaluation of regional model data through a synoptic environment assessment 34
5.2.4 Assessment of rare extremes using extreme value theory 36
5.3 Observed changes of extreme rainfall in the station record 36
5.4 Evaluation of CORDEX dynamical regional models using ETCCDI indices 37
5.5 Rare extremes in observed and model data 40
5.6 Synoptic circulations in the 20th century simulations 41
5.7 Evaluation of CORDEX RCMs using the synoptic environment approach 43
5.8 Rainfall characteristics of archetypal circulation environments 46
5.9 Synoptic environments associated with extreme rainfall 47
5.10 Discussion and conclusions 52

Chapter Six: Projected changes in the characteristics of extreme rainfall and their synoptic environments 54
6.1 Introduction 54
6.2 Trends of projected extreme rainfall at the station scale 54
6.3 Projected changes in synoptic environments associated with extreme rainfall 59
6.4 Projected changes in intensity of rare extreme rainfall events 62
6.4.1 Projected changes in regional model data using ETCCDI indices 62
6.4.2 Projected changes of rare extremes in the regional model data 62
6.5 Discussion and conclusions 64

Chapter Seven: Summary, conclusions and recommendations 66
7.1 Summary 66
7.2 Conclusions 67
7.3 Recommendations for future work 68

Chapter Eight: Capacity development, publications and conference proceedings 69
8.1 Capacity development 69
8.2 Journal papers 70
8.3. Conference papers 70

Chapter Nine: References 72

List of Figures

Figure 3.1. Regions of homogenous rainfall identified by Landman *et al*. (2001), based on Mason (1998).

Figure 3.2. Stations representative of homogeneous rainfall regions within South Africa

Figure 3.3. Kernel density plots of rainfall volumes at selected stations demonstrating different distributions of rainfall in different regions of South Africa. X-axis is rainfall amounts in mm.

Figure 3.4. Regions of homogenous extreme rainfall as produced by a k-means clustering technique.

Figure 4.1. Archetypal maps for Z500 in geopotential height (top) and T500 in degrees Kelvin (bottom).

Figure 4.2. Frequency mapping of extreme rainfall days associated with COLs in number of days.

Figure 4.3. Mapping of extreme rainfall days in summer (DJF), winter (JJA), autumn (MAM) and spring (SON). Numbers are the number of days mapping to each node.

Figure 4.4. Archetypal sea level pressure mid-latitude cyclone circulations (hPa) associated with extreme rainfall.

Figure 4.5. Mapping of MLC associated extreme rainfall days in MJJAS and each respective month. Numbers are the number of days mapping to each node.

Figure 4.6. Archetypal cloudiness patterns based on OLR (W/m^2). Darker shades indicate less OLR (colder cloud tops) and hence deeper convective systems (see Hart *et al.* 2012).

Figure 4.7. Seasonal frequency mapping of OND (left) and JFM (right) days that experienced extreme rainfall associated with TTTs in South Africa.

Figure 5.1. Bootstrapping methodology used to ascribe statistical significance to a trend.

Figure 5.2. Sub-regions used in the analysis that span tropical (region 1), sub-tropical (region 2) and mid-latitude region (region 3). From Kalognoumou *et al*. (2013).

Figure 5.3. Trends in the frequency of occurrence of extreme rainfall at observation stations data for (a) the full station record and (b) from 1979-2009. Blue (red) triangles indicate positive (negative) trends and solid triangles where the trend is statistically significant.

Figure 5.4. Comparison in rainfall (in mm per day) of the 8 RCA4-downscaled GCMs with the GPCC dataset.

Figure 5.5. Observed and downscaled ETCCDI indices over southern Africa. Indices are listed in Table 5.1

Figure 5.6. 20-year return values of annual maximum daily precipitation (P_{20}) for the period 1997-2006 for (a) GPCP, (b) TRMM (1998-2005), CCLM forced by (c) ERA-Int. and different GCMs (e-h), RCA forced by (d) ERA-Int. and different GCMs (i-l) and (m) multi-model ensemble mean of CCLM and RCA4 forced by GCMs. Pattern correlation coefficient relative to the GPCP (TRMM) is in the top (bottom) left corner of each map.

Figure 5.7. Self-organizing map produced using ERA-Interim MSLP, U-, V-wind and specific humidity at 850 hPa level.

Figure 5.8. Frequency mapping of the ERA-Interim, downscaled ERA-Interim (1989-2005) and downscaled GCM (1976-2005) data. Black bars are the raw ERA-Interim mapping frequencies, the light and dark grey bars the mappings of the downscaled ERA-Interim simulations for the CCLM and RCA4 respectively and the rest of the bars are the mapping frequencies of the downscaled GCM by each RCM as indicated.

Figure 5.9. Seasonal frequency mapping of the ERA-Interim, downscaled ERA-Interim (1989-2005) and downscaled GCMs (1976-2005) where (a) is DJF, (b) MAM, (c) JJA and (d) SON.

Figure 5.10. GPCP precipitation anomalies associated with each circulation type in the 5x4 SOM. Negative values indicate conditions drier than the 1979-2005 mean and positive values wetter than the mean.

Figure 5.11. Seasonal frequency of extreme precipitation days as observed (GPCP) from 1997-2005.

Figure 5.12. GPCP precipitation composites of days with extreme precipitation associated with each type in the 5x4 SOM.

Figure 5.13. Nodal mapping of extreme rainfall in region 1 (top row), 2 (middle row) and 3 (bottom row) for the GPGP data, downscaled ERA-Interim runs and the downscaled GCM runs.

Figure 6.1. Trends in the frequency of occurrence of extreme rainfall data for 12 downscaled CMIP5 GCMs under RCP 8.5 between 1951 and 2100. The GCM name is at the top of each image and more information is available in Table 4.1. Blue (red) triangles indicate positive (negative) trends and solid triangles where the trend is statistically significant.

Figure 6.2. Trends in the annual frequency of occurrence of extreme rain days for selected statistically downscaled GCMs under RCP 8.5 for three periods – the historical period (1951-2000, left), the full 150-year period (centre) and for the period 2000-2100 (right). Only statistically significant trends are displayed where red (green) represents decreasing (increasing) trends.

Figure 6.3. Projected changes in the mapping frequency of all days for the periods 2036-2065 and 2069-2098 relative to 1976-2005 under RCP4.5 and RCP8.5 for the full southern African domain.

Figure 6.4. As for Figure 6.3 but for extreme rainfall in regions 1 (top), 2 (middle) and 3 (bottom).

Fig. 6.5. Projected multi-model mean changes of moderate extreme events for the period 2069-2098 relative to a reference period 1976-2005 under RCP4.5 (top) and RCP 8.5 (bottom). Stippling indicates grid points that are not statistically significant (5% significance level using a t-test).

Figure 6.6. Projected multi-model mean changes 20-year return values over the time period 2036-2065 (a and b) and 2069-2098 (c and d) as differences relative to the reference period (1976-2005) for RCP4.5 (a and c) and RCP8.5 (b and d). Units are in mm/day.

List of Tables

Table 2.1. List of GCMs downloaded from the CMIP5 archive for statistical downscaling

Table 4.1. Variables used to develop the four synoptic-state based SOMs

Table 4.2. Node mapping of overlapping TTT and COL days. Bold dates indicate consecutive calendar days.

Table 5.1. ETCCDI extreme rainfall indices

Table 5.2. Total number of extreme precipitation days for each region for the period of 1997-2005 for RCA, CCLM, each respective ensemble average as well as the combined ensemble average.

List of Abbreviations

CCLM	COSMO-Climate Limited-area Modelling
CFSR	Climate Forecast System Reanalysis
CMIP	Coupled Model Intercomparison Project
COL	Cut Off Low
CORDEX	Coordinated Regional Downscaling Experiment
DJF	December/January/February
EI	Era-interim Reanalysis
ERA-Int	Era-interim Reanalysis
ETCCDI	Expert Team on Climate Change Detection and Indices
GCM	General Circulation Model
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
JJA	June/July/August
MAM	March/April/May
MLC	Mid-latitude Cyclone
MSLP	Mean sea-level pressure
NCEP	National Centre for Environmental Prediction
RCA	Rossby Centre Regional Atmospheric Model
RCM	Regional Climate Model
SAHP	South Atlantic High Pressure
SAWS	South African Weather Service
SD	Statistical Downscaling
SOM	Self-Organizing Map
SON	September/October/November
TRMM	Tropical Rainfall Measurement Mission
TTT	Tropical Temperate Trough
T500	Temperature at Z500 level
Z500	500 hecto Pascal geopotential height
Z850	850 hecto Pascal geopotential height

Chapter One: Introduction

1.1 The changing nature of extreme rainfall as a result of anthropogenic greenhouse gas warming

Extreme rainfall events are often associated with significant societal and infrastructural impacts. Events can lead to severe flooding and often cause massive displacement of communities and fatalities. Several studies have shown an historical increase in the intensity of extreme rainfall over many regions of South Africa as well as spatial heterogeneity in these changes (New et al., 2006; Kruger, 2006). The Intergovernmental Panel on Climate Change recently released a special report, 'Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation' (Seneviratne et al., 2012). In it they state: 'It is likely that the frequency of heavy precipitation or the proportion of total rainfall from heavy falls will increase in the 21st century over many areas of the globe.' Furthermore, many studies and assessments have found links between changes in global climate and changes in regional events such as heavy rainfall, heat waves and flooding (e.g. Min et al., 2011).

The only way to assess potential changes in the nature of extreme rainfall is to use climate data from general circulation models (GCMs) that can simulate global climate under greenhouse gas forcing. However, GCM data is at a relatively course spatial scale (typically ~ 200 km) and should be downscaled to a finer scale to capture local drivers of rainfall like topography – this is especially relevant for extreme rainfall (Giorgi and Mearns, 1999). The latest climate projections from the Coupled Model Intercomparison Project (CMIP5) have recently become available and data from these is freely available online and number of the CMIP5 GCM results have been downscaled over the African continent. Statistically downscaled rainfall data to the station scale has been produced at CSAG and as part of the coordinated regional downscaling experiment (CORDEX[1]) many regional modeling groups have downscaled the GCMs to a grid resolution of 0.44 degrees over Africa. These data are available to assess historical and projected changes in rainfall and extreme rainfall over South and southern Africa as well as their driving synoptic states.

In dynamical climate models, prognostic variables are those governed by prognostic equations that solve for conservation of mass, momentum and thermodynamic energy through integrations in time. Diagnostic variables, like rainfall, are derived variables computed from prognostic variables and

[1] www.cordex.org

other external parameters, which together form the synoptic environments that determine whether or not it will rain in a particular place.

1.2 Synoptic environments associated with extreme precipitation in South Africa

In South Africa synoptic systems associated with extreme rainfall in South Africa include tropical and extra-tropical cyclones (Taljaard 1996), cut-off lows (Taljaard 1985; Singleton and Reason 2007; Favre et al., 2012), closed mid-tropospheric low pressure systems in the sub-tropics (Taljaard 1995; Dyson and van Heerden, 2002), tropical temperate troughs (Hart et al., 2010, 2012, 2013) and tropical and sub-tropical convective activity from mesoscale convective systems to single-cell storms (Laing and Fritch, 1993; Blamey and Reason, 2009; Dyson et al. 2015). Establishing a robust relationship between the extreme rainfall surface and the driving synoptic environment facilitates an examination of future projections from GCMs and RCMs without having to rely purely on rainfall data from the models.

In this study, instead of examining only the diagnostic rainfall variable we examine the prognostic synoptic environment in which extreme rainfall occurs and characterize these. It is believed a more reliable and defensible rationale can be developed to explain projected changes in the nature of extreme rainfall over South Africa by understanding changes in extreme rainfall synoptic environments. We propose that examining the characteristics of prognostic variables (a) provides a more robust quantification of extreme rainfall changes and (b) provides an understanding of changes to the synoptic environment that facilitates the surface rainfall response statistics.

A recent WRC study (Lennard et al., 2013) adopted a synoptic quantification approach and demonstrated that summer extreme rainfall in South Africa is associated primarily with (1) the sub-tropical low pressure system over the interior of the country and (2) circulations that show a surface linkage between the sub-tropics and the passage of a trough in the mid-latitudes to the south of the country. Winter extreme rainfall was associated with the passage of mid-latitude cyclones. Moreover, significant trends in extreme rainfall identified in station data were associated with significant trends in synoptic circulations in summer rainfall cases.

1.3 Aims and Objectives

In this study, we build on the results of Lennard et al. (2013) to quantify more completely the types of synoptic states associated with extreme rainfall in South Africa and also investigate potential changes in the nature of extreme rainfall and their driving synoptic environments as a result of global warming. The primary objective of this project is to investigate projected changes in the nature of extreme rainfall by examining changes in the synoptic environments associated with

extreme rain. Four specific aims were developed to facilitate this:

1. Identify homogeneous extreme rainfall regimes across South Africa.
2. Produce a statistically downscaled dataset for stations across South Africa and identify historical and projected trends.
3. Explicitly examine known circulation features that are associated with extreme rainfall.
4. Identify changes in the historical and projected frequency of occurrence of synoptic circulations associated with extreme rainfall.

This report is constructed in the following way in order to most logically address the primary objective of the study. Chapter 2 introduces the data sources and methods employed to produce these data. Chapter 3 presents a regionalization procedure that produces homogeneous extreme rainfall regions over South Africa. Chapter 4 presents a characterization of known synoptic environments associated with extreme rainfall (cut-off lows, tropical temperate troughs and mid latitude cyclones) and evaluates a methodology to identify of these environments in climate data. Chapter 5 describes extreme rainfall in observed station and gridded data and evaluates the ability of regional climate models to reproduce the characteristics of rainfall and extreme rainfall over the observed period. Synoptic environments associated with extreme rainfall are also described. Chapter 6 assesses downscaled projection data to quantify future potential changes in extreme rainfall and their synoptic environments. Chapter 7 presents the summary and conclusion as well as suggested research avenues to pursue based on this work.

Chapter Two: Data and downscaling methods

2.1 Introduction

It is well understood that GCMs are the only tool available to assess potential changes to the global climate as a result of greenhouse gas warning. However, output from GCMs participating in the latest CMIP5 are available at horizontal grid scales as coarse as ~3x3 degrees to grids as fine as ~0.5 degrees. However, these fine resolution models are exceptional and most GCMs provide data at horizontal grid resolutions at between 1-2.5 degrees (IPCC, 2013). At these resolutions the large-scale circulation features like mid-latitude cyclones or subsidence-induced sub-tropical anticyclones may be captured but fine-scale processes associated with local and regional phenomena like topographically forced uplift or land surface forcing cannot be resolved. This is particularly important for rainfall, which often operates at much smaller spatial scales and especially for convective rainfall. This has implications for extreme rainfall as topographic and dynamic drivers that may cause extreme rainfall cannot be resolved at such scales. There is therefore a need to downscale low-resolution GCM climate data to the relevant scales. This downscaling may be performed dynamically using regional climate models (RCMs), which simulate the 3-dimensional atmosphere (Giorgi and Mearns, 1999) or statistically where empirical relationships between large-scale atmospheric variables and local climate variables are derived (e.g. Hewitson and Crane 2006). Dynamical and statistical downscaling methods are recognized as being of comparable skill (Christensen et al., 2007), however the methods can have markedly different attributes.

In this project, data from both downscaling methodologies are used. A description of the statistical downscaling methodology is presented below followed by information about the dynamically downscaled data obtained from the CORDEX project. Observational data is needed as a benchmark from which projections of change can be compared to and these are presented in the next section.

2.2 Observational data

Three sets of observation data were used in this study, namely station rainfall data, gridded rainfall data and reanalysis circulation data.

2.2.1. Station data

Deliverables from a previous WRC project (K5/1960) required that (a) station data be obtained from the South African Weather Service and (b) the development of a rainfall station data quality control system by the Climate System Analysis Group. In this project we have updated the station data record at CSAG with SAWS data until July 2012 and passed them through the quality control system developed at CSAG. This procedure produced data that could confidently be assessed to

fulfil the requirements of the project.

The station data are used in two general ways. Firstly, they form the predictand of the statistical downscaling (see section 2.3 below) and the evaluation data against which trends in the downscaled projected extreme rainfall data at the station scale can be compared (Chapters 5 and 6). Secondly, extreme rainfall thresholds are determined at each station and all days equal to or over this threshold are considered as extreme rainfall days. These dates are used to (1) extract gridded data used to develop homogeneous extreme rainfall regions in Chapter 3 and (2) the characterization of extreme rainfall synoptic environments in Chapters 4.

2.2.2. Gridded rainfall data

Gridded rainfall datasets were used for the regionalization of extreme rainfall in Chapter 3, the characterization of observed extreme rainfall statistics in Chapter 4 and the validation of regional model output in Chapter 6. In Chapter 3 data from the Tropical Rainfall Measuring Mission (TRMM 3B42 v7) are used to develop the homogenous extreme rainfall regions and is available from 1998 to present. In Chapters 4, which uses CORDEX downscaled data, the monthly Global Precipitation Climatology Centre (GPCC) and daily Global Precipitation Climatology Project (GPCP) rainfall data are used as they are the most widely adopted rainfall dataset within the CORDEX-Africa framework and are representative of station data in South Africa and Mozambique (Kalognoumou *et al*., 2013). Several observational datasets are available for model evaluation and have been reviewed to some extent in Nikulin *et al*. (2012) and Sylla *et al*. (2012). Nikulin *et al*. (2012) show large differences between satellite and gauge-based products and highlight the lack of reporting station gauges over large parts of Africa. However, Kalagnoumou *et al*. (2013) demonstrate that the differences between datasets are generally small (although they so show some temporal and spatial biases).

2.2.3. Reanalysis data

Two reanalysis datasets were utilized in this study. The ERA-Interim reanalysis (Dee *et al*. 2011) were used for the statistical downscaling as predictor variables (Chapter 5) as well as in the CORDEX simulations as boundary conditions for the evaluation simulations. The CFSR and ERA-Interim reanalysis were used in the SOM analyses (Chapter 4 and 5) for the characterization of synoptic states associated with extreme rainfall. The CFSR data are of a slightly higher resolution that the ERA-Interim data and it was thought the former would aid in characterizing the synoptic environments associated with extreme rainfall like cut-off lows and tropical temperate troughs more accurately.

2.3 Statistical downscaling

As part of this project, CMIP5 GCM data were statistically downscaled at CSAG to the station scale using the technique described in Hewitson and Crane (2006). Self-organizing map based downscaling (SOMD) is a leading empirical downscaling technique for Africa and can provides meteorological station level or gridded response to global climate change forcing. In essence this technique develops statistical relationships between rainfall at a number of stations within a specified proximity as well statistical relationships between these stations and synoptic scale circulation characteristics.

The downscaling of a GCM is accomplished by deriving the normative local response from the atmospheric state on a given day (predictors), as defined from historical observed data (predictants). The method recognises that the regional response is both stochastic as well as a function of the large scale synoptics. As such it generates a statistical distribution of observed responses to past large scale observed synoptic states. These distributions are then sampled based on the GCM generated synoptics in order to produce a time series of GCM downscaled daily values for the variable in question. Advantages of this method are that it is computationally much less expensive than dynamical downscaling and that the relatively unskilled grid scale GCM precipitation is not used by the downscaling but rather the relatively highly skilled large scale circulation (pressure, wind and humidity) fields are employed. This method can downscale to the point (station) scale and is used here to facilitate an assessment of projected changes in the frequency of extreme rainfall.

The SOMD methodology has undergone a number of key changes during the project. These include the modification of the code to allow it to handle the long continuous period provided by the CMIP5. A major change is associated with the stochastic element of how the code samples the local response under a given synoptic state. In the original version, there was strong auto-correlation in the sampling between grid cells. This produced an unrealistic spatial pattern of daily rainfall and resulted in an overestimation of the temporal variance when sampled over an area. The current version has a far more sophisticated method of sampling the local response for each grid cell; a random sampling field is generated so that each grid point has a random time series, but the correlation between each time series and each of the surrounding 8 grid points approximates the spatial correlation of observed precipitation. This project has therefore contributed to the advancement of the statistical downscaling methodology developed at UCT.

The SOMD is trained on the “observed” daily atmospheric states, or predictors, obtained from the ERA-Interim Reanalysis dataset for the period 1979-2012. The variables used include; de-trended

near surface temperatures and 10 meter winds, atmospheric temperature lapse rate, 700hPa relative humidity and winds. The local responses to these states, or predictants, are characterised using daily rainfall from weather stations across South Africa obtained from the South African Weather Service. These station data have been rigorously quality controlled according to procedures developed in Lennard *et al*. (2013). Once the relationship between each station and the synoptic environment is established, projected predictors from GCMs are used to develop projected downscaled rainfall at each station which are assessed in this study.

Data from 12 CMIP5 GCMs that span the period 1950-2100 were downloaded from the Earth Systems Grid (ESG) nodes for RCP4.5 and RCP 8.5 onto the CSAG storage system and checked for any errors (Table 2.1). Error checks related to the integrity of the data, which may become corrupted during the download owing to dropped packets that may lead to errors later in the process, and whether the simulations produced reasonable results.

Table 2.1. List of GCMs downloaded from the CMIP5 archive for statistical downscaling

General Circulation Model	Number of latitudes	Number of longitudes	Resolution (degrees)
BNU-ESM	64	128	2.8
CanESM2	64	128	2.8
CNRM-CM5	128	256	1.4
GFDL-ESM2G	90	144	2.5
GFDL-ESM2M	90	144	2.5
inmcm4	120	180	2 x 1.5
IPSL-CM5A-MR	143	144	2.5 x 1.3
MIROC5	128	256	1.4
MIROC-ESM	64	128	2.8
MIROC-ESM-CHEM	64	128	2.8
MPI-ESM-LR	96	192	1.8
MRI-CGCM3	160	320	1.1

Data from the statistical downscaling are used to examine long-term trends in the frequency of occurrence of extreme rainfall as projected by the models (Chapter 6).

2.4 Dynamical downscaling – CORDEX

The RCM data is sourced from the Co-ordinated Regional Downscaling Experiment (CORDEX[2]) and consists of downscaled circulation and rainfall data. The CORDEX protocol requires the downscaled data to be at a resolution of 0.44 degrees and the downscaling has been run over all of

[2] http://www.cordex.org

Africa. The downscaling programme consists of two streams, (1) an evaluation period (1979-2010) where ERA-Interim reanalysis data have been downscaled by a number of RCMs and (2) downscaled CMIP5 GCM projection data for RCP4.5 and RCP8.5 that spans 1951-2100. Data are freely available on the earth system grid federation (ESGF) network.

Data were retrieved from the ESGF – at the time of analysis this was restricted to two regional models that had each downscaled 4 GCMs. The two regional models are the Swedish Meteorological and Hydrological Institute (SMHI) RCA4 model, which has downscaled eight CMIP5 GCMs and the Consortium for Small Scale Modelling Climate Limited-area Modelling (CCLM) model, which downscaled four GCMs. The eight GCMs that were downscaled by the RCA4 regional model are the CCCma-CanESM2, MIROC-MIROC5, MPI-M-MPI-ESM-LR, NOAA-GFDL-ESM2M, CNRM-CERFACS-CM5, ICHEC-EC-EARTH, MOHC-HadGEM2-ES and the NCC-NorESM1-M. The four GCMs downscaled by the CCLM were the CNRM-CM5, EC-EARTH, MPI-ESM-LR, HadGEM2-ES which overlapped with the RCA4 downscaling and facilitated direct comparisons.

These CORDEX data are used in Chapter 5 to assess observed changes in extreme rainfall statistics and evaluate the eight RCA4 downscalings using a parametric and non-parametric assessment of rainfall and extreme rainfall. The data are also used in Chapter 6 where the overlapping RCA4 and CCLM downscalings are analysed to assess projected changes in the synoptic drivers of extreme rainfall.

Chapter Three: Regionalization of extreme rainfall over South Africa

3.1 Introduction

There have been a number of regionalizations of climatic zones in South Africa since Jackson (1951) and Preston-Whyte (1974), which have largely focused on rainfall and temperature (Harrison, 1984; Taljaard and Phil, 1996; Mason, 1998; Landman and Mason, 1999). More recently, Kruger (2002) described the climatic characteristics of 24 vegetation regions across the country and Engelbrecht and Engelbrecht (2015) focus on projected shifts in temperature climate zones over Southern Africa. However, there is no spatial regionalization of extreme rainfall over South Africa yet. In this chapter we attempt to address this using a multi-step, event-based assessment to develop regions across South Africa that experience similar types of extreme rainfall.

3.2 Methodology and data

In Lennard et al. (2013) eight rainfall regions developed by Mason (1998 – Fig. 3.1) were populated with 69 stations that were representative of rainfall within these regions (Fig.3.2). Extreme rainfall dates, where extreme rainfall is defined to be equal to or exceeding the 95th percentile of the station data record, were identified in the 69 stations and used to extract corresponding dates from the Tropical Rainfall Measuring Mission (TRMM – Kummerow et al. 2000) gridded rainfall dataset that has a high-resolution spatial coverage over South Africa. These gridded data are used to develop the homogenous extreme rainfall regions.

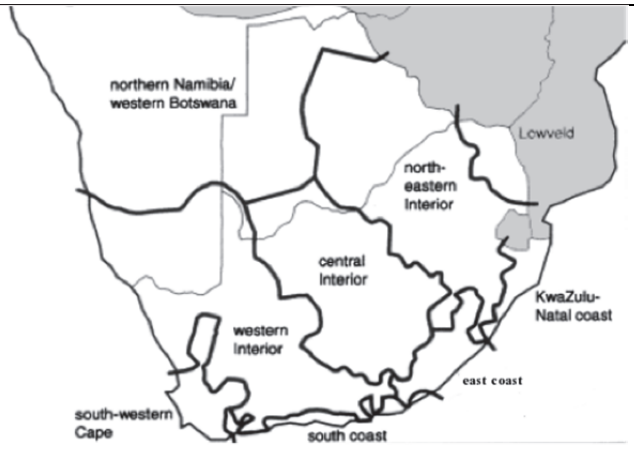


Figure 3.1. Regions of homogenous rainfall identified by Landman et al. (2001), based on Mason (1998).

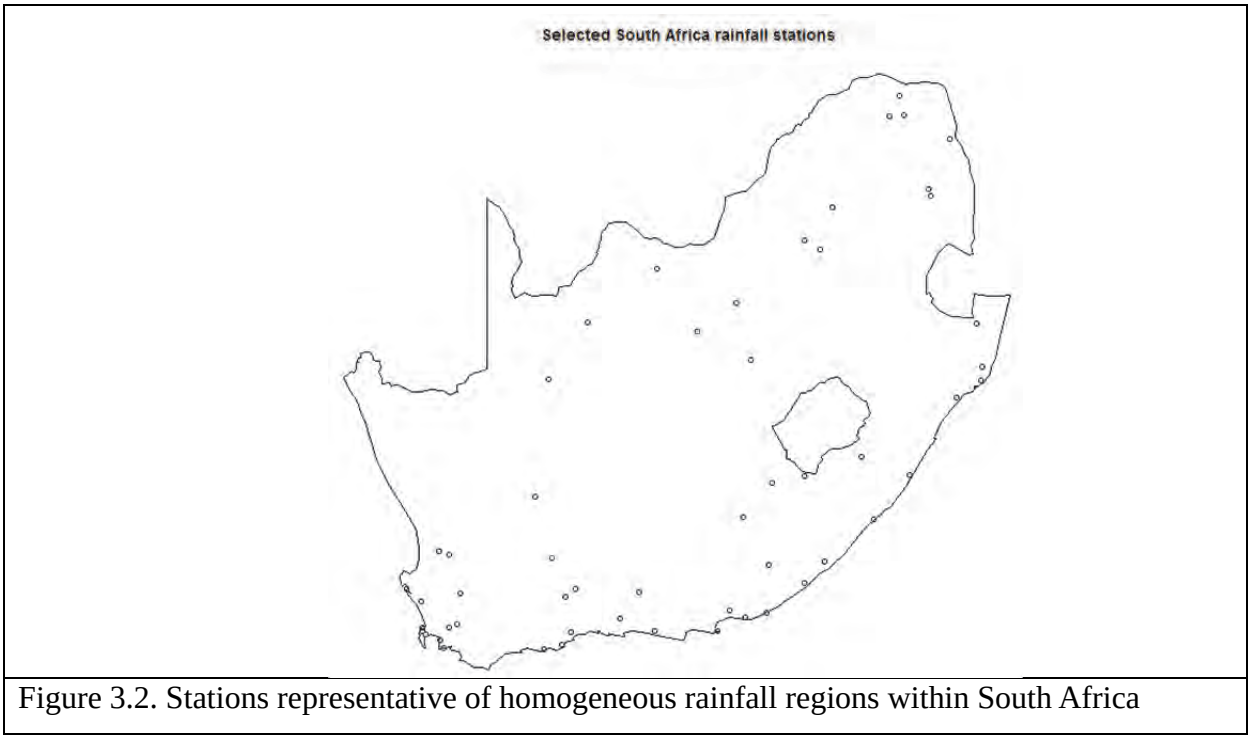


Figure 3.2. Stations representative of homogeneous rainfall regions within South Africa

We use the TRMM 3B42 (v7) daily rainfall data, which are available from 1998 to present. The TRMM observation system is designed to capture convective types of rainfall so is particularly suited to the region which experiences primarily convective rainfall. The resolution of the data is ~25 km. Many studies have used this data to investigate rainfall in tropical and sub-tropical regions (e.g. Kim et al., 2004; Futyan and Genio, 2007; L'Ecuyer and McGarragh, 2010; Diatta et al., 2010). In the mid-latitudes, TRMM data has been used to characterize rainfall, which may be convective or stratiform, in South America (Rasmussen et al., 2013), the north-western Pacific (Yamamoto et al., 2006), China (Ji and Chen, 2012), in a latitudinal bands up to 40 degrees north and south of the Equator (Kodoma and Tamaoki, 2002; Biasutti et al., 2012; Liu et al., 2013) and with a focus over Africa by Balogun and Adeyewa (2013). These and other studies have highlighted problems with the TRMM data that include an underestimation of deep convective rainfall and an underestimation of rainfall over regions of complex topography. However, the TRMM data captures the spatial pattern of rainfall even though it may not correctly capture the magnitude of rainfall (Han et al., 2009). Given that we are developing a spatial regionalization based on extreme rainfall dates obtained from station data, days that exhibit similar types of extreme rainfall in the TRMM data would be expected to fall in similar spatial rainfall extents. We therefore think the using TRMM data, despite the known shortcomings of the data, is justified.

These gridded data are clustered (spatially) into homogeneous extreme rainfall regions to produce a map of homogeneous extreme rainfall regions. This procedure is carried out in a 3-steps:

- *Step 1. Identification of extreme rainfall days*

Stations are not homogenous in their rainfall distributions, some have longer tails so extreme rainfall volumes at one station may be heavy rainfall in another (Fig. 3.3). Highest densities are found in lower rainfall volumes as expected but the length of the tails vary. Stations (a-c) are located in the southwestern Cape with station (b) in the mountains and (c) up the west coast. Stations (d, e and g) are on the south and east coasts and have much longer tails. Stations (f, h and i) are located over the interior of the country and have generally shorter tails. As a result of the heterogeneity of the distributions of the tails, we defined extreme rainfall days as equal to or above the respective stations 95^{th} percentile threshold of each station. In establishing the 95^{th} percentile value the full available station record was used and only days that recorded rainfall considered. Then for the regionalization, extreme rainfall days are identified in each of the 69 stations and the dates of each day extracted.

Step 2. Extraction of extreme rainfall data from gridded rainfall dataset

In the station data, 1278 days were identified as having experienced extreme rainfall between 1998 and 2009 (the period TRMM data were available). For each day identified in the station data that recorded extreme rainfall the corresponding TRMM rainfall data are extracted for the overlapping day(s). If more than one station reported extreme rainfall on a particular date data for that date was extracted only once.

- *Step 3. Regionalization of extreme rainfall*

Extreme rainfall is spatially regionalized using a k-means clustering methodology (Hartigan, 1975; Hartigan and Wong, 1979). The method divides *M* points in *N* dimensions into *K* clusters so that the within-cluster sum of squares is minimized. An input matrix of *M* points in *N* dimensions and a cluster matrix of *K* initial cluster points in *N* dimensions is required. Using Euclidian distance, the algorithm then searches for a K partition with local optimal (minimum) within-cluster sum of squares by moving points from one cluster to another.

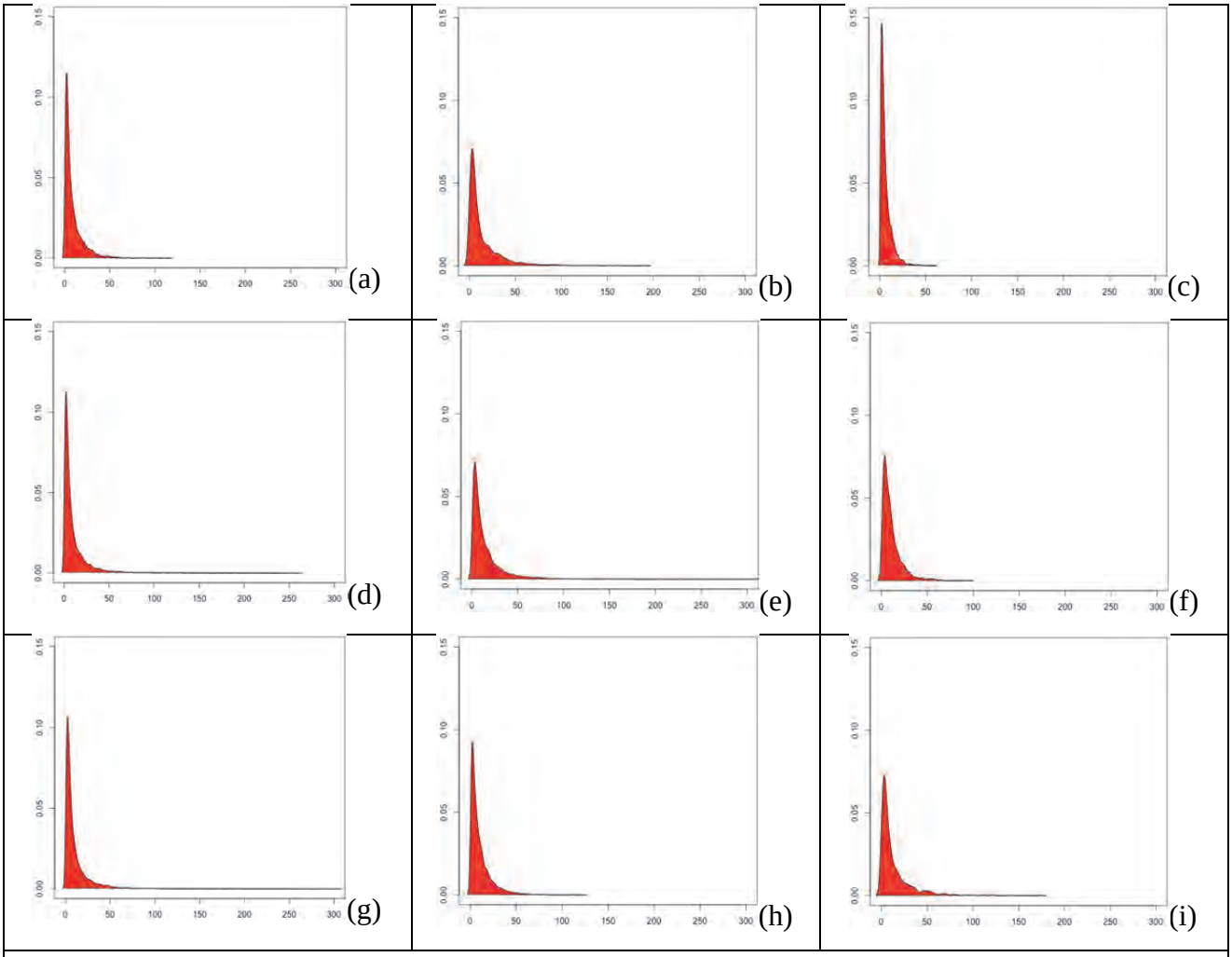


Figure 3.3. Kernel density plots of rainfall volumes at selected stations demonstrating different distributions of rainfall in different regions of South Africa. X-axis is rainfall amounts in mm.

It should be noted that data measured at stations are accumulated rainfall from 08h00 on the particular day to 08h00 the following day. The TRMM data available for this study are daily data from midnight to midnight so there is not complete overlap in the two measurement systems. This means that the following situation could exist: if an extreme rainfall event occurred, for example on 2 November between 00h00 and 08h00, this would be captured in the station data as having occurred on the 1st November. Therefore the TRMM data extracted 1st November would not reflect this extreme rainfall event. In a strongly convective region, where convection can occur late at night, this scenario is possible, however, the convective rainfall generally peaks in the late afternoon over continents (Yang and Smith, 2008; Yang et al., 2008). For this current analysis we use the daily TRMM data and are cognizant that there may be some mismatch in the timing between the observed and gridded data.

3.3 Results

The gridded data for each extreme rainfall day form the input matrix. A number of cluster regions were tested and eight were found to represent the types of extreme rainfall affecting each region, although one cluster is entirely over the ocean (Fig. 3.4). Seven homogeneous extreme rainfall regions are discriminated by the k-means methodology and an 8th region is placed offshore to the southeast of the country. Regions 1 and 3 are over typically summer convective rainfall areas and likely associated with extreme rainfall due to mid-tropospheric low-pressure systems and tropical temperate troughs. Likewise for region 2, although the influence of tropical cyclones is likely captured in this cluster as well. Region 6 covers the dry Northern Cape and northern parts of the western Cape and region 7 largely covers the Western Cape. Extreme rainfall in these regions is likely associated with extreme rainfall due to cold fronts and cut-off lows. In fact, Favre *et al*. (2012) show a high percentage of extreme rainfall in these two regions is due to COLs (up to 50 % in some places) and increasing trends in extreme rainfall due to COLs in region 7.

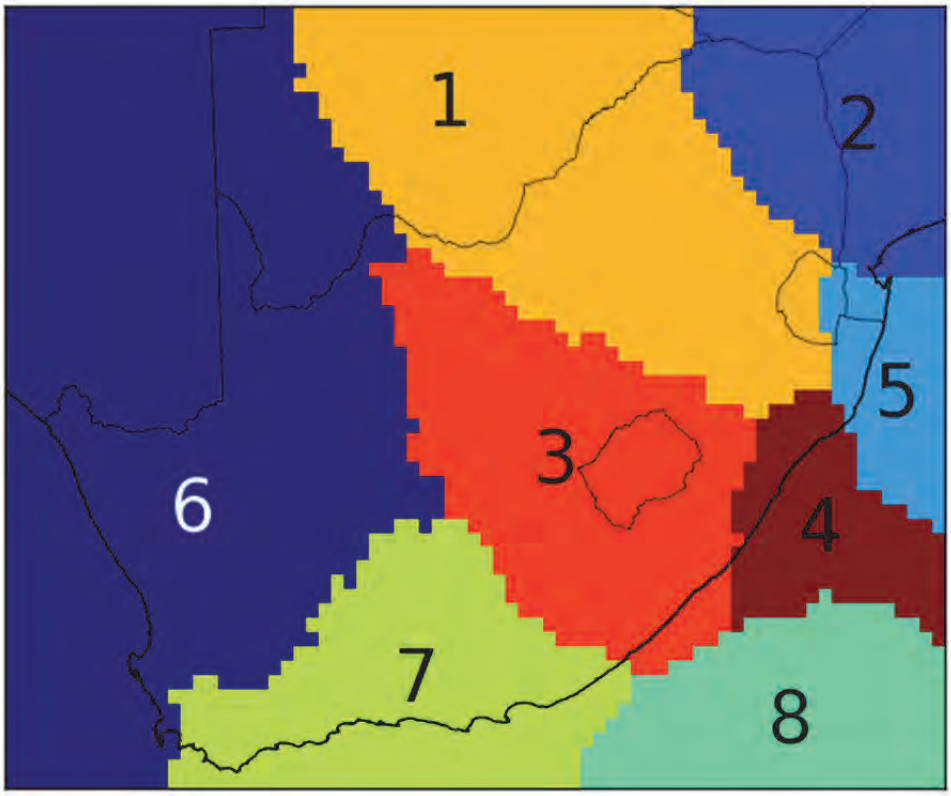


Figure 3.4. Regions of homogenous extreme rainfall as produced by a k-means clustering technique.

3.4 Discussion and conclusions

The regionalization produces spatial clusters that are similar to the homogeneous rainfall regions of Mason (1998) and Landman et al. (2001). Regions 1, 2, 3, 4 and 5 and 6 cover similar spatial extents as the Landman et al. (2001) north-eastern interior, Lowveld, central interior, Kwazulu-

Natal coast (although it splits the KZN region into two clusters) and western interior respectively. However, there are differences with the k-means clustering does not identify south coast all year rainfall region of Landman et al. (2001) nor does it capture the south-western Cape region. This may be a function of (1) the rainfall type and (2) the topography in this region. In the case of the former the TRMM system has been shown to not capture shallow, stratiform rainfall associated with cold fronts as well as it captures convective rainfall (Yang and Smith 2008). Also it should not be expected that an extreme rainfall regionalization produce the same results as a general rainfall regionalization as the extreme rainfall dataset is much smaller and the statistical characteristics are different. However, the regions presented here have a good physical basis for the regional characterizations of extreme rainfall and this study provides the first such characterization over South Africa.

Chapter Four: Characterizing known synoptic environments associated with extreme rainfall

4.1 Introduction

As mentioned in Chapter 1, certain synoptic environments have been associated with extreme rainfall in South Africa that include tropical and extra-tropical cyclones, cut-off lows, closed mid-tropospheric low pressure systems in the sub-tropics, tropical temperate troughs and tropical and sub-tropical convective activity from mesoscale convective systems to single-cell storms. It is believed a more reliable and defensible rationale can be developed to explain projected changes in the nature of extreme rainfall over South Arica by understanding changes in these synoptic environments that are associated with extreme rainfall.

The results from WRC project K5-1960 (Lennard *et al.*, 2013) demonstrated that a self-organizing map (SOM) could not discriminate closed low pressure circulations like cut-off lows or mid-tropospheric sub-tropical lows when applied to 30 years of daily synoptic data. The report suggested that SOMs instead be developed based on particular synoptic types, i.e. develop one map with circulation parameters associated with only cut-off lows and a separate SOM, e.g. mid-latitude cyclones, etc. This suggestion is implemented here and three types of known circulations that are associated with extreme rainfall above are examined. These are (1) mid-latitude cyclones (MLCs), (2) cut-off lows (COLs) and (3) tropical temperate troughs (TTTs). For each of these known circulation states a number of synoptic types are produced depending on the available number of observations and their seasonal frequency examined.

In this chapter we characterize three synoptic environments associated with extreme rainfall, namely cut-off lows, mid-latitude cyclones and tropical temperate troughs. We also attempted to use the SOM methodology to identify extreme rainfall environments in downscaled projection data.

4.2. Data and Methodology

4.2.1 Identification of synoptic states associated with extreme rainfall

Previous work has identified days in the historical station record that experienced extreme rainfall due to cut-off lows (Favre *et al.*, 2013) and tropical temperate troughs (Hart et al, 2010). These authors kindly provided the dates of days that recorded extreme rainfall as a result of the respective systems. For COLs a total of 522 the extreme rainfall days were identified and for TTTs 45.

Days that experienced extreme rainfall associated with MLCs were identified as part of this study. We assessed station records from May to September during 1979 to 2010 and examined the Climate

Forecast System Reanalysis data (Saha *et al.*, 2010) for evidence of frontal activity associated with the extreme rainfall. We excluded dates associated with COLs to avoid a conflation effect of multiple systems. However, as COL dates were only available between 1980-2006, to obtain a complete 30 year climatology it was also necessary to extract a subset of dates for the periods of May to Sept 1979, 1980, 2006-2009 and analyse each of the extreme rainfall days to check for the occurrence of COLs through examining the 500 hPa geopotential height data of the Climate Forecast System Reanalysis (CFSR).

Dates identified as being associated with extreme rainfall caused by frontal systems were cross-checked against the South African Weather Service (SAWS) Daily Weather Bulletin for the time period 1979-2009 (May to September). The SAWS reports were provided by the Department of Oceanography, University of Cape Town for 1979 to 2007 and SAWS online from 2008 (http://www.weathersa.co.za/climate/publications). The SAWS reports are typically given as snapshots either at 12:00 or 14:00 and they report rainfall as a total for a 24 hour period (i.e. mm between 08:00 yesterday to 08:00 today). These results were compared to the CFSR circulation data and good agreement was found with some discrepancy in only five days when frontal synoptic features observed in SAWS were not reflected in the CFSR data.

The SAWS reports were useful for the analysis of variables and conditions associated with the position of fronts and location of the extreme rainfall conditions across the region. This information was particularly valuable for days when multiple- and interacting- synoptic systems were responsible for the extreme rainfall conditions experienced across the country. An example that was common in the analysis, that involved some subjectivity in selection, was the occurrence of a ridging high-pressure system behind a cold front that extended across the country. In such a case, a subjective decision would be needed in order to determine whether the circulation around the high-pressure system (HP) behind the front was the result of moisture from the warm ocean, leading to orographic rain fall over the interior and/or whether the passage of the cold front had also caused extreme rainfall conditions over SE topographical regions of South Africa. While it was easily possible to glean the presence (or absence) of a frontal from the CFSR (through interpreting the contour spacing and values of sea level pressure and geopotential height isobars), the SAWS charts provided complementary information on variables (*e.g.* wind direction, cloud cover and the temperature) and the position of the rainfall in relation to these variables (e.g. the west coast rainfall was mostly frontal, if along the south coast mostly orographic around the HP). This would help determine if, for a given date, the region(s) that reported heavy rainfall, were likely under the influenced of the HP system advecting moisture or still under the influence of the frontal system

moving eastward. This analysis produced 670 extreme rainfall days associated with MLC extreme rainfall over South Africa for the period 1979-2009. The frontal circulation states were then confirmed by visual inspection.

Having identified dates that experienced extreme rainfall associated with COLs, MLCs and TTTs, days that experienced extreme rainfall due to each of these three synoptic types were grouped respectively. For each of the three states, circulation data associated with extreme rainfall days were extracted from the CFSR to develop self-organizing maps.

4.2.2. Self Organizing maps

Self-organizing maps (SOMs) are a form of artificial neural net (ANN) that reduce high dimensional data space to a lower dimensionality (usually two-dimensional), discretized representation of the input data and produce a map of discretized data archetypes. They extract and display the major characteristics of a multidimensional data distribution function to produce an array of generalized data archetypes (Hewitson and Crane, 2006). A number of climatological and oceanographical studies have used SOMs over southern Africa including Engelbrecht *et al.* (2014), Lennard and Hegerl (2014), Thomas *et al.* (2007), Hewitson and Crane (2006), Tadross *et al*. (2005) and Richardson *et al*. (2003). The reader is referred to Lennard and Hegerl (2014) and Lennard *et al*. (2013) and for a detailed description of the SOM methodology as applied in this analysis and to Kohonen (2001) for a detailed description of the SOM algorithm. However, in Section 5.2.4 a brief summary of the procedure is given.

The SOM software used in this study is SOM_PAK version 3.1[3] which has three distinct stages in the data mapping routine in which (1) the type and size of the SOM is set, (2) the training occurs and (3) error is evaluated and results are visualized.

4.2.3 Training Data

The reanalysis to be used for the training phase of the SOM is the Climate Forecast System Reanalysis data (Saha *et al.*, 2010) available from 1979 to 2010. We use CFSR rather than the ERA-Interim data as it has a higher spatial resolution of ~50km (vs ~75 km) even though the former is available until end-2012. As the procedure is a synoptic state-based assessment it is not necessary to have up to date reanalysis data. The data are extracted for each of the extreme rainfall days identified above for a domain that covers southern and central Africa. Variables extracted for the training of the SOM are specific to the synoptic state being examined as each state has unique

[3] Software and documentation available online at http://www.cis.hut.fi/research/som_lvq_pak.shtml

variables that determine its characteristics (Table 5.1). These data are extracted at 12Z for dates on which these particular systems were the cause of extreme rainfall as reflected in the station record. In the cases of mid-latitude cyclones and cut-off lows more than one variable was used in the SOM training and these were standardized to preserve the gradients in the data.

Table 4.1. Variables used to develop the four synoptic-state based SOMs

Mid-latitude cyclone	**Cut-off low**	**Tropical temperate trough**
Sea level pressure	500 hPa geopotential height 500 hPa temperature	Outgoing long-wave radiation

4.2.4 The SOM methodology

Self-organizing maps are used to classify synoptic types associated within each of the extreme rainfall producing systems above. For example, to identify the types of cut-off lows associated with extreme rainfall in the region, a SOM would be developed using reanalysis data only from days on which extreme rainfall was caused by cut-off lows. This provides information on the spatial extent, depth and seasonality of the synoptic archetypes.

The SOM methodology employed here has three distinct stages in the data mapping routine where (1) the type and size of the SOM is set, (2) the training occurs and (3) error is evaluated and results are visualized. Once the size of the SOM has been decided, i.e. the number of desired synoptic archetypes, a random distribution of reference vectors or nodes numbering the size of the SOM is set within the data space. Random numbers are then assigned to the reference vectors equally across the data space. The data space consists of training vectors, which are obtained from a set of predictor variables that are, in this case, variables selected from the CFSR reanalysis. The SOM is now ready for training.

Training the SOM is an iterative process in which the weights on a node (the reference vectors) are adjusted toward the training vectors such that they span the variance structure of the data space. During the training process, the node whose weight matches the training vector most closely (where there is the minimum Euclidian distance between node vector and data vector) is chosen as the 'winning' node for the particular vector. Not only is the winning node updated but the nodes surrounding the winning node also benefit from the learning process by adjusting their weights such that each vector converges to the input pattern. These form an update neighbourhood, at whose centre the winning node lies. The learning rate **determines how fast the weights move towards the data points. As** the computational needs of the SOM analysis are very modest a small learning rate is usually set and an arbitrarily large number of iterations are made. The reference vectors

adjust during the iterative training process such that they span the whole data space where the number of iterations is chosen such that final convergence (the fine adjustments of the nodes in the data space) can be achieved by the end of the training. The trained SOM now consists of the selected number of nodes (centroids) each surrounded by its data cloud of training vectors (see Sang *et al*., 2008, Fig 4).

The third stage of the process evaluates the quantization error (QE) of the trained SOM in order to produce the best SOM. For each input vector, the winning node is selected through an assessment of the Euclidean distance to the node vectors. The shortest distance between nodes, or quantization error, is calculated and the input vector is allocated to that node. In this way each day in the training period is associated with a particular node, however the mapping co-ordinate in the data space is not exactly the same as that of the node but is some Euclidean distance from the node. The best SOM is attained when the averaged minimum Euclidean distance across the SOM is obtained (the smallest average quantization error). A number of parameters can be set during the configuration of the SOM to minimize the QE and these are discussed in the next section.

4.2.5 Using quantization error as a threshold for identifying characteristic synoptic types

Each day in the SOM training data maps to a particular node some Euclidean distance from the node centroid. Based on the configuration of the training, this distance can be minimized to produce a SOM most representative of the training data. Using this trained SOM, it was hypothesized it would be possible to map other data to it and quantify the QE in the mapping of the second dataset to discriminate synoptic states similar to those of the trained map.

For example, a SOM trained on circulation data from cut-off lows (COLs) would produce archetypal COL circulation states and each day of the training data would map to a particular node but lie some Euclidean distance from the centroid of the node. These Euclidean distances would span a range for each node and this range would form a reference threshold distance. Then a second dataset, say for example 30 years of daily CFSR data or downscaled projection data, would be mapped to the trained SOM and each day of the second dataset would map to a particular node some Euclidean distance from the centroid. The Euclidean distance of the daily mapping can be compared to the respective reference threshold distance and if it falls within the reference range the day could be flagged as a potentially extreme rainfall causing system and the circulation and rainfall data extracted and assessed.

In this way the SOM method can be used to (1) identify extreme rainfall causing systems in the

reanalysis data that lie outside the training period as an evaluation of the method and (2) identify synoptic states in projection data associated with extreme rainfall and changes in the characteristics of these compared to historical traits.

The method relies on the minimization of the QE in the training phase of the reanalysis data associated with extreme rainfall. If QE values are too large in the training phase, data being mapped to the trained SOM may fall within the threshold Euclidean distance and produce a lot of "false positive" extreme rainfall causing synoptics which would render the process unusable. There are a number of ways to reduce the QE of the training data reported in the literature. Sun, (2000) shows that having a small update neighbourhood is the most important factor in reducing the QE. In verifying this we also tested SOM size, the number of training iterations (500000 to 10 million) learning rate the initialization of the reference vectors (random vs linear initialization). Like Sun, (2000) we found that a low update neighbourhood lowered the QE for a particular sized SOM. Additionally, we found making the training SOM size have the same number of nodes as the number of days whose circulation variables were being used as training data dramatically lowered the QE. However, the latter produces maps that are difficult to interpret.

Typically the dates used for creating the trained SOM did not completely span the full CFSR time period of 1979-2012. This presented the opportunity to evaluate the procedure. If the methodology is unable to identify the synoptic states of concern in the reanalysis data that lay outside the training period it is unlikely to do so in projected model data. So for each of the primary driving circulations, the full 30-year reanalysis data (CFSR) were mapped to the trained SOM and the Euclidean distances of the daily mappings compared to the threshold values of those from the trained SOM.

4.3 Results

4.3.1 Cut-off lows

Dates associated with extreme rainfall caused by COLs were obtained from Dr Alice Favre for the time period 1980-2006 (see Favre *et al.*, 2013). Two circulation variables that characterize these systems, 500 hPa geopotential height (Z500) and temperature at 500 hPa (T500) were extracted from the CFSR at 12z for each of 522 the extreme rainfall days. These data were then standardized to preserve gradients within the data after which a SOM was generated. A series of SOM sizes was tested and it was decided a 40-node SOM was optimal for this assessment given the relatively low number of events but high degree of specificity required to characterize the spatial distribution and depth of these systems.

The SOM categorized the available data according to spatial location and intensity of the system (Fig.4.1). On the left of the map are located deeper, well defined COLs and to the right less intense, shallower systems and coldest temperatures. It should be noted that although many of the nodes do not show closed low-pressure systems this is expected as a result of the SOM algorithm developing the most representative state of the data for each node.

Considering the rows of the SOM, the bottom row represents COLs with spatial locations to the west of the country. For each column, as one moves up the rows, the SOM represents systems whose spatial locations lie further east in the domain. For example, in the bottom row, node 4 represents COLs with centres to the west of the country and then in nodes 12, 20, 28, 36 the centre of the COL lies progressively more eastward until in node 36 the centre of the COL is over the Eastern Cape or KwaZulu.

The highest number of days in the record mapped to nodes 5 (23 days) and 16 (22 days), which represent COLs to the west of the country (with 500 hPa geopotential heights at about 5770 geopotential meters) and weaker COLs over the interior of the country respectively (Fig.4.2). Seasonally, a higher number of COLs associated with extreme rainfall occurred in the shoulder seasons (MAM-140; SON-152) than in summer and winter (DJF-140; JJA-110) and with the exception of SON, seasonal mappings are evident in the SOM.

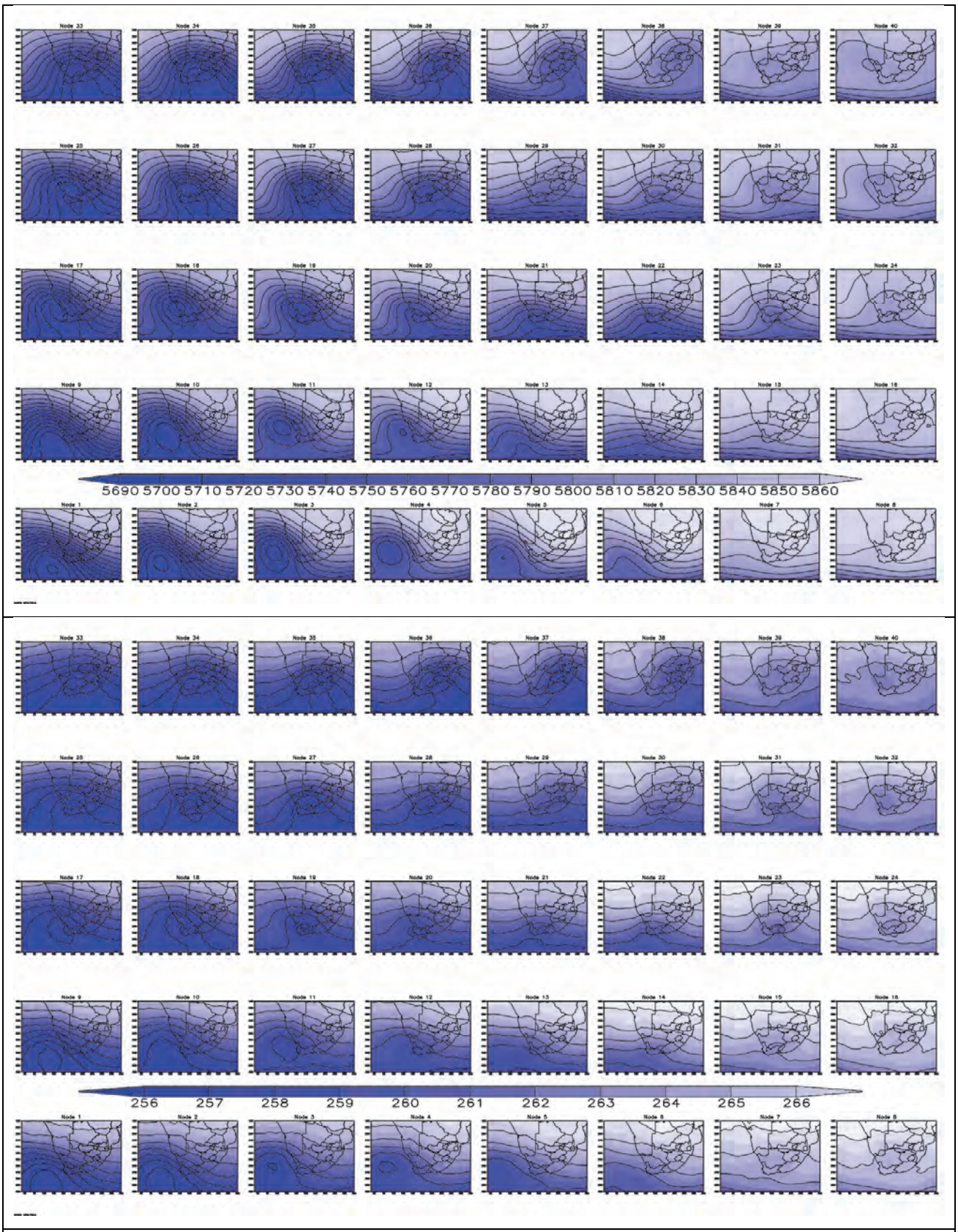


Figure 4.1. Archetypal maps for Z500 in geopotential height (top) and T500 in degrees Kelvin (bottom).

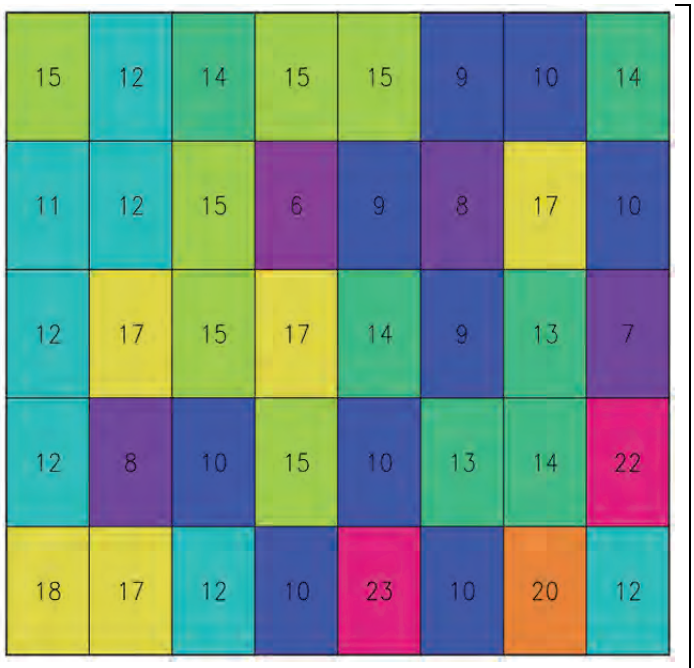


Figure 4.2. Frequency mapping of extreme rainfall days associated with COLs in number of days.

Extreme rainfall associated with COLs in summer (DJF) map frequently to the bottom right of the SOM which are representative of relatively, the shallowest and warmest systems (Fig.4.3, top left). This suggests the these types of COLs are associated with thermally driven convection and moisture transport from the sub-tropics as the Inter-tropical Convergence Zone (ITCZ) is located in its southerly-most position. The most frequently mapped to nodes (16, 7, 15 and 8, ordered by frequency) are adjacent nodes, which indicates the close proximity of these frequently mapped to states within the data space.

In winter (JJA) the most frequent mappings are to the top left and left of the SOM, which represent deep and relatively cold systems (Fig.4.3, top right). These circulation regimes are associated with the translation of westerly wave troughs. Nodes 33 and 25 are most frequently mapped to and represent COLs located over the interior of the country (as opposed to lying to the west of SA). There is very little overlap between summer and winter mappings indicating that systems are unique to each season and the associated seasonal dynamics associated can be further investigated. Another characteristic evident is that atmospheric thickness in JJA nodes (as indicated by the depth of blue in Figure 4.1) is much lower than in DJF nodes.

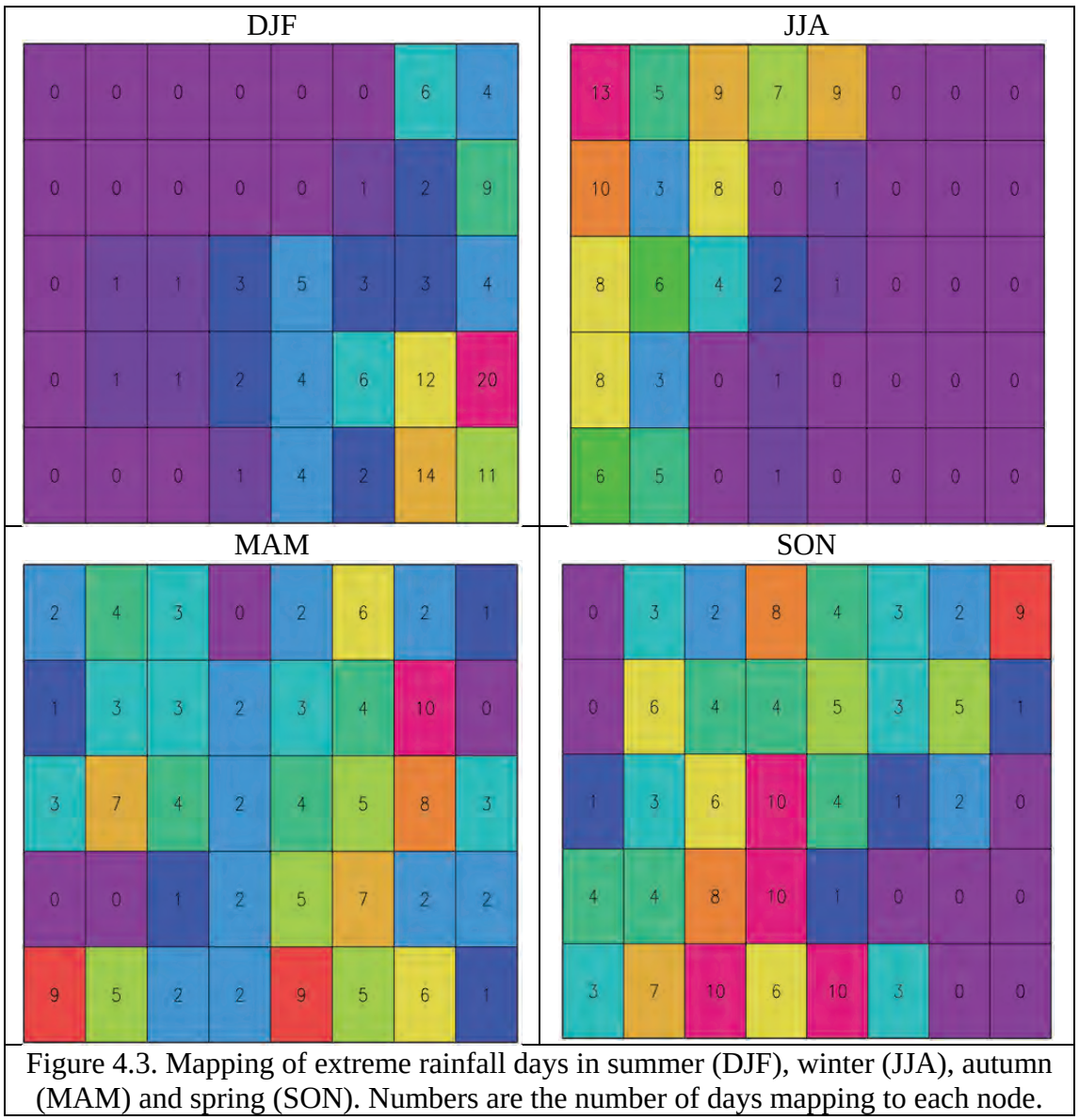


Figure 4.3. Mapping of extreme rainfall days in summer (DJF), winter (JJA), autumn (MAM) and spring (SON). Numbers are the number of days mapping to each node.

During autumn (MAM) there is no single location in the SOM that is most frequently mapped to and the most highly mapped to nodes (31, 1 and 5) are not often represented in summer and winter circulations (Fig.4.3, bottom left). The heterogeneous and unique SOM-space mapping is indicative of the transition from a thermally driven convective summer circulation regime with moisture sourced largely from the sub-tropics to winter states as a more westerly wave dominated regime. During spring (SON) a region near the bottom and centre of the SOM is frequently mapped to (nodes 3, 4, 5, 11, 12, 19 and 20; Fig.4.3, bottom right). These nodes represent circulations that are associated with westerly wave troughs of various depths over the western parts of the country, however, only 8 days overlap with core winter circulation patterns. An additional highly mapped to node in SON (40) is also associated with a COL over the eastern parts of the country, but the system is shallower and warmer. There is a relatively heterogeneous mapping to other nodes in SON, however, the core winter (33, 25) and summer nodes (7, 8, 15, 16) are not mapped to. This shows that in spring, which of the four seasons recorded the highest number of extreme rainfall causing

COLS, most of these systems are associated with the westerly wave troughs. However, these are generally weaker than core winter circulation patterns, likely a function of the southward movement of the westerly belt with the approach of summer, and located over the west of the country.

4.3.2 Mid-latitude cyclones

Six hundred and seventy days were associated with extreme rainfall caused by mid-latitude cyclones (MLCs) from 1979-2009. Sea-level pressure (SLP) characterizes these systems well and SLP data were extracted from the CFSR at 12z for each of the 670 extreme rainfall days. It was again decided a 40-node SOM was optimal for this assessment given the relatively low number of events but high degree of specificity required to characterize the spatial distribution and depth of these systems.

The SOM categorized the available data according to spatial location and intensity of the system (Fig 4.4). Generally, deeper systems are located on the left and shallower systems of the right of the SOM. As one progresses from the left to right side of the SOM frontal systems lie in a more easterly location and on the right of the map are either not evident or in the southeast corner of the maps. The two right-most columns of the SOM indicate a relatively higher pressure system ridging in behind the frontal system, however, the pressure associated with these is still low (103-104 hPa) and could either represent weaker fronts following the main front or a high-pressure system ridging in behind the front. In the first four columns of the SOM a tongue of relatively higher pressure is seen extending over the south and east coast. On inspecting the daily data a coastal low was often seen preceding the passage of the MLCs and it is likely this very shallow low-pressure system producing the pattern in the SOM.

The most frequently mapped to SOM node is node 33, which is associated with a deep MLC centre over the southeast of the domain and low-pressure region over the western parts of the country (Fig.4.5). The top row of the SOM is more frequently mapped to than the bottom, and moving from top to bottom the number of days mapping to each row are 149, 142, 134, 126, and 119 respectively. This indicates more extreme rainfall is associated with deeper MLCs in the southwest and south of the region and relatively higher pressure systems that follow these deep systems.

Different nodes of the SOM could be associated with extreme rainfall recorded at stations in particular regions of South Africa. For example, extreme rainfall recorded at stations in the Western Cape and south coast can be linked to the lower-right of the SOM, for instance nodes 7, 8, 15, 16 and 24. Similarly, extreme rainfall over the southern parts of the country and south-eastern coast

can be linked to the top-right of the SOM, including nodes 16, 23—24, 30—32, 38—40, while synoptic states associated with extreme rainfall over KZN and the Eastern Cape are linked to the top-left of the SOM, including for example, nodes 25-26 and 33-34. Thus changes in frequency mappings of particular nodes are likely to affect particular regions of the country.

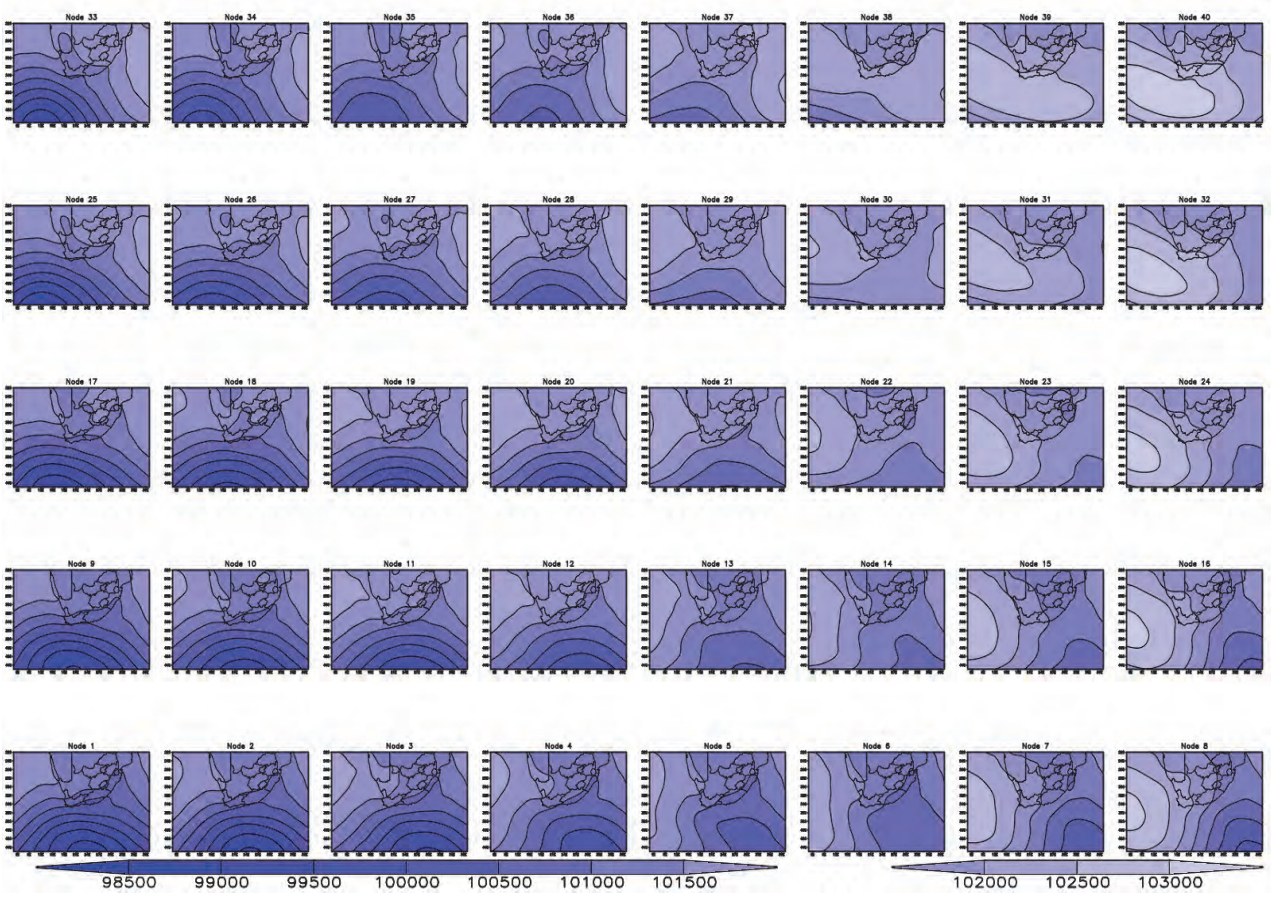


Figure 4.4. Archetypal sea level pressure mid-latitude cyclone circulations (hPa) associated with extreme rainfall.

The frequency mappings of each month were then computed. July experienced the highest number of extreme rainfall days (177) followed by June (137), August (130), May (119) then September (111). A higher number of July and August extreme rainfall days map to top three rows of the SOM and in May and September this is true for the bottom two rows of the SOM (Fig.4.5). With the exception of node 8, June events map very weakly to the right hand side of the SOM. This indicates in the core winter rainfall months (June, July and August) extreme rainfall is more often associated with deeper synoptic systems located on the top left of the SOM and in May and August with a larger, more heterogeneous range of synoptic states.

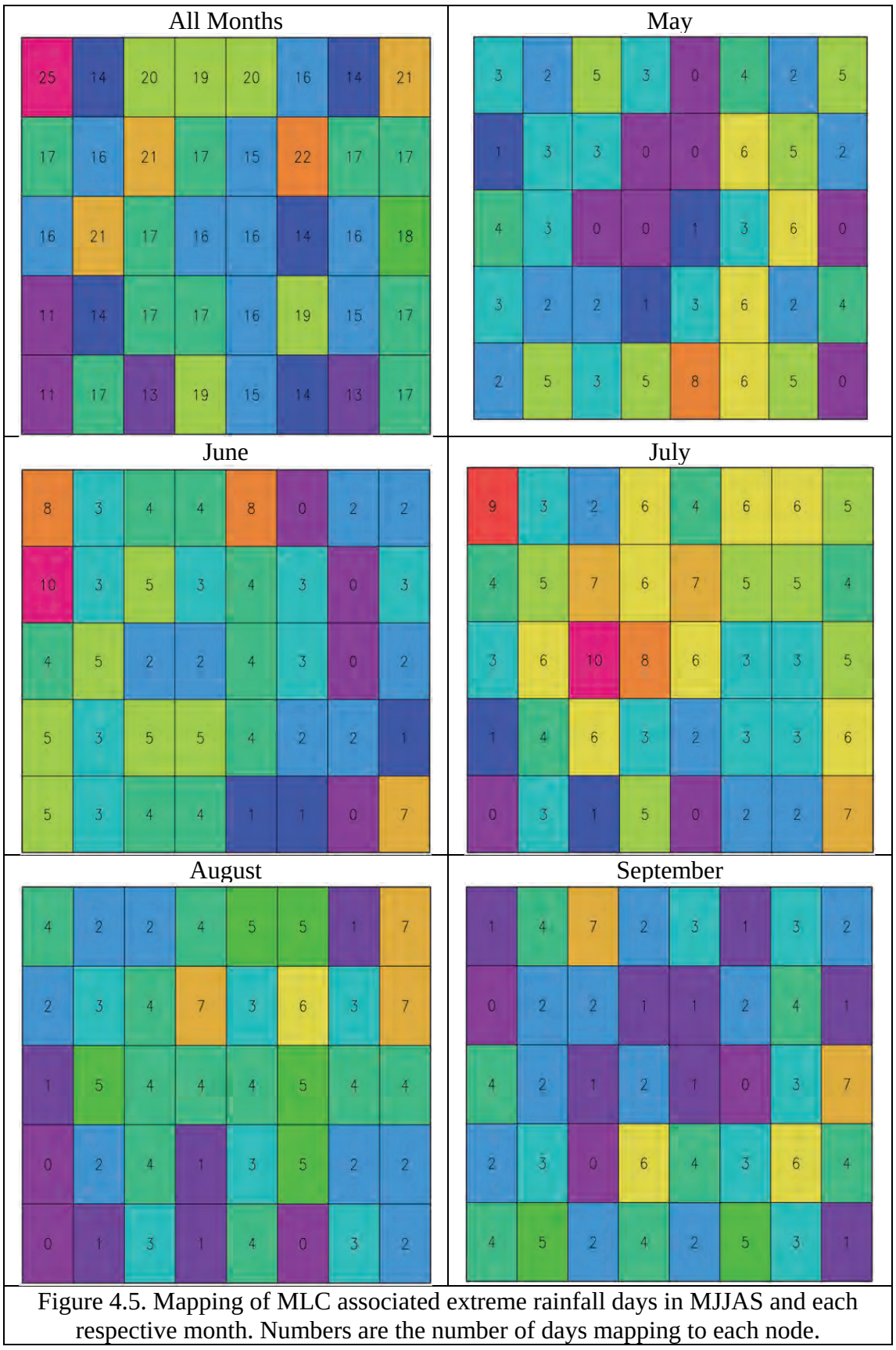


Figure 4.5. Mapping of MLC associated extreme rainfall days in MJJAS and each respective month. Numbers are the number of days mapping to each node.

4.3.3 Tropical Temperate Toughs

Only 45 TTTs associated with extreme rainfall in South Africa were obtained from Dr Neil Hart for the years 1980 to 1998, all of which occurred between the months of October and March. According to Hart *et al.* (2012) the most accurate and reliable discriminator of TTTs, especially those associated with extreme rainfall is cloudiness so the upward long wave radiation flux at the top of the atmosphere (or outward long-wave radiation- OLR) was extracted from the CFSR data and used to develop a SOM. The domain assessed stretched from Angola so that the effect of the Angola low could be captured to 45 degrees south to capture the interaction with the mid-latitude cyclones. A 12-node SOM was decided on as the small sample size meant if the SOM were too small, specific features of the characteristic patterns would be smoothed over. The SOM was developed using a learning rate of 0.1, a radius of influence of 2 and the number of iterations of 10000.

The SOM was able to characterize OLR patterns that show deep cloud bands stretching diagonally across almost the whole domain from north-west to south-east in most nodes (Fig. 4.6). The darker shades of grey indicate less OLR hence colder cloud tops implying deeper convective systems. On the right of the SOM, archetypal states are developed that indicate well defined tropical-temperate linkages in the cloud bands. Moving from right to left across the SOM bands become broader and less well defined and in nodes 11 and 12 the shape of the TTT is not well defined. In nodes 1 and 2 weaker linkages between regions of strong convection the tropics and mid-latitudes are apparent and moving up these columns shows stronger linkages. It should be noted here that the cloud bands are still evident, it is only the intensity that is varies.

The frequency mapping of the daily data was not homogenous across the SOM. Most days mapped to the top and bottom rows of the SOM, which indicate well-defined deep TTTs (top row) and weaker (nodes 1 and 2) and broader linkages (nodes 3 and 4) between the tropics and mid-latitudes (Fig. 4.7). Seasonally, 16 days were recorded in the three-month period October-December (OND) and 28 days in the season January-March (JFM). In OND 10 days mapped to the top row of the SOM and no days mapped to nodes 3 and 4. During JFM 16 days mapped to the bottom row of the SOM and 9 days to the top row. This suggests the SOM has been able to discriminate a seasonality in the type of TTT where mid- and late summer experiences wider types of cloud bands as the ITCZ reaches its most southerly location and is able to interact over a wider area with passing mid-latitude cyclones. This is also apparent in the monthly mappings where February, March and January recorded the highest number of TTTs per month in the 19-year period (11, 9 and 8 respectively) and of these had 6, 6 and 4 days map to nodes in the bottom row respectively.

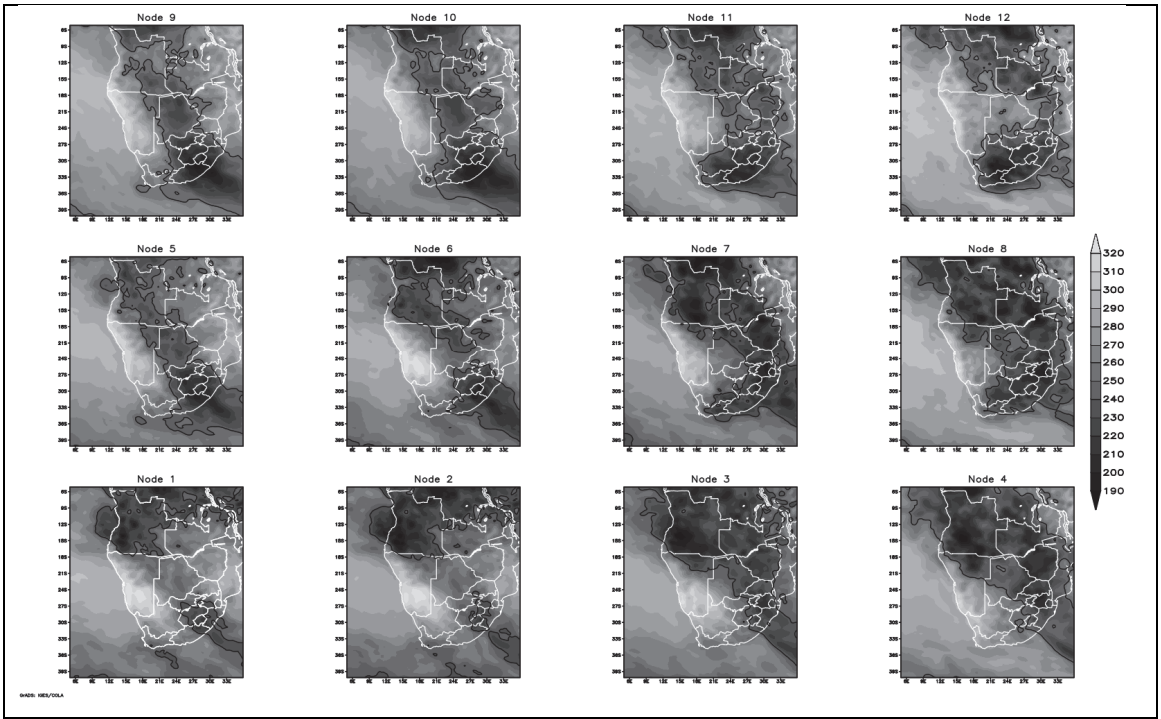


Figure 4.6. Archetypal cloudiness patterns based on OLR (W/m^2). Darker shades indicate less OLR (colder cloud tops) and hence deeper convective systems (see Hart *et al.* 2012).

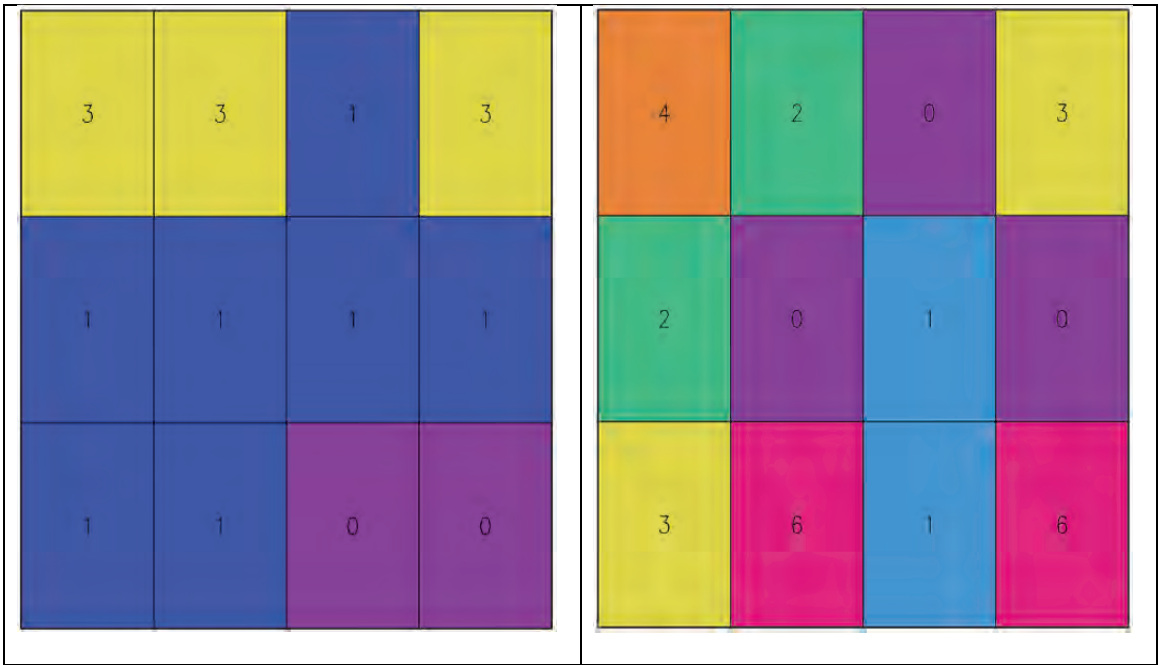


Figure 4.7. Seasonal frequency mapping of OND (left) and JFM (right) days that experienced extreme rainfall associated with TTTs in South Africa.

Hart *et al*. (2012) show that TTTs can interact with COLs to result in extreme rainfall. Dates from the COL and TTT records were examined and where they overlapped the node to which that date mapped to extracted. Seventeen days were identified and these mapped overwhelmingly to the top of the SOM, especially nodes 12 and 10 (Table 4.2). Three events were over consecutive days and these are indicated in bold. In one case the consecutive days mapped to the some node whereas the

other two mapped to different nodes. In the COL SOM ten of these days mapped to nodes with COLs to the west of the country (nodes 2, 5, 7, 12 and 14) and the remainder to nodes with COLs located over the western (nodes 18 and 26) or southern parts of the country. Although there was no particular area in the COL SOM these nodes mapped to, most TTT-COL patterns are associated with COLS to the west of the country with the convective activity being set up to the east of the COL.

Table 4.2. Node mapping of overlapping TTT and COL days. Bold dates indicate consecutive calendar days.

Date	TTT Node	COL Node
24Jan1981	2	22
25Jan1981	2	7
25Mar1981	1	14
28Nov1983	12	5
08Feb1985	2	31
09Feb1985	12	32
30Oct1985	10	26
31Oct1985	11	18
16Feb1988	12	24
19Feb1988	12	14
17Mar1991	7	24
16Dec1995	8	12
17Dec1995	10	12
25Mar1995	9	12
07Nov1996	9	2
26Mar1998	12	5
18Nov1998	10	12

The frequent mapping of TTT/COL days to node 12 likely explains the weak TTT shape in the OLR pattern represented in this node. The spatial location of the OLR minimum over the central-southern parts of SA matches the location of the COL location in nodes 31 and 39 in the COL SOM (see Fig. 2.1).

4.3.4 Quantization error method for identifying extreme rainfall synoptic environments

The Euclidean distances from the centroid of mapped days in the full 30-year period were comparable to the reference threshold distance established in each map. This meant that the

methodology could not discriminate the extreme rainfall causing synoptics from the respective variables used in the reanalysis data through an assessment of Euclidean distance. This was a disappointing result and although tested many times did not change. Tests included making the domain smaller to only encompass the region in which the particular synoptics occur, using coarser reanalysis (NEP2) and then using only one variable for each SOM. It seems that self-organizing maps trained on only particular types of synoptic circulations, cannot discriminate those types of circulations from a broader dataset such as reanalysis. It appears the noise in the data is too great to allow for the identification of any synoptic-specific signal.

4.4 Discussion and conclusions

Distinctive summer and winter COL circulations associated with extreme rainfall were identified – in the former deep cold westerly wave driven systems are characteristic whereas in the former systems are less deep and warmer suggesting a greater convective influence. In the shoulder seasons circulations are mapped to more heterogeneously although in spring extreme rainfall mappings are not associated with core summer circulations but with weaker mid-latitude systems. Extreme rainfall associated with MLCs generally mapped to deeper systems in the core winter season (JJA) than in the early and late season. In the latter, shallower systems and those following the passage of deep systems tended to be mapped to more frequently. Extreme rainfall associated with TTTs were summer bound but a much higher number of extreme rainfall events associated with TTTs occurred in the mid- to late summer (JFM) than in the early summer (OND). Early summer cloud patterns were characterized as both well-defined and ill-defined, the latter often being associated with COLs hence the heterogeneity.

Although the proposed quantization error method used to identify extreme rainfall causing synoptic environments was unsuccessful, the SOM is still a useful tool to disaggregate and identify particular synoptic states as well as their seasonality, e.g. deep and shallow cut-off lows. Given this, we propose that the more robust SOM methodology, similar to that of Lennard *et al.* (2013), can be adopted to assess projected changes in the synoptic drivers of extreme rainfall. In Chapter 5 a SOM is developed using the full set of reanalysis data and extreme rainfall days are mapped to the archetypal states of this SOM. We also use the SOM results to evaluate the available CORDEX regional climate models. Then in Chapter 6 we assess projected changes in the characteristics of extreme rainfall and relate these changes to changes in synoptic environments.

Chapter Five: Observed changes in extreme rainfall

5.1. Introduction

As mentioned in Chapter 2, statistically downscaled data (at the station scale) and dynamically downscaled CORDEX data are available to assess changes in the characteristics of extreme rainfall over southern Africa. However, before any assessment of projected change can be conducted, an assessment of historical change must be performed to develop a baseline from which any projected departure can be assessed against.

In this chapter we present trends in the frequency of occurrence of extreme rainfall events at 69 representative stations for the observational period obtained from the statistical downscaling. We also present the CORDEX RCA4 results, which consist of 8 downscaled GCMs and the downscaled ERA-Interim reanalysis to evaluate the ability of the models to replicate statistics of the observed climate.

5.2. Methods

5.2.1 Station data trend analysis

Trends in the frequency of occurrence of extreme rainfall days at each station are investigated using a bootstrapping methodology. For each station, the number of extreme rainfall events is counted for each year of the record and then a linear regression is fitted to the computed data to produce the observed slope of the trend (Fig. 5.1a). The yearly counts are then bootstrapped 1000 times and for each of the bootstrapped solutions a slope value is calculated (Fig. 5.1b). These 1000 replicates form a normal distribution for regression slopes from which the 5^{th} and 95^{th} percentiles are determined (Fig. 5.1c). The slope of the regression of the actual data is placed within the normal distribution and if it falls below the 5^{th} percentile or above the 95^{th} percentile the trend is considered significant (Fig 5.d – red dot is the actual regression slope value, so in this case the trend is not statistically significant).

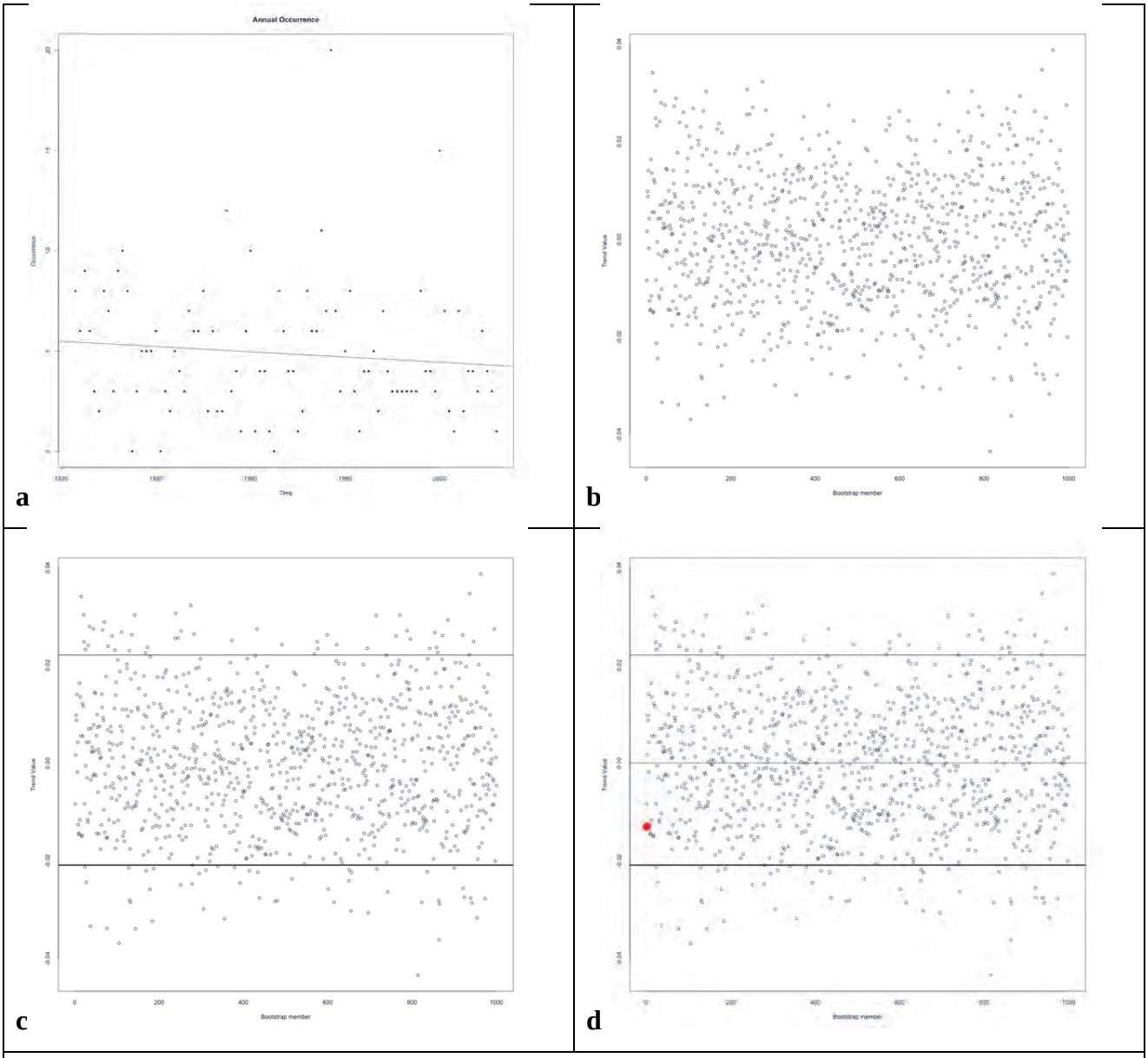


Figure 5.1. Bootstrapping methodology used to ascribe statistical significance to a trend.

This type of significance testing has been applied elsewhere (Chu and Wang, 1997; Sardeshmukh et al. 2000; Xu, 2006). The last paper cited also presents a number of other studies that use this type of methodology.

5.2.2 Evaluation of regional model data using ETCCDI indices

The results from the numerical downscaling first had to be evaluated to determine the reliability of the RCA4 model. The ensemble average of the 8 downscaled GCMs was compared with the gridded observation datasets of the monthly Global Precipitation Climatology Centre (GPCC) for the period 1981-2005. This record was used to evaluate the ability of the regional model to produce a realistic rainfall climatology in the downscaling.

However, extreme rainfall usually occurs at the daily to weekly scale therefore a daily observational dataset is necessary. The Global Precipitation Climatology Project (GPCP) is a daily dataset but only covers a relatively shorter period from 1997-2005. We used this dataset to test the ability of the RCM to capture extreme rainfall statistics over the common 1997-2005 period. A non-parametric approach, which uses indices of extreme rainfall defined by the Expert Team on Climate Change Detection and Indices (ETCCDI) of extreme rainfall, has been adopted to assess the characteristics of extreme rainfall. The non-parametric approach used extreme rainfall indices from the ETCCDI that consider absolute values, duration, thresholds and percentiles of extreme rainfall (Table 5.1).

Table 5.1. ETCCDI extreme rainfall indices

Index parameter	**Abbreviation**	**Units**
Total wet-day precipitation	prcptot	mm
Consecutive dry days	CDD	days
Maximum 5-day precipitation amount	RX5day	mm
Number of heavy precipitation days	R10mm	days
Number of very heavy precipitation days	R20mm	days
95th percentile of precipitation on wet days	R95pTOT	mm
99th percentile of precipitation on wet days	R99pTOT	mm
Simple daily intensity index	sdii	mm

5.2.3 Evaluation of regional model data through a synoptic environment assessment

Sea level pressure and wind and humidity fields at the 850 hPa level were used to assess the models' ability to reproduce a physically consistent regional circulation patterns. A SOM was developed using the ERA-Interim reanalysis data (Dee *et al.*, 2011) for the period 1989-2005 to explore the circulation patterns responsible for rainfall and extreme in Southern Africa. The ERA-Interim reanalysis (EI) is used instead of the CFSR because this is the reanalysis data used for the CORDEX evaluation period and as the SOM being developed is a more general SOM, the higher resolution of the CFSR is not essential. Variables used to develop the SOM were mean sea level pressure (MSLP), U- and V- wind components and specific humidity at the 850 geopotential height level. The MSLP is useful for the description of synoptic states and the latter 3 for information on moisture transport. The domain covers southern Africa from 9°S to 44°S to include synoptic weather system of both tropical and mid-latitude origins from 6°E to 45°E and encompasses processes from the oceans around the sub-continent. The choice of the number of SOM patterns is subjective; the fewer the number of nodes in the SOM array, the more general each pattern will be, while, with a greater number of nodes, a wider range of circulations can be represented. Different

SOM sizes were assessed in order to find a reasonable compromise between detail and interpretability of the SLP patterns. A 20-node SOM patterns (5 x 4) was selected to span the range of synoptic conditions for all seasons.

CORDEX data from two RCMs (RCA4.5 and the CCLM) which each downscaled four GCMs (CNRM-CM5, EC-EARTH, MPI-ESM-LR, HadGEM2-ES) as well as the EI were used in this analysis. This resulted in a 9-member ensemble whose data could be mapped to the EI-trained SOM. Through doing this is it possible to (1) evaluate the ability of the models to produce the historical circulation characteristics of between 1989-2005, (2) evaluate the downscaled GCM historical runs between 1976-2005 and (3) assess potential changes in circulation states in the mid- and late 21st Century.

The GPCP rainfall data were used to investigate rainfall over the region as it has been the most widely adopted rainfall dataset within the CORDEX-Africa framework and is representative of station data in South Africa and Mozambique (Kalognoumou *et al*., 2013). Extreme precipitation days are defined as days where at least 10 % of the grid points in the domain have precipitation above the 95th percentile of daily rainfall distribution from 1997-2005. Extreme precipitation days are extracted from GPCP gridded data set and mapped to the trained SOM to identify synoptic conditions associated with extreme precipitation. Three sub-regions are selected based on Global Precipitation Climatology Centre (GPCC, Rudolf *et al*., 2010) standardized annual precipitation cycles over Africa, which are distinguished into nine classes using the k-means method (Kalognoumou *et al*., 2013). The location of the sub-regions and their topography is shown in Fig. 5.2. The sub-regions are the centre and northern Mozambique, Zimbabwe, Zambia, northern Botswana and east parts of Angola (region 1), southern Mozambique, central and east coast of South Africa (region 2) and the south-western Cape of South Africa (region 3). Region 1 experiences rainfall in summer that results from mesoscale convective systems, tropical storms and tropical cyclones from Mozambique Channel (MC). Region 2 is also a summer rainfall that results from mesoscale convective systems, warm fronts, subtropical lows, mid- to upper-tropospheric troughs, and cloud bands. Region 3 is a winter rainfall region that receives the bulk of its annual rainfall from frontal systems.

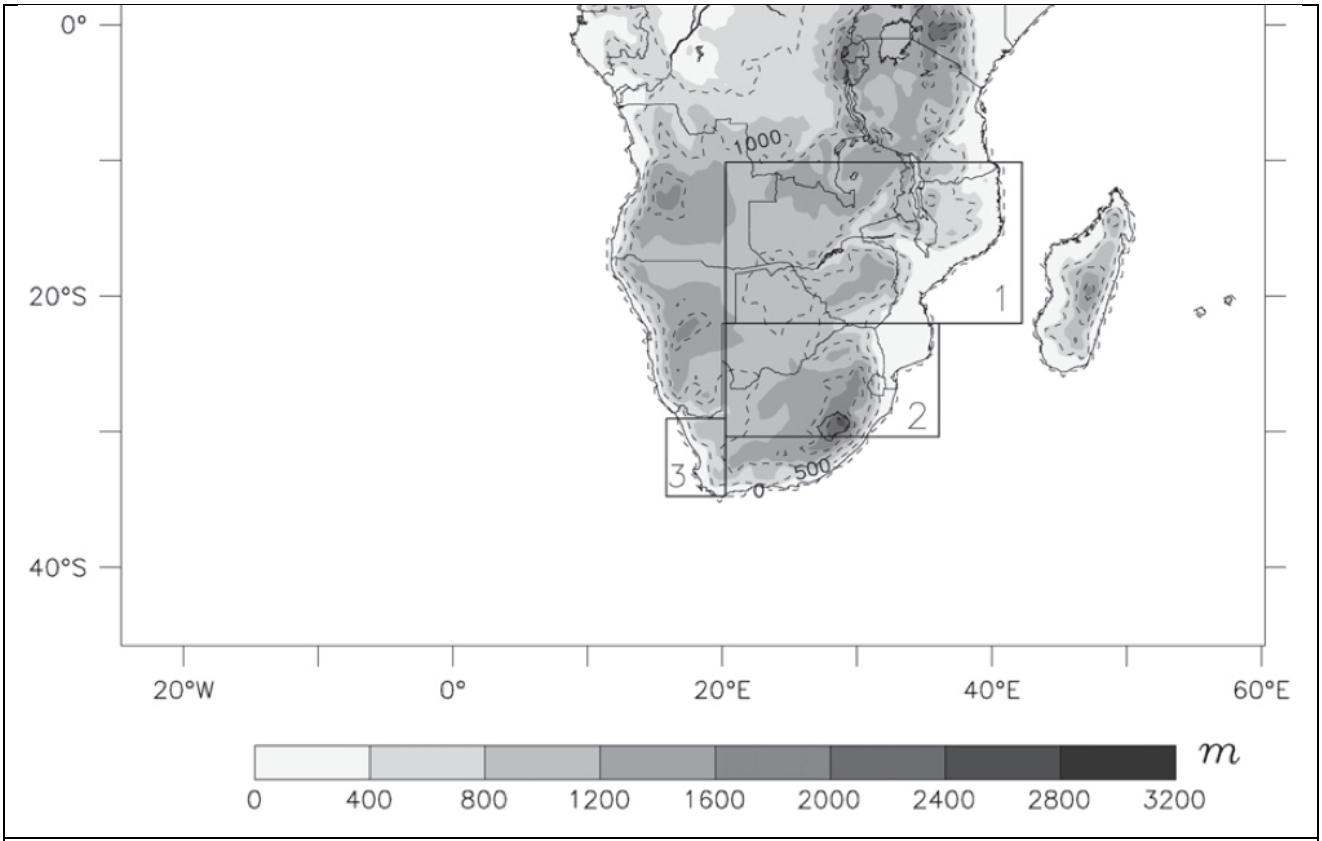


Figure 5.2. Sub-regions used in the analysis that span tropical (region 1), sub-tropical (region 2) and mid-latitude region (region 3). From Kalognoumou et al. (2013).

5.2.4 Assessment of rare events using extreme value theory

The extreme value theory (EVT) approach has been used in hydrology (e.g. Katz et al., 2002), atmospheric science (e.g. Palutikof *et al*., 1999; Buishand, 1989), finance and insurance (e.g. Embrechts *et al.*, 1997) and many other fields of applications. In this study the EVT approach follows that used in Zwiers and Kharin (1998) and Kharin and Zwiers (2000). The diagnostics of primary focus is to estimate extreme events in terms of T-year return values. The T-year return value is defined as the threshold that is expected to be exceeded on average, once every T years, where T is the return period or waiting time. These thresholds are estimated by the block maximum technique (Coles, 2001), where a generalized extreme value (GEV) distribution is fitted to a sample of annual/seasonal maxima at every grid point. It is important to test the goodness-of-fit (GOF) of the fitted distribution to examine whether the sampled annual maxima are realizations of a random process with the GEV distribution and the standard Kolmogorov-Smirnov (KS) GOF is applied for this.

5.3 Observed changes of extreme rainfall in the station record

The example presented in Fig. 5.1 uses a station with records from 1920, however, not all of the 69 representative stations used in this study have such a long record. Length of record varies from as long as 130 years to 34 years. Two trend assessments were therefore performed, one that used the

full record of each station and a second that only considered data between 1979-2009. Considerably different trends were obtained in terms of statistical significance and regionally in terms of direction of the trend (Fig. 5.3). The trends computed using the full station record show a greater number of statistically significant trends than the truncated dataset. However, with the exception of the Western Cape and the east coast, the patterns of trend are similar between datasets.

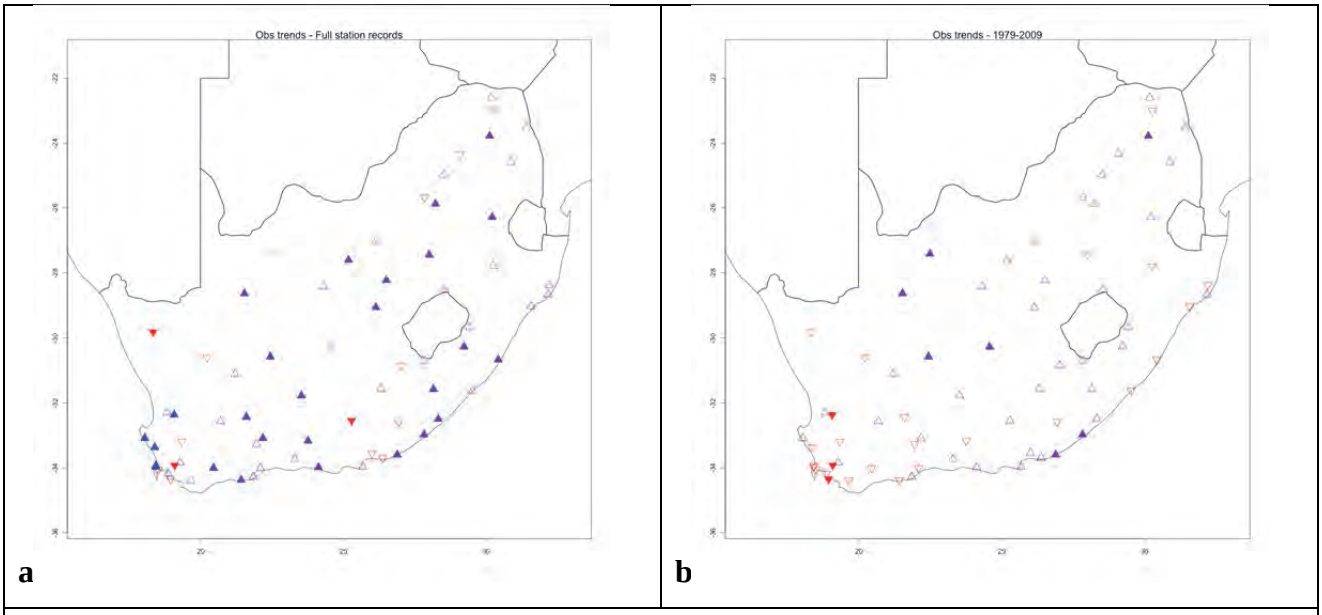


Figure 5.3. Trends in the frequency of occurrence of extreme rainfall at observation stations data for (a) the full station record and (b) from 1979-2009. Blue (red) triangles indicate positive (negative) trends and solid triangles where the trend is statistically significant.

We would suggest the results of the full-length station records are more appropriate to consider than the truncated period and that there has been a general increase in the frequency of occurrence of extreme rainfall as defined by the 95th percentile over most of the country.

5.4 Evaluation of CORDEX dynamical regional models using ETCCDI indices

The comparison between the GPCC data and the downscaled GCM data shows that although there is variability between the different GCM members, the downscaling captured the general spatial rainfall patterns in the ensemble mean between 1981-2005 (Fig. 5.4). The top map is the GPCC data, the 2nd row left-most figure the ensemble median and the remaining figures the individual downscalings. One exception is the Lesotho Highlands and the Drakensberg north of here which has a consistent wet bias in all ensemble members. However, it should borne in mind that this gridded dataset is based on gauge data and there are very few stations in many regions, especially in mountainous areas, so observation errors are likely high in these regions. This result, and others from CORDEX (e.g. Endris *et al.*, 2013; Kalagnoumou *et al*., 2013), demonstrate the downscaled rainfall results can be used for the assessment of potential change into the future.

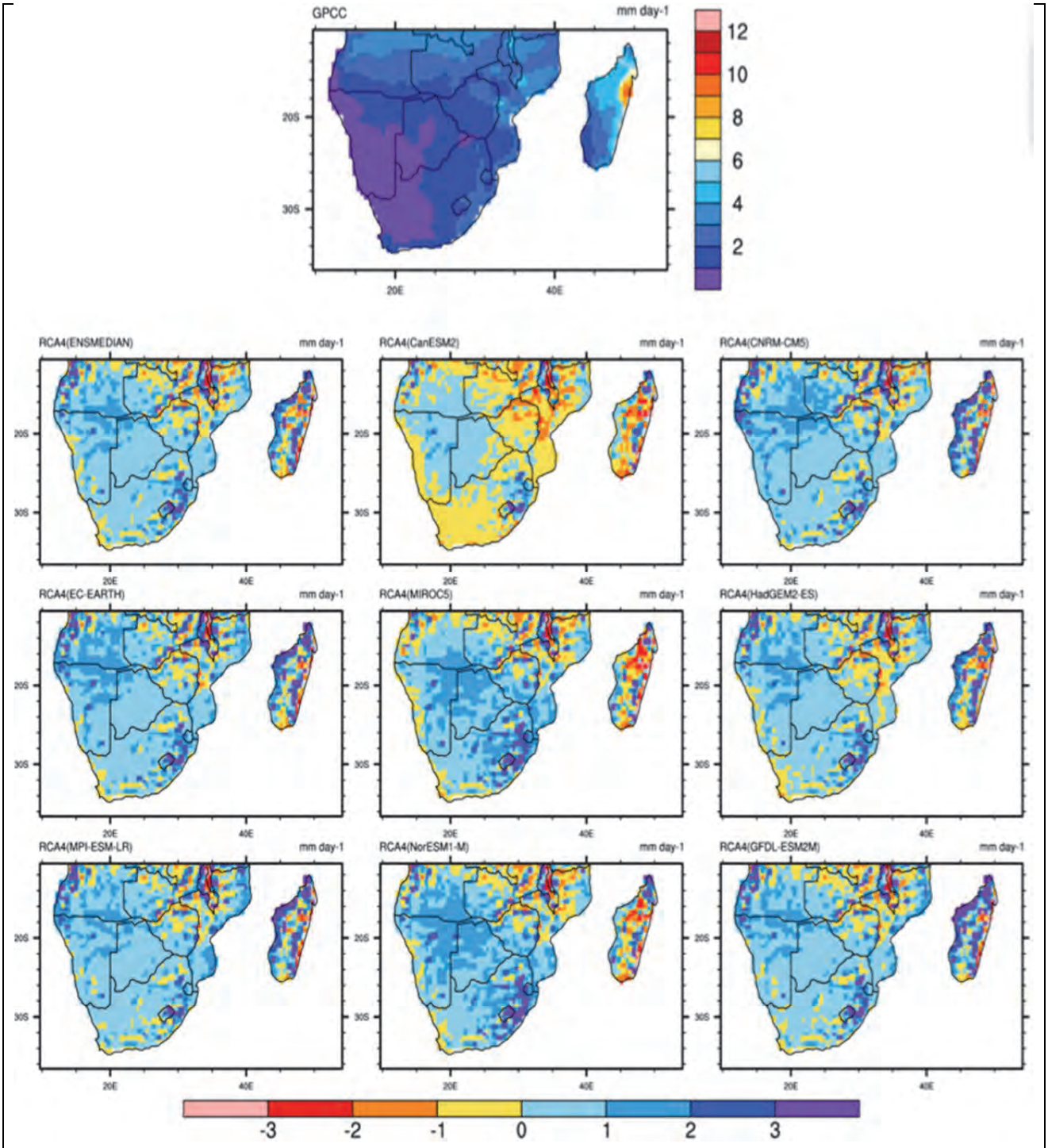


Figure 5.4. Comparison in rainfall (in mm per day) of the 8 RCA4-downscaled GCMs with the GPCC dataset.

The ETCCDI indices were computed for the GPCP and downscaled GCM data between 1997 and 2005. These indices were also computed for an evaluation run of the RCM which downscaled ERA-Interim reanalysis data to assess any bias of the RCM in the context of extreme rainfall. Figure 5.5 presents the indexed results for the evaluation period 1997-2005 of the GPCP data, the downscaled ERA-Interim simulations and the ensemble median of the eight downscaled GCMs. Each row is a particular index listed above. Some noteworthy points are:

1. The total precipitation (prcptot) of both downscaled results (ERA-Interim and GCM ensemble average) generally match the pattern of the GPCP data over the northern and western parts of the domain as well as over Madagascar. However, over Lesotho the model produces higher than observed values which was also seen in the GPCC results.
2. The model produces many more consecutive dry days in the drier north-western part of the domain in the ERA-Interim downscaling. The GCM ensemble mean does not reflect this, however, it does not capture the local maximum over northern Namibia.
3. The model produces too much extreme rainfall over the Mozambique Channel and Madagascar compared to the GPCP (see the rx5day, R95Ptot, R99Ptot). The GCM downscaling again has closer spatial representativity than the ERA-Interim downscaling, likely because this is an ensemble median which when computed results in error cancellation.
4. Heavy rain days (r10mm and r20mm) produced by both the ERA-Interim and GCM downscalings have similar spatial patterns to the GPCC data, although there are higher localised maxima around topography.
5. The downscaled simple daily intensity index (sdii) values are also similar to the GPCP data except over the Mozambique Channel.

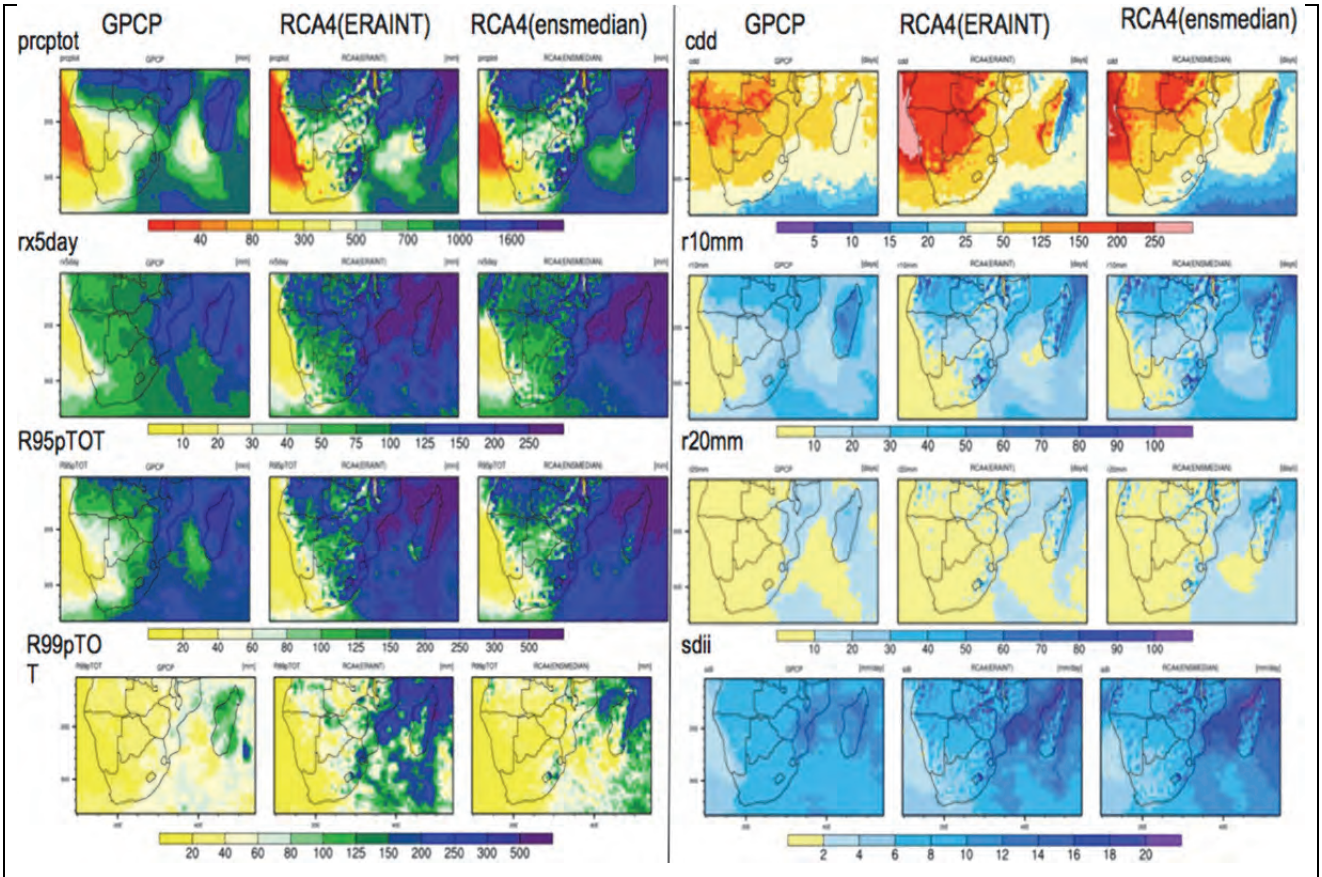


Figure 5.5. Observed and downscaled ETCCDI indices over southern Africa. Indices are listed in Table 5.1

5.5 Rare extremes in observed and model data

Spatial patterns of the estimated 20-year return value for GPCP, TRMM, both RCMs driven by the ERA interim and ensemble mean of both RCMs driven by GCMs is shown in Fig. 5.6. The pattern correlation coefficient of 0 indicates a no level of agreement between GPCP and TRMM datasets. The magnitude of the 20-year return value is generally higher in the TRMM than in GPCP. The spatial pattern from the downscaled models show a complex structure defined by local topographical conditions. The Kolmogorov-Smirnov goodness-of-fit test has been applied to verify the accuracy of the GEV fits. At the 5% significance level, no grid box GEV distributions are rejected. This indicates that the GEV distribution is a reasonable approximation for the distribution of annual precipitation maxima at each grid box. The RCA4 forced runs are in better agreement with the observed 20-year return value estimated from the GPCP and TRMM data compared to the CCLM runs and this is also seen in the GCM downscalings. The RCMs consistently simulate maximum 20-year return values to the east of southern Africa, whereas minima occur to the very dry region to the west. Each model shows a coherent spatial distribution of precipitation extreme and there is minimal spread between the members of the ensemble driven by the same RCM indicating that rare extremes are mostly controlled by model physics. The CCLM simulates maximum 20-year return values over the north of the continent, which is not seen in the GPCP and TRMM data or the RCA4 downscaling. The pattern correlations are generally low suggesting that the spatial distribution of rare extremes is not well captured by the models.

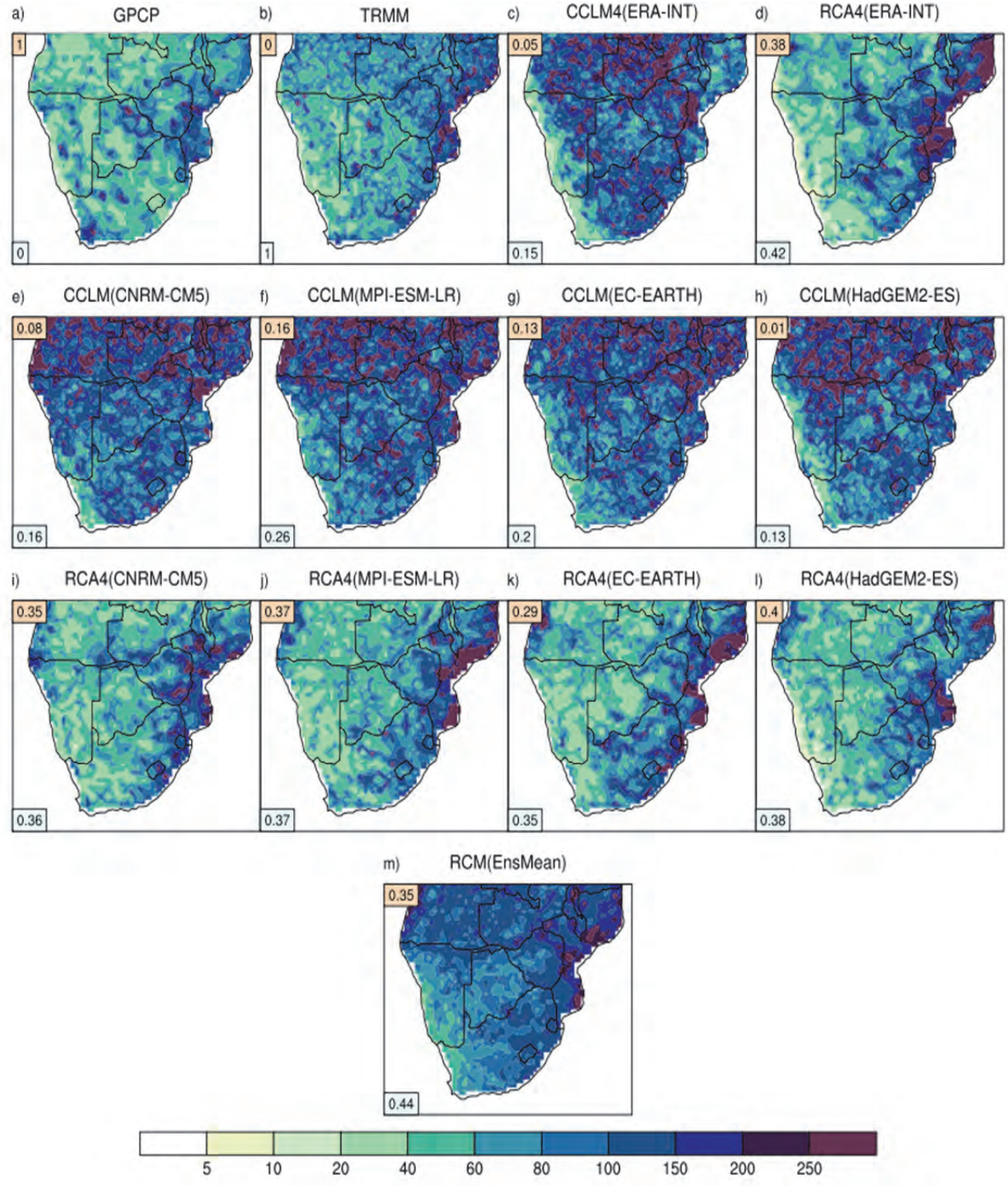


Figure 5.6. 20-year return values of annual maximum daily precipitation (P_{20}) for the period 1997-2006 for (a) GPCP, (b) TRMM (1998-2005), CCLM forced by (c) ERA-Int. and different GCMs (e-h), RCA forced by (d) ERA-Int. and different GCMs (i-l) and (m) multi-model ensemble mean of CCLM and RCA4 forced by GCMs. Pattern correlation coefficient relative to the GPCP (TRMM) is in the top (bottom) left corner of each map.

5.6 Synoptic circulations in the 20th century simulations

The 20 SOM patterns for the full 1989-2005 period of the EI reanalysis is shown in Fig. 5.7. The shaded region is MSLP and arrows indicate the moisture transport over the region and are derived

from the U-, V-wind components and specific humidity at 850 hPa. In order to construct this plot, each day in the period was mapped to the most similar SOM node to construct a composite of moisture transport. Similar synoptic circulation patterns are clustered together on the map while more distinct types are further apart as a consequence of the SOM mapping algorithm (Hewitson and Crane, 2002). Features of the low level circulation can be seen through the MSLP patterns – these include the South Atlantic and Indian anticyclones, mid latitude westerlies and the northeast monsoon. The types displaying the tropical troughs are in the lower portion of the SOM and types with broad regions of high pressure and mid-latitude troughs are near the top of the SOM. In the two bottom rows of the SOM, nodes represent strong anticyclones ridging in over the east or west coast with a dominant South Atlantic High Pressure (SAHP), and tropical troughs extending down from the equator. (Mid-latitude cyclones are displaced southward compared to nodes 1, 2, 6, and 7.) These nodes represent circulation patterns typical of summer (DJF) over the region. The moisture is transported into southern Africa from the southwest Indian Ocean (IO) and from the southeast tropical Atlantic Ocean (AO) (Reason et al., 2006) through the Angola low (Cook et al., 2004). The westerly moisture from the AO converges with a warm moist air from the IO in the heat low located over the interior of southern Africa, which result in ideal conditions for strong vertical uplift and the formation of the cloud bands. These features are commonly referred to as tropical-temperate troughs (TTTs) and contribute significantly to summer rainfall over southern Africa (Harangozo and Harrison, 1983) including heavy rainfall events (Hart et al., 2010). Circulations represented by node 12 and 17 suggest a linkage between the mid-latitudes and subtropical low-pressure systems and may represent conditions favourable for the formation of TTTs. The high-pressure system located in the south-west of the domain results in a poleward displacement of frontal systems and dry summers over south-western part of the Western Cape Province of South Africa.

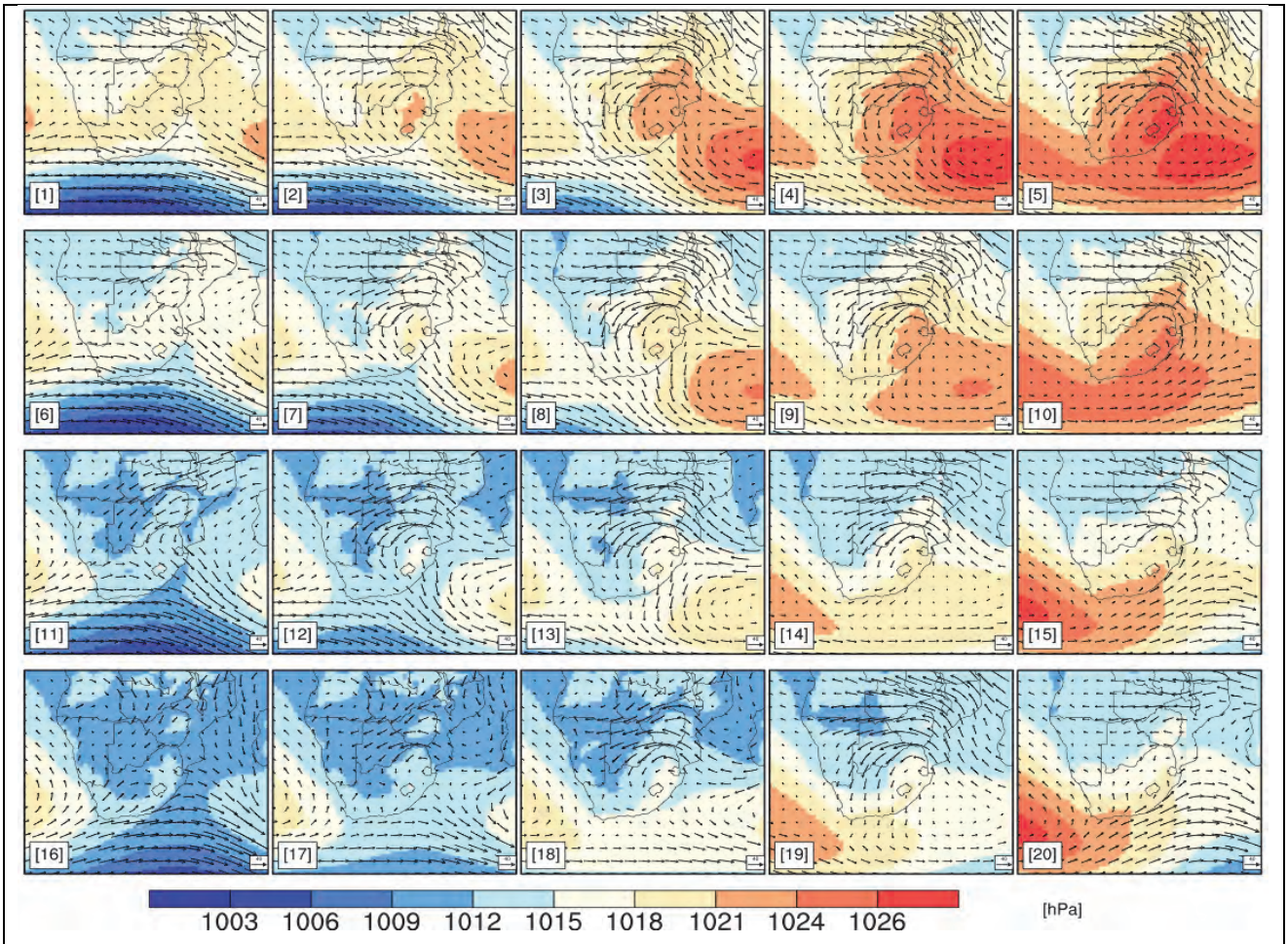


Figure 5.7. Self-organizing map produced using ERA-Interim MSLP, U-, V-wind and specific humidity at 850 hPa level.

Typical winter (JJA) circulation patterns for the region are located at the top rows of the SOM map. These nodes are associated with ridging highs from the AO, strong high pressure over the IO and the strong westerly wave with associated cold frontal systems. These cold fronts are the primary rain-bearing system along the southern coast and the south-western part of the Western Cape Province of South Africa (Tyson and Preston-White, 2000). The moisture flux in this region is westerly, in contrast to the easterly flux experienced across the bulk of the subcontinent. During autumn (MAM) and spring (SON) transitions periods, synoptic situations typical of either winter or summer can occur.

5.7 Evaluation of CORDEX RCMs using the synoptic environment approach

Having derived characteristic circulation patterns from the reanalysis data, the same variables that were used to train the SOM were extracted from each RCM simulation and mapped to the trained SOM. The evaluation downscalings, where each RCM downscaled the EI reanalysis data for the period 1989-2005, most node mapping frequencies are similar to that of the EI reanalysis (Fig. 5.8). There are four exceptions – node 5 is over-simulated and nodes 6 and 11 are under-simulated by both models.

In nodes where there are notable differences between frequency mappings of the two RCMs in the EI downscaling, the difference in pattern is repeated in the CNRM, MPI and HadGEM GCM downscalings but is switched in the EC-Earth (ECE) downscaling. For example, in node 18 the CCLM-EI downscaled mappings are higher than the raw EI mappings and the RCA4-EI downscaled mappings lower. This pattern is repeated in all the downscaled GCM results except the EC-Earth results where the pattern is reversed. This suggests that the regional models converge results toward their internal climate in the southern African region under this particular synoptic environment.

In nodes where there are not large differences between the downscaled mapping frequencies of each GCM-RCM combination, e.g. in nodes 8-10, 12-15, 19-20 the mapping frequency of a GCM downscaled by CCLM is similar to the same GCM downscaled by RCA4. This suggesting that in terms of the simulation of these particular synoptic states, the GCM state is propagated into the domain by the RCMs. These nodes represent relatively weaker synoptic environments over southern Africa compared to other nodes where there were large mapping frequency differences between the RCM-GCM downscaling combinations. The latter case is particularly evident in node 5, which represents a strong high-pressure system over most of the region, that is systematically more frequently mapped to by the CCLM-GCM downscalings than the RCA-GCM downscaling mapping. Mapping frequency differences between the RCM-GCM downscaling combinations can also be seen to lesser degrees in nodes 6, 7, 16, 17 and 18. However, the opposite pattern in the EC-Earth downscaling also suggests that the boundary conditions of the GCM can determine the trajectory a RCM will take.

In node 16 (and to a lesser degree nodes 17 and 18), downscaled GCM mapping frequencies are higher than for the EI data and the EI downscaled data for both RCMs, implying that the GCMs simulate this type of circulation too frequently. These are all nodes that are mapped to most frequently in summer (Fig. 5.9). Furthermore, in the left-most column of the SOM, evaluation (EI) downscaling frequency mappings are lower than the original EI frequency mappings but most of the GCM downscaling frequency mappings are equal to or higher than the EI frequency mappings. As in the case of the EC-Earth model above, this again suggests that the boundary conditions of the GCM can determine the trajectory a RCM will take, depending on the type of synoptic environment.

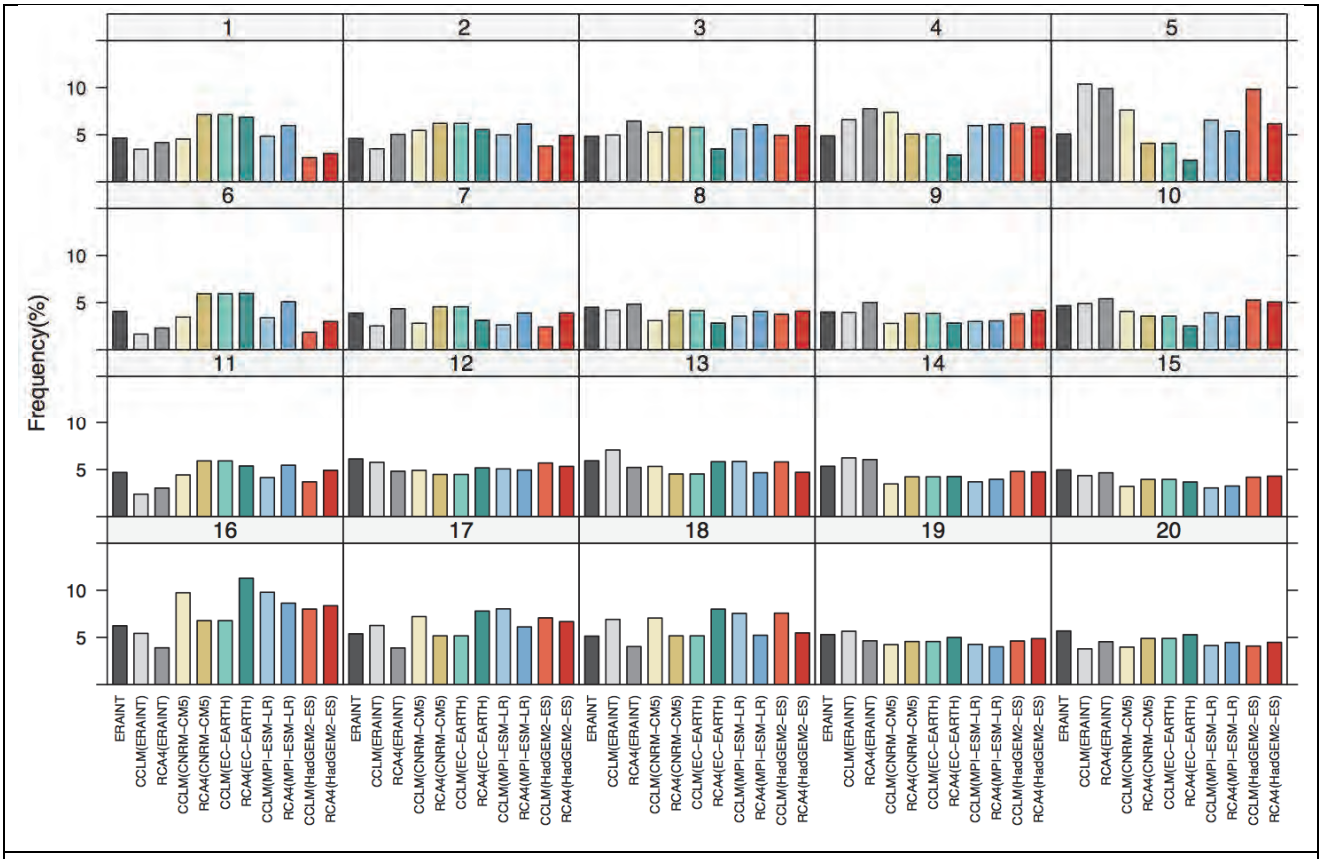


Figure 5.8. Frequency mapping of the ERA-Interim, downscaled ERA-Interim (1989-2005) and downscaled GCM (1976-2005) data. Black bars are the raw ERA-Interim mapping frequencies, the light and dark grey bars the mappings of the downscaled ERA-Interim simulations for the CCLM and RCA4 respectively and the rest of the bars are the mapping frequencies of the downscaled GCM by each RCM as indicated.

Seasonally (Fig. 5.9), nodes in the bottom row of the SOM (11-14, 16-19) are mapped to more frequently by the models during DJF (Fig. 5.9a) and the nodes in the top row and right hand side (1-6, 10-16) appear more frequent during JJA (Fig. 5.9c). During transitions periods synoptic circulation of both winter and summer are evident, however with lower frequencies (Fig. 5.9b and 5.9d). The bias in the EI downscaled results in node 5 is attributable to an over-simulation of winter-time circulations dominated by high-pressure synoptics which may be a function of the southernmost extent of the CORDEX-Africa domain. Here the minimum latitude of the domain is 45^{o} south and would likely inhibit the ability of the regional models to develop mid-latitude troughs that move northward through this latitude (see also Kalognoumou et al. (2013) who describe this problem). This problem may also be a factor in the left-most column of the SOM that have characteristically deep mid-latitude cyclones to the south of the domain. The lower frequency bias in nodes 11 and 16 seem to be present throughout summer, autumn and spring. During DJF the majority of the downscaled GCMs overestimate the frequency of nodes 16, 17, 18 and 19 and underestimate nodes 12,13 and 14. During SON and MAM the inter-model variability of the downscaled GCMs is smaller compared with JJA and DJF.

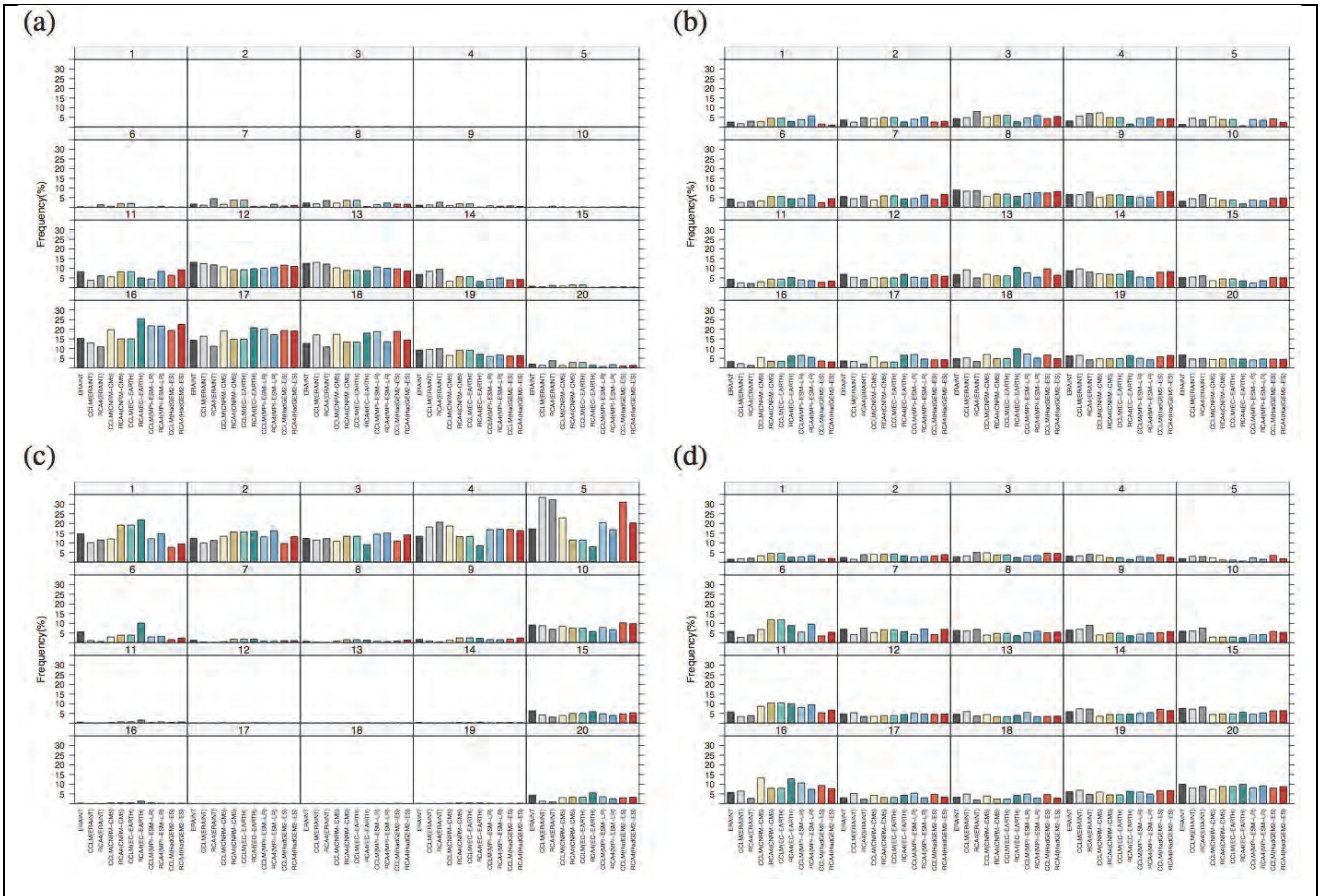


Figure 5.9. Seasonal frequency mapping of the ERA-Interim, downscaled ERA-Interim (1989-2005) and downscaled GCMs (1976-2005) where (a) is DJF, (b) MAM, (c) JJA and (d) SON.

5.8 Rainfall characteristics of archetypal circulation environments

In order to identify which circulation nodes could be associated with rainfall, each day of the observed period was mapped to the closest SOM node to construct a composite of rainfall for each SOM node (e.g. Hope et al., 2006). The daily mapping was performed for the period 1997-2005 and the corresponding rainfall composited from GPCP daily rainfall data. The rainfall anomalies from this composite mapping are shown in Fig. 5.10. Brown shading indicates regions of less rainfall compared to the 1997-2005 mean while green shading indicates more. The nodes 11-14 and 16-18 in the SOM are clearly responsible for most of the wet conditions over the bulk of the sub-continent, and most of this rainfall occurs during summer. Highest rainfall occurs over northern Mozambique and is attributable to the tropical low formed just off the north coast of Mozambique between 15° and 18° S that advects moisture into the region as demonstrated in nodes 17 and 18. The winter rainfall is largely associated with nodes 1-3 and would represent frontal rainfall.

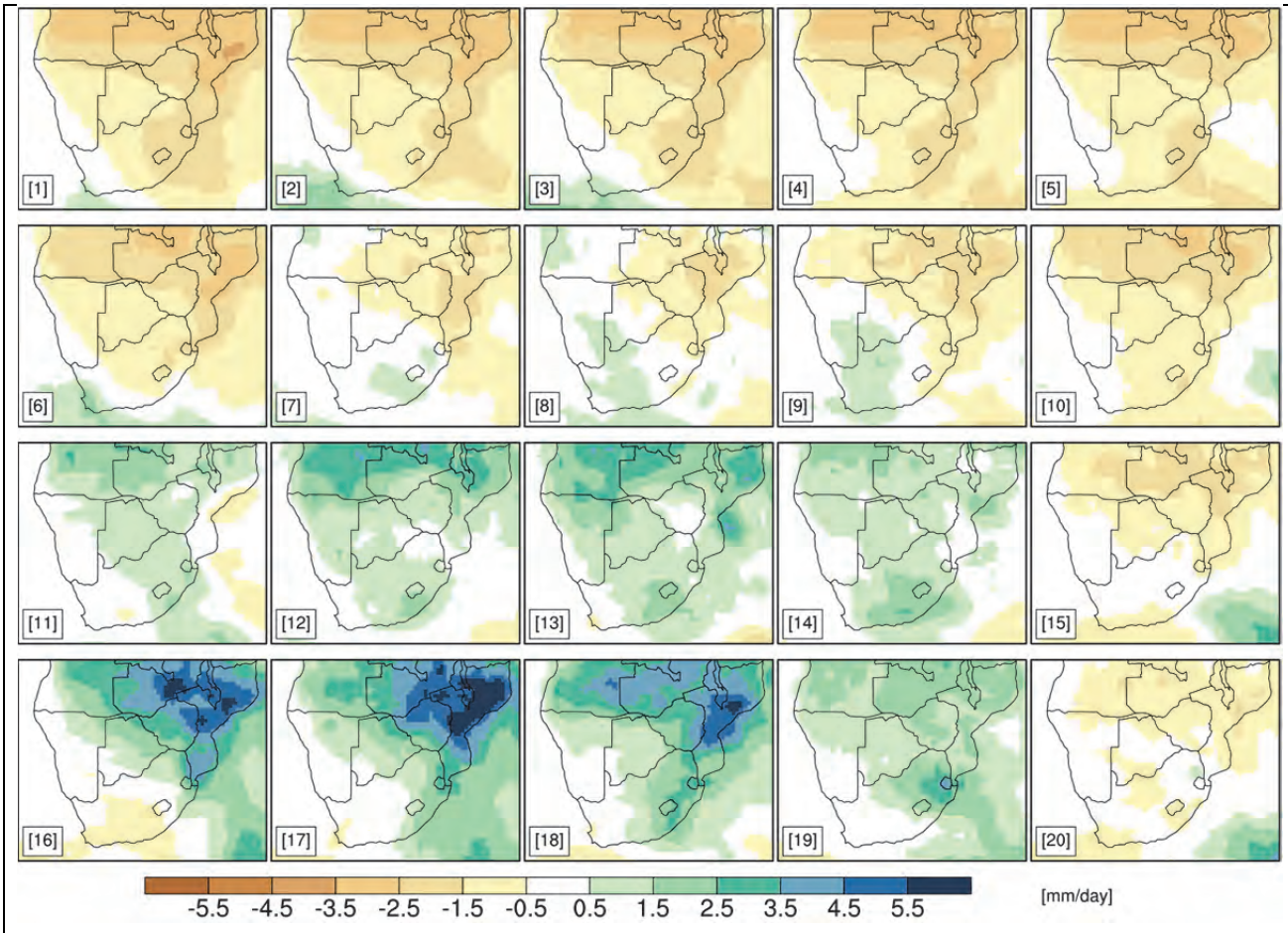


Figure 5.10. GPCP precipitation anomalies associated with each circulation type in the 5x4 SOM. Negative values indicate conditions drier than the 1979-2005 mean and positive values wetter than the mean.

5.9 Synoptic environments associated with extreme rainfall

It is also possible describe the driving synoptics of extreme rainfall days by mapping only these days to the trained SOM. For the period from 1997 to 2005, the total number of extreme precipitation days extracted from the GPCP data was 128 in region 1, 47 days in region 2 and 59 days in region 3 (Fig. 5.11). The highest number of extreme rain days occurred in region 1 during the summer, an observation also true for region 2 but here far fewer days were recorded. In region 3 the highest number of extreme rainfall days occurred in JJA and MAM. The composites of extreme precipitation days during this period associated with each synoptic circulation is shown in Figure 5.12.

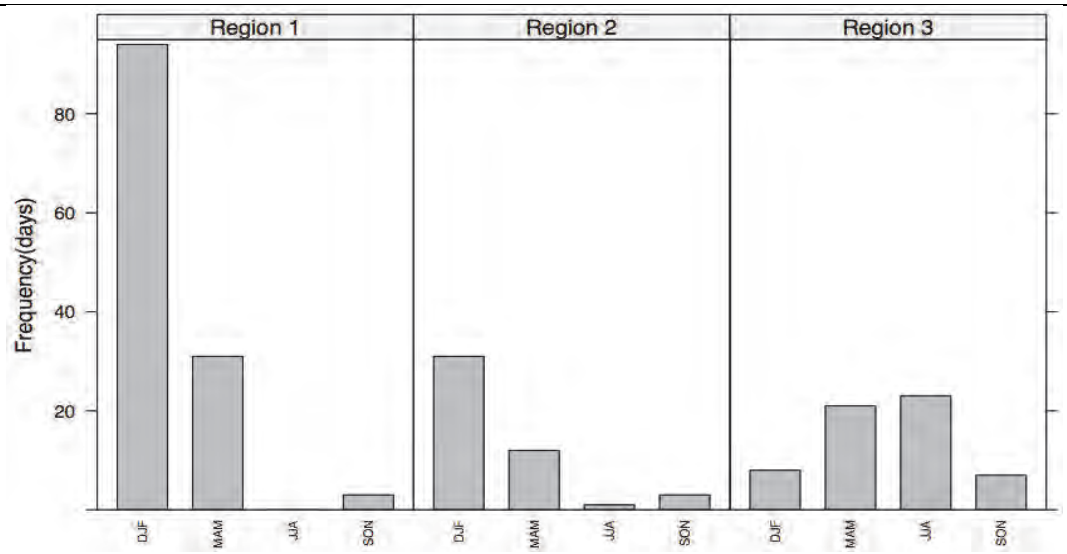


Figure 5.11. Seasonal frequency of extreme precipitation days as observed (GPCP) from 1997-2005.

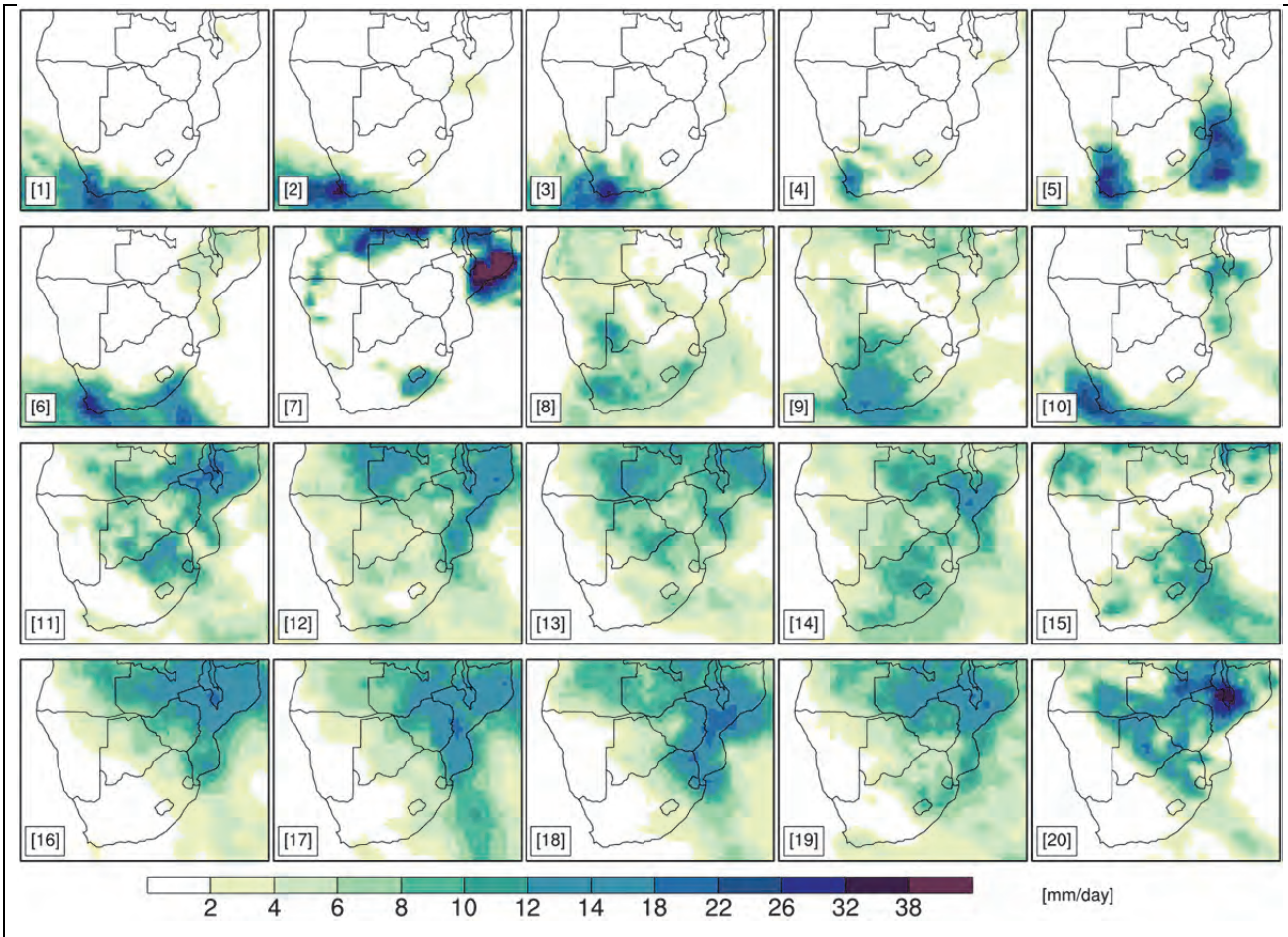


Figure 5.12. GPCP precipitation composites of days with extreme precipitation associated with each type in the 5x4 SOM.

In region 1 the dominant summer month nodes (16-19) are associated with extreme rainfall and of these most extreme rain days map to nodes 16 and 17 (Fig. 5.13 – top row). Spatially, extreme

rainfall mapping to these nodes is in the north east of the domain over Mozambique, Zimbabwe and Zambia where the extreme rainfall is associated primarily with westward moisture transport into the region from the Mozambique Channel. The downscaling does not capture the frequent mapping of extreme rainfall to these nodes in both the EI and GCM downscaling (Table 5.2). This suggests the regional models may have deficiencies in developing and advecting in the moisture necessary for such events, problems with the boundary layer and convective parameterization schemes and are perhaps not able to capture the effect of tropical cyclones, which are a primary cause of extreme rainfall in this tropical region. It is likely a combination of these factors. However, it should be remembered that the GCM simulations are not time-synchronous with the observed climate. Each GCM is initialized with atmospheric, sea and ice states in 1850 and left to develop its own climate – it is this non-real global climate that the RCM downscales. Despite this it would be expected that over the 9-year period the salient circulations responsible for extreme rainfall in the region would be captured. The relatively poor EI results suggest that even if the GCMs captured synoptic environments associated with extreme rainfall in the tropics, the downscaling would not necessarily produce the expected extreme rainfall response.

Spatially, extreme rainfall over Region 2 lies largely in the second from the bottom row of the SOM and also nodes 8 and 9. This is apparent in the frequency mapping of extreme rainfall in this region, however nodes 16-19 also record equivalent mappings (Fig. 5.13 – middle row). Being a summer rainfall region these results are expected. Although the downscaled mappings in this region appear better than in region 1, as was noted earlier the RCA4 model produced too many extreme rainfall events and the CCLM too few giving an overall smoothed result. This region is situated in a transitional zone from summer to winter rainfall and it is influenced by synoptic conditions typical from DJF and the transitional seasons MAM and SON.

Region 3 includes the winter rainfall region of South Africa and spatially extreme rainfall is associated with the mid-latitudes. Extreme rainfall days in this region map most frequently to nodes 2 and 11 (Fig 5.13 – bottom row) and these synoptic systems would likely take the form of deep mid-latitude cyclones and cut-off lows. However, extreme precipitation days are not only influenced by synoptic circulation typical from JJA but also from other seasons as evident especially nodes 13, 14 and 19. This is because the northern part of this region lies in a region that is also influenced by summer rainfall regimes. In this region the RCA4 again over-simulates the total number of extreme precipitation days and the CCLM simulates almost exactly the same number of days as the observed (Table 5.2).

Table 5.2. Total number of extreme precipitation days for each region for the period of 1997-2005 for the GPCP observed data, RCA and CCLM downscalings, each respective ensemble average as well as the combined ensemble average.

	Region 1	**Region 2**	**Region 3**
GPCP	**128**	**47**	**59**
RCA4(CNRM-CM5)	82	54	83
RCA4(EC-EARTH)	70	60	94
RCA4(MPI-ESM-LR)	79	73	80
RCA4(HadGEM2-ES)	56	72	51
Ensemble average – RCA	***71.75***	***64.75***	***77***
CCLM(CNRM-CM5)	52	16	41
CCLM(EC-EARTH)	16	19	76
CCLM(MPI-ESM-LR)	17	22	78
CCLM(HadGEM2-ES)	20	16	47
Ensemble average – CCLM	***25.25***	***18.25***	***60.5***
Ensemble average – RCA and CCLM	***49***	***41.5***	***68.75***

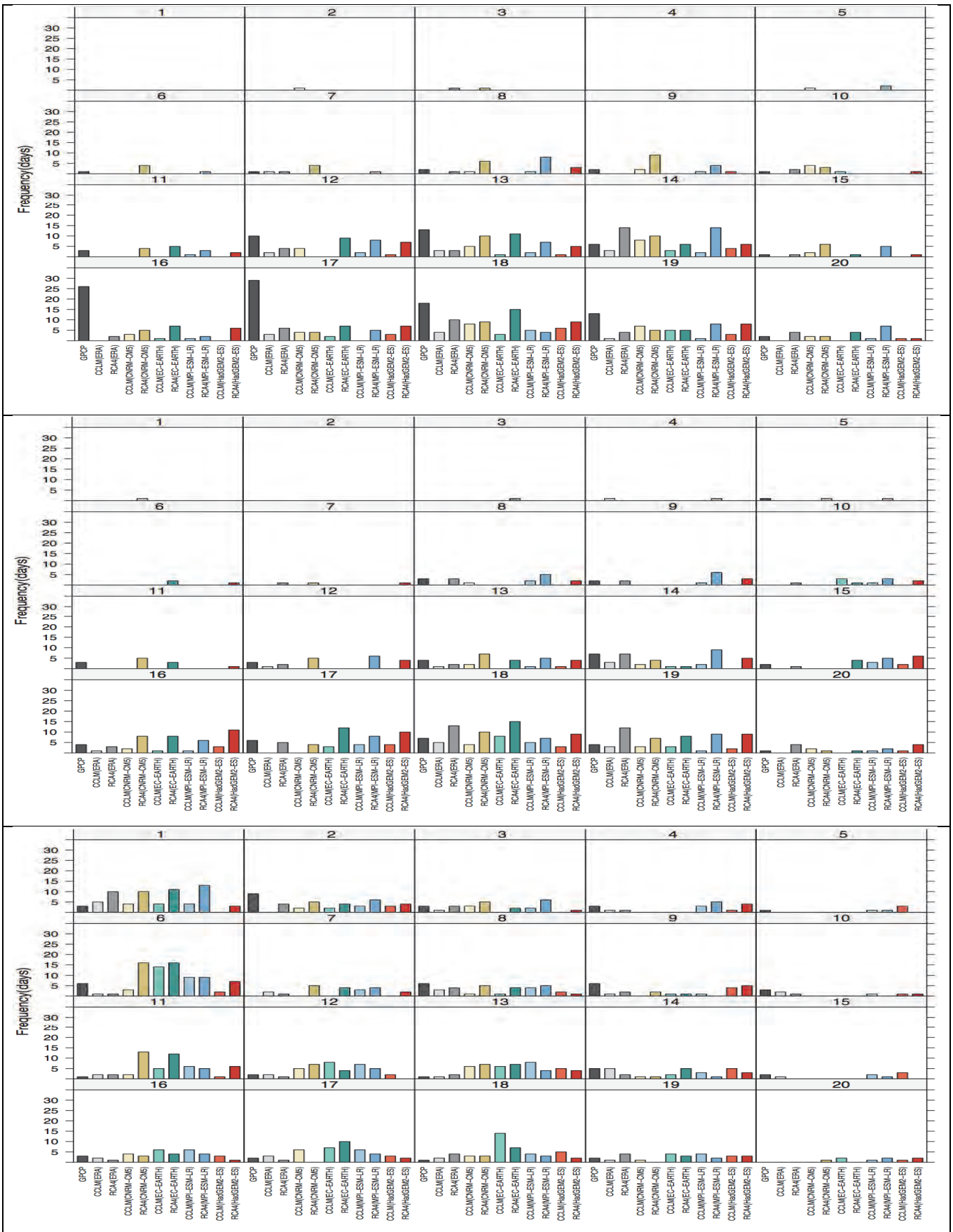


Figure 5.13. Nodal mapping of extreme rainfall in region 1 (top row), 2 (middle row) and 3 (bottom row) for the GPGP data, downscaled ERA-Interim runs and the downscaled GCM runs.

5.10 Discussion and conclusions

When considering data from the station records, many more stations show statistically significant increases in the annual count of extreme rainfall events in the full record compared to the truncated period is used. This result (using the full station record) is similar to what is reported in Mason et al. (1999) and Groisman et al. (2005) and Kruger (2006) who also use long records to identify increasing occurrence of heavy and extreme rainfall over many parts of South Africa. It should be noted that although they do not use the same metrics as used in this trend analysis, the theme of generally increasing trends in frequency or intensity of heavy or extreme rainfall in many parts of South Africa is common.

This result highlights the need for caution when performing rainfall trend analyses over South Africa. Nicholson, (1994) shows that rainfall over the region was generally above average in the 1970s and below average in the 1980s. This means that any rainfall study in southern Africa beginning in the late 1970s or early 1980s, as many do because 1979 is when many reanalysis and satellite products begin, is immediately influenced by the dry decade of the 1980s. This would be especially true for trend studies where decadal variability can easily skew results. In Chapter 6 we present trends of the statistically downscaled projection data that are computed over a 150-year period between 1951-2100 so suggest that in terms statistical significance, trends computed from the full observational dataset be used in this comparison.

Extreme rainfall as represented in the ETCCDI indices (95^{th} and 99^{th} percentile rainfall, rx5day) simulated by the regional model in both the ERA-Interim and GCM downscalings show the RCM produces too much extreme precipitation compared to the observed indices. However, the spatial patterns of these indices are reproduced and are similar to those in the shorter station record. The spatial correlation of rare rainfall events is low between two observational products, which suggests that it is difficult to make specific statements about these events, however, broader generalizations are possible.

The SOM assessment identifies characteristic circulation types that have distinctive seasonal features to which composite rainfall and extreme rainfall mappings align, further demonstrating the utility of method to describe synoptic environments associated with rainfall and extreme rainfall. The SOM was also used to evaluate the downscaled results from two RCMs (downscaling four GCMs each) and assess the projected frequencies of synoptic states associated with extreme rainfall in the short, medium and long term.

The evaluation downscaling demonstrated the models could not capture the frequency mapping characteristics of extreme rainfall. In the tropics both under-simulated these states and in the sub-tropics and mid-latitudes there are mixed results. The CCLM model produces much fewer extreme rainfall synoptic states than the RCA4 model, which has the effect of producing an ensemble average that is closer to the GPCP data in regions 2 and 3 through smoothing the differences between models. The models do, however, capture the features of low-level circulation reasonable well over southern Africa in the historical period as well as general rainfall states. This further demonstrates that an assessment of the synoptic environments produced by regional models provides much more understanding/information than simply assessing the rainfall products in isolation. The regional models poor representation of extreme rainfall is likely a combination of factors such as grid resolution, which at 0.44 degrees does not resolve convective activity nor completely the topographic amplification of rainfall, parameterizations of microphysics, boundary layer processes and convection as well as the model core.

The evaluation runs also demonstrated that the RCM internal climate seems to dominate for some archetypal states but for others the GCM climate propagates through into the downscaled domain. The type of synoptic environment determines whether the RCM internal climate emerges in the frequency mapping (evident when the forcing environment is stronger) or if the RCM translates the GCM boundary forcing into the downscaled domain (weaker forcing environment).

That the EC-earth GCM-RCM frequency mapping results are opposite to the other three GCMs demonstrates that (a) multi-model ensemble analysis is necessary to more completely sample the uncertainty space in projection data and (b) that GCM selection in any particular downscaling analysis is important. It is beyond the scope of this study to understand why the downscaled EC-Earth frequency mappings of the two RCMs are opposite to those of the other GCMs.

These results form a basis from which to approach the projection data. We are satisfied that the downscaled GCM data has replicated to an adequate degree the statistical properties of extreme rainfall and that circulation states characteristic of extreme rainfall events have been identified. We can therefore proceed to an analysis of projection data produced by the GCMs using these approaches and assess any changes in the characteristics and synoptic environments associated with extreme rainfall.

Chapter Six: Projected changes in the characteristics of extreme rainfall and their synoptic environments

6.1 Introduction

As mentioned in Chapter 1, the only way to assess potential changes to the climate as a result of global warming and potential impacts these changes may have on the natural and human systems is to use climate data from general circulation models (GCMs) that simulate global climate under greenhouse gas forcing. However, GCM data is at a relatively coarse spatial scale (typically ~ 200 km) and should be downscaled to a finer scale to investigate regional messages of concern. In the context of this study, the downscaling captures local drivers of rainfall like topography, that are especially relevant for extreme rainfall.

In Chapter 5 changes in the frequency of occurrence of extreme rainfall events in the observed record was presented and regional models evaluated for their representation of rainfall characteristics and synoptic environments. In this chapter we assess statistically downscaled projections of extreme rainfall at the station scale and dynamically downscaled results for two RCMs that downscaled four GCMs at the gridded scale. We also identify projected changes in the characteristics of synoptic environments associated with extreme rainfall.

The chapter is structured such that station-scale statistically downscaled projections of the frequency of occurrence of extreme rainfall are presented first after which changes in the synoptic circulations associated with extreme rainfall are investigated. A short section detailing projected changes in rainfall intensity is presented after with these results are discussed and some conclusions drawn.

6.2 Trends of projected extreme rainfall at the station scale

The 150-year trends in frequency of occurrence of extreme rainfall days at the 69 downscaled stations from each of the 12 GCMs is presented in Figure 6.1. Unlike the observed trends, in most GCMs the projected trends are negative at most stations with varying degrees of statistical significance. The negative trends most persistent across models in the north-eastern parts of South and also around Lesotho.

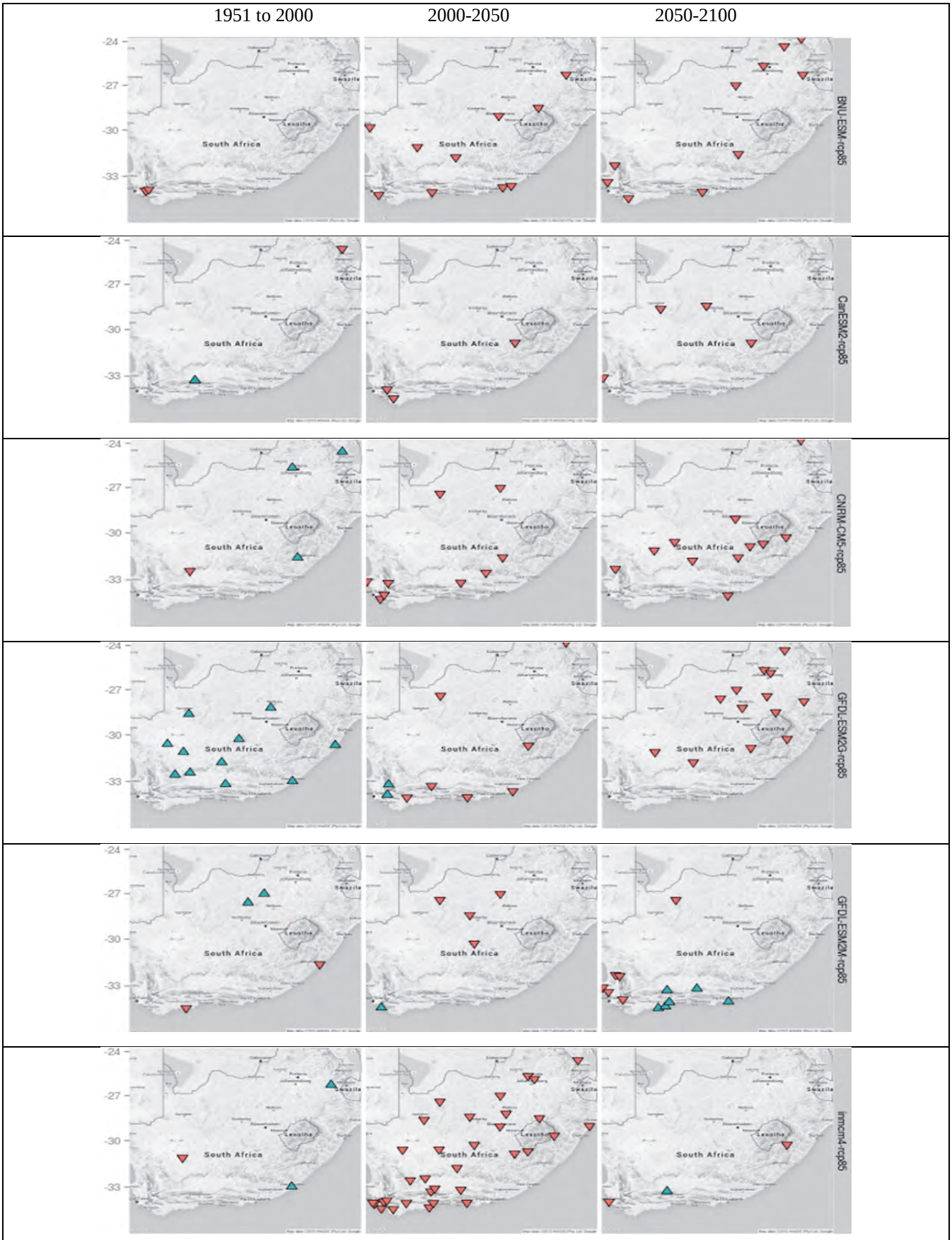
In the entire 12-model ensemble, only four GCMs produce stations that have statistically significant increasing trends and only 6 stations are affected. Downscaling two versions of the same GCM could produce different results. For example, of the two GFDL-based earth systems models (ESMs), the ESM2M version produces non-significant positive trends in the south-western Cape

and along the south coast (although one station this trend is significant) whereas the ESM2G shows negative trends in the region and many of these are significant. The difference in the two models is in the vertical coordinate system of the ocean model where the ESM2M uses pressure-based vertical coordinates in the ocean model and the ESM2G an isopycnal model[4] which may explain the coastal location of these differences. The three MIROC-based models are two ESMs, one with atmospheric chemistry (MIROC-ESM-CHEM) and one without (MIROC-ESM) and one atmosphere-ocean model (MIROC5). Their results are similar with negative trends in the north-eastern parts of South and around Lesotho. The ESMs show significant trends in the central interior whereas the atmosphere-ocean model has this region of significant trend shifted westward and also shows positive trends in the south-western Cape not evident in the ESMs. The results of one GCM (the inmcm4 model) shows statistically significant results at all stations which are likely spurious but are presented here to demonstrate the dangers of using data from one downscaling of one GCM. It is not possible to isolate a particular GCM as "most believable" as each is an equi-probable future nor do we in this study weight GCMs or bias correct their results according to their simulations of the historical period. We take all the GCM results and identify common meta-threads that are evident in the ensemble to understand what the strongest and most robust messages are.

The reversed trend in the projection data compared to the observed data is further investigated by splitting the 150-year period into three sub-periods, namely 1951-2000 and 2000-2050 and 2050-2100. This helps establish if similar trends were found in the historical period of the GCM runs and identifies when the trend emerges. For the historical period (1951-2000), trends in the annual frequency of occurrence of extreme rain days at the station were generally positive in most models, as was seen in the observed data (Fig. 6. 2), although the number of stations that display statistically significant changes vary between models. The decreasing trends are most evident post-2000 with a generally higher number of stations reporting significant trends in the latter 2050-2100 period. The downscaled trends for the historical period generally agree with the observed trends so the methodology captures the general response to the forcing synoptic environment provided by the GCMs. The general reversal between the observed and projected trends suggests the statistical downscaling is responding to a change in the drivers of extreme rainfall as a result of greenhouse gas (GHG) warming. In the next section results from dynamically downscaled GCM data are presented to compare against the statistically downscaled results.

[4] http://www.gfdl.noaa.gov/earth-system-model

Figure 6.1. Trends in the frequency of occurrence of extreme rainfall data for 12 downscaled CMIP5 GCMs under RCP 8.5 between 1951 and 2100. The GCM name is at the top of each image and more information is available in Table 4.1. Blue (red) triangles indicate positive (negative) trends and solid triangles where the trend is statistically significant.



.....continued on next page

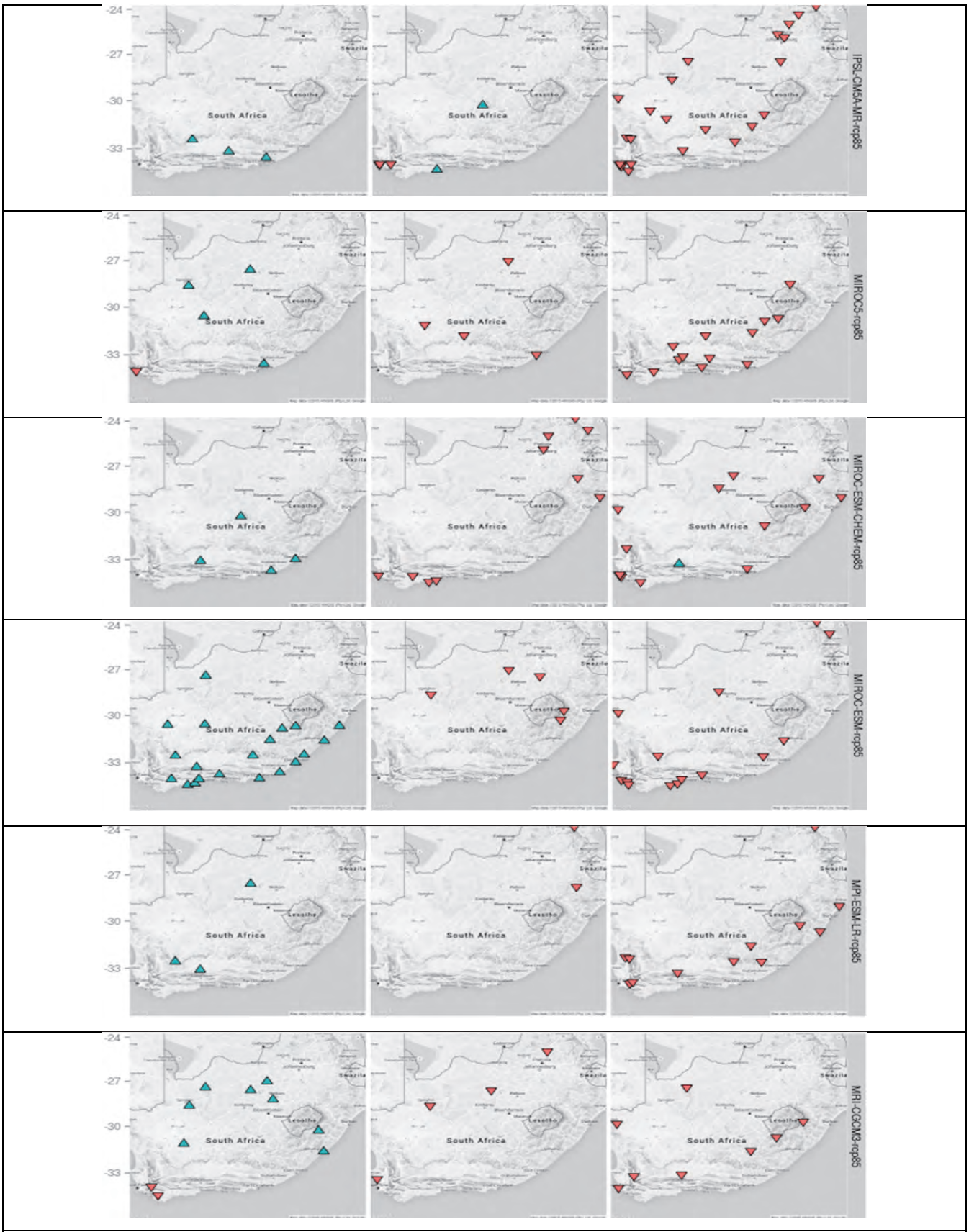


Figure 6.2. Trends in the annual frequency of occurrence of extreme rain days for selected statistically downscaled GCMs under RCP 8.5 for three periods – the historical period (1951-2000, left), the full 150-year period (centre) and for the period 2000-2100 (right). Only statistically significant trends are displayed where red (green) represents decreasing (increasing) trends.

6.3 Projected changes in synoptic environments associated with extreme rainfall

In Chapter 5 the dynamical downscaling results were evaluated and extreme rainfall was associated with particular synoptic states. Here, we present the dynamically downscaled results of the projected period from a synoptic circulation point of view by mapping the projected circulation states to the SOM developed in Chapter 5. Changes in frequency mappings provide information about which circulation states may become more or less frequent in the future and facilitate an investigation into the surface response in the context of projected extreme rainfall. We assess two representative concentration pathways (RCP 4.5 and 8.5) and two time slices. The stationarity hypothesis regarding climate dynamics is assumed, that is, the historical synoptic types remains valid throughout the 21st Century.

The change in synoptic circulation is defined as the difference between the simulated climate under future and current forcings. The difference in the number of days matching each synoptic type in the SOM are calculated for the middle (2035-2065) and late (2069-2098) 21st Century relative to 1976-2005. Box plots are developed to indicate the interquartile model spread (25th and 75th quantiles) with the horizontal line indicating the ensemble median and the whiskers showing the extreme range of the ensemble members. This indicates the primary spread of the projected change from the different models.

Over the full southern African domain, decreases in the frequency of occurrence of many nodes that are associated with wet conditions are projected (Fig. 6.3 – nodes 2, 3, 6, 11, 12, 16, 17). This is especially relevant for the Western Cape region where there is no corresponding projected increase in rainfall-associated nodes. Decreases in the frequency of nodes associated with rainfall over the interior (11, 12, 16, 17) is perhaps somewhat offset by increases in projected frequencies of nodes 13, 14, 18 and 19, however, the reduction of projected “wetter” state days is far higher than projected increase in “wetter” state days (medians of -50 to -230 days compares to medians of +20 to 100 days. This suggests a decrease in circulation patterns that favour the occurrence of precipitation over most of southern Africa. The increasing frequency of occurrence of nodes in the top of the map (dry nodes) implies an increase in circulation patterns that are associated with reduced precipitation events. In general, these changes represent a projected increased occurrence of the oceanic high-pressure systems, a more dominant high-pressure circulation poleward of the continent and a decreased occurrence of patterns of continental lows and mid-latitude lows. The magnitude of change of the frequency of the synoptic types of the ensemble median is higher for RCP8.5 for most of the nodes compared with RCP4.5 for the same period.

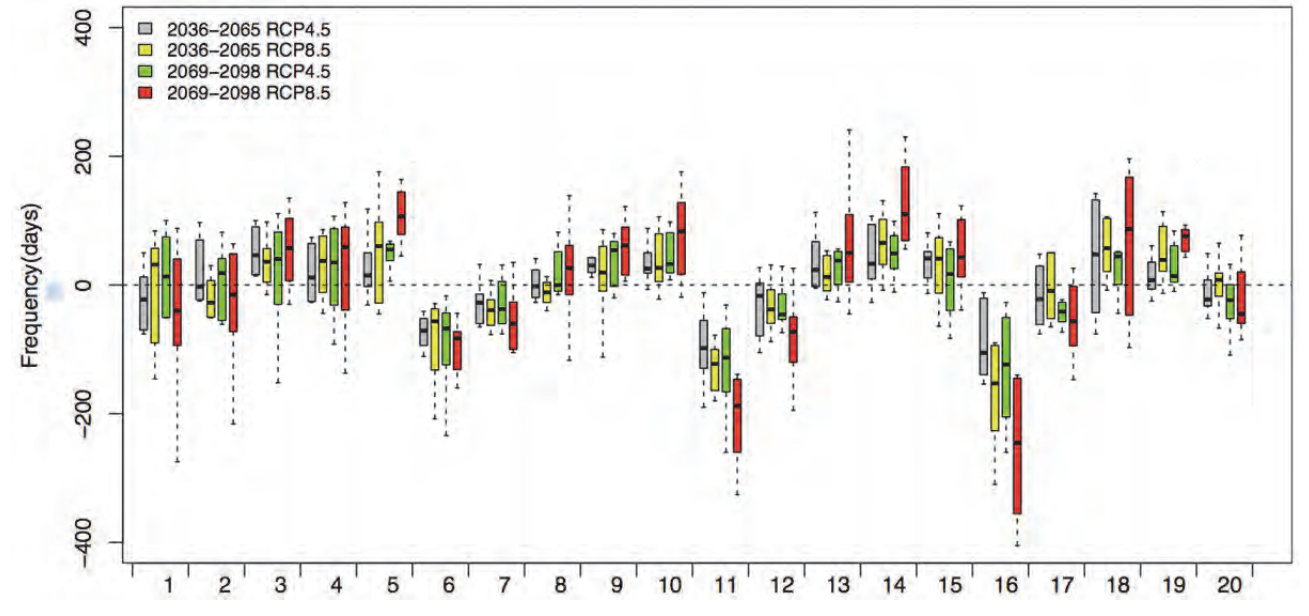


Figure 6.3. Projected changes in the mapping frequency of all days for the periods 2036-2065 and 2069-2098 relative to 1976-2005 under RCP4.5 and RCP8.5 for the full southern African domain.

In region 1, the circulation patterns associated with extreme precipitation days over region 1 (see Fig. 5.8 top row) are projected to increase in frequency in all scenarios (Fig. 6.4 – top row). However, the overall frequency mapping to these nodes is projected to decrease (Fig. 6.3) and suggests that when these types of circulation occur in the future, they are more likely to be associated for extreme precipitation days. For region 2, extreme precipitation days are projected to increase in association with nodes 18 and 19 but magnitudes are lower than in region 1 (Fig. 6.4 – middle row). For region 3, most of the nodes associated with extreme precipitation (nodes 1, 2, 11, 13, 14 and 19) are projected to decrease with the exception of node 1, which has a small projected increase in frequency of occurrence (Fig. 6.4 – bottom row). This suggests fewer extreme precipitation days for this region in the future, however it is not possible to make statements about extreme precipitation intensity with this type of assessment. In the next section, however, non-parametric and parametric analyses of projected extreme rainfall facilitates an inspection of changes in the intensity of extreme rainfall in addition to the frequency assessment performed above.

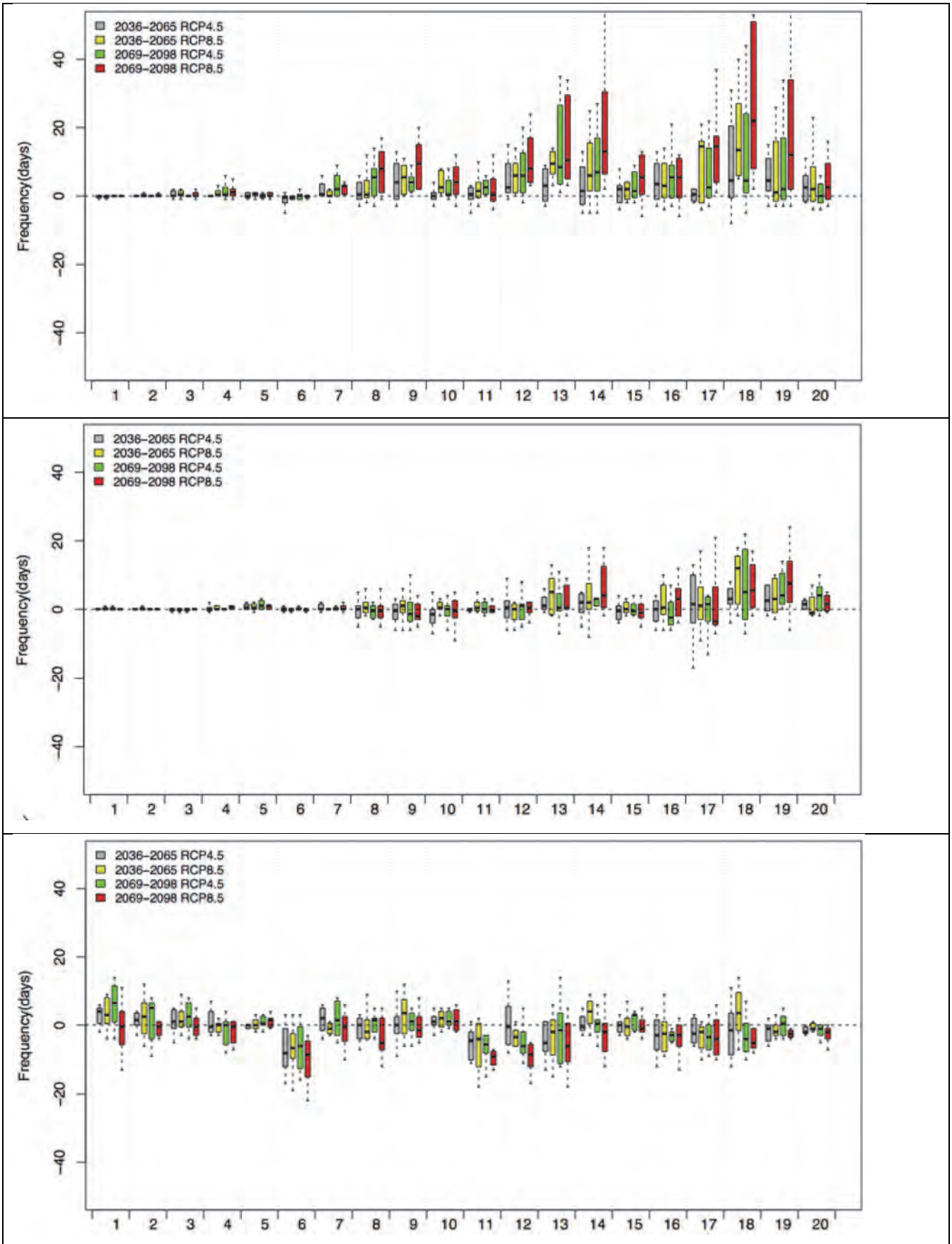


Figure 6.4. As for Figure 6.3 but for extreme rainfall in regions 1 (top), 2 (middle) and 3 (bottom).

6.4 Projected changes in intensity of extreme rainfall events

6.4.1 Projected changes in regional model data using ETCCDI indices

Projected rainfall data from the two RCMs were analyzed using extreme rainfall indices of the ETCCDI. The projected multi-model mean changes in these indices are largest in tropical and sub-tropical parts of southern Africa in all indices (Fig. 6.5). The spatial extent and intensity of change is generally higher in RCP4.5 compared to RCP8.5.

Over South Africa a strong change signal is seen over the eastern and south-western parts of the country. In the former region, strong positive changes (increases in rainfall intensity) are seen particularly over the Drakensburg in both RCP scenarios. In this largely summer rainfall region that lies in Region 2, there are projected increases in the frequency of occurrence of nodes 13, 14 18 and 19 (Fig. 6.4, middle row), circulation states that are associated with onshore flow of warm, moist air from the Agulhas region and pressures that are not very high. It is therefore likely that a combination of factors contribute to the increased extreme rainfall statistics in the region: the increased frequency of occurrence of this type of circulation, an enhanced moisture content in the region as a result of a warmer world and the topographic forcing that assists in triggering convection associated with extreme rainfall.

In the south-western parts of the country the decreasing statistics are more evident and statistically significant in RCP 8.5. The region is seen most clearly in the RCP 8.5 map of R95pTOT where there are significant decreases. This region is not in a purely winter rainfall region so the change is likely attributable to projected decreases in the frequency of occurrence of nodes 2, 11, 13, 14 and 19 that are associated with both mid-latitude systems and summer convective circulation types.

6.4.2 Projected changes of rare extremes in the regional model data

The ensemble mean of the projected changes in rare extreme precipitation (here defined as 20-year return values) for the late 21st Century for both RCP4.5 and RCP8.5 is shown in Figure 6.6. A general increase in the magnitude of the 1-in-20 year extreme precipitation event is projected over the central and eastern parts of southern Africa and a decrease over western parts of South Africa and central and southern Namibia. This pattern of change in extreme precipitation is projected consistently across both scenarios although the magnitude of the increases are generally higher under RCP8.5 in areas where the change is positive. The projected increases in return values imply more frequent recurrence of the current 20-year event in the future.

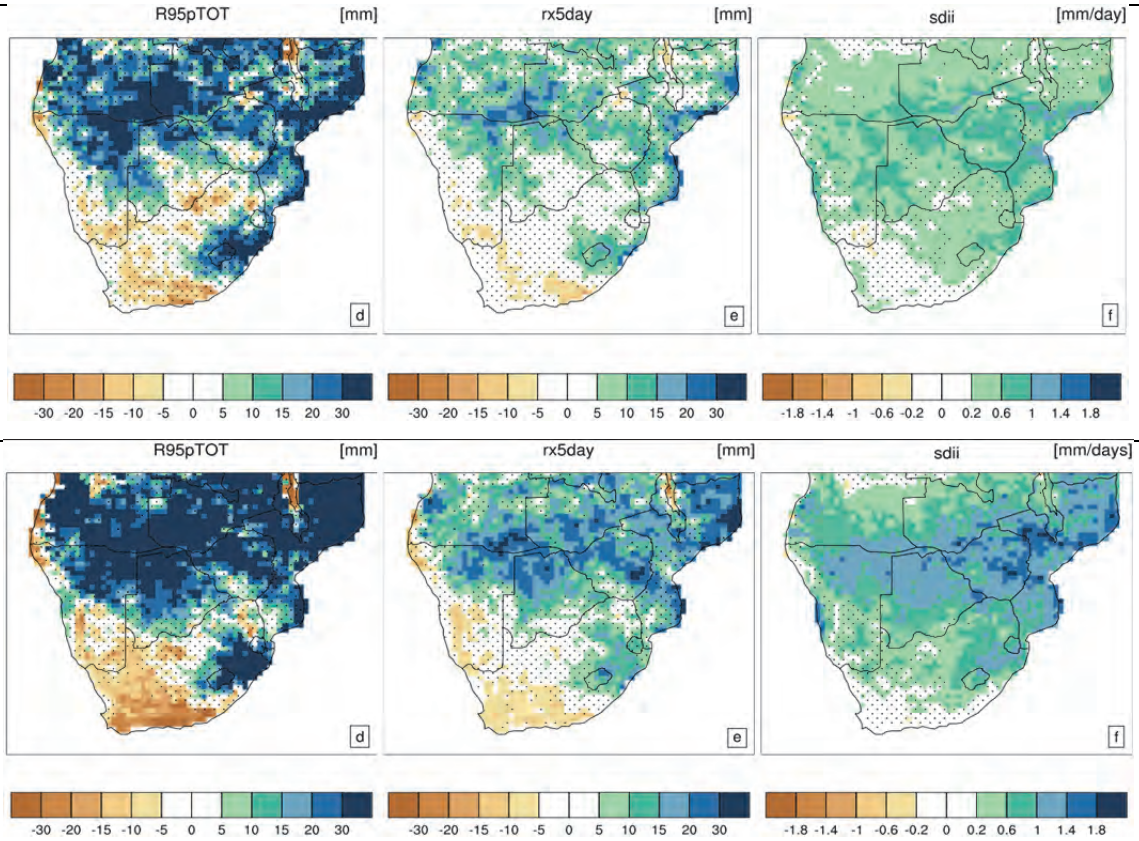


Fig. 6.5. Projected multi-model mean changes of moderate extreme events for the period 2069-2098 relative to a reference period 1976-2005 under RCP4.5 (top) and RCP 8.5 (bottom). Stippling indicates grid points that are not statistically significant (5% significance level using a t-test).

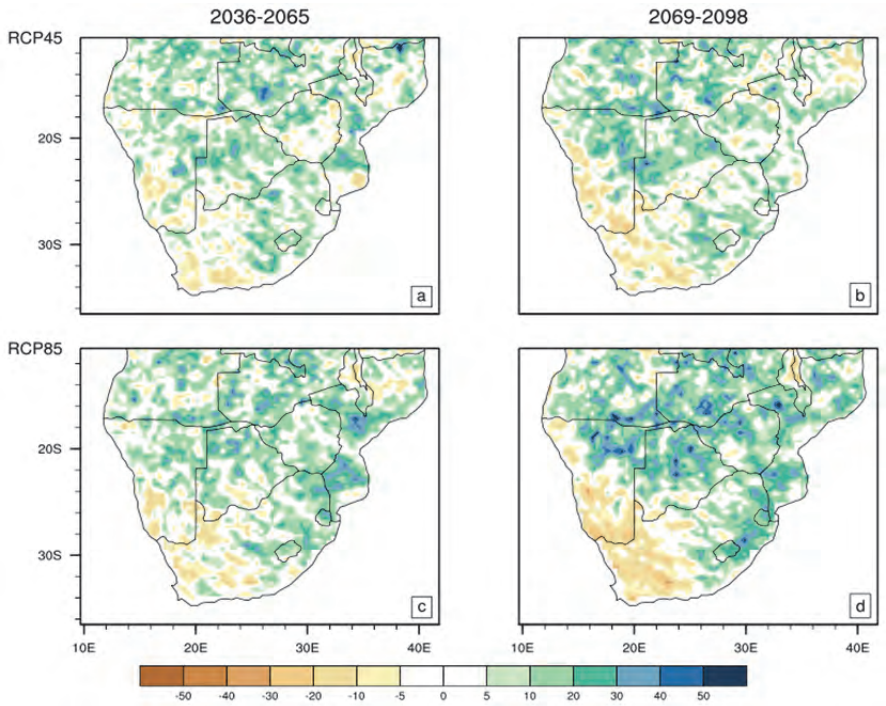


Figure 6.6. Projected multi-model mean changes 20-year return values over the time period 2036-2065 (a and b) and 2069-2098 (c and d) as differences relative to the reference period (1976-2005) for RCP4.5 (a and c) and RCP8.5 (b and d). Units are in mm/day.

6.5 Discussion and conclusions

An increase in the frequency of occurrence of synoptic states that are associated with drier conditions are projected for the future, while synoptic states that enhance precipitation are projected to decrease. This is especially evident in the winter rainfall region and the arid to semi-arid western parts of South Africa where the projected emergence of drier synoptic environments is evident in all scenarios. These changes are generally larger for RCP8.5 and in the far future. Synoptic states associated with extreme precipitation are projected to increase in the tropical and sub-tropical summer rainfall region with the magnitude of the change higher in the tropics. In the winter rainfall region and the arid to semi-arid western parts of the country the frequency of synoptic environments associated with extreme precipitation are projected to decrease. Changes in the extreme rainfall data from the dynamical downscaling reflect these circulation changes where increases in the statistics of extreme rainfall in the tropical and sub-tropical summer rainfall region are seen. Over South Africa, increases in extreme rainfall statistics over the eastern parts of the country, especially over the eastern Drakensburg, are projected for the future.

The changes over the eastern Drakensberg are likely a result of globally and regionally warmer temperatures in summer that, through the Clausius-Clapeyron relationship, facilitates greater moisture content in the atmosphere and increased rainfall potential. This relationship describes how a 1^{o}C rise in temperature leads to an almost exponential 7 % rise saturation vapour pressure. This has an effect on both the mean and extremes of rainfall, however the increase in precipitation extremes can be greater than changes in mean precipitation because extreme precipitation relates to increases in moisture content and thus the nonlinearities involved with the Clausius-Clapeyron relationship (Allen and Ingram, 2002; Kharin and Zwiers, 2005). It is very likely that because of this strong convective environments and associated precipitation would increase in a warmer world (Berg *et al*., 2013, Diffenbaugh *et al*., 2013, Kendon *et al*., 2014). This thermodynamic effect is greatest in subtropical regions. In tropical, mid-latitude and arid regions dynamical circulation changes also play an additional role in future changes in extreme precipitation (Emori and Brown, 2005; Meehl *et al*., 2005).

The statistical downscaling replicates the increasing trends in frequency of extreme rainfall events during the GCM historical period as seen in the observed data. However, decreasing trends in the frequency of extreme rainfall are generally downscaled for the 21^{st} Century at the examined stations under RCP 8.5. There is thus both agreement and disagreement between the dynamically downscaled and statistically downscaled projections. Over the western and southwestern parts of the country the statistically significant results from both downscaling methods agree and indicate a

decrease in the frequency and intensity of extreme rainfall. However, over the Drakensberg region, the results from the two methods are contradictory. A likely explanation of this discrepancy is that the statistical downscaling cannot capture the enhanced thermodynamic Clausius-Clapeyron relationship in the warmer RCP 8.5 world. As the statistical downscaling procedure is trained on an historical Clausius-Clapeyron relationship, it is perhaps unable to account for an enhanced thermodynamic effect. This “stationarity problem” is well known in the statistical downscaling community and acknowledged. Unfortunately, only three stations lie in the affected region, which makes it difficult to quantify the extent of the difference and it would be desirable to repeat the analysis with as many stations in the region.

In summary, over the western and southwestern parts of the country, which is semi-arid to arid and also affected by mid-latitude systems, the results from both downscaling methods are in general agreement. This region has a stronger dynamical driving component and projections indicate a decrease in the frequency of synoptic environments associated with extreme rainfall, likely a consequence of the southward displacement of mid-latitude cyclone tracks by an enhanced descending limb of the Hadley circulation. This implies fewer storm tracks are likely to pass over the region which would likely result in fewer extreme rainfall events. In the sub-tropical Drakensberg region the increase in frequency and intensity of extreme rainfall is likely a consequence of an enhanced thermodynamic driver, the Clausius-Clapeyron relationship, as well as an increased frequency of circulation states that indicate onshore advection of moisture into the region. Here the results of each downscaling method diverge significantly and, as discussed above, this is likely because the statistical downscaling cannot capture an enhanced thermodynamic effect.

In conclusion, many downscaling evaluations only consider surface diagnostic fields like rainfall or temperature and do not investigate the prognostic variables behind these. Similarly, projected changes in diagnostic fields are often not explained within the context of projected prognostic environments. Through relating characteristic rainfall and extreme rainfall fields to particular synoptic environments, we can understand projected changes in rainfall and extreme rainfall through an assessment of the circulation and thermodynamic environment. As the dynamical downscaling is able to capture both of these, and we have understood the changes in the synoptic environments associated with extreme rainfall produced by these models, we suggest the message in the extreme rainfall context from the dynamical downscaling is more defensible at the moment. It is important to have this physical understanding of projected changes in high-demand, largely diagnostic variables in order to develop robust change messages.

Chapter Seven: Summary, discussion, conclusions and suggestions for further work.

7.1 Summary

Data from general circulation models consists of prognostic variables – those governed by prognostic equations that solve for conservation of mass, momentum and thermodynamic energy through integrations in time and diagnostic variables, which are computed from prognostic variables and other external parameters. Prognostic variables include pressure, temperature, humidity, wind speed and direction and diagnostic variables include cloud, two-meter temperature and rainfall. We relate the prognostic synoptic environment to responses in the (diagnostic) extreme rainfall field using downscaled atmospheric and rainfall data.

Synoptic environments associated with extreme rainfall in South Africa have been characterized here using the self-organizing maps methodology. We find cut-off lows associated with extreme rainfall are deeper and westerly-wave driven in austral winter whereas in summer they are shallower and warmer suggesting a greater convective influence. Similarly, mid-latitude cyclones are deeper in the core winter months (June-July-August) and shallower in spring and autumn months. Extreme rainfall associated with tropical temperate troughs display clear intra-seasonal characteristics during summer where a much higher number occur in late summer (January-February-March) than in the early summer (October-November-December). Seven homogenous extreme rainfall regions have been identified over South Africa that include extreme rainfall associated with the three above synoptic environments as well as those influenced by tropical cyclones.

Historical trends in the frequency of occurrence of extreme rainfall events show a general increase across South Africa. Greater statistical significance is found in longer station records and suggests that trend analysis in South Africa beginning in 1979, which is the year most reanalysis products begin, is likely to be skewed as the 1980s were anomalously dry over much of the country. Statistically and dynamically downscaled results replicate the observed trends. Dynamically downscaled future projections indicate synoptic environments associated with extreme precipitation over southern Africa are likely to increase in the tropical and sub-tropical summer rainfall region with the magnitude of the change higher in the tropics. Over South Africa, increases in frequency and intensity of extreme rainfall are projected over the eastern parts of South Africa, especially over the Drakensberg while decreases are projected over the western and southern parts of the country. Dynamically and statistically downscaled projections indicate decreases in the frequency (and in

dynamical models intensity) of extreme rainfall over the western and southern parts of South Africa as a consequence of fewer circulation states associated with extreme rainfall. However, over the eastern Drakensberg the results diverge – statistical downscaling suggests a decreased frequency of extreme rainfall events and the dynamical downscaling an increase in both intensity and frequency of extreme rainfall.

7.2 Conclusions

Three major conclusions have emerged from this study. Firstly, self-organizing maps have been shown as be a useful tool to investigate the prognostic synoptic environments associated with the surface extreme rainfall response over South(ern) Africa. They have teased out seasonally characteristic systems and provided information on projected changes in the frequency of these systems. However, they cannot be used as a tool to identify extreme rainfall associated in the way they were used in this analysis. Further experimentation of the Euclidean distance method of identifying particular synoptic environments. One idea here is to first run a principle component analysis on the training data to reduce the degrees of freedom and rerun the SOM analysis. It is envisaged this would be performed on higher and lower resolution reanalysis data.

Secondly, dynamically and statistically downscaled results concerning changes in the frequency of extreme rainfall events both agree and disagree. We have understood the reasons for these similarities and differences based on the prognostic synoptic environment and the thermodynamic changes that are likely to occur in the future as a result of a warming world. It is crucial to understand these reasons as this contributes to the robustness of the message being conveyed. Here we demonstrate a likely decrease in the frequency of occurrence of extreme rainfall events in the western regions of South Africa as a result of circulation changes, and increase in the frequency of occurrence of extreme rainfall over the eastern escarpment because of circulation and thermodynamic changes. Only the downscaling using regional climate models can capture the enhanced thermodynamic effect in the future so we have given preference to this message.

Lastly, we used two both dynamical and statistical methods to downscale data from multiple GCMs and the dynamically downscaled results were from two RCMs. In this way we have sampled a large portion of the uncertainty space in the projection data to develop the messages above. Using only one GCM or downscaling method or RCM limits the amount of information available.

In conclusion, high-demand, diagnostic variables (like rainfall) are particularly difficult to simulate as a result of the way they are derived, and have relatively large uncertainties associated with them.

We have used multiple methods to quantify the prognostic synoptic environments associated with extreme rainfall over South Africa and developed a physical understanding of projected changes that has facilitated the emergence of robust and defensible change messages.

7.3 Recommendations for future work

Identifying particular synoptic environments associated with extreme rainfall in future projection data is an important and desirable outcome. Currently this is possible with closed low-pressure systems such as cut-off lows, tropical cyclones and cold fronts but not with other systems that result in extreme rainfall over South Africa like tropical temperate troughs, deep easterly wave lows and onshore flows over the southern coastal parts of the country. It is suggested that the development of a methodology to identify prognostic synoptic environments associated with extremes (extreme rainfall or drought) be pursued as more projection data becomes available through CMIP5 (and CMIP6) as well as CORDEX and other more local studies.

Secondly, we recommend the differences in the results between the statistical and dynamical downscaling over the eastern parts of the country be investigated. Although the reasons put forward for the differences by this study are valid, a proper examination is necessary to more thoroughly understand the differences and the implication these may have in application models and for decision-making. It would also be desirable to have additional results from other statistical downscaling models, which should be possible through CORDEX in 2016.

Chapter Eight: Capacity Development, publications and conference proceedings.

8.1 Capacity development

Mr Izidine Pinto

This project funded a proportion of Mr Izidine Pinto's Ph.D. He submitted his Ph.D. for examination and it is expected he will graduate in either December 2015 or July 2016. His thesis title is "*Future changes in extreme rainfall events and circulation patterns over southern Africa"*.

Additionally, the project funded an opportunity for Mr Pinto to attend the "WCRP-ICTP Summer School on Attribution and Prediction of Extreme Events" at the International Centre for Theoretical Physics in Trieste, Italy between 21 July and 1 August 2014. The goal of the school was to train students with outstanding research potential in the techniques that are required to better understand observed and future changes in extremes climate events. A short report from Mr Pinto is included below and it will be seen a paper has emerged from this summer school. Regrettably, because of the large author team and journal page restrictions, the WRC could not be acknowledged in this paper for having made it possible for Mr Pinto to attend the summer school.

WCRP-ICTP Summer School on Attribution and Prediction of Extreme Events – I. Pinto

I was invited to participate at the "WCRP-ICTP Summer School on Attribution and Prediction of Extreme Events". The summer school was held at ICTP, Trieste-Italy, from 21 July to 1 August 2014. The goal of the school was to train students with outstanding research potential in the techniques that are required to better understand observed and future changes in extremes climate events. This Summer School was part of the activities of the WCRP Grand Challenge on "Understanding and Predicting Weather and Climate Extremes".

The School was organized around three broad topic areas:

• Statistical theory underpinning extreme value analysis;

• Detection and attribution of observed changes in the frequency and/or intensity of extremes;

• Event attribution, and the physical mechanisms that are involved in amplifying and/or extending the duration of some specific extreme events such as heat waves.

The two week course consisted of morning lectures and afternoon small group sessions, focused on

the practical application of the material covered in the lectures. A team of 5 students, including myself from different regions of the world (Spain, England, Argentina, Brazil and South Africa) worked on one of the projects for the duration of the school and presented preliminary results at the end of the school. After we got back to our home research institutions we continued to work as a team and published the results of the project in a peer-reviewed journal. The publication is shown below:

Bellprat, Omar, Fraser C. Lott, Carla Gulizia, Hannah R. Parker, Luana A. Pampuch, Izidine Pinto, Andrew Ciavarella, and Peter A. Stott. (2015). Unusual past dry and wet rainy seasons over Southern Africa and South America from a climate perspective. Weather and Climate Extremes. *http://dx.doi.org/10.1016/j.wace.2015.07.001*.

I would like to thank the Water Research Commission, specifically Project K5/2240, for funding my attendance of the summer school.

Izidine Pinto

Miss Myra Nyak

Miss Nyak joined the project relatively late but was key in identifying extreme rainfall associated with extreme rainfall. She graduated with a M.Sc. in July 2015 although this was not directly related to the project. Her exposure to the methods employed in the analysis has prompted her to consider reading for a Ph.D. in 2016.

8.2 Journal papers

Pinto, I., C. J. Lennard, M. Tadross, B.Hewitson, A. Dosio, G. Nikulin, H-J. Panitz and M. E. Shongwe (In Review): Simulated changes in extreme precipitation over southern Africa from CORDEX models. *Climatic Change:*

8.3 Conference papers

Lennard C.J., Pinto, I, (2015): Projected changes in synoptic circulation states associated with extreme rainfall in Southern Africa. *EGU General Assembly Conference Abstracts* (Vol. 17, p. 14737).

Pinto, I., Lennard, C., Hewitson, B., Tadross, M., (2014). Assessment of the ability of CORDEX models to capture observed extreme rainfall over southern Africa. WCRP-ICTP Summer School on Attribution and Prediction of Extreme Events, 21 July-1 August, Trieste, Italy.

Pinto, I., Lennard, C., Hewitson, B., Tadross, M., (2013). Projections of extreme rainfall events in CORDEX simulations for southern Africa. 2nd International Conference on Coordinated Regional Downscaling Experiment (CORDEX),4-7 November, Brussels, Belgium.

Pinto, I., Lennard, C., Tadross, M., Hewitson, B., 2013. Climate extremes indices in the CORDEX multimodel ensemble: Future climate projections. 29th Annual conference of South African Society for Atmospheric Sciences, 26-27 September, DURBAN, South Africa. ISBN 978-0-620-56626-1

Pinto, I., Lennard, C., Hewitson, B., 2013. Intercomparison of precipitation extremes over southern Africa in CORDEX simulations. Geophysical Research Abstracts, Vol. 15, EGU2013-8872, EGU General Assembly 2013

Chapter Nine: References

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

Balogun, A. R., & Adeyewa, Z. D. (2013). Analysis of Storm Structure over Africa Using the Trmm Precipitation Radar Data. *Atmospheric and Climate Sciences*, **3**, 538.

Berg, P., C. Moseley, and J.O. Haerter, (2013). Strong increase in convective precipitation in response to higher temperatures. *Nature Geoscience*, **6**(3), 181-185.

Biasutti, M., Yuter, S. E., Burleyson, C. D., & Sobel, A. H. (2012). Very high resolution rainfall patterns measured by TRMM precipitation radar: seasonal and diurnal cycles. *Clim. Dyn.*, 39(1-2), 239-258.

Blamey R.C. and C.J.C. Reason (2009). Numerical simulation of a mesoscale convective system over the east coast of South Africa. *Tellus* 61A:17-34

Buishand, T. (1989). Statistics of extremes in climatology. *Statistica Neerlandica*, 43(1):1-30.

Christensen, J.H., B. Hewitson, A. Busuioc, A. Chen, X. Gao, I. Held, R. Jones, R.K. Kolli, W.-T. Kwon, R. Laprise, V. Magaña Rueda, L. Mearns, C.G. Menéndez, J. Räisänen, A. Rinke, A. Sarr and P. Whetton, (2007). Regional Climate Projections. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Chu, P-S and Wang J., (1997). Tropical cyclone occurrences in the vicinity of Hawaii: are the differences between El Nin˜o and non-El Nin˜o years significant? *J. Clim*. 10:2683-2689.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics, London, UK.

Cook, C., Reason, C. J.C., and Hewitson, B. C. (2004). Wet and dry spells within particularly wet and dry summers in the South African summer rainfall region. *Climate Research*, *26*(1), 17-31.

Diatta, S., Hourdin, F., Gaye, A. T., & Viltard, N. (2010). Comparison of Rainfall Profiles in the West African Monsoon as Depicted by TRMM PR and the LMDZ Climate Model. *Mon. Wea. Rev.*, *138*(5), 1767-1777.

Dee, D. P., and Coauthors, 2011: The ERA-Interim reanalysis: Configuration and performance of the data assimilation sys- tem. *Quart. J. Roy. Meteor. Soc*., 137, 553-597.

Diffenbaugh, N. S., Scherer, M., & Trapp, R. J. (2013). Robust increases in severe thunderstorm environments in response to greenhouse forcing. *Proceedings of the National Academy of Sciences*, *110*(41), 16361-16366.

Dyson LL and Van Heerden J (2002). A model for the identification of tropical weather systems over South Africa. *Water SA* 28(3):249-258

Dyson, L. L., Van Heerden, J., and Sumner, P. D. (2015). A baseline climatology of sounding-derived parameters associated with heavy rainfall over Gauteng, South Africa. *Int. J. Climatol.*, *35*(1), 114-127.

Embrechts, P., Klüppelberg, C., and Mikosch, T. (1997). *Modelling extremal events for insurance and finance*. Springer Verlag.

Emori, S., and S. J. Brown, 2005: Dynamic and thermodynamic changes in mean and extreme precipitation under changed climate. *Geophys. Res. Lett.*, **32**, L17706.

Endris, H. S., P. Omondi, S. Jain, C. Lennard, B. Hewitson, L. Chang'a, J. L. Awange, A. Dosio, P. Ketiem, G. Nikulin, H-J. Panitz, M. Büchner, F. Stordal and L. Tazalika, (2013) Assessment of the Performance of CORDEX Regional Climate Models in Simulating East African Rainfall. *J. Clim.*, **26**, 8453-8475.

Engelbrecht C.J., and F. A. Engelbrecht, (2015). Shifts in Koppen-Geieger climate zones over Southern Africa in relation to key global temperature goals. *Theoretical and Applied Climatology*,

Engelbrecht, C. J., F. A. Engelbrecht, and Dyson, L. L. (2013). High-resolution model-projected changes in mid-tropospheric closed-lows and extreme rainfall events over southern Africa. *Int. J. Climatol.*, *33*(1), 173-187.

Engelbrecht, C. J., Landman, W. A., Engelbrecht, F. A., & Malherbe, J. (2014). A synoptic decomposition of rainfall over the Cape south coast of South Africa. *Clim. Dyn.*, 1-19.

Favre A, Hewitson BC, Tadross M, Lennard C and Cerezo-Mota R (2012). Relationships between cut-off lows and the semiannual and southern oscillations. *Clim. Dyn.*, 38:1473-1487

Favre, A., Hewitson, B., Lennard, C., Cerezo-Mota, R., & Tadross, M. (2013). Cut-off Lows in the South Africa region and their contribution to precipitation. *Clim. Dyn.*, 41(9-10), 2331-2351.

Futyan, J.M., and A.D. Del Genio, (2007). Deep convective system evolution over Africa and the tropical Atlantic. *J. Climate, 20, 5041-5060.*

Giorgi, F., and Mearns, L. O. (1999). Introduction to special section: Regional climate modeling revisited. *Journal of Geophysical Research: Atmospheres (1984-2012), 104*(D6), 6335-6352.

Groisman, P., Knight, R., Easterling, D., Karl, T., Hegerl, G., and Razuvaev, V. (2005). Trends in intense precipitation in the climate record. *J. Clim.*, 18(9):1326-1350.

Harangozo, S. A., and M.S.J Harrison, (1983). On the use of synoptic data in indicating the presence of cloud bands over Southern-Africa. *South African Journal of Science, 79*(10), 413-414.

Harrison, M. S. J., (1984). A generalized classification of South African summer rain-bearing synoptic systems. *J. Climatol.*, **4**:547-560, 1984. ISSN 1097-0088. doi: 10.1002/joc.3370040510.

Hart N.C.G., C.J.C. Reason and N. Fauchereau (2010). Tropical-extratropical interactions over southern Africa: three cases of heavy summer season rainfall. *Mon. Wea. Rev.* 138:2608-2623.

Hart, N. C., Reason, C.J.C., & Fauchereau, N. (2012). Building a tropical-extratropical cloud band metbot. *Mon. Wea. Rev.*, *140*(12), 4005-4016.

Hart, N. C., Reason, C. J.C., & Fauchereau, N. (2013). Cloud bands over southern Africa: seasonality, contribution to rainfall variability and modulation by the MJO. *Clim. Dyn.*, *41*(5-6), 1199-1212.

Hartigan, J. A. (1975). Clustering algorithms. John Wiley and Sons, New York.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Applied statistics*, 100-108.

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

Hewitson, B. C., & Crane, R. G. (2006). Consensus between GCM climate change projections with empirical downscaling: precipitation downscaling over South Africa. *Int. J. Climatol.*, *26*(10), 1315-1338.

Hope, P. K., Drosdowsky, W., and Nicholls, N. (2006). Shifts in the synoptic systems influencing southwest Western Australia. *Clim. Dyn.*, 26(7-8):751-764.

Jackson, S., (1951). Climates of Southern Africa. *South African Geographical Journal*, **33**(1):17-37. doi: 10.1080/03736245.1951.10559278.

Ji, X., and Chen, Y. (2012). Characterizing spatial patterns of precipitation based on corrected TRMM 3B43 data over the mid Tianshan Mountains of China. *Journal of Mountain Science*, *9*(5), 628-645.

Kalognomou, E., C. Lennard, M.Shongwe, I. Pinto, M. Kent, B. Hewitson, A. Dosio, G. Nikulin, H. Panitz, M. Buchner (2013). A diagnostic evaluation of precipitation in CORDEX models over southern Africa. *J. Clim.* 26(23):9477-9506. doi:10.1175/JCLI-D-12-00703.1

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., & Senior, C. A. (2014). Heavier summer downpours with climate change revealed by weather forecast resolution model. *Nature Climate Change* **4**(7), 570-576. doi:10.1038/nclimate2258

Kharin, V. V. and Zwiers, F. W. (2000). Changes in the Extremes in an Ensemble of Transient Climate Simulations with a Coupled Atmosphere-Ocean GCM. *J. Clim.*, 13(21):3760-3788.

Kim, M.J., J.A. Weinman and R.A.Houze, (2004). Validation of Maritime Rainfall Retrievals from the TRMM Microwave Radiometer. *Journal of Applied Meteorology*: Vol. 43, No. 6, pp. 847-859.

Kistler R et al. (2001). The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, 82:247-267.

Kodama Y. M. and A. Tamaoki. A re-examination of precipitation activity in the subtropics and the mid-latitudes based on satellite-derived data. J. Meteorol. Soc. Jpn, 2002, 80: 1261-1278.

Kohonen, T., (2001). *Self-Organizing Maps*. Third, extended edition. Springer, Berlin.

Kruger, A. C., (2002). Climate of South Africa: surface winds. South African Weather Service Technical Report, 2002.

Kruger, A. C., (2006). Observed trends in daily precipitation indices in South Africa: 1910-2004. *Int. J. Climatol.*, *26*(15), 2275-2285.

Kummerow C, Simpson J, Thiele O, Barnes W, Chang ATC, *et al*. 2000. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *J. Appl. Meteorol*. 39:1965-82.

Laing A.G. and J.M. Fritsch (1993). Mesoscale convective complexes in Africa. *Mon. Wea. Rev.* 121:2254-2263

Landman, W. A., and S.J. Mason (1999). Change in the association between Indian Ocean sea-surface temperatures and summer rainfall over South Africa and Namibia. *Int. J. Climatol.*, **19**(13):1477-1492.

Landman W., A, S. J. Mason, P. D. Tyson and W. J. Tennant, (2001). Retro-active skill of multi-tiered forecasts of summer rainfall over southern Africa. *Int. J. Climatol.* 21: 1-19.

L'Ecuyer, T. S., & McGarragh, G. (2010). A 10-year climatology of tropical radiative heating and its vertical structure from TRMM observations. *J. Clim.*, *23*(3):2345-2352.

Lennard, C.J, D. Morrison, L. J. Coop and Grandin (2013). Extreme events: Past and future changes in the attributes of extreme rainfall and the dynamics of their driving processes. Water Research Commission report No 1960/1/12, pp. 105.

Lennard, C., & Hegerl, G. (2014). Relating changes in synoptic circulation to the surface rainfall response using self-organising maps. *Clim. Dyn.*, 1-19. doi:10.1007/s00382-014-2169-6

Liu, P., Li, C., Wang, Y., & Fu, Y. (2013). Climatic characteristics of convective and stratiform precipitation over the Tropical and Subtropical areas as derived from TRMM PR. *Science China Earth Sciences*, *56*(3), 375-385.

Malherbe J, Engelbrecht FA, Landman W and Engelbrecht C (2011) Tropical systems from the southwest Indian Ocean making landfall over the Limpopo River Basin, southern Africa: a historical perspective. *Int. J. Climatol.* 32:1018-1032

Mason, S. J., (1998). Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *Int. J. Climatol.*, **18**(2):147-164.

Meehl, G., J. Arblaster, and C. Tebaldi, 2005: Understanding future patterns of increased precipitation intensity in climate model simulations. *Geophys. Res. Lett.*, **32**, L18719.

Mei, H., S.A. Braun, P. Ola, G. Persson and J.W. Bao. (2009). Along front Variability of Precipitation Associated with a Midlatitude Frontal Zone: TRMM Observations and MM5 Simulation *Mon. Wea. Rev.*, 137(3):1008-1028.

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

New, M., *et al*. (2006), Evidence of trends in daily climate extremes over southern and west Africa, *J. Geophys. Res*., **111**, D14102, doi:10.1029/2005JD006289

Nicholson, S. E., (1994). Recent rainfall fluctuations in Africa and their relationships to past conditions over the continent. *Holocene*, **4,** 121-131.

Nikulin G, C. Jones, P. Samuelsson, F. Giorgi, B. Sylla, G. Asrar, M. Buchner, R. Cerezo-Mota, O. Christensen, M. De´que´, J. Fernandez, A. Hansler, E. van Meijgaard and L. Sushama, (2012). Precipitation climatology in an ensemble of CORDEX-Africa regional climate simulations. *J. Clim.*, **25**, 6057-6078.

Palutikof, J., Brabson, B., Lister, D., and Adcock, S. (1999). A review of methods to calculate extreme wind speeds. *Meteorological Applications*, 6(02):119-132.

Preston-Whyte, R. A., (1974). Multivariate Approach. Climatic Classification of South Africa. *South African Geographical Journal*, **56**. doi: 10.1080/03736245.1974.10559527.

Rasmussen, K. L., S.L. Choi, M.D. Zuluaga, and R.A. Houze, (2013). TRMM precipitation bias in extreme storms in South America. *Geophysical Research Letters*, 40(13), 3457-3461.

Reason, C. J. C., Landman, W., & Tennant, W. (2006). Seasonal to decadal prediction of southern African climate and its links with variability of the Atlantic Ocean. *Bulletin of the American Meteorological Society*, *87*(7), 941-955.

Richardson, A. J., Risien, C., & Shillington, F. A. (2003). Using self-organizing maps to identify patterns in satellite imagery. *Progress in Oceanography*, *59*(2), 223-239.

Saha, S., et al. (2010). The NCEP climate forecast system reanalysis. *Bulletin of the American Meteorological Society*, *91*(8), 1015-1057.

Sang, H., Gelfand, A. E., Lennard, C., Hegerl, G., & Hewitson, B. (2008). Interpreting self-organizing maps through space-time data models. *The Annals of Applied Statistics*, 1194-1216.

Sardeshmukh, P.D., G.P. Compo and C. Penland, (2000). Changes of probability associated with El Nino. *J. Clim.* 13:4268-4286.

Seneviratne, S.I., N. Nicholls, D. Easterling, C.M. Goodess, S. Kanae, J. Kossin, Y. Luo, J. Marengo, K. McInnes, M. Rahimi, M. Reichstein, A. Sorteberg, C. Vera, and X. Zhang, (2012). Changes in climate extremes and their impacts on the natural physical environment. In: *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation (SREX). Special Report of the Intergovernmental Panel on Climate Change (IPCC).* [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi *et al.* (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA, pp. 109-230.

Singleton A.T. and C.J.C. Reason, (2007) Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. *Int. J. Climatol.* 27:295-310

Sun, Y. (2000) On quantization error of self-organizing map network. *Neurocomputing*, *34*(1), 169-193.

Sylla, M.B., A.T. Gaye, and G.S. Jenkins, (2012) On the fine-scale topography regulating changes in atmospheric hydrological cycle and extreme rainfall over West Africa in a regional climate model projections. *International Journal of Geophysics,* doi:10.1155/2012/981649

Taljaard JJ (1985) Cut-off lows in the South African region. South African Weather Bureau technical paper, vol 14, Pretoria, South Africa, 153 pp

Taljaard JJ (1995) Atmospheric circulation systems, synoptic climatology and weather phenomena of South Africa: Atmospheric circulation systems in the Southern African region. South African Weather Bureau Technical paper no. 28, Pretoria, South Africa, 65 pp

Taljaard JJ (1996) Atmospheric circulation systems, synoptic climatology and weather phenomena of South Africa: Synoptic climatology and weather phenomena of South Africa, rainfall in South Africa. South African Weather Bureau Technical paper no. 32, Pretoria, South Africa, 98 pp

Tadross, M. A., Hewitson, B. C., & Usman, M. T. (2005). The interannual variability of the onset of the maize growing season over South Africa and Zimbabwe. *J. Clim.*, *18*(16), 3356-3372.

Tennant, W. (2004). Considerations when using pre 1979 NCEP/NCAR reanalyses in the southern hemisphere. *Geophys. Res. Lett.* 31:L11112

Thomas, D. S., Twyman, C., Osbahr, H., & Hewitson, B. (2007). Adaptation to climate change and variability: farmer responses to intra-seasonal precipitation trends in South Africa. *Climatic Change*, *83*(3), 301-322.

Tyson, P. D., & Preston-Whyte, R. A. (2000). *Weather and climate of southern Africa*. Oxford University Press.

Xu, K.M. (2006). Using the bootstrap method for a statistical significance test of differences between summary histograms. *Mon. Wea. Rev.* 134:1442-1453 .

Yamamoto, M. K., Higuchi, A., and Nakamura, K. (2006). Vertical and horizontal structure of winter precipitation systems over the western Pacific around Japan using TRMM data. *Journal of Geophysical Research: Atmospheres (1984-2012)*, *111*(D13).

Yang, S., & Smith, E. A. (2008). Convective-Stratiform Precipitation Variability at Seasonal Scale from 8 Yr of TRMM Observations: Implications for Multiple Modes of Diurnal Variability. *J. Clim.*, *21*(16).

Yang, S., Kuo, K. S., & Smith, E. A. (2008). Persistent nature of secondary diurnal Codes of precipitation over oceanic and continental regimes. *J. Clim.*, *21*(16).

Zwiers, F. W. and Kharin, V. V. (1998). Changes in the Extremes of the Climate Simulated by CCC GCM2 under CO2 Doubling. *J. Clim.*, 11(9):2200-2222.