

East African Rainfall: Classification of Rain Producing Systems – a Modelling and Observation Study



Helen Elizabeth Pearce

**Keble College
University of Oxford**

A thesis submitted for the degree of
Doctor of Philosophy
Michaelmas Term 2013

East African Rainfall: Classification of Rain Producing Systems – a Modelling and Observation Study

Helen Elizabeth Pearce, Keble College, University of Oxford

A thesis submitted for the degree of Doctor of Philosophy, Michaelmas Term 2013.

Abstract

The study of anthropogenic climate change is a research area of vital importance for the coming decades, with rainfall change and variability expected to be keenly felt in vulnerable regions of the world, including Africa. The focus of this study is daily rainfall during the short rains season over East Africa from October to December, which has one of the most coherent increasing rainfall projections in the Coupled Model Intercomparison Project (CMIP) dataset. This thesis aims to examine the fidelity of coupled climate models over an East African domain, with the question approached through focus on the daily (rather than monthly) rainfall field.

The self-organising map is used as a clustering tool to establish contemporary characteristics of daily rainfall events in reanalysis (ERA-40) and satellite (TRMM) rainfall datasets between 1971-2000 and 1998-2010 respectively for the East African short rains. Moisture flux divergence is found to be the circulation parameter that is most closely related to the presence of rainfall events or dry conditions over East Africa. Coupled climate models are poor at replicating the daily rainfall field over the domain. A key result of the analysis is the consistent overestimation of daily rainfall by climate models for days where dry conditions of suppressed convection should prevail. In contrast, the moisture flux divergence field maps well to dry nodes for days of the self-organising map array for the models. Dry days are associated with widespread anomalous moisture flux divergence and rainfall events with co-located anomalous moisture flux convergence. This is in agreement with the moisture flux divergence field in the ERA-40 dataset; it is the rainfall field where there is disagreement for days of suppressed convection.

Twenty to thirty-five percent of the projected rainfall increase towards the end of the twenty-first century results from an increase in the proportion of days assigned to nodes of suppressed convection in six of the seven models and the ensemble mean. There is an accompanying projected rainfall increase associated with days assigned to these nodes. Such days in the 2090s are characterised by projected increased strength moisture flux divergence over East Africa. Given that the moisture flux field was more successfully simulated in the coupled models under contemporary conditions than the daily rainfall field, this suggests that rainfall projections under a high emissions scenario at the end of the twenty-first century are overestimated and that an important part of the key increase in the projected rainfall may not be real.

Acknowledgements

I would like to thank everyone who has provided me with support and encouragement throughout my DPhil, whilst not all named, I am grateful to all. First, I would like to thank my supervisor, Professor Richard Washington, for providing guidance, support and feedback throughout my DPhil. In particular, I would like to thank Richard for suggesting such an interesting topic to study and for his continued support as my work evolved throughout the years, through valuable intellectual input and suggestions. I would also like to thank him for the opportunities he provided me with outside of my DPhil; the Radcliffe Met Station and additional projects, which made it possible for me to undertake my thesis in the first place.

On the technical and computing side of my work, I would like to thank Gil Lizcano for his help in developing my skills in working on the Linux system, untangling my scripts when they failed to run properly and pointing me in the right direction to keep them going. Also thanks to Sebastian Engelstaedter for taking on this role for the last eighteen months.

Thank you to the IT staff at the School of Geography and the Environment for continually providing me with more space on the system for all of my data and analysis.

I would like to thank my family; my siblings, Catherine and Rob and, in particular, my parents for their unfailing support, in every way, throughout the years. Finally, to Ivan, thank you.

Contents

Abstract	i
Acknowledgements	iii
Abbreviations	ix
1. Introduction and Thesis Aims	1
1.1 Introduction	1
1.2 Thesis Focus – East African daily rainfall regimes during the short rains	3
1.3 Thesis Aims	9
1.4 Thesis outline	10
2. Data and Methods	15
2.1. ECMWF ERA Data	16
2.1.1. ERA-40 Dataset	16
2.1.2. ERA-Interim Dataset	21
2.2. NCEP/NCAR Dataset	23
2.3. H08 – a statistical precipitation series	24
2.4. UEA CRU gridded dataset	25
2.5. TRMM satellite precipitation series	27
2.6. Meteorological station precipitation data	28
2.6.1. Introduction to the station data	28
2.6.2. Interpolation of the station data	30
2.6.3. Evaluation of the gridded precipitation series	30
2.6.3.1. CRU vs. interpolated station series	31
2.6.3.2. TRMM vs. interpolated station series	35
2.6.4. Masking of the interpolated station series	39
2.7. Mapping of datasets to research questions	40
2.8. AOGCM data from the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4)	43
2.9. Clustering techniques	48
2.10. The self-organising map	49
2.10.1. Benefits of the SOM over alternative clustering methods	49
2.10.2. Previous applications of the SOM in climatology	50
2.10.3. Self-organising map methodology	50
2.11. Normalisation	55
2.12. Statistical Testing	56
2.12.1. Chi-squared goodness of fit	56
2.12.2. Pearson’s product moment correlation coefficient	57
2.12.3. Significance testing of projected rainfall changes – Z-test	57
2.13. Moisture flux divergence	58
2.14. Chapter summary	58

3. Twentieth century climatology of East Africa and potential changes to the twenty-first century climate	60
3.1. Background climate of East Africa	61
3.1.1. East African climate in the research literature	62
3.1.2. Basic climatology of East Africa	63
3.2. East African rainfall climatology – station data series	70
3.3. Trends in East African climate from 1950 to the present day	71
3.3.1. East African climate trends in the literature	71
3.3.2. Trends in East African climate	73
3.4. Extremes in the East African climate	81
3.5. Evaluation of AOGCMs – how well can they replicate the twentieth century East African Climate?	85
3.5.1. Use of AOGCMs in climate research	85
3.5.2. Climatology of the IPCC AR4 models over East Africa	86
3.6. AOGCM projections of East African climate during the twenty-first century	94
3.6.1. East African climate projections in the literature	94
3.6.2. Projections over East Africa from the IPCC AR4 model subset	95
3.7. Chapter Summary	105
4. Contemporary characteristics of daily rainfall regimes during the short rains in the ERA-40 dataset	108
4.1. Evaluation of the SOM classification method: A simple test case	110
4.1.1. Absolute data	112
4.1.2. Evaluation of the classification scheme	112
4.2. Application of the SOM classification to three realisations of the ERA-40 rainfall series	118
4.2.1. ERA-40 rainfall anomaly SOM	119
4.2.2. ERA-40 rainfall normalised SOM	123
4.2.3. ERA-40 rainfall normalised SOM reduced domain	125
4.3. Most closely matched nodes	155
4.4. Least closely matched nodes	156
4.5. Progression of circulation fields over the days preceding a rainfall event	161
4.6. Chapter summary	169
5. Contemporary characteristics of daily rainfall regimes during the short rains – an ERA-40 and TRMM comparison	172
5.1. Classification of TRMM satellite rainfall	173
5.1.1. Daily rainfall states in the TRMM SOM	174
5.1.2. Comparison of daily rainfall states in the TRMM and ERA-40 SOMs	175
5.1.3. Atmospheric circulation parameters associated with TRMM	177

daily rainfall types	
5.2. Comparison of ERA-40 and TRMM daily rainfall events	192
5.2.1. Daily rainfall events; correlation between ERA-40 and TRMM	192
5.3. Daily rainfall events 1998-2000; comparison between ERA-40 and TRMM	195
5.3.1. Daily states in the ERA-40 and TRMM rainfall datasets	195
5.3.2. Date matching of ERA-40 and TRMM daily rainfall states	200
5.3.3. Closely matched nodes in the TRMM and ERA-40 datasets	202
5.4. A 5 by 4 rainfall self-organising map over the East Africa array	204
5.4.1. A common period: 1998-2000 5 by 4 SOM realisations	204
5.4.2. Matching the ERA-40 rainfall data to the 5 by 4 TRMM rainfall SOM	208
5.5. Proportional occurrence of rainfall events in TRMM and ERA-40	209
5.6. Chapter summary	214
6. East African daily rainfall regimes: a comparison of CMIP3 models with ERA-40 and TRMM	217
6.1. Daily rainfall in AOGCMs during the twentieth century	219
6.1.1. Daily rainfall characteristics of eight CMIP3 AOGCMs	219
6.1.2. Proportion of occurrence of different daily rainfall states during the control period	227
6.2. Application of AOGCM data on to the ERA-40 rainfall self-organising map	231
6.2.1. Quantisation error evaluation	232
6.2.2. Distribution of AOGCM nodes to the ERA-40 SOM	233
6.3. Trends in East African rainfall during the 1971-2000 period	240
6.4. Daily rainfall per day (mm) during the East African short rains in AOGCMs in relation to ERA-40	242
6.4.1. Mean daily node rainfall in reanalysis and satellite rainfall datasets over East Africa	243
6.4.2. Mean daily node rainfall in model datasets over East Africa	249
6.5. Atmospheric conditions associated with rainfall nodes within the ERA-40 self-organising map	254
6.6. Chapter summary	259
7. Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario – Part I – Rainfall projections	261
7.1. Mean changes to East African rainfall under the SRES A2 scenario	262
7.1.1. Projected mean change to East African rainfall during the short rains under the SRES A2 scenario	263
7.1.2. Projected spatial changes to East African rainfall during the short rains under the SRES A2 scenario	264

7.2. Application of 2090s data to the climatology ERA-40 SOM, fit to the SOM	267
7.2.1. Quantisation error of 2090s daily data when applied to the control period self-organising map	269
7.2.2. Distribution of daily states in the 2090s under the SRES A2 scenario	271
7.3. Quantification of projected daily rainfall changes between 1971-2000 and 2090-2099	280
7.4. Chapter summary	290
8. Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario – Part II – Spatial changes and moisture flux projections	291
8.1. Ensemble rainfall – 2090-2099 under the SRES A2 high emissions scenario	292
8.2. Individual AOGCM rainfall – 2090 to 2099 under the SRES A2 high emissions scenario	299
8.2.1. CCCMA	300
8.2.2. CNRM	303
8.2.3. CSIRO	305
8.2.4. MPI	307
8.2.5. GFDL, MIUB and MRI	310
8.3. Moisture flux changes	315
8.3.1. Group B – events of suppressed convection over the East African domain	320
8.3.2. Projected moisture flux divergence changes to other nodes within the SOM array	326
8.4. Chapter summary	337
9. Conclusions and directions for further research	340
9.1. Justification for study and thesis aims	340
9.2. Conclusions for each research question	343
9.2.1. Question 1 and methodology sub question	343
9.2.2. Question 2 and sub question	350
9.2.3. Question 3	353
9.2.4. Question 4	354
9.3. Final conclusions	356
9.4. Directions for future research	358
9.4.1. A thorough examination of the convection scheme in AOGCMs	359
9.4.2. Investigation of daily rainfall regimes of East African short rains in regional climate models	359
9.4.3. Extending investigation of daily rainfall regimes to other seasons over East Africa	361
9.5. Final words	363
References	364

Abbreviations

AMV	Atmospheric Motion Vectors
AOGCM	Atmosphere-Ocean General Circulation Model
AR4	The Fourth Assessment Report of the IPCC
AR5	The Fifth Assessment Report of the IPCC
A2	SRES High Emissions Scenario
CCCMA	cccma_cgcm3_1, Canadian Centre for Climate Modelling and Analysis
CMIP3	The Third Climate Model Intercomparison Project
CMAP	CPC Merged Analysis of Precipitation
CNRM	cnrm_cm3, Centre National de Recherches Météorologiques
COADS	Comprehensive Ocean-Atmosphere Data Set
CPC	National Weather Service Climate Prediction Center (USA)
CSIRO	csiro_mk3_0, CSIRO Atmospheric Research (Australia)
EAT	East African Time (UTC +3 hours)
ECMWF	European Centre for Medium-Range Weather Forecasts
EOF	Empirical Orthogonal Function
ERA	ECMWF Reanalysis Datasets
ERS	European Remote Sensing Satellite
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
GARP	Global Atmospheric Research Program
GFDL	gfdl_cm2_0, US Dept. of Commerce, NOAA Geophysical Fluid Dynamics Laboratory
GISS	giss_model_e_r, NASA/Goddard Institute for Space Studies (USA)
GOME	Global Ozone Monitoring Experiment
GPCP	Global Precipitation Climatology Project
GTS	Global Telecommunications System
HIRS	High-resolution Infrared Radiation Sounder

IFS	Integrated Forecasting System (of the ECMWF)
IPCC	Intergovernmental Panel on Climate Change
JAXA	Japan Aerospace Exploration Agency
MIUB	miub_echo_g, Meteorological Institute, University of Bonn, Meteorological Research Institute of KMA, and Model and Data Group
MPI	mpi_echam5, Max Planck Institute for Meteorology (Germany)
MRI	mri_cgcm2_3_2a, Meteorological Research Institute (Japan)
NASA	National Aeronautics and Space Administration
NCAR	The National Center for Atmospheric Research (USA)
NCEP	National Centers for Environmental Prediction (USA)
NCL	The NCAR Command Language
OND	October, November and December rainy season
PCA	Principal Component Analysis
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PREC/L	NOAA Precipitation Reconstruction over Land
RCM	Regional climate model
SOM	Self-organising map
SRES	Special Report on Emissions Scenarios
SST	Sea surface temperature
SSTA	Sea surface temperature anomaly
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
TWERLE	Tropical Wind Energy Conversion Reference Level Equipment
UEA CRU	University of East Anglia Climate Research Unit
UTC	Coordinated Universal Time
VTPR	Vertical Temperature Profile Radiometer
20C3M	IPCC Historical Twentieth Century Run

Chapter 1: Introduction and Thesis Aims

1.1 Introduction

It is now well established that ‘the observed increase in global average temperatures...is *very likely* due to the observed increase in anthropogenic greenhouse gas concentrations’ (Solomon et al, eds, 2007). With projections of average global temperature increases of between 2.4°C and 6.4°C by the end of the twenty-first century under the high emissions SRES (IPCC Special Report on Emissions Scenarios) A2 scenario, the potential impacts of anthropogenic climate change over the coming decades signals a research area of vital importance. In particular, future climate projections depend heavily on the fidelity of climate models leading to the need for climate models and their performance to be carefully scrutinised. The climate orientated scientific community is facing a significant challenge in terms of the amount of data accumulating in databases from both observed and model sources which has not yet been analysed (Gebril et al, 2011). Sylla et al (2013) assert that reliable daily data is crucial to assess model performance, although it has been underused in model evaluation thus far. Given that climate model output is now routinely interpreted at the regional and sub-regional level, in order to put in place adaptation policies, it is particularly important to analyse models at this level, equating to the spatial scale for which practical information for policy making is utilised. Globally, least developed countries are prone to a scarcity of high quality climate data, a situation exacerbated by a lower research capacity and through receiving less attention from the global climate community, with the continent of Africa being a prime example (Shongwe et al, 2011).

At the same time, Africa is especially vulnerable to the effects of rainfall variability and change due to diverse factors including poverty, an overdependence on rain fed agriculture and an inequitable distribution of land (Zhao et al, 2005). Within the continent, East Africa and East African coastal regions and rangeland systems specifically, have been highlighted as prospective ‘hotspots’ that are markedly vulnerable to climate change socially and environmentally (Thornton et al, 2009). Africa is the subject of fewer studies on climatological variability and climate extremes within the research community, even though it is the poorest countries which are disproportionally affected by such events (Shongwe et al, 2011). The least studied continent in terms of climate variability on a global level is Africa even though changes to the frequency, intensity and predictability of rainfall could negatively impact on water availability in East Africa, with potentially devastating effects (Anyah and Qiu, 2012).

Therefore, research into African rainfall systems and their possible changes under anthropogenic climate change signals an important research agenda for the coming years, with East Africa the focus of this study. Coupled climate models (AOGCMs, coupled atmosphere-ocean general circulation models), are the sole tool available to look at projections of future climate and for these to be effectively employed in research confidence in their ability to simulate current conditions is required. A key part of the challenge is that crucially important variables from a social or economic perspective, such as rainfall, are highly parameterised in models. Simulating realistic climatologies is therefore challenging. Giannini et al (2008) highlight the differing rainfall projections emerging from AOGCM modelling groups with no universally accepted way to prioritise their results. Modelling uncertainties also increase over time through twenty-first century projections and dominate the level of overall uncertainty in climate projections by the 2090s (Rowell, 2012). Rowell (2012) concludes that there is the potential to decrease this

uncertainty through targeted model improvement, especially in the Africa where doubts surrounding local changes to tropical rainfall peak due to model uncertainty. The tropics are the region where many of the impacts of anthropogenic climate change will be felt, providing additional impetus for research in this region. Key to addressing this uncertainty is to understand the weaknesses within AOGCM simulations over East Africa, which will be scrutinised here.

In AOGCMs the spatial scale remains relatively coarse, continuing the necessity for downscaling methods to be employed to obtain regional level information. This can be done using two methods, both employing large scale AOGCM output as a base. First, utilising AOGCM data as a boundary condition to drive dynamical regional climate models (RCM) and second, through creating statistical relationships between large and small scale climatic features which are then applied to AOGCM output. Consequently, the reliability of any local scale information inferred through statistical downscaling or from RCMs depends on the validity of the large scale information driving these methodologies (Herrera et al, 2006). This reinforces the value of evaluation of AOGCM output over climatologically vulnerable regions to ensure that conclusions drawn from these projections or utilised in other methodologies can have known confidence. This is further emphasised through the assertion that current inter-variable biases make it very difficult to accurately assess the realism of climate change projections (Wilby et al, 2002).

1.2 Thesis focus – East African daily rainfall regimes during the short rains

The focus of this study will be daily East African rainfall events, employing a study region of 20°E to 60°E, 15°S to 20°N. The focus countries of Tanzania, Kenya, Uganda and

Ethiopia are contextualised in Figure 1.1 (a) with the domain shown in detail in Figure 1.1 (b).

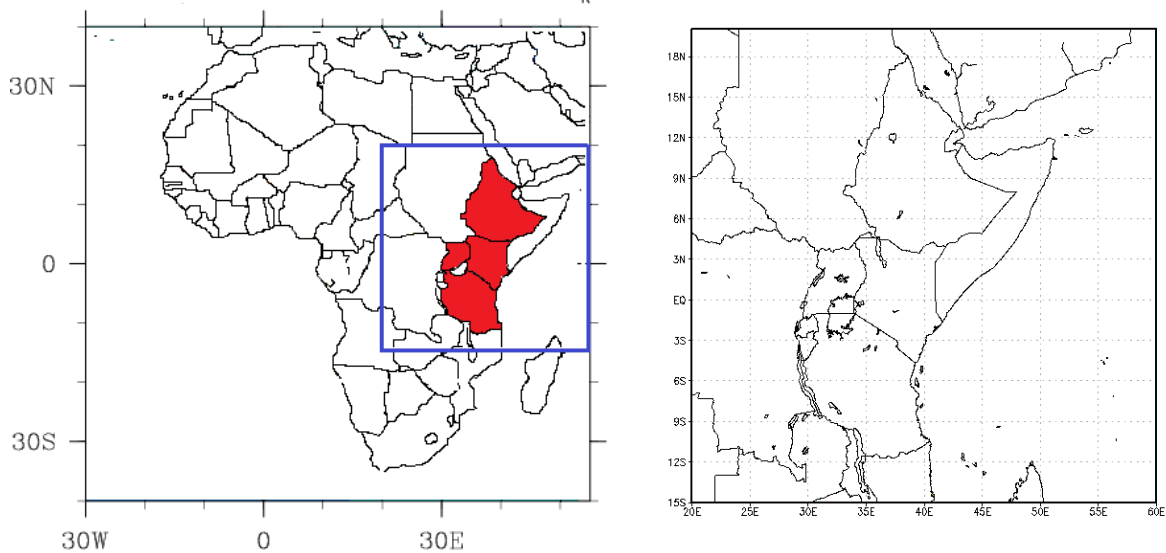


Figure 1.1 (a) Study domain, with focus countries highlighted and (b) Detailed view of the East African domain.

The base climate in East Africa is governed by convective rainfall events linked to the migration of the inter-tropical convergence zone (ITCZ), trailing the region of greatest solar insolation by approximately one month throughout the year (Black et al, 2003). Over the majority of East Africa the annual rainfall cycle features two rainy seasons; the long rains of the boreal spring (March, April and May; MAM) and the short rains of the boreal autumn (October, November and December; OND), co-located with the movement of the ITCZ northwards and subsequently southwards during the year (Mutai and Ward, 2000). In comparison to the long rains, the short rains are more spatially coherent across East Africa but also show greater interannual variability (Manatsa et al, 2012). Annual rainfall across East Africa totals less than would be expected for the latitude, as seen in Figure 1.2, showing the GPCP (Global Rainfall Climatology Project) global annual mean daily rainfall in millimetres.

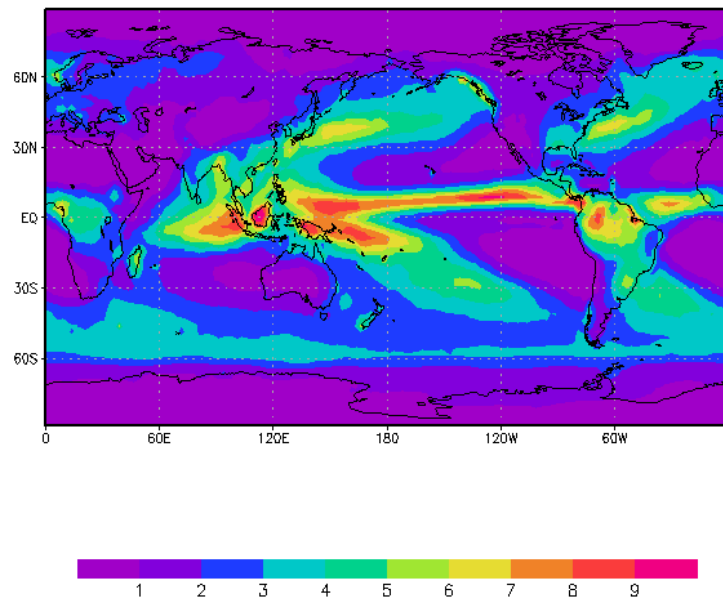


Figure 1.2 Global mean daily GPCP rainfall in mm, 1996-2009.

There are several known large scale drivers of interannual variability of the rainy seasons in East Africa. The El Nino Southern Oscillation (ENSO) has two distinct teleconnections to the region, with cool La Nina events associated with drought during the short rains and warm El Nino events linked to pluvials over East Africa (Mutai and Ward, 2000). Strengthened moist South Atlantic westerly flows are associated with above average East African rainfall and often occur in conjunction with El Nino events (McHugh, 2006). The Indian Ocean Dipole sea surface temperature (SST) variability mode is linked to East African rainfall anomalies (McHugh, 2006). Quantified through the dipole mode index (DMI), the difference in Indian Ocean sea surface temperature anomalies (SSTAs) between the western and eastern Indian Ocean regions, a positive DMI is known to be associated with increased East African rainfall, with the link strongest in the September to November (SON) season (Conway et al, 2007). This type of event is characterised by a

combination of warm SSTAs in the western equatorial Indian Ocean and cool SSTAs around Indonesia (McHugh, 2006).

It is the short rains season over East Africa that forms the basis of this thesis. The short rains are defined as the months of October to December each calendar year, in line with other studies over this region. This research follows the example set by, amongst others, Clark et al, 2003, Anyah and Semazzi, 2007 and Shongwe et al, 2011, in excluding September from all analysis of the short rains period. Whilst the short rains could be said to begin in September each year, atmospheric dynamics in this month are distinct from the other three months as are teleconnections to large scale features such as ENSO (Shongwe et al, 2011).

The short rains are selected over the long rains as the focus of this study for the following reasons. The projected increase in the East African short rains rainfall under rising anthropogenic greenhouse gas emissions by the end of the twenty-first century is one of the most coherent results within the Intergovernmental Panel on Climate Change, Fourth Assessment Report (IPCC AR4), using the CMIP3 models (Solomon et al, eds, 2007). The IPCC Fifth Assessment Report (AR5), released in 2013, reinforced this, stating that it ‘*likely*’ (66-100% probability) that the short rains will see increased rainfall by the end of the twenty-first century (Christensen and Kanikichana, eds, 2013). Large rainfall anomalies in the short rains, either positive or negative, have a greater impact on local communities in the region due to the changes they impart on the regional hydrological cycle (Manatsa et al, 2012).

However, there has been an emergence of research that has suggested possible mechanisms resulting in potential rainfall changes of opposite signs over East Africa through the two rainy seasons (for example Williams and Funk, 2010, Copsey et al, 2006,

Williams et al, 2012, Lyon and Dewitt, 2012, Schreck and Semazzi, 2004) which gives rise to two questions; namely, are the climate models that are projecting increases in short rains rainfall by the end of the twenty-first century over East Africa doing so for sound and physically consistent reasons? Or are the parameterised rainfall values at odds with the dynamics emerging within the models in comparison with observational datasets?

Hence, this investigation focuses on the examination of daily rainfall events within the short rains over four key focus countries; Tanzania, Kenya, Uganda and Ethiopia. The majority of studies over East Africa that address model evaluation do so either through the mean monthly rainfall field or the simulation of known teleconnections to East Africa in AOGCMs (for example Shongewe et al, 2011, Anyah and Semazzi, 2007, Camberlin et al, 2004, Black et al, 2003, Mutai and Ward, 2000). The monthly rainfall field is an accumulation of daily rainfall and therefore the monthly totals are dependent on the fidelity of the systems that deliver rainfall on daily timescales. The structure of daily rainfall regimes in AOGCMs during this season has been neglected, compromising evaluation of model rainfall over East Africa. Work on daily timescales in Africa has so far been limited, with studies focused on southern Africa (Rocha et al, 2008 and Williams et al, 2010), the Sahel (Sylla et al, 2010) and West Africa (Sylla et al, 2012). However, it has been suggested that in Africa the diversity of topography and land-sea contrasts necessitates model evaluation of fine temporal scales of rainfall due to the impact of daily rainfall statistics on natural systems and human impacts (Sylla et al, 2013). This setting is complex and deserves a closer examination than can be achieved by the analysis of monthly totals alone.

In view of these considerations, this study clusters rainfall events in East Africa making use of a methodology not previously applied to the convection dominated tropical circulation. A series of circulation related parameters are examined in association with the

classification of daily rainfall types in order to determine, first from reanalysis data, which of the circulation parameters are consistent with daily rainfall distributions and which are not. Variables include moisture flux divergence, moisture flux, specific humidity, zonal winds and geopotential height. Subsequently, the circulation variable most closely associated with the rainfall types can be used as a basis for assessing the AOGCMs.

Information obtained from the daily classification of rainfall can be utilised in evaluating the daily occurrence of rainfall in AOGCMS over East Africa in comparison to reanalysis and satellite datasets and to investigate the origins of these events in models relative to such datasets. The classification also allows for a detailed examination of anthropogenically induced climate change projections of daily rainfall over East Africa and a greater insight can be gained into the confidence with which climate change projections over the region can be used. An underlying aim throughout this study is that a fuller understanding of the coupled climate model parameterised rainfall field is obtained through the thorough examination of the daily rainfall events in a CMIP3 model subset over East Africa.

Methodologically, self-organising maps (SOMs) are heavily used as a classification tool of daily rainfall and other circulation variables in this thesis. The SOM is a neural network based clustering algorithm which clusters similar rainfall events together in an objective and reproducible way, allowing the examination of associated circulation patterns. The method facilitates visualisation of the distribution of the complexity of synoptic states within a dataset, whilst retaining the continuum of the data (Hewitson and Crane, 2002). The SOM is similar to other clustering techniques with the SOM placing the selected number of nodes within the data space in such a way that they represent the multi-dimensional input data. Nodes are placed further apart in regions of sparse data and closer together where data density is high, instead of being evenly distributed (Hewitson and

Crane, 2002). There are two main benefits guiding the choice of the SOM in this study. First, the SOM is not liable to overtraining and so can be trained through multiple iterations to obtain the ‘best’ visualisation in the array (Hewitson and Crane, 2006). Second, the SOM does not make assumptions (and so is not dependent) on the data belonging to a particular underlying distribution (Hewitson and Crane, 2002).

1.3 Thesis Aims

Taking into consideration current research directions and requirements for study on East Africa during the short rains, the overall aim of this thesis is to evaluate the fidelity of model rainfall projections for the East African short rains from the perspective of changes to daily rainfall regimes. The strategy is to establish whether AOGCMs are good at simulating contemporary rainfall regimes over East Africa during the short rains. This thesis proposes an original approach for achieving this aim, through the daily rainfall field. To reiterate, the majority of studies over East Africa that address model evaluation do so either through the mean monthly rainfall field or the simulation of known teleconnections to East Africa in AOGCMS (for example Shongewe et al, 2011, Anyah and Semazzi, 2007, Camberlin et al, 2004, Black et al, 2003, Mutai and Ward, 2000).

The monthly field often used in research is an accumulation of daily rainfall during the months of the short rains. However, the structure of the rainfall regimes in AOGCMs, comprised of daily events over East Africa, has been neglected in research, with this study aiming to redress the balance. Here, the focus is on whether the daily rainfall fields in AOGCMs are a feasible representation of observed conditions in order to evaluate the overall question of the fidelity of model rainfall projections over East Africa during the short rains. This has been addressed through four distinct research questions:

1. What is the contemporary structure of daily rainfall regimes over East Africa during the October to December short rains season?
 - Is the self-organising map an appropriate tool for this investigation?
2. To what extent can coupled climate models replicate the contemporary structure of the East African daily rainfall regimes of the short rains season?
 - To what extent is the mean rainfall climatology replicated in AOGCMs?
3. What mean changes to the East African short rains occur under the high emissions SRES A2 scenario at the end of the twenty-first century in coupled climate models?
4. What patterns of daily rainfall are responsible for the change in future modelled rainfall?

1.4 Thesis outline

This thesis utilises the following structure through answering the research questions in order to establish an answer to the overall aim of this thesis, to evaluate model rainfall projections for the East African short rains from the perspective of changes to daily rainfall regimes. Each chapter contributes to one of the four research questions; these contributions are outlined below to illustrate how this thesis fulfils its overall aim.

Chapter 2: Data and Methods

This chapter documents the datasets, reanalysis, satellite and model, which are employed in the thesis. Key methodologies, including the rationale and implementation of the self-organising map, are highlighted and explained. Chapter 2 contributes to the aims of the thesis through providing a thorough grounding of the datasets and tools which facilitated the investigation.

Chapter 3: Twentieth century climatology of East Africa and potential changes to the twenty-first century climate

Chapter 3 evaluates the climate of East Africa during the second half of the twentieth century including the climatology, 95th percentile extremes, interannual variability and trends in reanalysis, observational and model datasets. Model projections over East Africa for the 2030s, 2050s and 2090s are also examined.

This chapter forms the foundation for the addressing the thesis aims in later chapters through establishing the mean and extreme climate of East Africa and the mean projections for the region under a high emissions scenario at the end of the twenty-first century. It provides the frame of reference for work undertaken in later chapters. First, it provides the basis to establish results for research question 1, the contemporary structure of daily rainfall regimes within the short rains of East Africa. Second, the preliminary model evaluation in the subset of eight CMIP3 models answers the sub question of the second research question; the extent to which the mean rainfall climatology can be replicated in AOGCMs.

Finally, contextualisation, through projected mean model changes under a high emissions scenario by the end of the twenty-first century in AOGCMs, is provided to look at projections to daily rainfall regimes over East Africa during the short rains.

Chapter 4: Contemporary characteristics of daily rainfall regimes during the short rains in the ERA-40 dataset

Chapter 4 addresses the first research question. First, the sub question of establishing the validity of the self-organising map as a classification method on the convection dominated climate of the region is answered. The ability of the SOM to cluster the rainfall patterns seen in the short rains effectively and assign individual days to groupings within the SOM

is established. In relation to the primary aim of the first research question, daily rainfall events during the last decades of the twentieth century are investigated, with associated circulation indicators also analysed to determine how they characterise rainfall events or dry conditions within the domain.

Chapter 5: Contemporary characteristics of daily rainfall regimes during the short rains – an ERA-40 and TRMM comparison

Chapter 5 addresses the first research question of this thesis. The absence of station data with sufficient spatial and temporal coverage over the domain necessitates the introduction of a second observed approximation of daily rainfall conditions over the East African domain. This chapter focuses on the use of the TRMM satellite derived rainfall series, covering the period 1998-2010. Evaluation of the two differently sourced datasets answers the first research question, establishing a detailed picture of the contemporary characteristics of the short rains of East Africa in the daily field. This information can be used to evaluate the other three research questions.

Chapter 6: East African daily rainfall regimes: a comparison of CMIP3 models with ERA-40 and TRMM

This chapter answers research question 2, the extent to which the contemporary structure of the daily short rains regimes in East Africa can be replicated in AOCGMs during the control period of 1971-2000. Daily rainfall characteristics during the short rains in a subset of CMIP3 models are evaluated for their ability to replicate the rainfall states clustered in Chapter 4 and 5. The data is then constrained to the ERA-40 rainfall SOM established under control conditions for a direct comparison of the models to this dataset. Circulation parameters of daily rainfall events are evaluated to establish if model rainfall events are characterised by the same features as the observed datasets.

Chapter 7: Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario - Part I – Rainfall projections

Chapter 7 addresses the third and part of the fourth research questions. It examines the projected mean changes to the short rains under the SRES A2 scenario at the end of the twenty-first century. Mean change statistics for daily rainfall events are examined and spatial patterns of change across the domain are evaluated. The origins of the projected change in terms of the type of daily event are quantified to begin to address the final research question to establish which daily rainfall patterns are responsible for the changes in modelled rainfall.

Chapter 8: Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario - Part II – Spatial changes and moisture flux projections

Chapter 8 continues to investigate the fourth research question through evaluating the circulation fields in the coupled models which are associated with changes in the daily rainfall types. In particular moisture flux divergence is evaluated because this field was shown in the analysis of ERA-40 and TRMM data (Chapters 4 and 5) to have close spatial association with daily rainfall maxima and conditions of suppressed convection. To address the fourth research aim of this work and the overall aim, the coherence of projections between different variables in the model is evaluated; the daily rainfall field alone is not sufficient for confidence in model projections.

Chapter 9: Conclusions and directions for further research

Chapter 9 is the synthesis of this work; combining results from all chapters in order to converge on an answer to the overall question of the reliability of model rainfall

projections over East Africa during the short rains. The four research questions are considered individually to assess the overall aim. Three directions for future research are suggested.

Chapter 2: Data and Methods

This chapter introduces the datasets and key methods employed within this research.

Datasets originating from observational, satellite, reanalysis, statistical and model sources are employed in this work, namely the ERA-40 reanalysis dataset, ERA-Interim reanalysis dataset (ECMWF Data Server), a set of meteorological station daily rainfall series from the East African region (McSweeney, *personal correspondence*) and TRMM satellite rainfall data (<https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission>). For work on the climatology of East Africa the NCEP reanalysis dataset (<http://www.esrl.noaa.gov/psd/>), the statistically created H08 Japanese rainfall series (Hirabayashi et al, 2008) and the University of East Anglia's CRU gridded dataset are also used (Climate Research Unit, University of East Anglia). For the study of climate projections at the end of the twenty-first century, a subset of eight individual CMIP3 models, alongside an ensemble mean of these models, are utilised. These datasets are described in more detail below.

Methods used in this research are also documented, including the interpolation of the station data to a gridded dataset and a comparison of the interpolated data to both CRU and TRMM datasets to evaluate the rainfall characteristics of the dataset resulting from the interpolation, leading to the methodology used to mask this dataset. Application of the self-organising map methodology which facilitates the analysis of daily rainfall events through clustering similar events together is described. The chapter concludes with an outline of the normalisation methodology, standard statistical tests and the derivation of moisture flux divergence.

2.1 ECMWF ERA Data

2.1.1 ERA-40 dataset

The primary dataset used for work on the twentieth century climate of East Africa is the ERA-40 reanalysis dataset. All of the ERA data (both ERA-40 and ERA-Interim) is provided by the European Centre for Medium-Range Weather Forecasting (ECMFW) data services (<http://www.ecmwf.int/products/data/archive/>).

A reanalysis dataset incorporates observations from several sources using the same data assimilation system and forecast model throughout the time series in order to yield a globally complete and physically consistent dataset for all meteorological variables without the climate ‘jumps’ associated with variations in assimilation systems that would occur if archived analyses were used (Kalnay et al, 1996). The premise is that the only remaining source of discontinuity in a reanalysis record is that which arises from changes to observation methods and coverage (Trenberth et al, 2001). The ERA-40 dataset is derived from a reanalysis of observations produced by ECMWF in collaboration with other institutions and covering the period September 1957 to August 2002. This particular dataset uses observations from the archives of the ECMWF, NCEP and the Japan Meteorological Agency as well as data specifically supplied for this project, including in situ and satellite data from NCAR and EUMETSAT satellite wind products.

Instrumentally, there is a heavy reliance on radiosonde data and SYNOP (surface synoptic observation) reports at the beginning of the ERA-40 record before data collected from aircraft and satellite records became of much greater importance from the 1970s onwards (Uppala et al, 2005). A key requirement of a reanalysis dataset is that the estimated parameters are consistent with the laws of physics in addition to observations (Dee et al, 2011).

ERA-40 data is available on monthly and six-hourly timescales, with output for 0000, 0600, 1200 and 1800 UTC. This work utilises the six-hourly series over the time period 1971-2000, focusing on the months of October to December, with data processed to obtain daily values for the time period 0600 UTC to 0600 UTC. The methodology is to calculate cumulative values for rainfall and the mean value at each grid point for other variables over each twenty-four hour period. The 0600 to 0600 daily period is selected as most likely to be concurrent with the daily range of the station rainfall data based on correlation calculations with station rainfall over the East African domain.

With very limited information available about the station data (see Section 2.2), and initial intentions for it to be closely utilised with the ERA-40 dataset in this thesis, the decision is taken to establish which twenty four hour period each daily rainfall value corresponds to since model rainfall is logged instantaneously against time whereas observed gauge data is read each morning and is an integration of the previous 24 hours but recorded against the date of measurement.

Establishing the optimal combination of four six-hourly reanalysis time slots to match the twenty-four hour daily rainfall record is done by calculating correlations between the ERA-40 rainfall and selected stations in Tanzania, Kenya, Uganda and Ethiopia for three different time periods. All time periods use the ERA-40 dataset 1200 UTC to 1200 UTC, but with offsetting to the station series by twenty four hours in each direction and also with no offset. Results are shown in Table 2.1. Stations in Ethiopia and Uganda show very little correlation with the ERA-40 rainfall in any combination. For stations in Kenya and Tanzania there are strong positive correlations present. The highest correlations are found when the second day of the month in the ERA-40 rainfall data is correlated with the first day of the month in the station data series and so on. Although this is not the case for every station, for half of the stations where a different correlation is higher there is also a strong

positive correlation for the -1 offset. It is concluded that the highest positive correlations correspond with a single daily reading taken at a manual meteorological station each morning and so 0600 UTC to 0600 UTC (equivalent to 0900 to 0900 East African Time, EAT) has been chosen as the twenty four hour daily period from which to calculate ERA-40 cumulative rainfall and mean values for additional variables from the six-hourly measurements.

Table 2.1 Correlation between station rainfall series and ERA-40 rainfall series, strongest positive correlations (r values) are highlighted in green

Country	Station	Correlation (+0)	Correlation (+1)	Correlation (-1)
Ethiopia (July 1988)	Bahirdar	-0.1099	-0.1093	-0.2460
	Gore	-0.1275	-0.1231	-0.1118
	Kulumsa Rc	0.0240	-0.0893	0.1782
Kenya (November 1976)	Elburgon F	0.2571	0.1508	0.6633
	Nyeri, M. O	0.0528	0.1166	0.1681
	Kiambu, D. C.	0.1983	0.1581	0.3332
	Limuru Mab	0.0398	0.3207	-0.0190
	Lamu Met S	0.2931	0.1017	0.0988
	Wajiir Met	0.4785	0.2766	0.0088
	Malindi Me	-0.0093	-0.2814	0.3677
Tanzania (November 1990)	Amani	0.4062	0.4203	0.3625
	Dar	0.5776	0.7381	0.6393
	Bagamoyo	0.5364	0.3107	0.6085
	Lushoto	0.0300	0.2369	0.4323
	Mafia	0.4245	0.5891	0.6610
	Morogoro	0.6973	0.4782	0.5953
Uganda (November 1978)	Buluwula G	-0.2002	-0.0609	0.0602
	Bushenyi A	-0.0544	-0.1517	-0.2618
	Entebbe In	-0.0990	0.0121	0.0232

Two pressure levels of circulation fields are chosen as a starting point to investigate circulation states physically consistent with rainfall events; a near surface level and the 850hPa level, located at a level equivalent to approximately 1.5km in the atmosphere over East Africa. There is evidence in the literature to support using the 850hPa level to represent circulation patterns over East Africa; for example, Mpeta and Jury (2001),

Mapande and Reason (2005), Camberlin and Planchon (1997), Kijazi and Reason (2005) and Camberlin et al (2010) and Philippon et al (2002).

Variables used are zonal (U) winds, meridional (V) winds, specific humidity, relative humidity, moisture flux, moisture flux divergence and geopotential height at both the surface and 850hPa levels. A suite of variables is used to aid the overall aim of evaluating the model rainfall field on the daily temporal scale to understand rainfall projections for East Africa. Model rainfall is parameterised and thus difficult to simulate in models; finding a variable with a field that is physically consistent with the presence of daily rainfall events in the reanalysis and satellite datasets allows this variable to be utilised in the examination of the model rainfall field, both for contemporary and late twenty-first century conditions.

ERA-40 is a second generation product (following the ECMWF ERA-15 product) with the goal of producing the best possible set of analyses, given constraints of changing observing systems and limited computational resources (Uppala et al, 2005). The fit of the dataset to other climatological series for individual variables and measurement instruments broadly improves as the record continues (Uppala et al, 2005). This steers the decision to use the latest possible period of the ERA-40 data, even though this means compromising on the number of recording stations with data available for those years.

The ERA-40 data can be evaluated through the size of analysis increments. An analysis increment measures the amount the background forecast is adjusted by the observations assimilated at a particular time. A smaller analysis increment means that the observations influence the assimilation through lesser adjustments that provide a better background forecast for future analyses rather than through very large corrections to poor background forecasts (Uppala et al, 2005). Analysis increments are, on the whole, smaller in ERA-40

than ERA-15 showing improvement between the two systems. Another improvement over ERA-15 is in the representation of African easterly waves from source near the Ethiopian Highlands in the ERA-40 dataset (Uppala et al, 2005).

Whilst the ERA-40 series is a reasonable source of data to use over East Africa (in combination with a satellite sourced rainfall series), given the lack of temporally and spatially comprehensive station series and the known positive characteristics of its climatic representation, there are recognised problems with the series which must be considered in conjunction with its use.

Uppala et al (2005) identify two main issues with the ERA-40 series; excessive tropical oceanic rainfall in short term forecasts run from the analyses and a Brewer-Dobson circulation that is too strong. The Brewer-Dobson circulation is the atmospheric meridional overturning circulation, as occurs in the stratosphere (Bunzel and Schmidt, 2013). Of particular relevance to this study, exaggerated rainfall occurs where rainfall is observed to be present. This results in realistic looking spatial patterns of rainfall when actually rainfall is enhanced due to a cycle of events set off by the way humidity analysis spreads increments in the horizontal plane. This spreading both adds moisture in areas that are too dry and also to nearby regions that are already close to the point of saturation.

Uppala et al (2005) suggest that this then increases the tropical circulation through excess latent heating in moist regions with excess drying occurring in regions of descent, creating an enhanced rainfall pattern, though one with many realistic looking characteristics. The positive rainfall bias in ERA-40 is most pronounced in the tropics, especially over oceans, when compared with the Global Rainfall Climatology Project (GPCP), with the difference larger than can be explained by uncertainty in the GPCP data. In extratropical regions both mean and interannual values compare more favourably with the GPCP (Betts and Beljaars, 2003). Hagemann et al (2005) conclude that because the biases within the hydrological

cycle of ERA-40 differ during different periods of the series, being strongly influenced by inputted observational data, that hydrological trends should not be derived directly from this series. ERA-40 rainfall values after the eruption of Mount Pinatubo in 1991 are higher than actually occurred globally. This is because the effects of volcanic aerosols need to be included in the bias corrections, something that was problematic for some data sources including from the NOAA-12 satellite (Uppala et al, 2005).

There are a limited number of studies evaluating the ERA-40 data over East Africa, which provide more specific information about the series over the domain of interest. Diro et al (2009) compare the ERA-40 and NCEP reanalysis rainfall estimates over Ethiopia. Both products have deficiencies but ERA-40 is found to capture the annual cycle of rainfall better than NCEP over much of the country although the Kiremt (JJAS) peak is exaggerated in the northwest and west and the Belg rainfall season is also overestimated in west and central parts of Ethiopia. ERA-40 captures the spatial variability of the rainfall well. Over Uganda, ERA-40 underestimated rainfall by approximately 10mm per dekad (Maidment et al, 2012). The authors also found that ERA-40 underestimates the seasonal pattern, in particular during peak rainfall periods.

2.1.2 ERA-Interim dataset

The ERA-Interim dataset is used in conjunction with the TRMM satellite derived rainfall series. This reanalysis product follows ERA-40 as part of the preparation of a product to replace ERA-40 with a new reanalysis dating back to the early twentieth century, but this particular product begins from January 1989 (Dee et al, 2011). The aim of this product was to address issues from the ERA-40 series including depiction of the hydrological cycle, quality of stratospheric circulation and consistency temporally of reanalysed geophysical variables.

Both ERA datasets use the ECMWF IFS (Integrated Forecasting System); ERA-40 with the IFS Cy24r3 release and ERA-Interim with the Cy31r2 version. Improvements made between the release of these two versions (and through the intervening versions) incorporate changes to convection schemes, cloud representation, radiation, ozone chemistry and an increase in vertical levels (Dee et al, 2011). Up to 2002, most observations inputted are as those initially prepared for ERA-40. After 2002, the data input is from ECMWF's operational forecasting archive. Additionally, some datasets were obtained specifically for ERA-Interim including AMV wind data from geostationary satellites, reprocessed ozone profiles from the GOME instrument and reprocessed altimeter wave height data from ERS-1 and ERS-2 (Dee et al, 2011).

Dee et al (2011) conclude that the hydrological cycle is improved in ERA-Interim in comparison to ERA-40. Rainfall estimates are one of the most challenging aspects of a reanalysis series, involving many parameters only indirectly constrained by observations, with approximations used in the representation of moist processes having a large impact on how realistic the hydrological cycle is. In ERA-40 there was excessive rainfall over tropical oceans and a poor representation of global transport of moisture in the atmosphere (Uppala et al, 2005). In ERA-Interim, two major factors for these problems were overhauled; the humidity analysis and the radiance data bias correction method. Whilst ERA-Interim rainfall does not have the high overestimation of ERA-40 after the 1991 Mount Pinatubo eruption there is a sudden decrease in the global mean rainfall of approximately 1mm day^{-1} at the beginning of 1992 with continued smaller reductions until after 2006 (Dee et al, 2011). These reductions are only seen over oceans and so require less consideration here when the key focus is continental rainfall over East Africa.

2.2 NCEP/NCAR dataset

In addition to the ERA-40 data, two further reanalysis datasets are employed in this work, one being ERA-Interim. The second is the result of collaboration between the National Center for Environmental Prediction (NCEP) and the National Center for Atmospheric Research (NCAR) and will be referred to as the NCEP dataset from this point. This dataset is similar to ERA-40, with the main premise being to remove the ‘climate jumps’ associated with changes to assimilation systems (Kalnay et al , 1996).

The data used in the NCEP reanalysis dataset comes from several different sources (Kalnay et al, 1996). Global rawinsonde data is sourced from the NCEP GTS (Global Telecommunications System) data and also from national archives of countries including South Africa, Australia, Canada, Argentina, Brazil, the United Kingdom, France and the United States. Data from China and the USSR is obtained through bilateral agreements with academic sources and data from Japan obtained from the Japanese Meteorological Agency is also used. Surface marine data is obtained from the COADS (Comprehensive Ocean-Atmosphere Data Set) dataset, containing information from ships, fixed buoys, drifting buoys, pack-ice buoys and near surface data from ocean surface reports. Aircraft data is obtained from the NCEP GTS and other sources including data direct from New Zealand and from experiments encompassing the GARP (Global Atmospheric Research Program) Atlantic Tropical Experiment and the First GARP Global Experiment (FGGE). The NCEP dataset additionally contains data from the Tropical Wind Energy Conversion Reference Level Equipment (TWERLE) constant pressure balloons for the southern hemisphere. Surface land synoptic data is used from the NCEP GTS and air force sources. Satellite sounder data utilised in the series is available from 1979 for the TOVS sounders, with data before that coming from VTPR and HIRS data. Special Sensing Microwave/Imaging data is used for surface wind speeds. Finally, satellite cloud drift

winds from the Geostationary Meteorological Satellite (GMS) cloud drift winds are included in the series from 1978-1991.

The NCEP reanalysis data is used in the evaluation of East African climate in Chapter 3 to record mean states of temperature and rainfall, building a profile of the East African climate during the twentieth century. Trends over the second half of the twentieth century are examined, and the NCEP data is also used to evaluate representations of the East African climate in the CMIP3 AOGCM subset. More than one reanalysis dataset is used in this thesis because of the known differences between the datasets.

As with ERA-40, there are known issues with NCEP. Trenberth and Guillemot (1998) evaluate NCEP rainfall using the Xie-Arkin global rainfall estimates, noting several biases. They highlight a double ITCZ in the central Pacific, which results in the South Pacific Convergence Zone not being well captured in the NCEP series. The NCEP series has a bias for excessive rainfall over southeastern USA and southeastern Asia but rainfall associated with ENSO in the central tropical Pacific Ocean is underestimated. The Hadley circulation in the NCEP reanalysis series is much weaker than in the ERA-40 series (Trenberth and Guillemot, 1998). NCEP has not been evaluated extensively over East Africa. However, Diro et al (2009) did find that NCEP was not as good at capturing the annual Ethiopian rainfall cycle as ERA-40.

2.3 H08 - a statistical rainfall series

This global dataset covers the period 1948-2006 at a spatial resolution of 0.5° by 0.5° with temporal resolution ranging from three-hourly to daily depending on the variable (Hirabayashi et al, 2008). The H08 series differs from a reanalysis series in that it is created statistically, from a gridded product whereby the statistical data set is interpolated to higher resolution, rather than through a numerical prediction model. Employed in this

work is the rainfall series, which was created using the global monthly 0.5° rainfall product PREC/L (NOAA Rainfall Reconstruction over Land) as a base which was divided into daily values using a gamma distribution algorithm. A previous product (Hirabayashi, dating to 2005) did not include spatial characteristics of the rainfall (that is, each grid box was created independently from the adjoining ones), something that the authors try to correct here through relating rainfall in adjacent grid boxes. This is done by utilising information from two products; NCEP/NCAR (corrected by CRU) for the period 1948-1977 and the GTS (Global Telecommunications System) from 1978 onwards with the rainfall intensity of the H08 data series, obtained by the gamma distribution, placed in the same order as either the NCEP or GTS product (Hirabayashi et al, 2008). For East Asia and India, the daily rainfall values were replaced by regional products, but these regions are not used in this study.

The spatial pattern of rainfall in the H08 product is improved over the earlier product (published in 2005 by Hirabayashi et al) due to the spatial distribution technique applied. The authors feel that this gives their dataset a strength that is not seen in reanalysis datasets where the number of rain days is often overestimated considerably in contrast to their statistical method where the number of rain days is constrained to that which was observed (Hirabayashi et al, 2008). In this study, the H08 is used in evaluation of the AOGCMs during the control period, in conjunction with the NCEP reanalysis dataset.

2.4 UEA CRU gridded dataset

The University of East Anglia's Climate Research Unit has created gridded climatic datasets. Utilised in this study is the 0.5° by 0.5° resolution (equivalent to approximately 55km at the equator) gridded monthly temperature series version CRU TS 2.1 (Mitchell and Jones, 2005). This dataset follows on from earlier versions (New et al, 1999 and 2000,

and Mitchell et al, 2004). CRU TS 2.1 includes data from seven different sources and where a particular station is included in more than one dataset, priority is given to the version likely to be more reliable. Each individual station record was checked and corrected for inhomogeneities before being merged into the final database. To create the gridded series anomalies are interpolated rather than absolute values, this is done by the station time series being expressed as anomalies relative to a 1961-1990 control period which are then interpolated to a grid and combined with the associated grid of normal for the same control period (Mitchell and Jones, 2005). Stations with missing values can also be used under this method which is particularly important in creating a gridded series where station data availability is already sparse. Given that a normal is necessary to use the anomaly method, for those stations without one (including stations that did not exist in the control period) an estimated normal was created. This was done by using neighbours to create a reference series including the 1961-1990 period from which the mean of this reference series during 1961-1990 was used as the normal (Mitchell and Jones, 2005). Interpolation of the anomalies was initially on to a continuous surface before a regular 0.5° latitude longitude grid was derived. ‘Dummy’ stations, of zero anomalies, were located where no station existed within the correlation decay distance to prevent extrapolation of data outside realistic possibility. Anomaly values were then combined with the 1961-1990 normals to obtain absolute values. Impossible values, for example negative rainfall values, were adjusted to the nearest possible value and a ratio was used to match the 1961-1990 mean to the normal (Mitchell and Jones, 2005).

The process used to construct the CRU TS 2.1 does have limitations. The method largely detects abrupt inhomogeneities and will not pick up on widespread gradual inhomogeneities. Being an automated method is essential for the quantity of data involved, although this removes all advantages of a manual method. Nonetheless, the automated

approach is suitable for a technique designed to provide an estimate of interannual climate variability as opposed to the detection of climatic trends (Mitchell and Jones, 2005). For this study, the CRU series is not used for detection of anthropogenic climate change and so it is an appropriate series to use to examine the representation of twentieth century climate.

2.5 TRMM satellite rainfall series

TRMM (Tropical Rainfall Measuring Mission) satellite data covers the broad region of 35°S to 35°N (extended to 50° either side of the equator in recent versions) and is used in this study to evaluate the mean characteristics of the interpolated station rainfall series and as a second rainfall series to establish contemporary characteristics of daily regimes during the short rains. The TRMM project aims to address the accurate spatial and temporal measurement of tropical rainfall and is a joint mission between NASA (National Aeronautics and Space Administration) and JAXA (Japan Aerospace Exploration Agency) (<http://trmm.gsfc.nasa.gov>). Instruments on the satellite include rainfall radar, TRMM microwave imager, a visible and infrared scanner, clouds and the Earth's radiant energy system and a lightening imaging sensor (<http://disc.sci.gsfc.nasa.gov>).

The B342 daily algorithm is used to produce TRMM merged high quality (HQ) and infrared (IR) rainfall and root-mean-square (RMS) rainfall-error estimates. The gridded estimates of rainfall have a 3 hour temporal resolution and 0.25° by 0.25° spatial resolution, from which a daily accumulated rainfall product is derived. The estimates are produced through the following stages. First, the microwave estimates of rainfall are calibrated and combined. These calibrated microwave rainfall estimates are then used to create infrared rainfall estimates and finally the microwave and infrared estimates are combined (<http://trmm.gsfc.nasa.gov>). The TRMM rainfall series is a relatively recent product and so only three years of overlap with the control period for this work exist, from

1998 to 2000; these three years and a climatology period of 1998 to 2010 are used in this study.

There are caveats related to the use of the TRMM data. Occurrence of rainfall over land tends to be underestimated in the series due to satellites sometimes failing to pick up light rainfall and that which is enhanced by flow lifting over mountains. Some, although not all, coastal areas also see an underestimation of rainfall. Conversely, arid coastal areas show persistent rainfall artefacts. Sometimes there is a low bias in areas of particularly complex topography due to the bias of gauge location towards lower elevation (<http://disc.sci.gsfc.nasa.gov>).

2.6 Meteorological station rainfall data

2.6.1 Introduction to the station data

Access was provided to a dataset of station rainfall records (McSweeney, personal correspondence). This dataset consists of 353 stations in total distributed over Ethiopia (21 stations), Kenya (211 stations), Tanzania (20 stations) and Uganda (101 stations). The distribution of these stations can be seen in Figures 2.1 to 2.4, with the shading indicating topography of each country.

This dataset is extremely disparate both in terms of spatial coverage over the domain and the temporal completeness of data within each series so a considerable amount of pre-processing is necessary. The temporal range of each individual series is from October 1909 to December 2000, although only data collected from 1950 onwards is utilised in the interpolation of the individual stations to a gridded dataset. One station in Kenya (reference K9036022, E. A. R. C. N., located at -0.717 latitude and 36.433 longitude) is

excluded from the dataset because it only contains data before 1950. Use of the daily station series is discussed in Chapter 3.

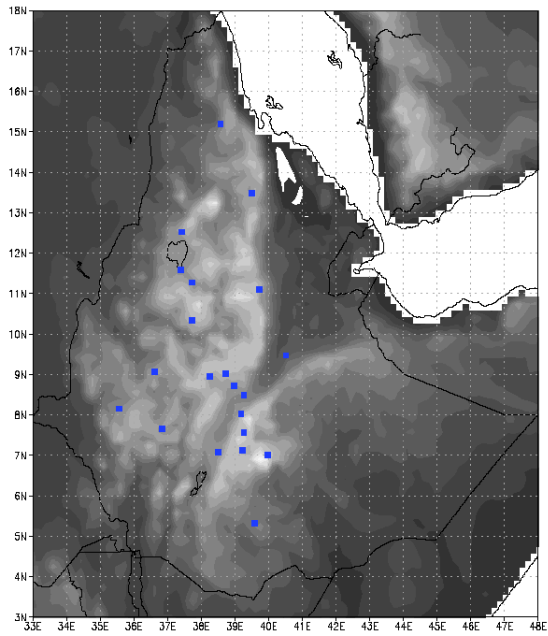


Figure 2.1 Location of rainfall stations in Ethiopia, topography as height in metres

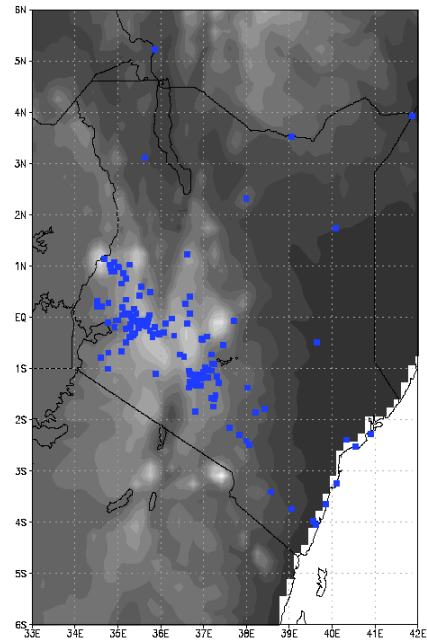


Figure 2.2 Location of rainfall stations in Kenya, topography as height in metres

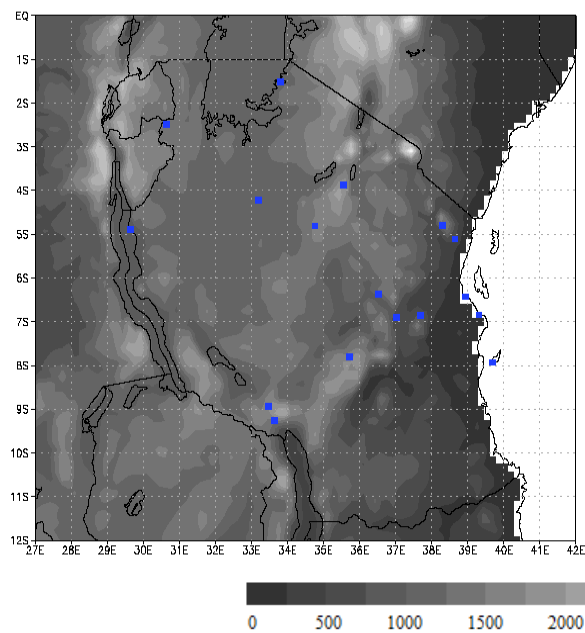


Figure 2.3 Location of rainfall stations in Tanzania, topography as height in metres

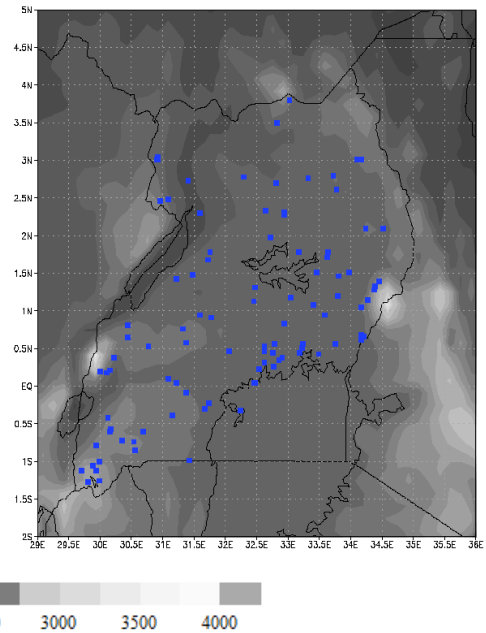


Figure 2.4 Location of rainfall station in Uganda, topography as height in metres



2.6.2 Interpolation of Station Data

Due to the spatial and temporal incompleteness of the station data series, it has been decided to use all data series that have data during the study period, 1971 to 2000, in order to get the widest possible coverage of rainfall over the domain. This results in the 352 stations containing data between 1950 and 2000 being used to create a daily gridded rainfall series using an interpolation method to a 0.5° by 1° grid. Interpolation is a method that utilises existing values within an area of observations in order to estimate values that are missing within the same area (Yang and Hodler, 2000). The initial grid chosen for the interpolated series is 28° to 42°E , 9°S to 16°N ; a rectangular domain around the region of station locations. The methodology uses a triangulation interpolation method, based on Delaunay triangulation. This process involves connecting together vertices to form triangles that do not overlap and is appropriate for this dataset due to its ability to cope with a dataset containing areas of both high and low density observations as is the case here (Taylor et al, 2004). The triangular plane is then sampled onto a regular grid. This interpolation was automated using NCL.

2.6.3 Evaluation of the gridded rainfall series

Following the interpolation of the station data it is necessary to evaluate the resulting gridded series, in particular due to the spatially uneven input data. The data is compared to two other gridded rainfall series, namely the University of East Anglia's CRU data and the TRMM satellite data (see Sections 2.4 and 2.5 respectively). For the CRU dataset evaluation is carried out on the period 1971-2000, October to December. However, for the TRMM dataset, more limited data availability meant that the comparison is done for the years 1998 to 2000, the only overlapping time period.

2.6.3.1 CRU vs. interpolated station series

First, monthly mean rainfall totals and daily mean rainfall figures for October, November and December are compared. Figures 2.5 (a) and (b) to 2.10 (a) and (b) show the monthly and daily mean rainfall totals for the CRU (a) and interpolated (b) series. For these plots and throughout the thesis, rainfall plots are shaded so that blue/purple colours show drier conditions and orange/red show heavier rainfall events.

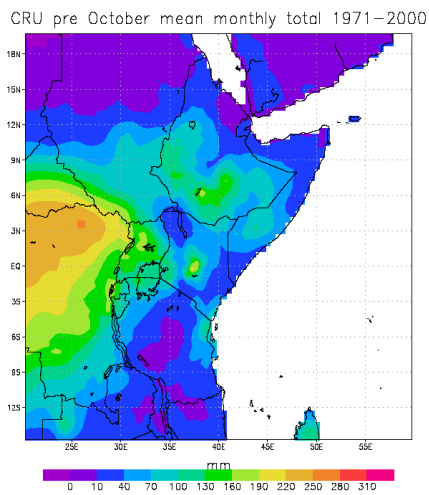


Figure 2.5 (a) CRU October mean monthly rainfall total in mm

Stat pre October monthly total 1971–2000

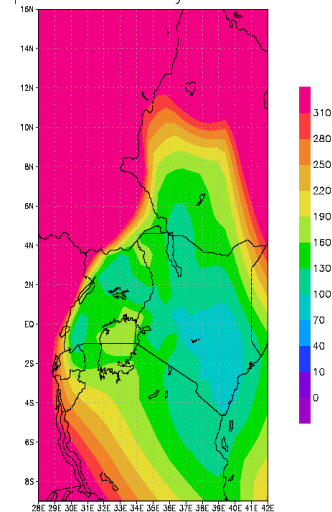


Figure 2.5 (b) Station October mean monthly rainfall total in mm

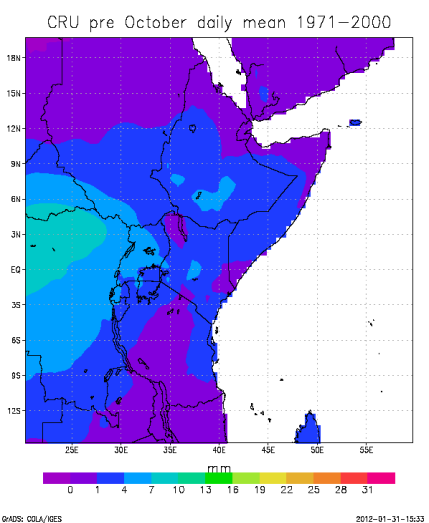


Figure 2.6 (a) CRU October daily mean rainfall total in mm

Stat pre October mean 1971–2000

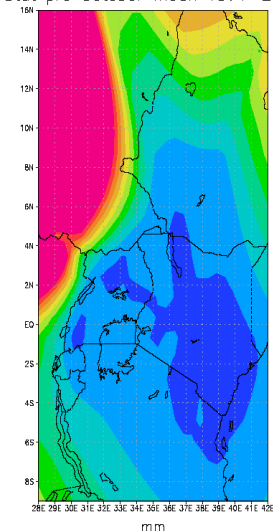


Figure 2.6 (b) Station October daily mean rainfall total in mm

CRU pre November mean monthly total 1971–2000

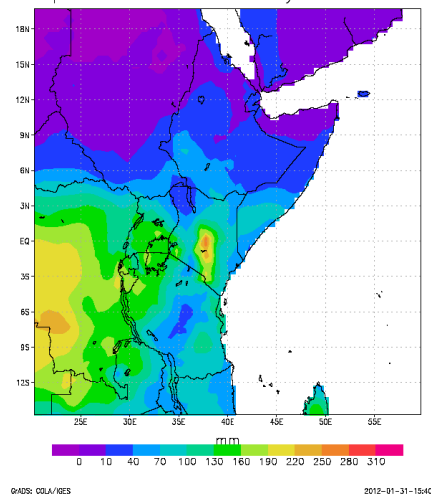


Figure 2.7 (a) CRU November monthly total rainfall total in mm

Stat pre November monthly total 1971–2000

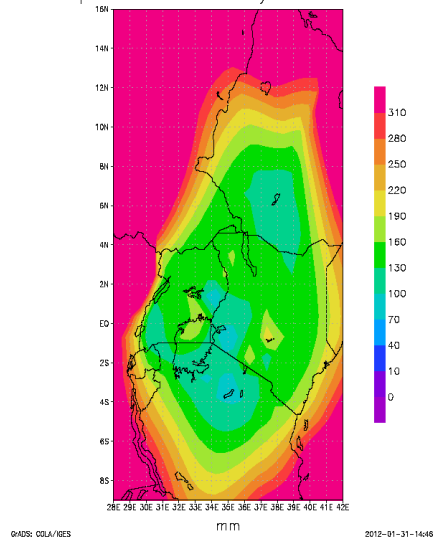


Figure 2.7 (b) Station November monthly total rainfall total in mm

CRU pre November daily mean 1971–2000

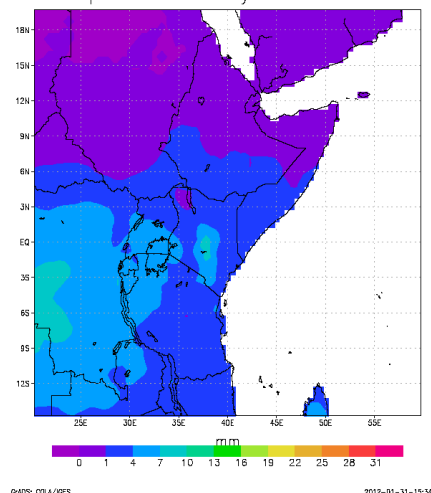


Figure 2.8 (a) CRU November daily mean rainfall total in mm

Stat pre November mean 1971–2000

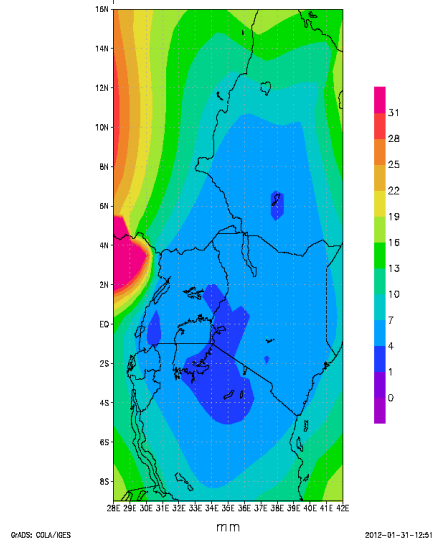


Figure 2.8 (b) Station November daily mean rainfall total in mm

CRU pre December mean monthly total 1971–2000

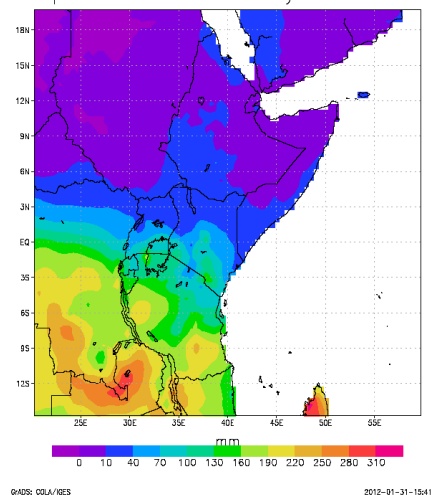


Figure 2.9 (a) CRU December monthly total rainfall total in mm

Stat pre December monthly total 1971–2000

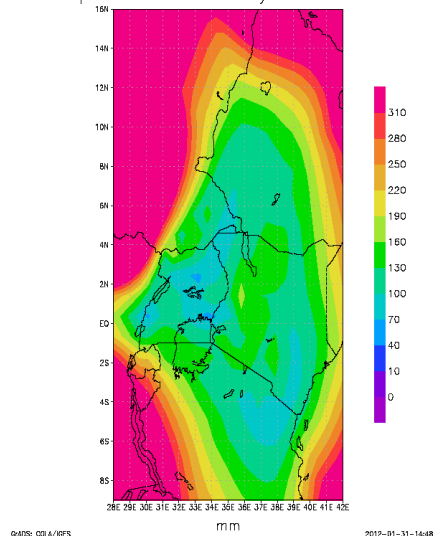


Figure 2.9 (b) Station December monthly total rainfall total in mm

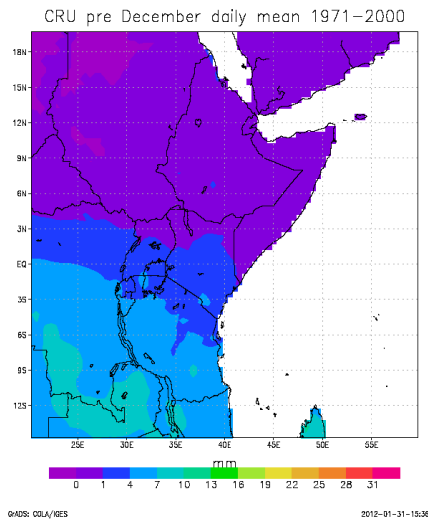


Figure 2.10 (a) CRU December daily mean rainfall total in mm

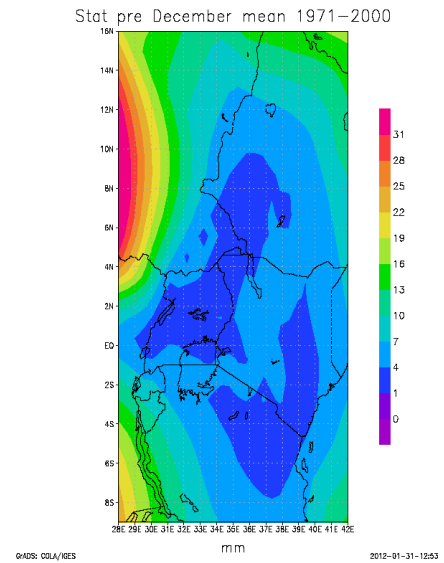


Figure 2.10 (b) Station December daily mean rainfall total in mm

Problems with the interpolated station series are immediately apparent. For the mean monthly rainfall total plots, values around the edges of the domain are unrealistically high where interpolation has occurred between the outermost stations and the edge of the rectangular domain. For example, over most of Sudan and South Sudan the interpolated series has monthly rainfall totals in excess of 310mm whereas the CRU data shows very little rainfall falling over this country with monthly totals during the OND period generally measuring less than 10mm. This can be explained due to the combination of method used and the lack of any station data from Sudan and South Sudan.

Taking each month individually, there are both similarities and differences between the CRU rainfall and the interpolated station rainfall. During October, the area outside of the spatial locations of the input stations for the series has rainfall that is much higher, in particular over South Sudan and northern Ethiopia. Over the four focus countries where the stations are located, the daily mean rainfall measures are in approximately the correct range of up to 220mm/month, although there are differences in the spatial distribution of the monthly rainfall with the peak located along the East African coast in the interpolated

series in comparison to a central location in the CRU series. Both datasets show raised mean monthly October rainfall total over Lake Victoria, measuring 190mm in the interpolated dataset and 130mm in CRU, and also over Uganda and southern Ethiopia compared to the rest of the domain. During November the outer edges of the domain show rainfall that is much too high in the interpolated series. Compared to October, the monthly rainfall over the focus countries is higher in November, and this is reflected in the interpolated station rainfall series with both datasets showing monthly rainfall totals of up to 280mm. There are also differences between the two series in November. Lowest mean rainfall in November is located over central to southern Tanzania in the CRU dataset and measures only 40mm whereas in the interpolated station series it is further north in Tanzania, stretching into Kenya, and also records a higher value of 70mm. For December the situation is very similar to that described for the previous two months. The mean rainfall is in approximately the correct range over the focus countries of between 70 and 220mm. Although the specific patterns of rainfall are not distinguished clearly in the interpolated series it is possible to pick out lighter rainfall over northern Uganda in both rainfall series.

Comparison of the mean rainfall patterns of CRU and the interpolated series demonstrate that the interpolated series is particularly weak around the edges of the domain where no input stations are located. This also serves to mask the progression of convective rainfall from north to south during the October to December rainy season although it is partially evident. However, over southern Ethiopia, Kenya, Tanzania and Uganda the interpolated series has some success in replicating rainfall within expected ranges for each month. In order to gain further understanding of the interpolated data, the two series for each individual month of the period 1971-2000 OND are compared (plots not shown). From this analysis, several conclusions can be made:

- The interpolated station series is very poor at accurately representing East African rainfall outside of the area where stations are located, especially over Sudan and South Sudan.
- Northern Ethiopia (from 10°N) has only seven stations, only two of which contain data after 1988, to input into the interpolated series; coverage of this northern part of East Africa is very sparse and rainfall here is often overestimated.
- Through the central band of the domain where the input stations are concentrated (Uganda and Kenya), the interpolated series is good. Rainfall amounts are consistently in the correct range and spatial patterns for each month are also frequently correct.
- If a rainfall centre is located over Ethiopia it frequently becomes exaggerated in magnitude due to the scarcity of stations.
- Where coastal rainfall occurs, caution is needed in interpreting the intensity, which is sometimes overestimated offshore where there is no station data.
- Rainfall centres over Tanzania become too high due to interpolation between only a few input stations.

The main issue emerging from this comparison is the overestimation of rainfall outside the region of high station density, with the problem particularly acute where no stations are located. The next stage is to compare the interpolated station series to a second rainfall dataset, the TRMM satellite rainfall series.

2.6.3.2 TRMM vs. interpolated station series

The same process of comparing the station interpolation series to an established rainfall dataset is repeated using the TRMM satellite dataset. This dataset begins in 1998 so the three overlapping years from 1998-2000 are used from TRMM and the interpolated series

in this comparison. For this reason, the comparison is discussed in this separate section. Figures 2.11 (a) to (b) to 2.16 (a) to (b) show the monthly mean rainfall totals and daily mean rainfall for October, November and December respectively.

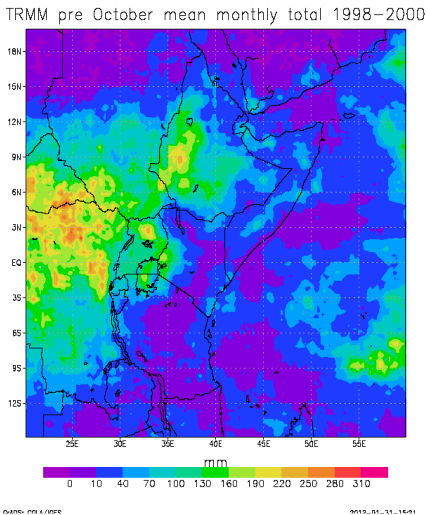


Figure 2.11 (a) TRMM October mean monthly rainfall total in mm

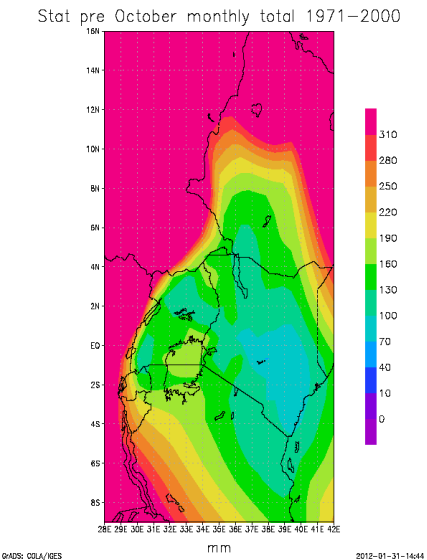


Figure 2.11 (b) Station October mean monthly rainfall total in mm

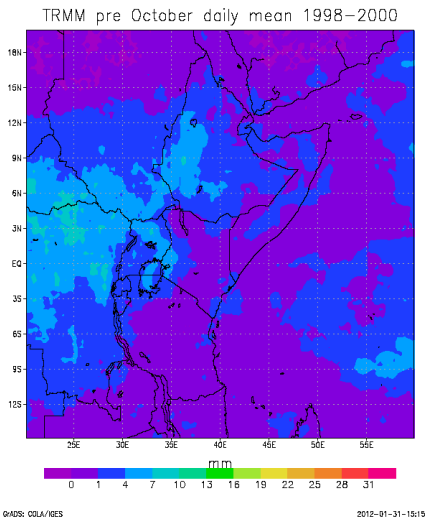


Figure 2.12 (a) TRMM October daily mean rainfall mm

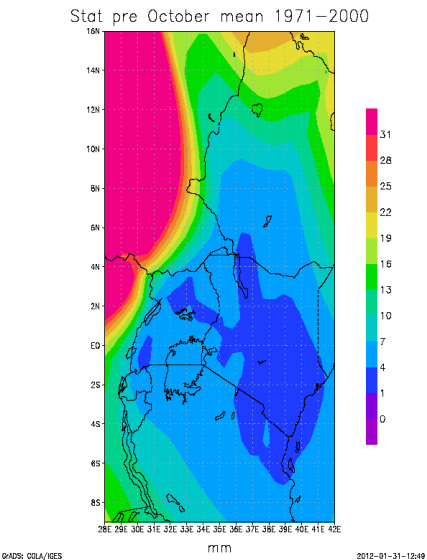


Figure 2.12 (b) Station October daily mean rainfall mm

TRMM pre November mean monthly total 1998–2000

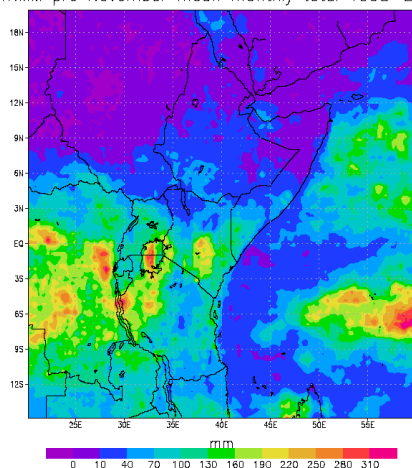


Figure 2.13 (a) TRMM November mean monthly rainfall total in mm

Stat pre November monthly total 1971–2000

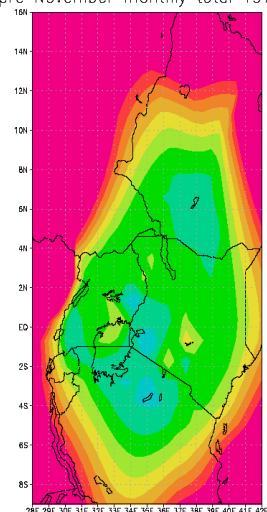


Figure 2.13 (b) Station November mean monthly rainfall total in mm

TRMM pre November daily mean 1998–2000

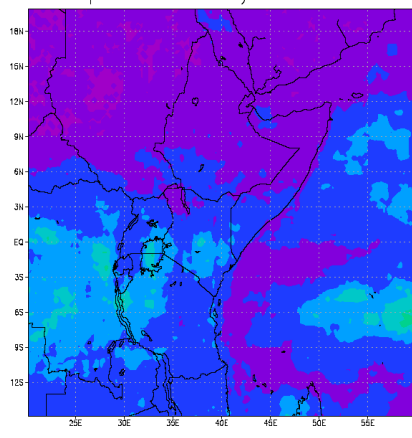


Figure 2.14 (a) TRMM November daily mean rainfall in mm

Stat pre November mean 1971–2000

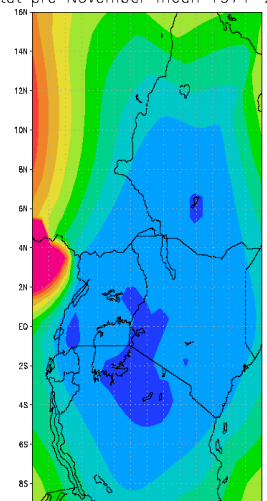


Figure 2.14 (b) Station November daily mean rainfall in mm

TRMM pre December mean monthly total 1998–2000

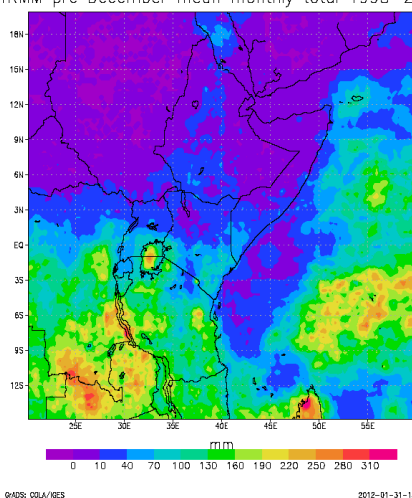


Figure 2.15 (a) TRMM December monthly mean rainfall in mm

Stat pre December monthly total 1971–2000

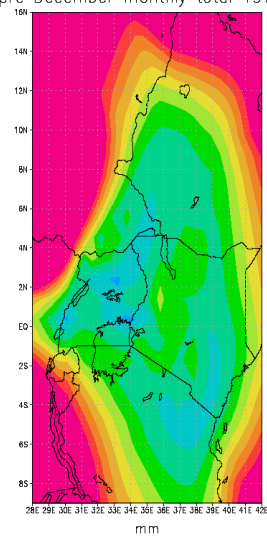


Figure 2.15 (b) Station December monthly mean rainfall in mm

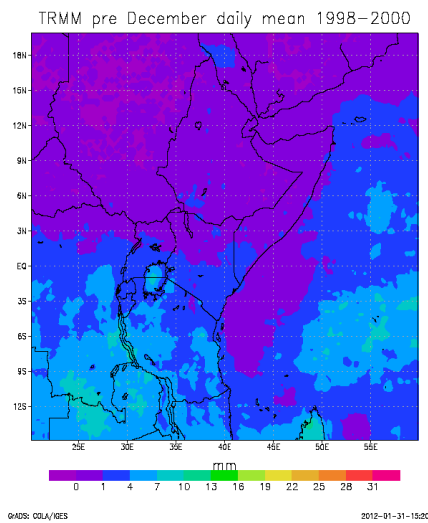


Figure 2.16 (a) TRMM December daily mean rainfall in mm

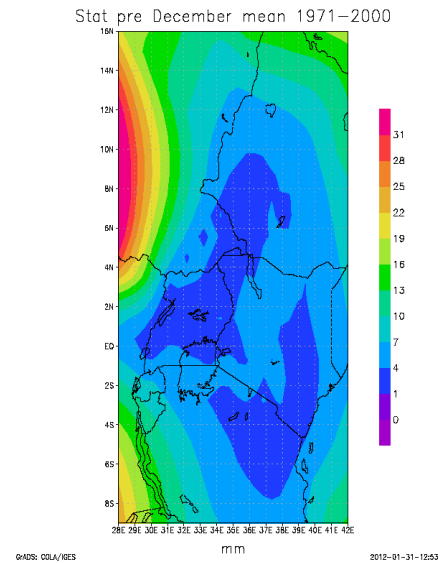


Figure 2.16 (b) Station December daily mean rainfall in mm

For the monthly mean plots of TRMM data in comparison to the station rainfall series, similar conclusions to the work between CRU and the interpolated series emerge. It is confirmed that rainfall outside the station locations is grossly overestimated. Much of northern Ethiopia, Sudan and South Sudan see monthly rainfall totals in excess of 310mm whereas values in the TRMM dataset measure less than 40mm. Some of the agreements taken from the first evaluation also hold true. The centre of rainfall noted over Lake Victoria in the CRU and station rainfall series during October is also present in the TRMM data, where it continues into November. This is matched in the station rainfall series, with monthly rainfall figures recorded at between 130mm and 190mm in both datasets in October.

When comparing the nine overlapping individual months of the TRMM data with the station rainfall series there are good similarities between the spatial patterns and magnitude of rainfall, especially over the focus countries. For example, centres over Kenya, Uganda and Tanzania consistently appear in both datasets. In a similar way to the CRU, this further

analysis confirms that rainfall is overestimated both over Ethiopia where station data is particularly scarce and Sudan and South Sudan where there are no stations.

2.6.4 Masking of the interpolated station series

Due to the results of the evaluation of the gridded rainfall series, it has been decided to mask the domain so that the areas of the series where characteristics of the annual rainfall series are unsatisfactory can be disregarded from analysis. The chosen mask is shown in Figure 2.17.

The primary areas masked out are located in the western regions of the domain. No station rainfall data is available for Sudan or South Sudan and the interpolation method led to unrealistically high rainfall values in this area so this region is masked out to the Ethiopian border at approximately 33°E. Ethiopia, north of 12°N, is also masked out; this region contains only three stations and this proves inadequate to get useful rainfall values for the northernmost part of the domain. The south west of the domain is masked out in an area extending north-eastwards into Tanzania. Again this masking is due to the interpolation method calculating unrealistically high values in this region.

Although most of Ethiopia is left in the gridded rainfall series, caution should be exercised from 9° northwards, where available station data is sparse and so there is a tendency for interpolated rainfall to be too high in this region. It is important to note that this mask applies only to the interpolated station series; all other datasets maintain the larger rectangular shaped East African domain.

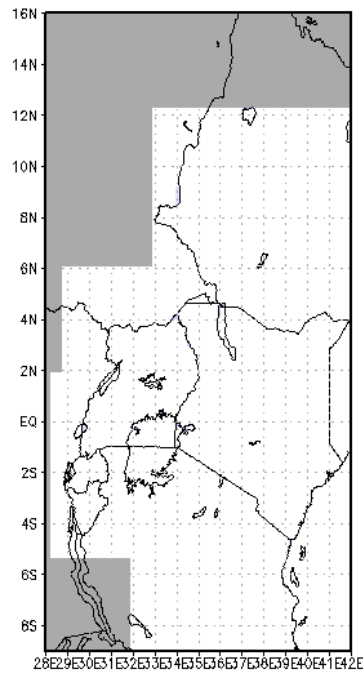


Figure 2.17 Mask of the East African domain for the interpolated station rainfall data. Masked out region is shaded in grey.

2.7 Mapping of datasets to research questions

As outlined in Sections 2.1 to 2.6, a number of reanalysis, satellite, statistical and station datasets are employed in this research to characterise the contemporary climate in East Africa, both mean and daily regimes. Additionally these datasets are used to evaluate the replication of contemporary conditions in a CMIP3 coupled climate model subset. Table 2.2 highlights the different datasets, in which chapters they are used and for which of the research questions their use relates.

To reiterate, the research questions for this thesis are as follows:

1. What is the contemporary structure of daily rainfall regimes over East Africa during the October to December Short Rains season?

- Is the self-organising map an appropriate tool for this investigation?
2. To what extent can coupled climate models replicate the contemporary structure of the East African daily rainfall regimes of the Short Rains season?
 - To what extent is the mean rainfall climatology replicated in AOGCMs?
 3. What mean changes to the East African Short Rains occur under the high emissions SRES A2 scenario at the end of the twenty-first century in coupled climate models?
 4. What patterns of daily rainfall are responsible for the change in future modelled rainfall?

Table 2.2 Summary of datasets employed for which parts of the research

Dataset	Dates	Variables	Chapter	Research Question
ERA-40	1971-2000, daily	Rainfall, specific humidity, zonal winds, meridional winds, geopotential height	4-8	All. ERA-40 is used to establish the contemporary conditions of the short rains and then the replication of these conditions in the AOGCMs. Model rainfall is constrained to the ERA-40 SOM under climate change conditions.
ERA-Interim	1989-2010, daily	Specific humidity, zonal winds, meridional winds, geopotential height	5	One. ERA-Interim is used to establish circulation patterns associated with TRMM rainfall events under contemporary conditions.

NCEP	1970-1999, monthly	Rainfall	3	Two, sub question. To examine the replication of the mean climatology in AOGCMs.
H08	1970-1999, monthly	Rainfall	3	Two, sub question. To examine the replication of the mean climatology in AOGCMs.
CRU	1970-1999, monthly	Temperature	3	Two, sub question. To examine the replication of the mean climatology in AOGCMs.
TRMM	1998-2010, daily	Rainfall	2, 5	One and two. TRMM is used to establish daily rainfall regimes under contemporary conditions.
Station	Variable, daily	Rainfall	2, 3	None. After evaluation the station data was not suitable to examine daily regimes due to poor spatial and temporal coverage.

For this work, ERA-40 was selected ahead of the alternative NCEP reanalysis data after a comparison of available literature for the region, as discussed in Sections 2.1 and 2.2.

ERA-Interim is a more recent reanalysis product and includes improvements over the earlier ERA-40 series but with the series only starting in 1989 this did not allow for a long

enough period to evaluate the CMIP3 20C3M contemporary short rains against, with the 20C3M data only covering the period up to the year 2000. This steered the choice of ERA-40 ahead of ERA-Interim. The TRMM satellite data was introduced as a second, alternatively sourced, rainfall product to use to examine daily rainfall regimes.

2.8 AOGCM data from the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4)

Climate model data is a key component in climatological research, being used to unravel cause and effect in the climate system under constantly changing conditions. For this thesis, a selection of models utilised in the Intergovernmental Panel on Climate Change Fourth Assessment Report (IPCC AR4, 2007) are used. These model runs originated as part of the Climate Model Intercomparison Project 3 (CMIP3). This project, following on from earlier CMIP phases, revolutionised climate model data through a coordinated effort between seventeen separate modelling groups to produce consistent model runs, which were made available for analysis through the Program for Climate Model Diagnosis and Intercomparison (PCMDI) (Meehl et al, 2007). Since this work began, CMIP5 model data from the next phase of the model intercomparison project has been made available with the intention that these new coordinated model runs will address questions that arose during the IPCC AR4 (<http://cmip-pcmdi.llnl.gov/cmip5/>). However, the IPCC AR5 found no conclusive evidence that the small increases in resolution and improvement in the simulation of land surface processes have led to marked increases in the reliability of simulations, or confidence in results, over Africa in CMIP5 models over CMIP3 models (Christensen and Kanikichana, eds, 2013).

The AOGCMs used in this work are selected based on access to all of the required daily data fields (rainfall, specific humidity, zonal and meridional winds and geopotential

height) which led to eight models being chosen. The characteristics of these models are summarised in Table 2.3 (Gleckler, et al, 2008 and Legutke and Voss, 1999).

A multimodel ensemble mean of the eight AOGCMs is created through calculating the mean field for each variable over the relevant time series and domain. A multimodel ensemble mean adds significant value to climate research engaging with AOGCM data for several reasons. First, the mean and median of a CMIP3 model ensemble has a lower error under climatology conditions than individual models for the majority of variables (Gleckler et al, 2008). Second, it is hoped that combining models with different parameterisations will lead to the suppression of projected changes with no physical basis whilst retaining robust climate change responses to greenhouse gas forcing (Giannini et al, 2008). Third, it is well understood that models are not equal in skill at simulating particular global regions and the use of an ensemble can be helpful in highlighting outliers amongst the individual models for further investigation.

The eight individual AOGCMs employed in this work come from eight different modelling groups in North America, Europe, Asia and Australia. Notable is the absence of a model within the CMIP3 grouping that was developed in Africa. Given that it is usual for AOGCMs to be tuned for the region in which they are developed it is useful to use multiple models to study the East African region as each will have differing dynamics and qualities for this domain. The performance of the individual models in reproducing the climatology conditions of the 1971-2000 control period is evaluated in Chapter 3.

Table 2.3 Key characteristics of AR4 AOGCMs used in this study

Modelling Group	Modelling Designation	AGCM horizontal/vertical resolution	OGCM horizontal/vertical resolution
Canadian Centre for Climate Modelling and Analysis	cccma_cgcm3_1 ‘CCCMA’	T63 L31	1.4 x 0.94 L29
Météo-France, Centre National de Recherches Météorologiques	cnrm_cm3 ‘CNRM’	T63 L45	182 x 152 L31
CSIRO Atmospheric Research (Australia)	csiro_mk3_0 ‘CSIRO’	T63 L18	1.875 x 0.84 L31
US Dept. of Commerce, NOAA Geophysical Fluid Dynamics Laboratory	gfdl_cm2_0 ‘GFDL’	T45 L24	1 x 0.33-1 L50
NASA/Goddard Institute for Space Studies (USA)	giss_model_e_r ‘GISS’	72 x 46 L17	4 x 5
Meteorological Institute, University of Bonn, Meteorological Research Institute of KMA, and Model and Data Group	miub_echo_g ‘MIUB’	T30 L19	T42 L20
Max Planck Institute for Meteorology (Germany)	mpi_echam5 ‘MPI’	T63 L32	1 x 1 L42
Meteorological Research Institute (Japan)	mri_cgcm2_3_2a ‘MRI’	T42 L30	2.0 x 0.5-2.0 L23

Two main different model runs are utilised; the 20C3M run for the control period, and the high emissions SRES A2 scenario for the 2090s period. The 20C3M run follows a level of emissions that evolve as recorded during the twentieth century and is used for all model runs covering the period 1971-2000. This run considers solar irradiance, sulphate aerosol

concentration, volcanic aerosol optical depth, greenhouse gas concentrations and ozone concentrations as ascertained through observational records (Gerdes and Koberle, 2007).

For the twenty-first century, the evolution of greenhouse gas emissions is uncertain; arising from a dynamic combination of factors including oscillating population demographics, socio-economical change and developing technology (Nakicenovic and Swart, 2000). This led the IPCC to develop a wide range of scenarios encompassing levels of greenhouse gas emissions dependent on hypothetical storylines of differently developing worlds. The IPCC does not assign a probability or likelihood to any of the individual storylines. This work primarily employs the high emissions A2 scenario; a heterogeneous world with maintenance of local identity and a high population growth with a global population of 15 billion by 2100. In this scenario, international cooperation does not develop greatly, leading to divergent economic development and the continued polarisation between the industrialised and less developed worlds (Nakicenovic and Swart, 2000). Table 2.4 shows the emissions for this scenario by 2100 (Nakicenovic and Swart, 2000).

Table 2.4 Emissions totals for the IPCC SRES A2 scenario by 2100

	A2 Scenario 2100
Cumulative total carbon dioxide GtC	1862.0
Sulphur Dioxide MtS/yr	60.0
Methane MtCH ₄ /yr	889.30
Nitrous Oxide MtN/yr	16.5
CFC/HFC/HCFC Mtc equiv/yr	753.0
PFC MtC equiv/yr	178.4
SF ₆ MtC equiv/yr	164.6
CO MtCO/yr	2326.0
NMVOC Mt/yr	342.0
NO _x MtN/yr	109.0

Chapter 3, the initial evaluation of the model climatologies over East Africa, additionally employs the SRES A1B and B1 scenarios (mid-range and low-range emissions respectively) in order to contextualise a range of potential futures for East Africa under increasing emissions. The A1B scenario is for a future of rapid economic growth, with population growth peaking mid twenty-first century followed by decline. This scenario envisages a balance of energy usage across sources. The B1 scenario follows the same population projections as the A1B storyline, but with convergence towards a service and information economy with the proliferation of clean, resource-efficient technologies (Nakicenovic and Swart, 2000). In comparison to the cumulative carbon dioxide total by 2100 of 1862GtC in the A2 scenario, the A1B and B1 scenarios have cumulative carbon dioxide totals by 2100 of 1499GtC and 983GtC respectively (Nakicenovic and Swart, 2000).

The SRES A2 scenario has been selected to be used in this work in Chapters 7 and 8, focusing on daily projections over East Africa, due to the expectation that projected changes for the twenty-first century will be of increasing magnitude through time and under higher greenhouse gas emissions. It has been suggested that since the year 2000 carbon dioxide emissions have been increasing at a rate above that of the highest SRES emissions scenario (Raupach et al, 2007). Whilst this trend will not necessarily continue in future, this provides justification for focusing on potential changes under a high emissions scenario.

In terms of daily data, key to this study, there are three ten year ‘timeslices’ available for analysis. These are the 2030s (2030-2039 inclusive), 2050s (2050-2059) and 2090s (2090-2099). Therefore the combination of a high emissions scenario and the final timeslice of the twenty-first century (2090-2099) provides the most theoretically sound combination for the most extreme rainfall projections for East Africa to be evaluated.

2.9 Clustering techniques

This thesis draws on clustering of daily rainfall events over East Africa to facilitate analysis of the characteristics of the daily scale within the short rains season. Much research on East African climate variability in the literature has drawn on the EOF (Empirical Orthogonal Function) technique, otherwise known as PCA (Principal Component Analysis) to extract dominant modes of variability. Examples include Basalirwa et al (1999), Schreck III and Semazzi (2004), Anyah and Semazzi (2006), Bowden and Semazzi (2007), Manatsa et al (2012) and Omondi et al (2012). EOFs are one of the most extensively utilised methods in atmospheric science which can be used to reduce dimensionality and extract patterns (Hannachi et al (2007).

The EOF methodology is applied to a single field, such as SST, finding spatial patterns of variability, their time variation and a measure of importance in relation to the data as a whole (Björnsson and Venegas, 1997). EOFs reduce the dimensionality of the data, constraining it to a small number of dominant nodes (Reljin et al, 2002). In meteorological applications, remaining data outside of the most dominant modes is most frequently judged to be noise (e.g. Schreck III and Semazzi, 2004) and so this methodology is particularly suitable for examining modes of interannual variability where the expectation is that a limited number of modes will be the dominant controls.

For this work, focusing on characterising daily rainfall regimes during the short rains of East Africa and how reliable future climate projections are, EOFs are not the most suitable methodology, given that focusing on a limited number of dominant nodes of variability could lead to the neglect of important, but less frequently occurring spatial rainfall patterns. It has been decided to utilise the self-organising map as an alternative clustering method to the EOF. Benefits of this method over the EOF are outlined in Section 2.10.1.

2.10 The Self-organising Map (SOM)

Self-organising maps have been proposed as an alternative method of classification in synoptic climatology given that they facilitate visualisation of the distribution of the complexity of synoptic states within a dataset, whilst retaining the continuum of the data (Hewitson and Crane, 2002). The methodology was proposed by Kohonen (1990) and has been widely applied across many research sectors; in finance (Xiao and Lei, 2008, Hsu et al, 2010, Severin, 2010), medical research (Liszka-Hackzell and Martin, 2005, Lessman et al, 2007, Faisal et al, 2010) and biology (Wang et al, 2002, Griebeler and Seitz, 2006, Abe et al, 2006).

2.10.1 Benefits of the SOM over alternative clustering methods

Self-organising maps can be viewed as an alternative way of examining synoptic climatology (Hewitson and Crane, 2002). A key feature of the SOM is that it allows visualisation of the complexity of atmospheric states whilst retaining the continuum of the data. The SOM is somewhat similar to other clustering techniques with the SOM placing the selected number of nodes within the data space in such a way that they represent the multi-dimensional input data. Nodes will be placed further apart in regions of sparse data and closer together where data density is high, instead of being evenly distributed (Hewitson and Crane, 2002). Hewitson and Crane (2002) identify two ways in which the SOM is different to other clustering techniques. First, the fact that the SOM is not explicitly trying to ‘cluster’ but instead to find a state representative of nearby observations, and second, that the results can be visualised with ease. An advantage over other clustering methods is that the SOM does not make assumptions (and so is not dependent) on the data belonging to a particular underlying distribution (Hewitson and Crane, 2002). Reljin et al (2002) suggest that this marks a key advantage of the SOM over

EOFs. EOFs are a linear method and whilst it is possible to address some nonlinearities in data using it, this may not be sufficient for the climate system (Reljin et al, 2012).

Moreover, the SOM will converge on the same results regardless of the SOM dimensions employed, revealing the same patterns but with differing levels of detail depending on the size of the SOM array. There is no issue of overtraining with a SOM because a function is not being fitted and the SOM can also deal with missing data without issues (Hewitson and Crane, 2006). Using EOFs introduces the risk of analysing an artificial node, statistically induced. This possibility is removed when using the SOM; no dates would be assigned to an artificial node (Tozuka et al, 2008).

2.10.2 Previous applications of the SOM in climatology

SOMs were first presented as part of discussions on neural networks to the broader physical geography circle in 1994 (Hewitson and Crane, 1994, eds), but they have not been widely utilised. Some research that does use this technique includes studies of South Africa by Hewitson and Crane (2006), research on Puerto Rico climate from 1960-1990 (Malmgren, and Winter, 1999), studies by Cavazos (1999 and 2000) on rainfall events in Mexico/Texas and the Balkans respectively and also work on cloud classification using the self-organising map technique (Ambroise et al, 2000) (Hewitson and Crane, 2002). Liu et al (2006) compared the performance of EOFs and SOMs in a climatic setting, concluding that the SOM outperformed the EOF for some measures and had the potential for further geophysical applications including for climate change scenarios and on numerical model data (Liu, 2006).

2.10.3 Self-organising map methodology

Self-organising maps are created through a series of stages visualised in Figure 2.18 (Crane and Hewitson, 2003). The SOM is created through an algorithm that maps highly

complex, multi-dimensional datasets onto a regular array of units, making it a useful method to visualise the complexities of climatological data and classify different synoptic types within the dataset (Nishiyama et al, 2007). Unlike other clustering methods, the SOM methodology is not susceptible to overtraining, and so the stages can be repeated through multiple iterations to obtain the ‘best’ visualisation in the array (Hewitson and Crane, 2006).

For the majority of SOMs created, the input data is first normalised to the thirty year climatology to have a mean of 0 and standard deviation of 1. This normalisation has a dual purpose. First, by standardising the different units of the various input variables it will allow the variables to be combined to create a single SOM that is representative of a state inclusive of key climate variables of the region. Second, this normalisation allows examination of interannual variability and changes to the occurrence of patterns emerging in different variables during the 1971-2000 period relative to the mean over the whole time period during each of the months October, November and December.

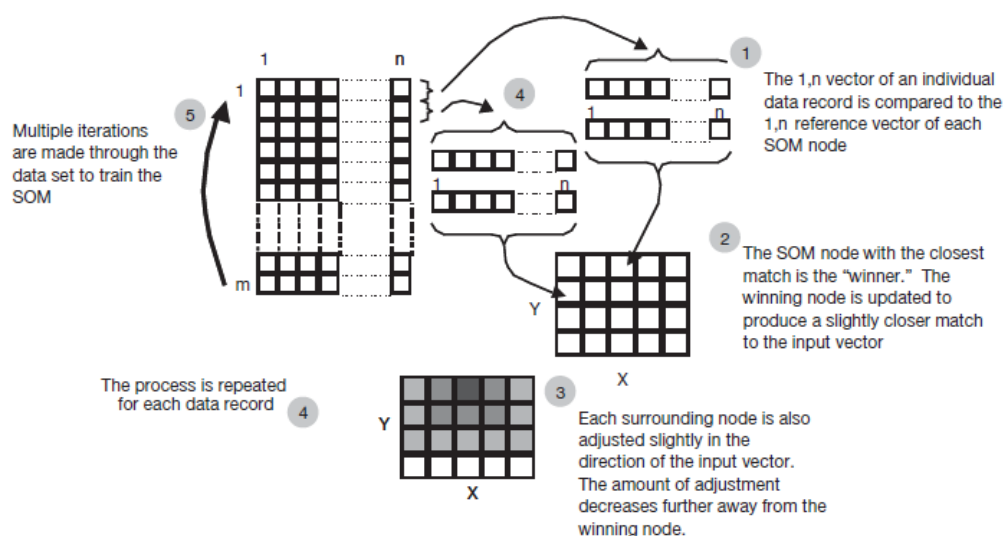


Figure 2.18 Visualisation of SOM methodology, from Crane and Hewitson (2003)

The stages of the training and visualising a self-organising map can be summarised as follows (drawn from Hewitson and Crane (2006) and Nishiyama et al (2007)). All training and visualisation of SOMs in this thesis is undertaken using the SOM_PAK self-organising map package programme (SOM Programming Team of the Helsinki University of Technology, 1995) and NCL:

1. Input data for the self-organising map takes the form of a ASCII file with one line for each observation, in this case daily data for the months of October, November and December 1971-2000; a total of 2760 daily inputs. Each line has a value for each grid point, for a grid sized 20°E to 60°E, 15°S to 20°N, resulting in a total of 1435 observations per line. The array of nodes for the outputted SOM is user specified and a balance has to be struck between visualising a realistic number of states to facilitate full analysis whilst retaining the ability of the SOM to capture the full continuum of states of the input variable without excessive generalisation. After experimenting with a number of different arrays it is concluded that the best balance between detail and analysis is struck with a 7 by 5 SOM, containing 35 nodes in a hexagonal arrangement.
2. The next stage is to initialise the 35 nodes, which is done on a random basis, after specifying the map dimensions (7 x 5), topology (hexagonal) and neighbourhood function type (bubble).
3. Taking each input data line in turn, the Euclidean distance between the input vector and the vector associated with each individual reference node is calculated until the node that is closest to the observational input data is declared the 'winning' node.
4. The reference vector of the winning node is then adjusted in the direction of the input vector by a user determined learning rate. This learning rate decreases through the iterations as the SOM converges on a solution. Surrounding nodes are, to a lesser extent

with distance from the winning node, also adjusted towards the input vector. The radius of nodes adjusted with each winning node is determined by an update kernel that decreases through the iterations.

5. Stages 3 and 4 are repeated through all days of the input data and through multiple iterations. The SOM_PAK programme undertakes the training specified in stages 3 and 4 in two parts. During the first phase, the neighbourhood radius is equal to the y dimension of the array (5) and decreases to one during training with the learning rate (initially 0.05) decreasing to 0. This stage orders the reference vectors within in the array. The second phase of training is used to finalise the values of each reference vector and repeats the same stages but with a smaller radius (3) and learning rate (0.02).

6. Once the SOM has been finalised through the two stages of training it is necessary to evaluate the quantisation error of the input data to the finished map. This is a measure of how closely the input data fits to the states depicted in the chosen number of nodes in the SOM for each, with the lowest mean quantisation error (QE) indicating a SOM that is best fitted to the dataset.

7. The random initialisation of the map nodes at the start of the process results in a slightly different end SOM each time training was completed through the stages. Therefore, it was decided to repeat the initialisation and training stages multiple times before selecting the self-organising map with the lowest quantisation error to continue to the visualisation and date matching stages. The selected SOM is referred to by the number of the initialisation of the reference vector, for example if it was the second SOM to be initialised that has the lowest, the resulting SOM is referred to as being associated with reference vector 2. The differences between the individual reference vectors are quantitatively very small and are not visible to the naked eye. The numbering system for the reference vectors was used to

select the SOM with the smallest quantisation error, but this selection does not impact on the results.

8. The trained SOM can now be used to analyse the data. Each individual input day of the dataset is matched to a particular SOM node and also the QE associated with this matching is calculated. The finished SOM is then visualised using NCL. This allows examination of which states or types in any particular variable have been extracted from an input dataset over the October- December months, and how the states could change between months or years, or were linked to other variables.

Figure 2.19 shows the numbering of a 7 by 5 self-organising map, as typically utilised in this work. In a SOM, similar states in the data space are placed close together within the organisation of SOM nodes. For example, the four corners of the SOM, 0 0, 6 0, 0 4 and 6 4, represent the states within the data which are the most extreme, and nodes 0 0, 1 0, 0 1 and 1 1 states close together in data space. However, within the continuum of data available for input the nodes are not evenly spaced, with more nodes being placed where density of data is higher (Hewitson and Crane, 2002).

0 0	1 0	2 0	3 0	4 0	5 0	6 0
0 1	1 1	2 1	3 1	4 1	5 1	6 1
0 2	1 2	2 2	3 2	4 2	5 2	6 2
0 3	1 3	2 3	3 3	4 3	5 3	6 3
0 4	1 4	2 4	3 4	4 4	5 4	6 4

Figure 2.19 Depiction of the numbering of a self-organising map (SOM)

2.11 Normalisation

There are different ways of normalisation which achieve the result of removing units from the dataset and resulting in the dataset statistics of mean, 0, and variance, 1, the two prerequisites for a successfully normalised data series. The calculation for normalisation is as follows:

$$\text{normalised value} = \frac{\text{anomaly}}{\text{standard deviation}}$$

The anomaly is defined as:

$$x - \text{mean}$$

Therefore, the type of normalisation output depends on how the mean is calculated, for which there is more than one option in the context of this study. This study employs daily data, months October, November and December, for the period 1971-2000. Each month should be treated independently from the other two which leaves two options for the calculation of the mean. First, the mean of each individual month over the thirty-year period could be calculated and used to establish the anomaly. Second, the climatology value for October, November and December over the thirty year period could be obtained and used in anomaly calculations.

Both methods are evaluated, but in order to retain interannual variability and not flatten the emerging patterns, the second option, calculating the anomaly based on the climatology value, is used and from here on in is referred to as ‘normalisation to climatology’. The first method, by using each individual month of the whole series as the basis of the anomaly calculations removes the interconnections between the climate as a whole over the control period and so does not contain any of the interannual variability that is a key component of

the East African climate. Therefore, all data series are normalised to the climatology value of the months of October, November and December and checked to ensure that they have the statistics of mean, 0, and variance, 1. This allows greater scope for analysis, retaining interannual variability over East Africa.

2.12 Statistical Testing

Statistical tests are used during this work in order to quantify changes to rainfall characteristics over East Africa. Examples include testing the statistical significance of trends over East Africa during the twentieth century and differences in the distribution of daily rainfall types in the different datasets and between time periods. Some of the statistical tests are carried out in Excel, for others the test statistic is calculated manually and compared to the relevant statistical table.

2.12.1 Chi-squared goodness of fit

The chi-squared goodness of fit test is used to calculate if there are any significant differences in the distribution of different types of daily rainfall over the domain between the observed and model datasets and additionally between the control and 2090s time periods. The chi-squared test compares ‘observed’ and ‘expected’ counts for different categories, with the null hypothesis being that the observed and expected values belong to the same distribution. Degrees of freedom are equal to the number of categories minus one (Dytham, 2003). This test is carried out in Excel, with the output statistic being the probability that the two datasets belong to the same distribution. If the probability is less than 0.05 the null hypothesis is rejected and it is concluded that the two distributions of node occurrence are different.

2.12.2 Pearson's product moment correlation coefficient

Pearson's product moment correlation coefficient is employed to examine the relationship between the daily rainfall states within the equivalent SOM nodes of two arrays calculated from two datasets; ERA-40 and TRMM. Correlation is used as a measure of the association (although not one that implies a causal relationship) between two datasets with the strength of the correlation represented through an r value which ranges from -1 (perfect negative correlation) through 0 (no correlation) to +1 (perfect positive correlation) (Dytham, 2003). The statistical significance of the correlation can be established through a t -test, using the following formula, where r is equal to the correlation coefficient and n is equal to the number of observations.

$$t = r \sqrt{\frac{n - 2}{1 - r^2}}$$

The value for t can then be compared against t_{CRIT} , found through a statistical table containing the critical values for testing the null hypothesis that the correlation coefficient is not significant. The degrees of freedom are equal to n minus 2. If the calculated t value is greater than the t_{CRIT} value then the null hypothesis can be rejected and it can be concluded that the correlation is statistically significant. This test is calculated in Excel.

2.12.3 Significance testing of projected rainfall changes – Z-Test

The Z-Test is utilised to see if there is statistical significance to the projected rainfall changes to East Africa under a high emissions scenario between the climatology period and the 2090s. The null hypothesis for this test is that the two means are equal. If the Z statistic is greater than 1.96, the difference is statistically significant at 95%, with 2.326 the critical value for significance at the 99% level. The Z statistic is calculated using the following formula (Chu et al, 2010).

$$z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Where $\bar{x}$ and $\bar{y}$ are the two daily rainfall means, s_1^2 and s_2^2 are the variances of the two series and n_1 and n_2 are the lengths of the two series. This test is calculated manually.

2.13 Moisture flux divergence

Moisture flux divergence is used in this thesis to analyse the presence of conditions leading to convective rainfall over East Africa. Moisture flux divergence is a combination of meridional winds, zonal winds and specific humidity that is known to be a good predictor of convergence in the tropics. Specifically the measure represents the mass convergence of moisture (and conversely, divergence of moisture) (Waldstreicher, 1989).

The equation for moisture flux convergence can be defined as follows:

$$q\nabla \cdot V + V \cdot \nabla q$$

where $q\nabla \cdot V$ represents the mass convergence, the wind convergence multiplied by the mixing ration, and $V \cdot \nabla q$ represents the moisture advection as the moisture field changes with time (Waldstreicher, 1989). In this thesis, moisture flux (vector plots, in g kg⁻¹ m s⁻¹) and moisture flux divergence (shaded plots, in g kg⁻¹ s⁻¹) plots are calculated using GrADS functions. These work from a file containing two variables, qu (specific humidity multiplied by zonal winds) and qv (specific humidity multiplied by meridional winds), for each time step. The GrADS functions allow this data to be displayed as a ratio for moisture flux, through a vector display. For moisture flux divergence a finite differencing function

is used to calculate the horizontal divergence at one atmospheric level; 850hPa in this case.

2.14 Summary of chapter

Chapter 2 introduces the datasets, originating from observational, reanalysis and model sources, which are utilised in this work. The origins, key information and caveats for each dataset are stated. With regards to methods, the premise of utilising the self-organising map to facilitate investigation into contemporary and future structures of the daily rainfall regimes over East Africa over other clustering methods employed in climate studies is examined. An overview of statistical tests used is provided. This chapter provides information on the grounding of data and methods which facilitate research towards achieving the overall aim of this thesis.

Chapter 3: Twentieth century climatology of East Africa and potential changes to the twenty-first century climate

This chapter examines the background climate of the East African domain including the climatology, 95th percentile extremes, interannual variability and trends over the second half of the twentieth century in reanalysis, statistical and observed datasets. The ability of AOGCMs to replicate this twentieth century East African climate is also evaluated in a subset of eight of the IPCC AR4 global coupled atmosphere-ocean models, addressing the sub question of the second research question. Finally, climate projections of these models over East Africa during the twenty-first century in an increased greenhouse gas atmosphere are examined through three time slices, the 2030s (2030-2039), 2050s (2050-2059) and 2090s (2090-2099).

The analysis in this chapter is undertaken on monthly totals rather than the daily timescale which is the focus of subsequent chapters. The intention is to provide the background state, on which the subsequent analysis of higher temporal resolution in Chapters 4 to 8 inclusive, can be placed. This contextualisation research therefore provides the frame of reference to achieving the thesis aims in later chapters.

The sub question of the second research question, to establish the extent to which mean contemporary rainfall climate is replicated in AOGCMs is addressed in Section 3.4. The East African climate is complex and this complexity is represented to differing degrees in the CMIP3 AOCGMs. Anyah and Qiu (2012) state the necessity to undertake a comparative analysis of global climate models in order to understand uncertainties between model and observed datasets prior to considering downscaling of this data. This underlines the importance of the premise of this study to assess the fidelity of global climate model rainfall projections. The sub research question will identify positive

elements of simulation and model biases in each model from the CMIP3 subset, information that could help to contextualise evaluation of daily rainfall regimes in the climate model and associated circulation patterns in each model in Chapter 6.

This work is taken from a project undertaken for the CCAFS report ‘Climate Change in East African Agriculture: Trends, current projections, crop-climate suitability, and prospects for improved climate model information’ (Washington and Pearce, 2012). All of the data manipulation, plotting and analysis presented within this chapter, that was also a part of the CCAFS report, was undertaken by the author.

3.1 Background climate of East Africa

The climate of East Africa is dominated by convective rainfall events linked to the migration of the inter-tropical convergence zone (ITCZ) in conjunction with the region of greatest solar insolation. The annual rainfall cycle lags the insolation cycle by approximately one month (Black et al, 2003). Over the majority of this study domain, the annual cycle is dominated by two rainy seasons; the long rains of the boreal spring (March, April and May; MAM) and the short rains of the boreal autumn (October, November and December; OND) with these convective events being co-located with the movement of the ITCZ northwards and southwards during the year (Mutai and Ward, 2000). In the northernmost parts of East Africa the rainfall cycle is unimodal, with the main rains occurring from approximately June to August (JJA). Much of northern Ethiopia also sees preliminary rainfall each year from March, followed by a dry pause before the JJA rains begin (Verdin et al, 2005).

For this study, the domain ‘East Africa’ relates to the region covering 20°E to 60°E, 15°S to 20°N. The key countries are those for which daily rainfall station series are available; Kenya, Ethiopia, Tanzania and Uganda. This region can be seen in Figure 3.1.

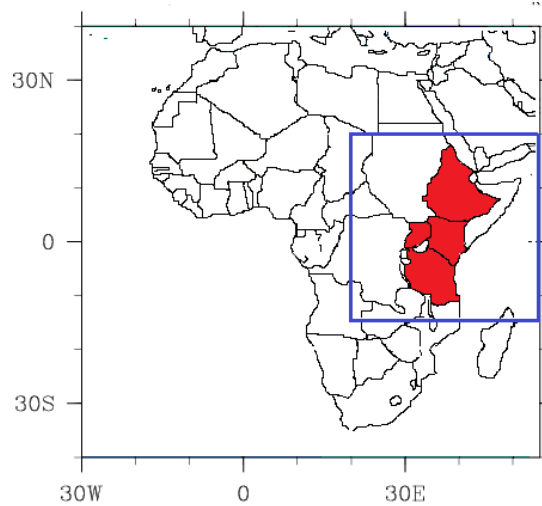


Figure 3.1 Study domain (20°E to 60°E, 15°S to 20 °N) with focus countries highlighted

3.1.1 East African climate in the research literature

The East African climate regime is dominated by convection, with a series of complex interactions governing interannual rainfall variability in the region. These include the interconnections between sea surface temperature anomalies (SSTAs), large scale atmospheric patterns, synoptic scale weather disturbances, tropical cyclones, subtropical anticyclones, extratropical weather systems, wave perturbations and free atmosphere variations (Mutai and Ward, 2000). Relative to the long rains of March to May, the October to December short rains have a greater interannual variability but are more spatially coherent across East Africa. There are established connections between East African rainfall and the El Nino Southern Oscillation (ENSO); increased rainfall during the OND season is associated with warm El Nino events, with cool La Nina episodes linked to decreased rainfall over East Africa (Mutai and Ward, 2000). Occurrences of moist South Atlantic westerly flows are linked to anomalous East African rainfall and tend to be strengthened during El Nino events and weakened during La Nina events (McHugh, 2006). This link is evident over all of the four focus countries.

The Indian Ocean dipole sea surface temperature (SST) variability mode also exerts control on East Africa. Increased East African coastal rainfall is linked to warm SSTAs in the western equatorial Indian Ocean and cool SSTAs around Indonesia (McHugh, 2006). Within the East African domain a significant control on the distribution of rainfall is topography (Oettli and Camberlin, 2005). Literature written on the climate of East Africa has a heavy focus on interannual drivers of climate rather than controls on daily states or intraseasonal timescales. This focus highlights the novel nature of the approach undertaken in this thesis through the examination of daily regimes within the short rains, from Chapter 4 onwards.

3.1.2 Basic climatology of East Africa

The basic climatology of East Africa is evaluated using the UEA CRU gridded dataset for temperature and the NCEP and Japanese H08 datasets for rainfall.

3.1.2.1 Temperature

Figure 3.2 shows the CRU annual mean temperature climatology and seasonal cycle over East Africa for the period 1970-1999.

Annually, the highest mean values are found to the north of the equator. Here, climatological mean temperatures measure between 24°C and 30°C over much of the region north of the equator, with the exception of Ethiopia where the mountainous topography mutes the mean temperatures to between 18°C and 24°C. South of the equator, East Africa has mean temperatures of between 20°C and 26°C during the year. Of the focus countries of this thesis, the warmest mean temperatures are found in Kenya, remaining largely stable throughout the yearly cycle at approximately 24° to 26°C inland and up to 4°C warmer on the Indian Ocean coastline.

The annual cycle has greatest amplitude over Sudan. Mean temperatures exceed 32°C during the boreal summer from May to September but reach just 14°C during the winter months. In contrast, mean temperatures remain in the range of 22°C to 26°C over much of the East African coastline, including Kenya and Tanzania, indicating a less defined annual cycle. The high topography in Ethiopia means temperatures remain relatively cool year round in comparison to other countries at similar latitudes. Consistent and moderate mean temperatures of between 20°C and 26°C are found in the landlocked equatorial countries of Uganda, Rwanda and Burundi.

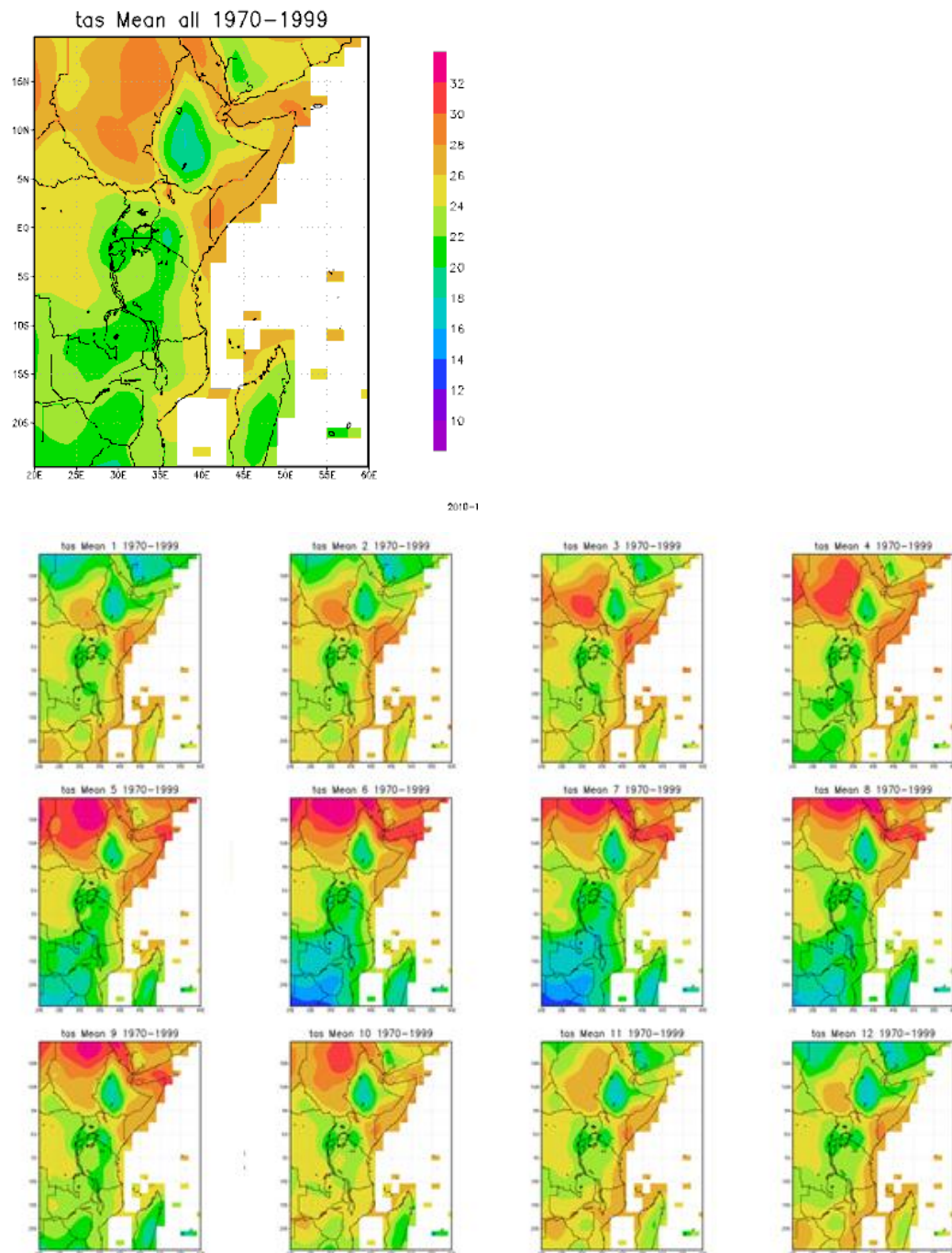


Figure 3.2 CRU annual mean temperature and seasonal cycle over East Africa, temperatures shown in °C, with all months following the same scale, as shown next to the annual plot

3.1.2.2 Rainfall

Figures 3.3 (a) and (b) show the annual climatology and monthly climatology over East Africa for the period 1970 to 1999 in the NCEP reanalysis and H08 rainfall datasets.

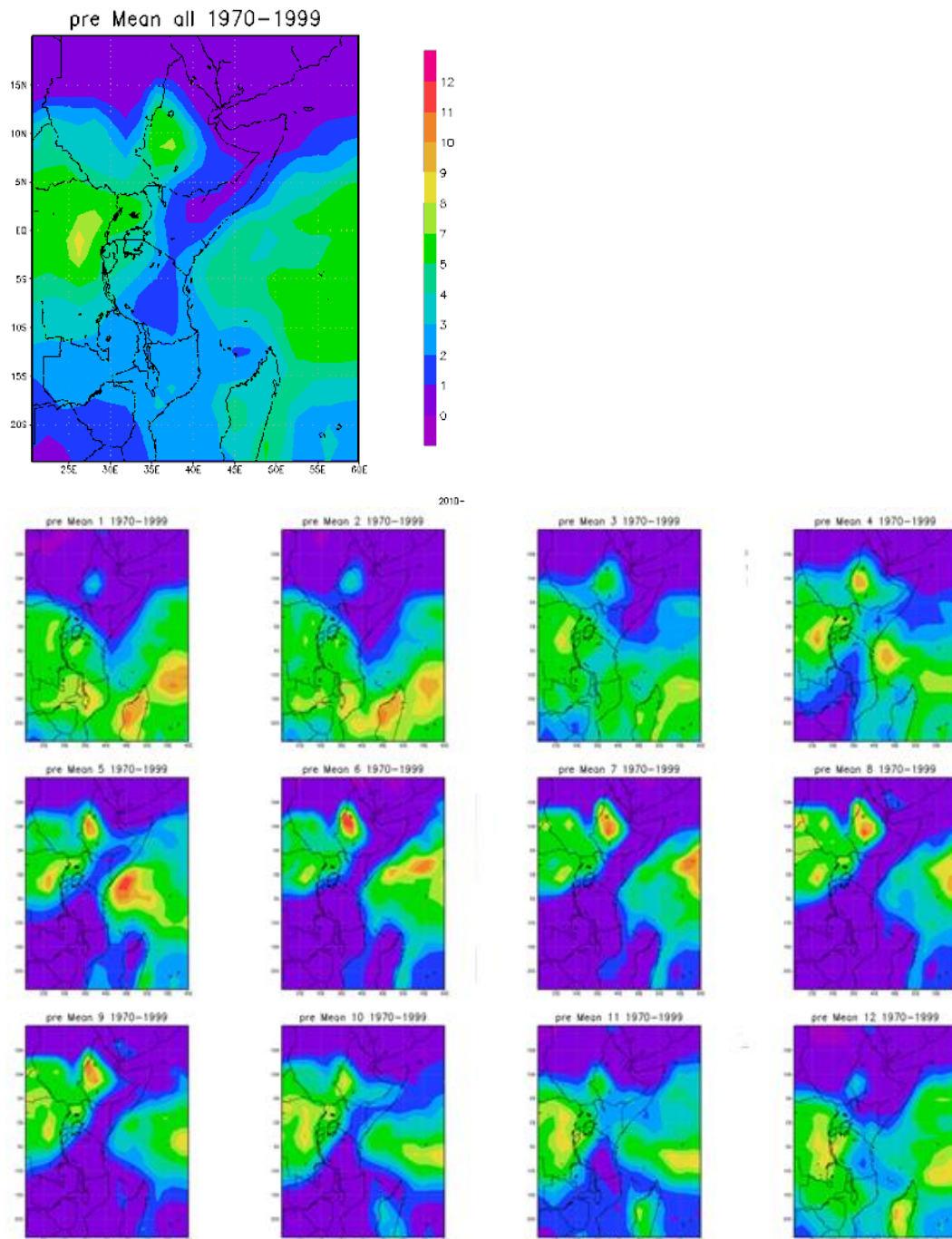


Figure 3.3 (a) NCEP annual and monthly rainfall climatology for East Africa in mm/day, all months following the same scale as shown next to the annual plot

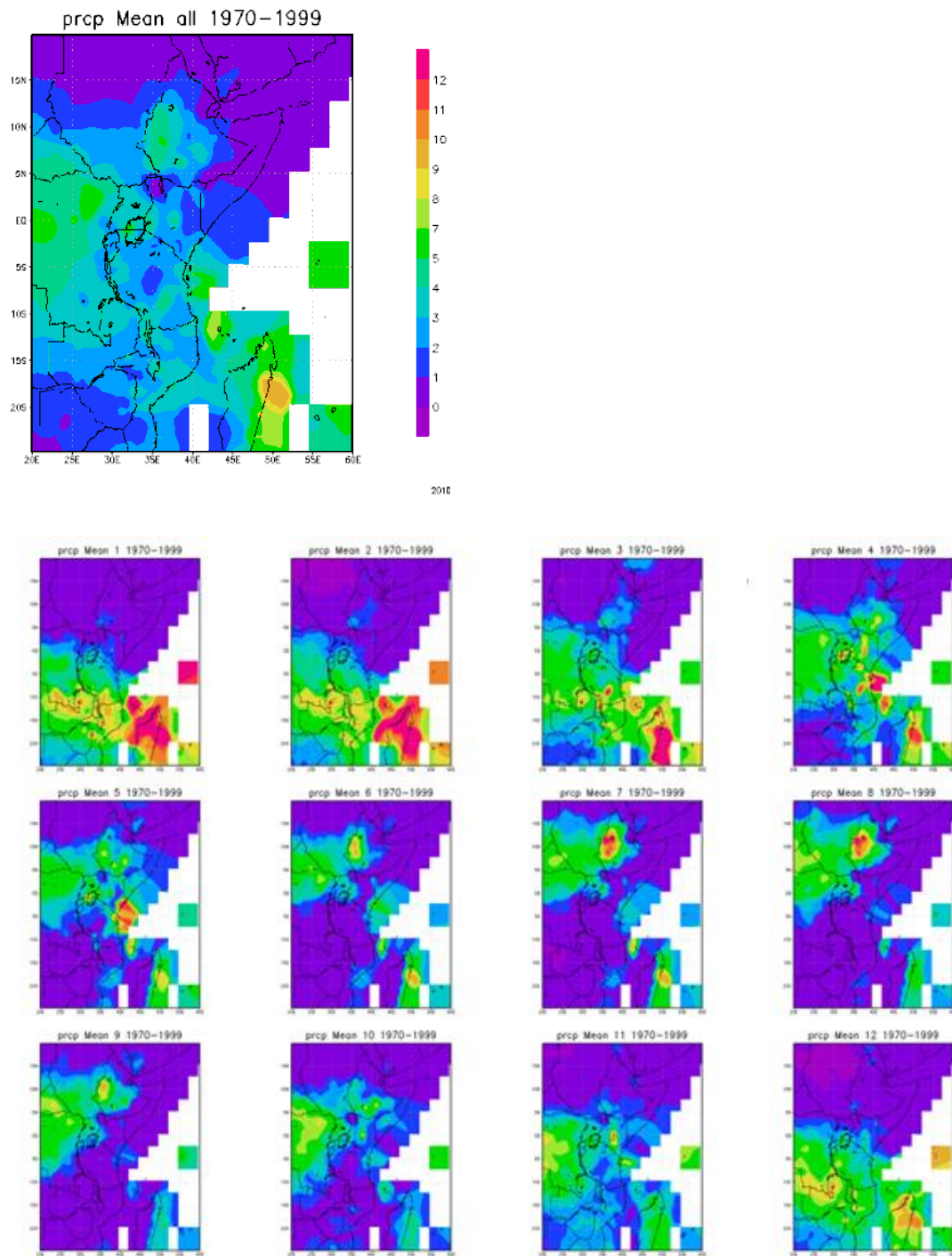


Figure 3.3 (b) H08 annual and monthly rainfall climatology for East Africa in mm/day, all months following the same scale as shown next to the annual plot

The seasonal cycle of rainfall, controlled by the migration of the ITCZ and accompanying convective rainfall as it follows the co-location of the most intense solar heating, lagging behind by approximately one month, is clear to see in Figures 3.3 (a) and (b). The ITCZ

moves from a position of 15°S in January to a northerly maximum in East Africa of 10-15°N from June to August and back again from September to December.

Within East Africa, different rainfall regimes are identifiable. Northern Tanzania, Uganda, Kenya, Burundi, Rwanda and southern Ethiopia have a bimodal rainfall regime, with two distinct rainfall seasons, March to May (the long rains) and October to December (the short rains) associated with the movement of the ITCZ. The NCEP reanalysis rainfall dataset has lower daily rainfall maximums than the H08 dataset, peaking at between 5 and 9 mm/day for NCEP in comparison with more intense rainfall of between 11 and 12 mm/day in the H08. These differences between the datasets are consistent throughout the domain and annual cycle, with the NCEP dataset showing lower daily mean rainfall than H08. The H08 dataset has a more northerly peak of rainfall in association with the movement of the ITCZ.

Northern East Africa, namely Sudan and South Sudan, northern Ethiopia and Eritrea, has just one main rainfall season that occurs during boreal summer from June to September with peak intensity found from July to August in conjunction with the maximum northerly location of the ITCZ. This region also sees a large rainfall gradient; Sudan is comprised of desert zones in the north with an annual rainfall of less than 75mm whereas South Sudan has a much higher annual rainfall total of between 1200 and 1800 mm (Osman and Shamseldin, 2002).

The cycle in the four focus countries can be described as follows:

Ethiopia- Much of northern Ethiopia has a unimodal rainfall regime with ITCZ associated rainfall falling between May and September. This pattern is apparent in both the NCEP and H08 datasets. Rainfall peaks during July and August in both datasets, at a rate of up to 10mm/day in NCEP series and up to 12mm/day in the H08 series. Southern Ethiopia has

the bimodal rainfall regime characteristic of much of East Africa. During the short rains southern Ethiopia is on the fringes of the rainfall region, averaging from 1-5mm/day in the H08 dataset and up to 8mm/day in the NCEP series where the region of rainfall is, on the whole, more widespread than in the H08 dataset.

Kenya- Kenya has two distinct rainy periods annually, which can be seen in both rainfall datasets. These are the short rains (October to December) and the long rains (March to May). Rainfall amounts in Kenya are relatively low in comparison to those expected at this latitude. In the NCEP reanalysis dataset daily rainfall rates over Kenya during the short rains measure up to 7mm/day, and during the long rains the rate is approximately the same. In the H08 dataset, the equivalent rates are up to 9mm/day during both the long and short rains, continuing the pattern of the H08 dataset having slightly higher rainfall peaks than the reanalysis NCEP database.

Tanzania- Tanzania has two rainy seasons each year, from March to May and October to December. In the two datasets examined here, H08 is more successful in capturing this regime. In the NCEP reanalysis dataset the representation of the movement of the ITCZ means there is no rainfall over Tanzania in both May and October/November. For the short rains, the two datasets have mean rainfall rates of 0-9mm/day and 0-7mm/day for the NCEP and H08 series respectively. This is in contrast to the annual pattern where the H08 dataset has higher rainfall values.

Uganda- Uganda has a bimodal rainfall regime in both datasets. Due to weaknesses in the representation of rainfall in both the NCEP and H08 datasets, this pattern is less clear over Uganda than Kenya. Uganda in these datasets sees rainfall to a lesser extent every single month. The NCEP reanalysis dataset has rainfall rates of approximately 5-7mm/day during

the short rains and 5-9mm/day during the long rains. The H08 has equivalent values of 5-8mm/day and 4-7mm/day in the long and short rains respectively.

3.2 East African rainfall climatology - station data series

More than 350 station gauge daily rainfall series were obtained for this work (McSweeney, 2009, *personal correspondence*). These stations are located over Tanzania, Kenya, Uganda and Ethiopia, with the main common characteristic between them the disparate nature of the series in terms of length and missing data. As shown in Chapter 2, the distribution of the series over the four focus countries is unequal and disproportionately focused over southern Kenya, in particular around Nairobi.

It was intended that the station rainfall series would be utilised in the interpolated format for examination of daily rainfall regimes over East Africa. However, as seen in Figure 3.4, the number of stations with data available (excluding missing periods), decreases drastically in the latter part of the twentieth century. This decrease in data availability happens in two stages. First, it occurs in the mid 1970s and then again after 1985, with a continual decrease in the number of recording stations so that by the last few years of the twentieth century there are only 50 stations recording data over this geographically diverse region. Such a pattern is evident in all of the countries for which data was available except Tanzania, for which there are only twenty available stations to begin with. The decision to use 1971-2000 as the control period, based on the prevalence of this period in the climate literature and the increasing quality of the ERA-40 reanalysis series towards the end of the twentieth century, exacerbates the implications of the decrease in station data available at the end of the twentieth century. Moreover, the amount of missing data and scarcity of coverage in much of East Africa means it is a significant challenge to adequately interpolate the series with coverage to a gridded daily rainfall series for the control period.

These issues reinforce the assertions of Sylla et al (2013), that gauge series in Africa are not adequate for the description of rainfall characteristics, in particular at the local level, but also to an extent at the regional and large scales. Considering this assertion in conjunction with evidence amassed as to the difficulty of utilising these daily series to examine daily rainfall events over East Africa, it has been decided to omit the daily station gauge series from further analysis in this thesis and to instead focus on gridded reanalysis and satellite rainfall datasets.

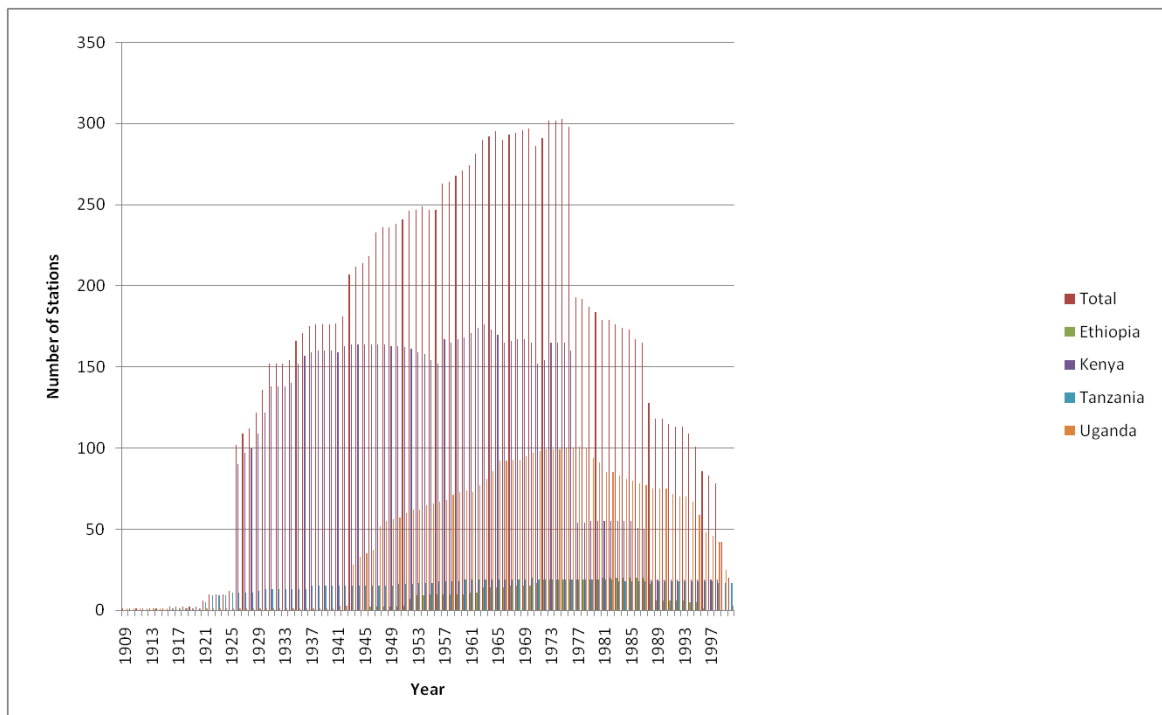


Figure 3.4 Number of stations recording rainfall data during the twentieth century over East Africa

3.3 Trends in East African climate from 1950 to the present day

3.3.1 East African climate trends in the literature

There has been some work undertaken to evaluate trends in East African climate in the literature. Analysis of twentieth century meteorological station temperature data from

Tanzania and Kenya conclude that between 1946 and 2004 the trend in maximum temperature in these countries is near zero. Minimum temperature is seen to increase significantly over the same time period, decreasing the diurnal temperature range (Christy et al, 2009). Over Sudan and Ethiopia, it has been noted that the diurnal temperature range has decreased by between 0.5°C and 1.0°C since the 1950s (Caminade and Terray, 2006). Studies looking at potential rainfall trends during the twentieth century have also been undertaken. An emerging trend during the OND short rains is identified by Schreck and Semazzi (2004). This comprises a northwest to southeast orientation, from the border of Uganda and the Democratic Republic of the Congo to the Indian Ocean coast of Zanzibar, of two distinct centres of rainfall anomalies with increasing rainfall seen to the north of eastern Africa and drying to the south. Although this trend is not mechanistically attributed to increasing anthropogenic greenhouse gases a possible link is inferred (Schreck and Semazzi, 2004). Moreover, a large positive increasing trend in twentieth century Lake Victoria rainfall is identified. It is noted that the short rains have a significant influence on annual rainfall totals, with most stations having a significant annual rainfall trend also having a significant coincident trend during the October to December period (Kizza et al, 2009).

Not all observational studies over the domain arrive at the same conclusions for twentieth century rainfall trends, highlighting the complexity of climate in the region and the sensitivity of results to datasets and methods. There have been studies that have identified negative rainfall trends during the twentieth century. Funk et al (2008) attribute a rainfall decline over Kenya, Burundi and Tanzania during the 1979-2005 period to Walker Circulation anomalies arising due to increased tropical Indian Ocean rainfall. This increased Indian Ocean rainfall, which occurs due to enhanced warming of SSTs, is thought to be linked to atmospheric ridging and anticyclonic moisture circulation over East

Africa which disrupts the usual onshore moisture flows. CMIP3 ensemble experiments connect anthropogenic greenhouse gas increases over the late twentieth century with the Indian Ocean SST warming inducing these changes to East African rainfall (Funk et al, 2008). An especially strong decrease in mean rainfall is found, to the magnitude of 35-45% of the 1950-1979 mean, through central and eastern Kenya and Ethiopia during the long rains (Williams and Funk, 2010). Data that suggests that drying has been amplified at high elevations, including the summit of Mount Kilimanjaro emphasises these trends (Mölg et al, 2009).

For East African agriculture, the onset of the wet season is as important as trends in mean rainfall. A mixed picture emerges with only isolated locations showing an earlier rainy season onset (Kniveton et al, 2009). The contrasting results that have emerged from previous studies over East Africa highlight the complexity of the East African region and in particular the difficulties in analysis of rainfall.

3.3.2 Trends in East African climate

3.3.2.1 Annually averaged trends

Figures 3.5 and 3.6 show the time series of the annually averaged CRU temperature, NCEP rainfall and H08 rainfall over the East African domain. Mean annual temperatures over the second half of the twentieth century show considerable interannual variation, with a standard deviation of 0.302. There is evidence of a general trend towards warmer mean temperatures which is particularly apparent from 1970 onwards, although this trend is not statistically significant. Both rainfall datasets show no statistically significant trend for this period but contain significant interannual variability, reinforcing assertions in the literature over East Africa (Mutai and Ward, 2000). NCEP has mean daily rainfall rates fluctuating

between 2.5mm/day and 3.15mm/day, with the H08 series having fluctuations of mean daily rainfall of 2.3mm/day to 3.0mm/day.

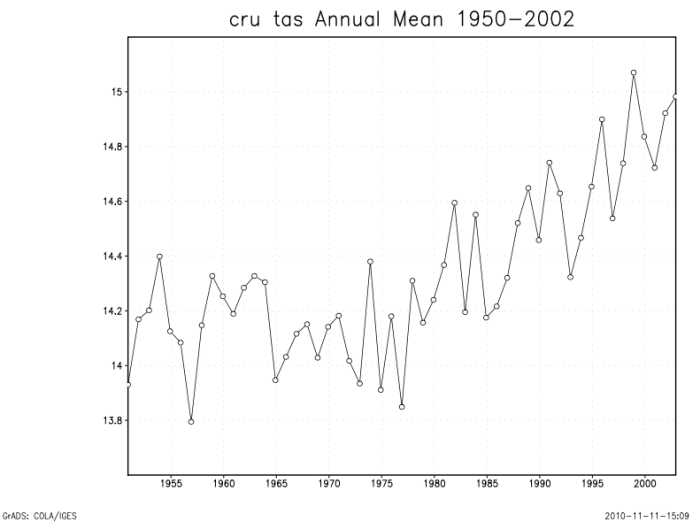


Figure 3.5 Annually averaged CRU East African temperature, x axis shows the year, y axis shows mean temperature in °C

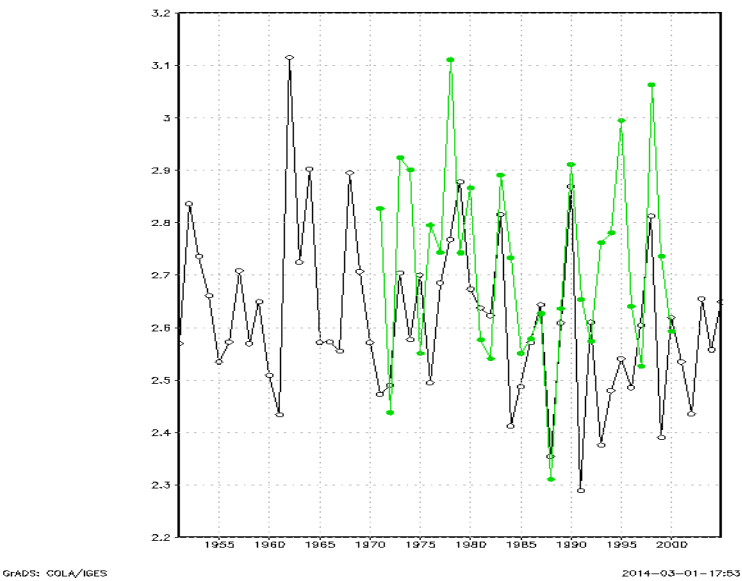


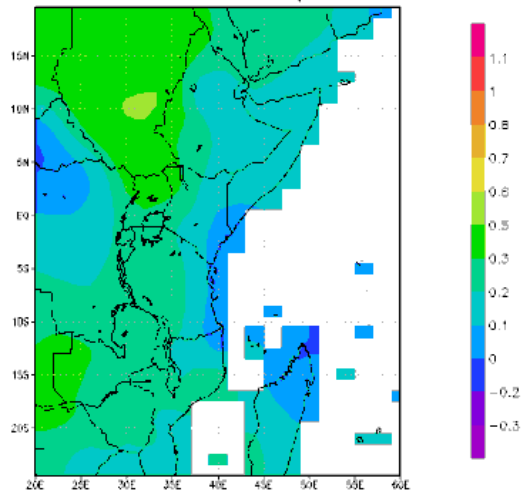
Figure 3.6 Annually averaged NCEP (green line) and H08 (black line) East African rainfall, x axis shows the year, y axis shows mean daily rainfall in mm

3.3.2.2 Spatial trends in East African climate

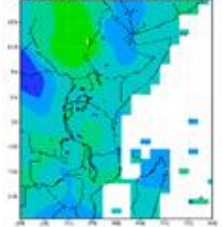
With increasing importance placed globally on the impact of rising anthropogenic greenhouse gas levels on climate, it was decided to consider trends which may already have occurred in East Africa during the second half of the twentieth century. The IPCC states that global mean temperatures have increased by $0.74^{\circ}\text{C} \pm 0.18^{\circ}\text{C}$ over the last century, 1906-2005, with an accelerated warming rate of $0.07^{\circ}\text{C}/\text{decade}$ on average over the past fifty years (Solomon et al, eds, 2007).

Figures 3.7 to 3.9 show the decadal changes in CRU temperature and NCEP and H08 rainfall respectively from 1961 to 2000. However, caution should be taken in interpreting spatial trends in the temperature dataset; the CRU series has relatively few individual stations contributing to it across the domain.

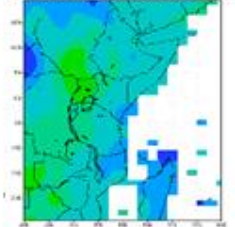
cru tas 1961–2000 trend/decade all



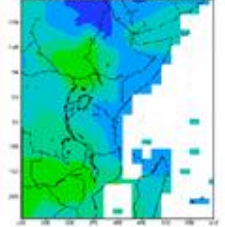
cru tas 1961–2000 trend/decade 1



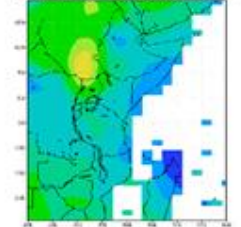
cru tas 1961–2000 trend/decade 2



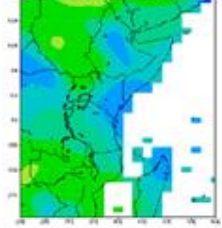
cru tas 1961–2000 trend/decade 3



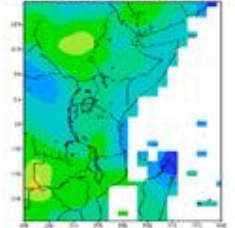
cru tas 1961–2000 trend/decade 4



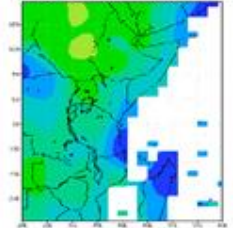
cru tas 1961–2000 trend/decade 5



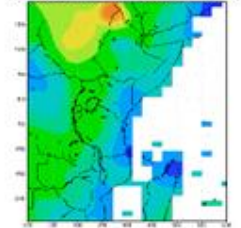
cru tas 1961–2000 trend/decade 6



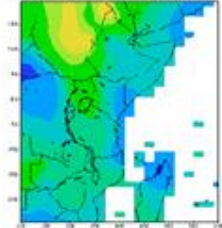
cru tas 1961–2000 trend/decade 7



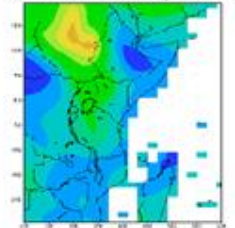
cru tas 1961–2000 trend/decade 8



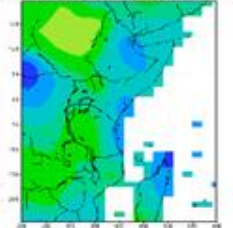
cru tas 1961–2000 trend/decade 9



cru tas 1961–2000 trend/decade 10



cru tas 1961–2000 trend/decade 11



cru tas 1961–2000 trend/decade 12

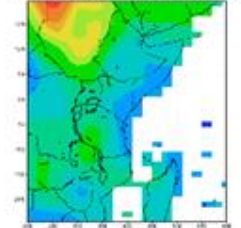
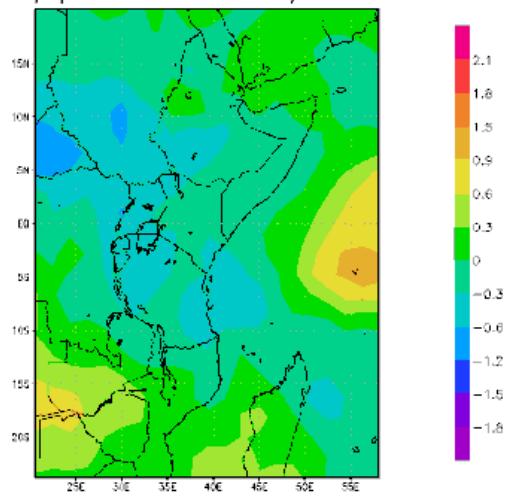
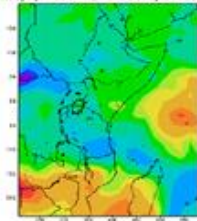


Figure 3.7 CRU annual and monthly temperature trend 1961–2000, in °C/decade, all plots follow the same scale as the annual plot

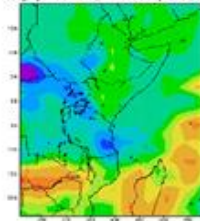
ncep pr 1961-2000 trend/decade all



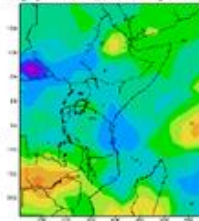
ncep pr 1961-2000 trend/decade 1



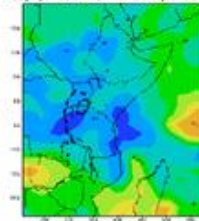
ncep pr 1961-2000 trend/decade 2



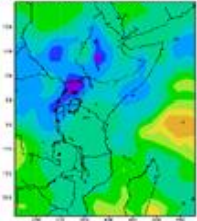
ncep pr 1961-2000 trend/decade 3



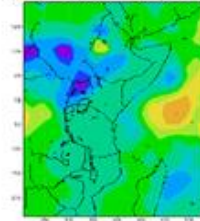
ncep pr 1961-2000 trend/decade 4



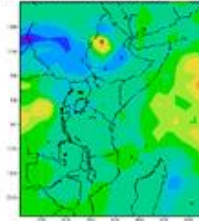
ncep pr 1961-2000 trend/decade 5



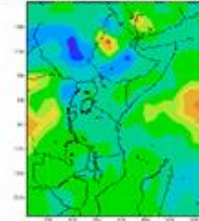
ncep pr 1961-2000 trend/decade 6



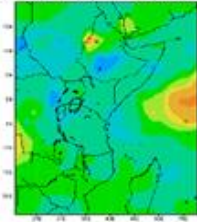
ncep pr 1961-2000 trend/decade 7



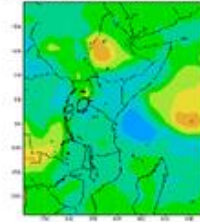
ncep pr 1961-2000 trend/decade 8



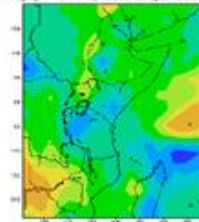
ncep pr 1961-2000 trend/decade 9



ncep pr 1961-2000 trend/decade 10



ncep pr 1961-2000 trend/decade 11



ncep pr 1961-2000 trend/decade 12

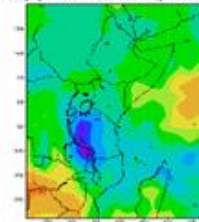


Figure 3.8 NCEP annual and monthly rainfall trend 1961-2000, in mm/day/decade, all plots follow the same scale as the annual plot

jap pr 1961–2000 trend/decade all

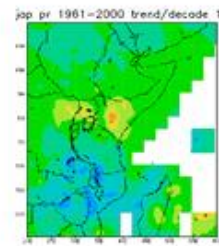
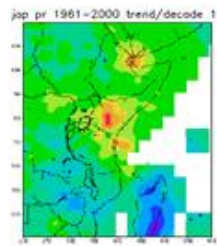
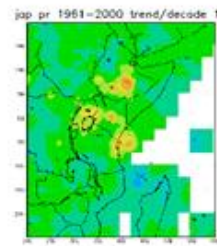
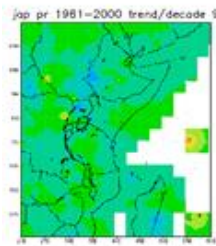
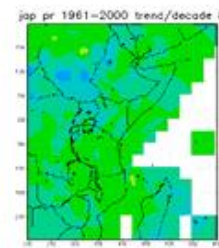
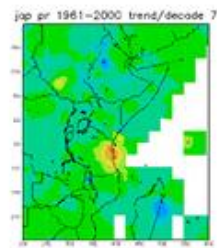
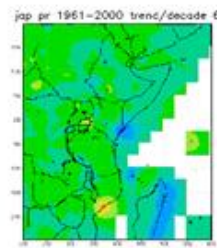
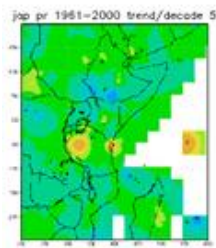
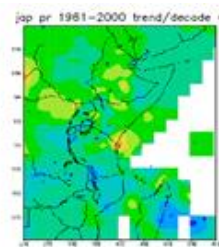
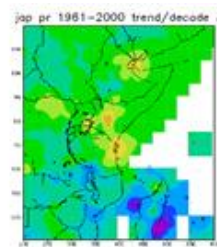
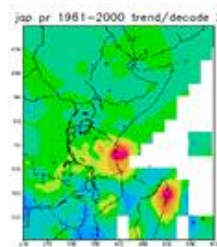
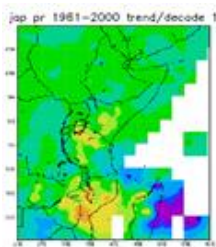
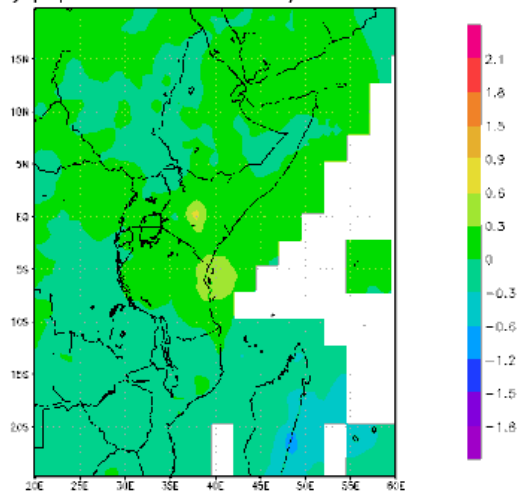


Figure 3.9 H08 annual and monthly rainfall trend 1961–2000, in mm/day/decade, all plots follow the same scale as the annual plot

Over the 1961-2000 period, much of the domain has an annual warming trend of 0.1°C to $0.6^{\circ}\text{C}/\text{decade}$. Outside of the focus countries there is a small region with no discernible change over the Democratic Republic of Congo. Additionally, a small region on the Indian Ocean coast, co-located with Tanzania and Kenya has very little change in mean temperatures during this time period. The positive trend of greatest magnitude is seen over northern parts of the domain, focused over Sudan and South Sudan and extending into Uganda and Ethiopia. This trend measures between 0.3°C and $0.6^{\circ}\text{C}/\text{decade}$. Elsewhere, including inland Kenya and Tanzania, Burundi, Rwanda and Eritrea a positive temperature trend of between 0.1°C and 0.3°C is seen.

On a monthly basis, the spatial pattern of warming trends is similar to the annual mean decadal trend with the magnitude of the trend varying seasonally. The greatest positive trend of up to $1.1^{\circ}\text{C}/\text{decade}$ is seen over northern parts of the region. This pattern reaches a peak in August but high warming rates continue through to December in this area. The positive warming trend is also evident throughout East Africa across these months, although to a lesser extent. January and February are the only two months where the trend is not of greatest magnitude in the north. At this time the largest positive temperature trend is seen over South Sudan and Uganda and measures $0.6^{\circ}\text{C}/\text{decade}$. North of this latitude during January and February there is a very slight cooling trend of up to $0.2^{\circ}\text{C}/\text{decade}$. For much of the southern part of the domain, the temperature trend remains consistent with the annually averaged trend. The time period shown here is relatively short (forty years), which Henson et al (2010) suggests is the minimum time period for which trends could be detected. A statistically significant trend was not detected over the domain as a whole.

Rainfall trends from 1961-2000 show much less coherence than the temperature trends, in particular between the two rainfall datasets; NCEP and H08. Equally, Hagemann et al (2005) suggest that for reanalysis datasets changes in measuring systems through the series

mean that great care should be taken in deriving trends in hydrological variables.

Therefore, whilst for this work the trends per decade are examined, no claim for statistical significance or otherwise is made. For NCEP reanalysis data, a key feature is an annually averaged drying trend of up to 1.2mm per day/decade over continental East Africa. In the H08 dataset, there is no clear trend on an annually averaged basis with a figure of between -0.3mm and +0.3mm per day/decade depending on the location.

From month to month, the difference in the size and direction of the rainfall trends between the two series continues. The greatest drying trends in NCEP are found during the long rains (March to May) and continuing into June. Here a decrease in rainfall of up to 1.8mm per day/decade is found centred over South Sudan, Ethiopia, Uganda, Rwanda and Burundi. This is in direct contrast to the H08 dataset where the short rains have a slight positive trend of up to 0.9mm per day/decade. For the short rains that are the focus of this study, the NCEP dataset shows a drying trend to the magnitude of 0.0 to -0.3mm per day/decade in October, increasing to up to 1.6mm per day/decade in December over Tanzania. This is the dominant trend for Tanzania and Kenya but for the other two focus countries (Uganda and Ethiopia) the trend is less pronounced and at times is for increased rainfall, such as the +1.5mm per day/decade trend present over the Ethiopian Highlands in October. This is closer to the pattern seen in the H08 data series during the short rains. Here, the dominant pattern through Kenya, Uganda and Ethiopia during the short rains is for increased rainfall of up to 2.1mm per day/decade during November, with similar patterns seen in October and December to a slightly lesser extent. To the south of this pattern there is a tendency to drier conditions of up to 1.2mm per day/decade. This creates a dipole pattern in rainfall trends during the October to December period.

3.4 Extremes in the East African climate

Extreme events are both difficult to model and key to the climatological variability of a region. Analysis of extreme events is more challenging than for mean conditions with the lack of availability of daily climatic datasets, particularly in the tropics, rendering such analysis even more troublesome (Kiktev et al, 2003). It has been said that model differences are a primary source of uncertainty in simulation of extremes at the current time, with the representation of extreme tropical rainfall particularly difficult (Kharin et al, 2007). Despite this, it has been proposed that the impact of increased anthropogenic greenhouse gases will be realised most strongly in changes in the intensity and frequency of extreme events and therefore the study of these events is key for understanding future climates of the region (Kharin et al, 2007).

Figures 3.10 and 3.11 show the mean and trend for NCEP and H08 95% rainfall events respectively. Figure 3.12 shows the mean and trend for NCEP 95% maximum temperature events. Previous sections have highlighted the difficulty of establishing rainfall characteristics over East Africa with the NCEP and H08 rainfall series showed varying attributes of mean values over East Africa. For the heaviest rainfall events, the spatial pattern of the climatology is similar for both datasets with the most intense 95% maximum rainfall events occurring in western Ethiopia and lower maximum rainfall events occurring in the coastal regions of East Africa through coastal Ethiopia, Kenya and Tanzania. Despite this spatial similarity, the magnitude of the heaviest rainfall events is different between the two datasets. Following the patterns in the climatology the H08 rainfall has much stronger heavy rainfall events with the 95% events peaking at 30-35mm/day; equivalent events in the NCEP dataset peak at values 10mm lower for the same regions.

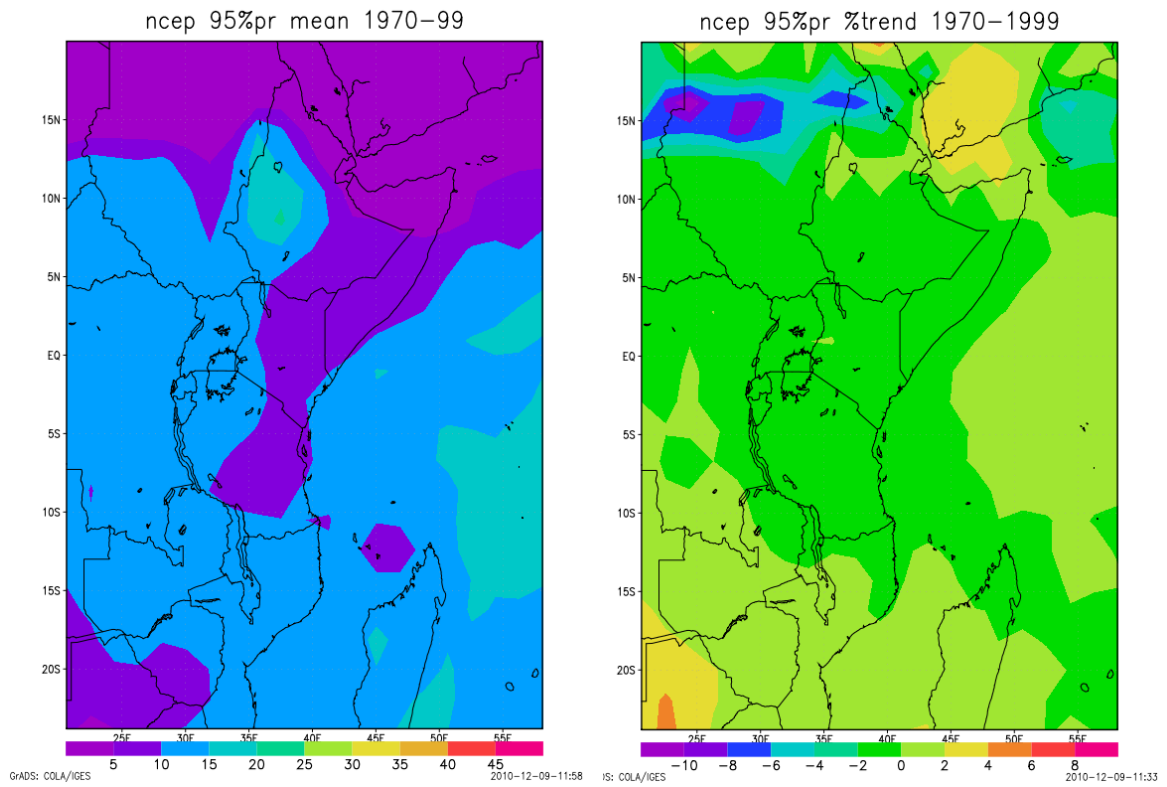


Figure 3.10 NCEP mean and trend 95% rainfall events, mean in mm/day and trend in percent change relative to the mean

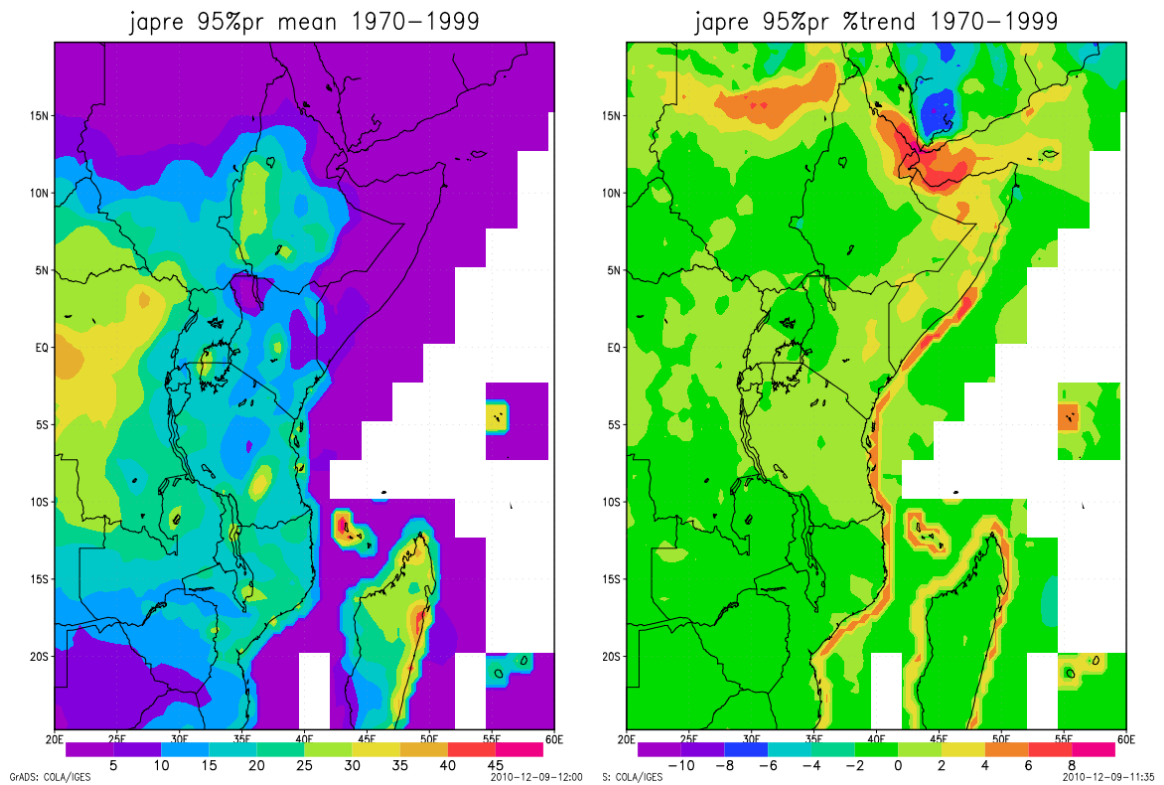


Figure 3.11 H08 mean and trend 95% rainfall events , mean in mm/day and trend in percent change relative to the mean

Differences in trends between the two series are pronounced. Over Sudan trends are of opposite signs in the two datasets with 95% rainfall events becoming heavier by up to 6% in the H08 dataset, and lighter by 10% in the NCEP series. Elsewhere, the NCEP dataset has a slight negative trend (2%) in heavy rainfall events over the focus countries of Kenya, Ethiopia, Uganda and Tanzania indicating that the heaviest rainfall events have become lighter in the last three decades of the twentieth century. The H08 dataset has no single trend covering such a large region. Here, coastal areas show evidence of an increase in the intensity of heavy rainfall events of up to 2% with the opposite inland. It is likely that the apparent trend in the coast Indian Ocean is an artefact in the dataset. It is noted that the very short time period may have produced artificial trends in the most extreme events of the region where interannual variability could be considerable and hence statistical significance of the trends is not tested in this case, the same applies to temperature.

For temperature episodes, maximum (95%) and minimum (5%), the CRU temperature is examined. To restate, caution must be taken with deriving trends from CRU due to the low density of the station network contributing to the series. The Ethiopian Highlands are the location for both the coldest minimum episodes and the lowest maximum temperature episodes, with climatology values of 0°C to 6°C and 24°C to 30°C respectively. Elsewhere, the minimum temperatures do not reach such low values; typical 5% minimum values over much of the East Africa domain measure 6°C to 18°C. The hottest temperature maximums are found in Sudan, with a peak of 48°C. South Sudan also sees hot maximum temperature episodes of up to 42°C. Over Tanzania, Kenya and Uganda the maximum temperature episodes measure 27°C to 36°C. In line with what is expected with increasing anthropogenic greenhouse gases the 95% maximum temperature episodes largely see an increase in temperatures occurring over the latter part of the twentieth century. The greatest increases are found over Uganda and along the border between South Sudan and

the Democratic Republic of the Congo, of up to $0.18^{\circ}\text{C}/\text{decade}$. Over the other three focus countries increases in the maximum temperature of up $0.09^{\circ}\text{C}/\text{decade}$ are seen, although the change is smaller ($0\text{-}0.03^{\circ}\text{C}/\text{decade}$) over Tanzania. Over much of the domain there is a slight cooling trend in the minimum temperatures. This is especially pronounced over Sudan and South Sudan where decreases in the 5% minimum temperatures of up to $-0.06^{\circ}\text{C}/\text{decade}$ are seen. Over the focus countries a very small decrease in minimum temperatures of 0.02°C to 0.04°C is seen.

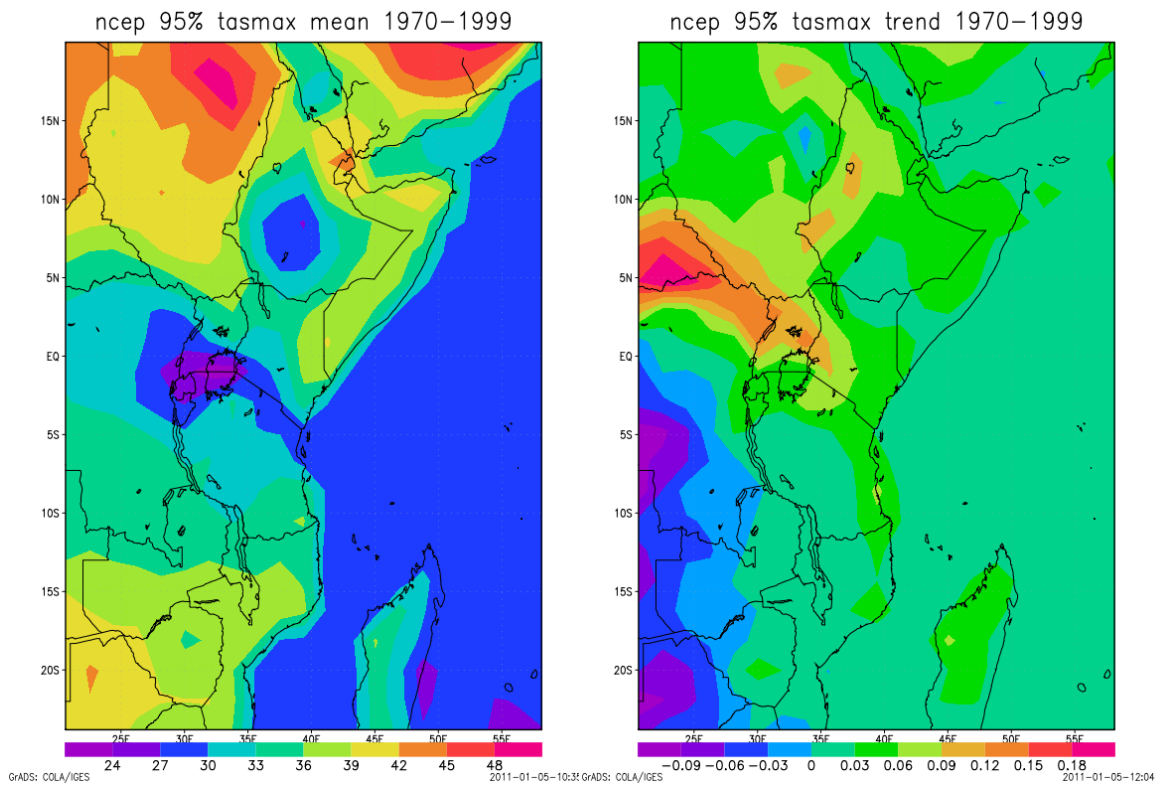


Figure 3.12 NCEP 95% maximum temperature mean and trend, mean in $^{\circ}\text{C}$ and trend in $^{\circ}\text{C}/\text{decade}$.

3.5 Evaluation of AOGCMs – how well can they replicate the twentieth century East African climate?

3.5.1 Use of AOGCMs in climate research

Whilst similarities between observed datasets and climate model simulations of the twentieth century are a necessity, they are not sufficient to have confidence in future projections (Caminade and Terray, 2010). This is despite both the improvements made to GCMs over the past decades with the models evolving from a solely atmospheric model to including oceans, land surface processes, ocean-ice interactions and numerous parameterisations and that AOGCMS remain the only way to project potential future climates (Herrera et al, 2006). Giannini et al (2008) assert that the uncertainty relating to different model simulations can be reduced through using a multimodel ensemble; they hypothesise that the averaging of different model parameterisations will neutralise differences with no physical basis whilst reinforcing and retaining information from robust responses. Additionally, individual models can be analysed to see if any particular response to increasing anthropogenic greenhouse gases occur; both approaches are employed here.

Despite limitations to the use of AOGCMs, they are a vital part of research over East Africa to evaluate potential climate changes and also in attribution studies. Moreover, whilst accurate simulations of the present day climate are not sufficient for confidence in twenty-first century projections, it is still crucial that they are able to reproduce the background climate state. This stage of evaluation increases understanding of mechanisms and influences on East African climate through which patterns are formed and without this stage the ability to verify future projections decreases further (Giannini et al, 2008).

3.5.2 Climatology of the IPCC AR4 models over East Africa

For this study a subsection of eight IPCC AR4 models are used. This selection is based on data availability with the selected models documented in Chapter 2, Section 2.8. An ensemble mean of all of the CMIP3 models is also used; this is beneficial because the use of an ensemble can decrease uncertainty. Globally, for almost all variables, the mean and median of all CMIP3 models has a lower error under climatology conditions than the individual models (Gleckler et al, 2008). Not all models are equally skilful in simulating the climatological annual cycle and variance of monthly anomalies, with the CCCMA, MPI and GFDL models found to have more accurate performance over the tropics to the magnitude of overall errors across all variables of up to 5% less than is typical, and up to 30% less than the most poorly performing models in the tropics. There is still uncertainty as to how a good mean statistics performance translates to simulations of interannual variability (Gleckler et al, 2008). In addition to mean characteristics and interannual variability, teleconnections, for example with the El Nino Southern Oscillation, play a large role in influencing the East African climate especially on an interannual basis. Many models do not simulate this connection accurately (Hulme et al, 2001). One model that does successfully simulate the East African climate, including the Indian Ocean Dipole and ENSO relationships, is SINTEX-F1 model (Behera et al, 2005).

Aside from AOGMS, RCMs (regional climate models) simulate a chosen region in more detail with boundary conditions from a global AOGCM used to force a high resolution RCM. There are several studies which employ RCMs over an East African domain. The earliest of these studies successfully customised a RCM over East Africa, utilising the NCAR RegCM2 model and reproducing interannual rainfall variability of the OND short rains for most years between 1982 and 1993 (Sun et al, 1999). Building on this result, the regional model CLM is employed in various configurations to evaluate the seasonal cycle

of rainfall. All configurations of the model are adequate although absolute rainfall values were influenced by the parameterisation schemes selected, highlighting how even a single model and single domain can produce vastly different results depending on human choice in setup (Kaspar and Cubasch, 2008). The importance of different configurations can be further reinforced through the need to incorporate a three-dimensional lake model to be coupled with the RCM in the Lake Victoria region in order to obtain a realistic climate simulation (Song et al, 2004).

The model MM5 successfully simulated the rainy season onset over the Greater Horn of Africa in comparison to three satellite and gauge based observational datasets (Riddle and Cook, 2008). A further regional model based study employed the ICTP-RegCM3 model to evaluate the OND short rains. This model has a simulation that is fairly consistent with observational datasets for the monthly and seasonal climatology and interannual variability. The most accurately matched simulations occur over northern Uganda, central Tanzania and the coastline. Rainfall is underestimated over north eastern Kenya and overestimated over the Lake Victoria basin. Over the Kenyan Highlands there are also issues associated with the model resolution. Despite these areas of poorer simulation for the region as a whole RegCM3 maintains teleconnections to global variability modes including ENSO and the Dipole Model Index (Anyah and Semazzi, 2007).

The following sections refer to evaluation of mean fields of the subset of CMIP3 AOGCMs and an ensemble of these models. Detail of the selected AOGCMs is located in Chapter 2, Section 2.8. One method of model evaluation which is often undertaken is to subtract the AOGCM field from the equivalent field in the observed datasets. This method is not used here due to the known issues with the available gridded datasets over East Africa and in particular the differences between gridded products over the same domain.

3.5.2.1 Temperature

Figure 3.13 shows the ensemble temperature cycle. Individual temperature cycles for each individual model are not shown here, but the patterns seen are summarised in Table 3.1.

Generally the models simulate the spatial distribution of temperature adequately, although absolute temperature values are less accurately represented in some models. The CMIP3 ensemble shows a good approximation of the CRU data during the thirty year climatology period both spatially and temporally for the highest temperatures throughout the year. The magnitudes of high temperatures tend to be underestimated, for example over South Sudan in May. Some cooler regions tend to be presented as even colder in the ensemble indicating that this model ensemble has a slight negative temperature bias.

Table 3.1 Summary of AOGCM simulations of the control period

Model	Temperature	Rainfall
Ensemble	Good spatial and temporal climatology but high temperatures underestimated	Spatial cycle well simulated but rainfall intensity muted
CCCMA	Mostly too cold across the East African domain in particular over Uganda, Rwanda, Burundi and Tanzania	Good placement of the most intense rainfall but the northern fringes of the rainfall region consistently extended too far north
CNRM	This model was consistently too cool by between 2°C and 4°C across the whole East African domain	Most intense rainfall was well represented but rainfall was too extensive for most of the year
CSIRO	Good spatial depiction of temperature patterns but slightly too cool	Annual cycle was well presented in this model and the intensity of the rainfall was accurately simulated
GFDL	The months of July to November were especially sound for temperature values, but this applied generally for the whole year	Good modelling of the annual cycle but October to May saw rainfall over land overestimated
GISS	One of the few models where temperatures were too	Rainfall was highly unrealistic with values much

	warm (in this case inland over much of the domain) This was coupled by temperatures too cool along the East Africa coast	too high over the East African domain
MIUB	Temperature was well represented for both magnitude and spatial patterns	The annual cycle moved northwards too soon and also to too high a latitude, which was coupled with rainfall intensity which was too low
MPI	Spatial patterns of temperature were good, but the magnitude was too extreme in both directions, that is, some months were colder and some warmer than actually occurred	Rainfall intensity was underestimated over land areas of the domain although the spatial pattern was well represented
MRI	This was one of the weakest models for spatial patterns of temperature but the values of temperature were well represented	A good model for East African rainfall. Both spatial location of the annual cycle and rainfall intensity were well represented

ensemble 20c3m tas Mean all 1970–1999

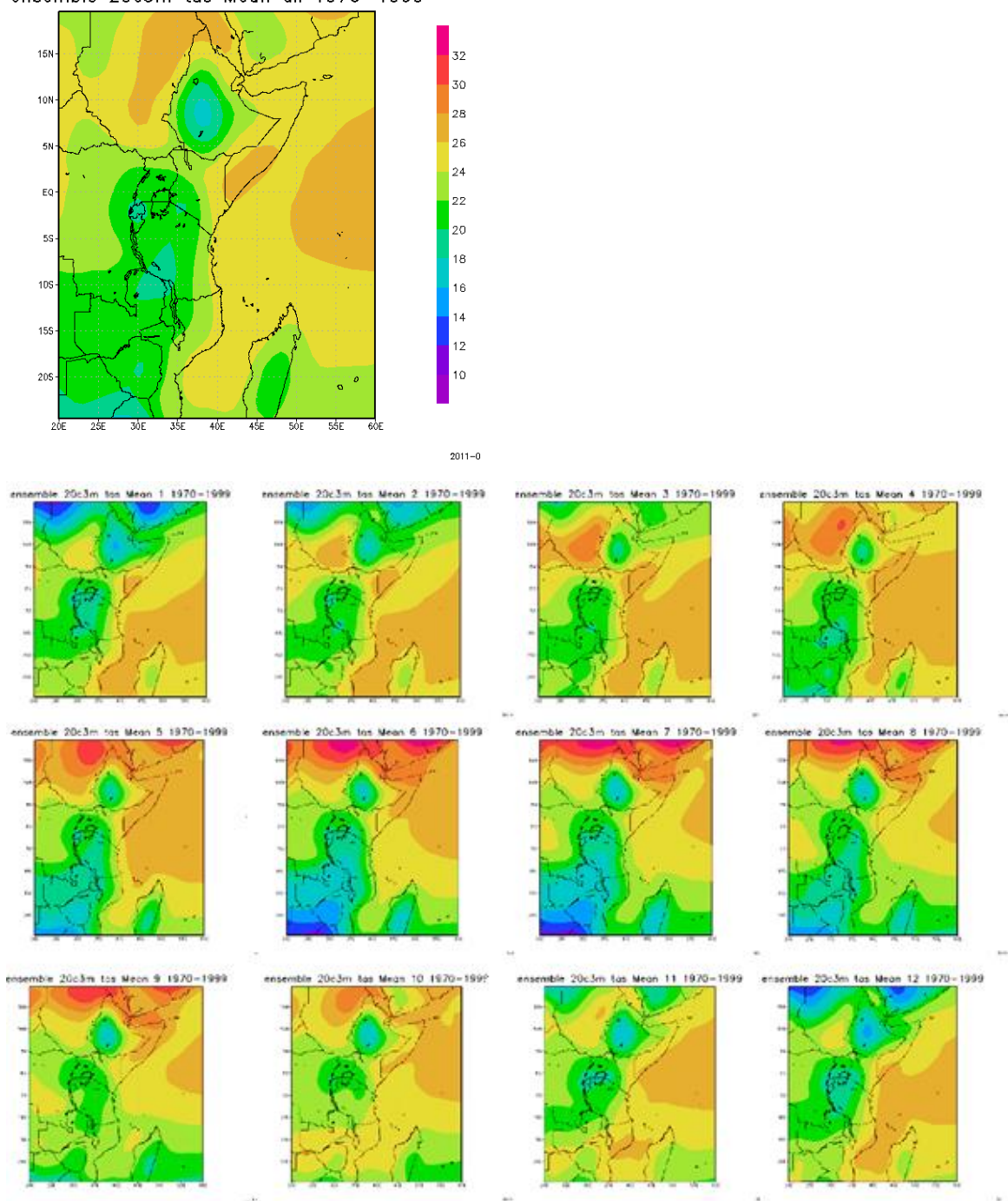


Figure 3.13 Ensemble mean temperatures, annual and monthly for the climatology in °C, the scale shown next to the annual plot applies to all months

3.5.2.2 Rainfall

The rainfall cycle for the ensemble is shown in Figure 3.14. Individual model cycles were also computed but are not shown. Over the year as a whole the ensemble depicts a reasonable approximation of the annual rainfall cycle over East Africa. The seasonal cycle of rainfall, including movement associated with the ITCZ is well illustrated although the rainfall intensity seen cannot be matched with the ensemble showing more muted rainfall values throughout the year. In the ensemble the southwards movement of the ITCZ related band of rainfall occurs too early, illustrated by the centre of rainfall remaining over Ethiopia until October to November in the NCEP and H08 rainfall datasets but only until September in the model ensemble.

Figure 3.15 shows the climatological cycle of rainfall averaged zonally across the domain for the NCEP and H08 datasets and the eight selected models for the climatology period. Whilst there are differences between the two datasets representing the ‘observed’ climate the contrast with the majority of the AOGCMs was evident. Most of the models have an annual cycle that moves too far northwards with rainfall spread over too wide a latitude for much of the year. This causes the rainfall amounts at many locations to be overestimated. Two models produce an annual cycle that is a close approximation to the reanalysis and statistical datasets; MPI and MRI. In these models, the daily rainfall rate and latitudinal spatial coverage is in good qualitative agreement with the NCEP and H08 cycles.

ensemble 20c3m pr Mean all 1970–1999

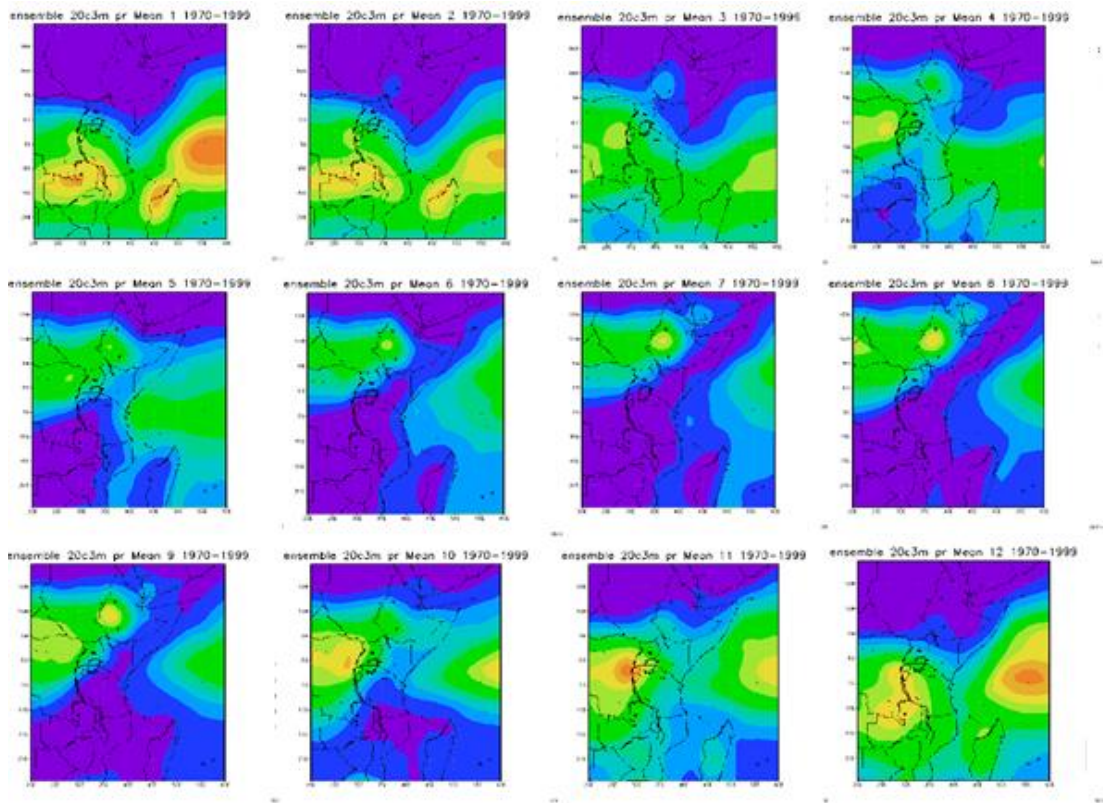
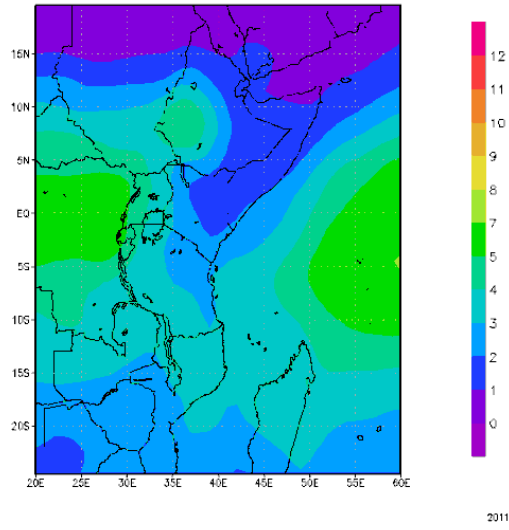


Figure 3.14 Model ensemble annual and monthly rainfall cycle for the climatology period in mm/day, the scale shown next to the annual plot applies to all months

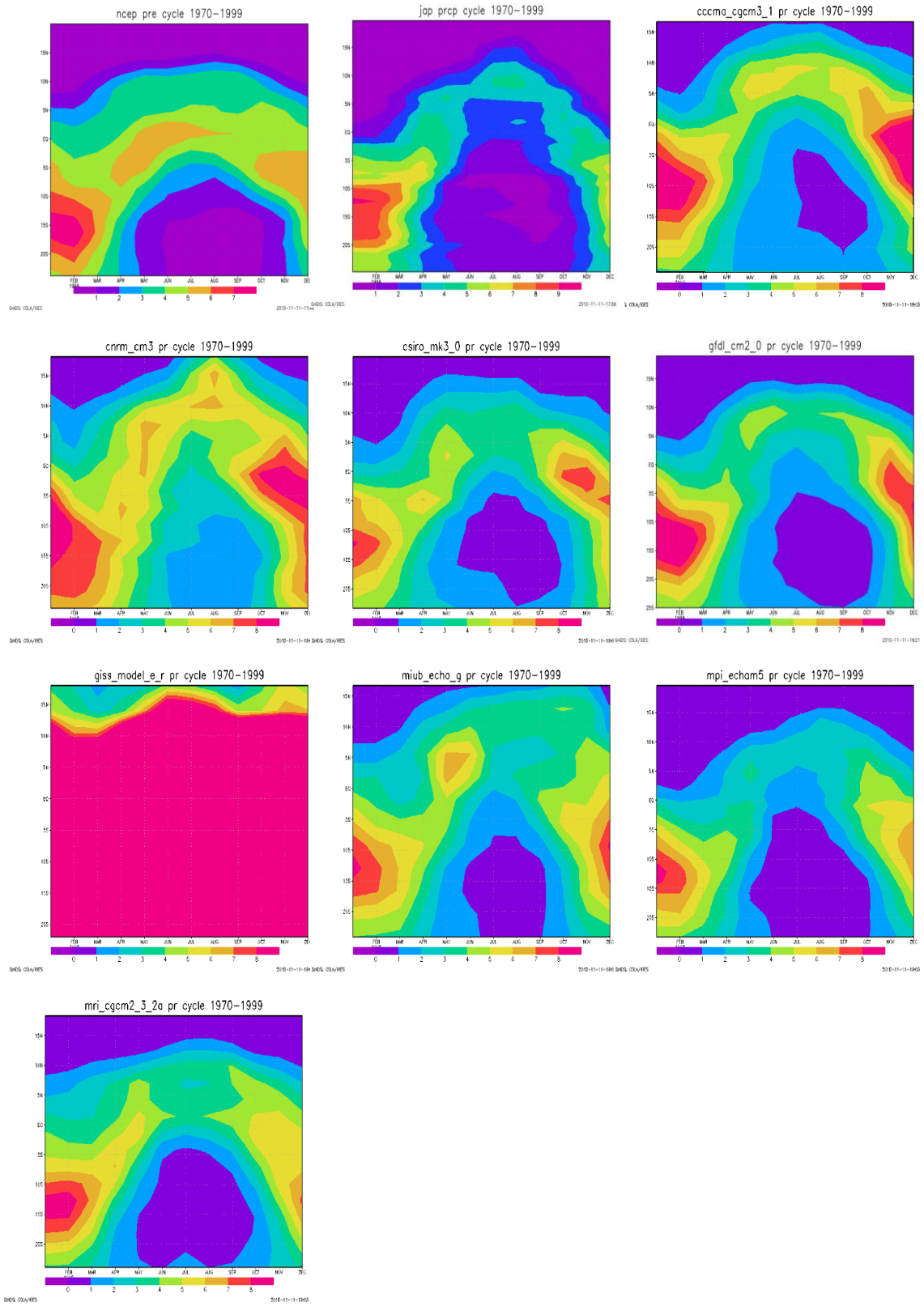


Figure 3.15 Zonally averaged climatological rainfall cycle for the two contemporary datasets and eight models in mm/day, x axis shows months (January to December from left to right), y axis shows latitude

3.6 AOGCM projections of East African climate during the twenty-first century

3.6.1 East African climate projections in the literature

There are previous studies over East Africa examining AOGCM projections during the twenty-first century. The IPCC Fourth Assessment Report, published in 2007, projects a warmer future under non-mitigation scenarios for all models. The multimodel global mean temperature increase is 3.4°C (2.0°C to 5.4°C) under the A2 high emissions scenario by 2090-2099 relative to 1980-1999 (Solomon et al, eds, 2007). Over East Africa, all models used in the IPCC AR4 report under the A1B scenario show increased temperatures over the domain during all seasons that are of greater magnitude than the global temperature response (Doherty et al, 2010). The response of rainfall to increased greenhouse gases is less certain but it has been hypothesised there could be an intensification of the hydrological cycle which will result in both increases and decreases in rainfall at the local level and an intensification of heaviest rainfall events (Trenberth et al, 2003).

The IPCC has proposed that one of the most robust rainfall responses in AOGCMs is for increased rainfall over East Africa. Eighteen of the twenty-one models used over this region in the IPCC AR4 simulate increased rainfall during the twenty-first century over a core region over East Africa to the east of the Great Lakes (Solomon et al, eds, 2007).

Whilst very few regions of Africa see projected changes outside natural variability under a low emissions B1 scenario, East Africa is an exception to this. Under this low emissions scenario, rainfall is projected to rise by up to 30% during the December to February boreal winter months. This pattern is reinforced under the high emissions A2 scenario where rainfall is projected to increase by up to 50%, and even up to 100% over some parts of East Africa. Caution is urged about projecting changes to ENSO associated rainfall due to the

inherent difficulties in the simulation of ENSO and its teleconnections to the region (Hulme et al, 2001).

There are also AOGCMs that support projections of a drier future for East Africa, especially for seasons aside from the OND short rains. During the June to August period there have been projections of decreases of between 5% and 10% under the conservative B1 scenario over the domain and also for significant decreases over parts of the Horn of Africa (Hulme et al, 2001). One particular study hypothesises that AOGCMs that project increases over East Africa are the same models as those projecting increased El Nino events over the twenty-first century with these events linked to increased rainfall during the short rains over East Africa (Williams and Funk, 2010). These authors suggest instead that the dominant trend will be the continuation of a drying trend that emerged at the end of the twentieth century over East Africa during the long rains. This trend is linked to tropical Indian Ocean warming, a feature which will continue during the twenty-first century irrespective of changes to ENSO, and which could therefore be more significant for the region (Williams and Funk, 2010).

Opposing hypotheses such as these highlight the importance of considering individual rainfall seasons over East Africa and not just annual conditions with potential trends not necessarily being universal due to the complex controls on the domain during the different seasons.

3.6.2 Projections over East Africa from the IPCC AR4 model subset

There is considerable uncertainty surrounding global emissions scenarios over the coming century. In order to provide different possible outcomes under anthropogenic climate change in East Africa this chapter employs three different emissions scenarios (see Chapter 2, Section 2.8) and three ‘time slices’ in evaluating potential changes to the

climate of the domain over the twenty-first century. The three scenarios used are the B1, A1B and A2 IPCC SRES scenarios, representing low, medium and high emissions futures respectively. Time periods used are the 2030s (2030-2039 inclusive), the 2050s (2050-2059) and the 2090s (2090-2099). The expected evolution of the combination of scenario and time slice is that the change will increase in magnitude over time and through the increased emissions scenarios. In this case, it is expected that the smallest change to climatology values will occur under the low emissions B1 scenario during the 2030s and the greatest changes will occur at the end of the twenty-first century under the high emissions A2 scenario.

3.6.2.1 Temperature

Figure 3.16 shows temperature anomalies in the IPCC AR4 model ensemble relative to the 1970-1999 climatology for the three scenarios and the three time slices. Plots for individual models have been computed but are not shown here.

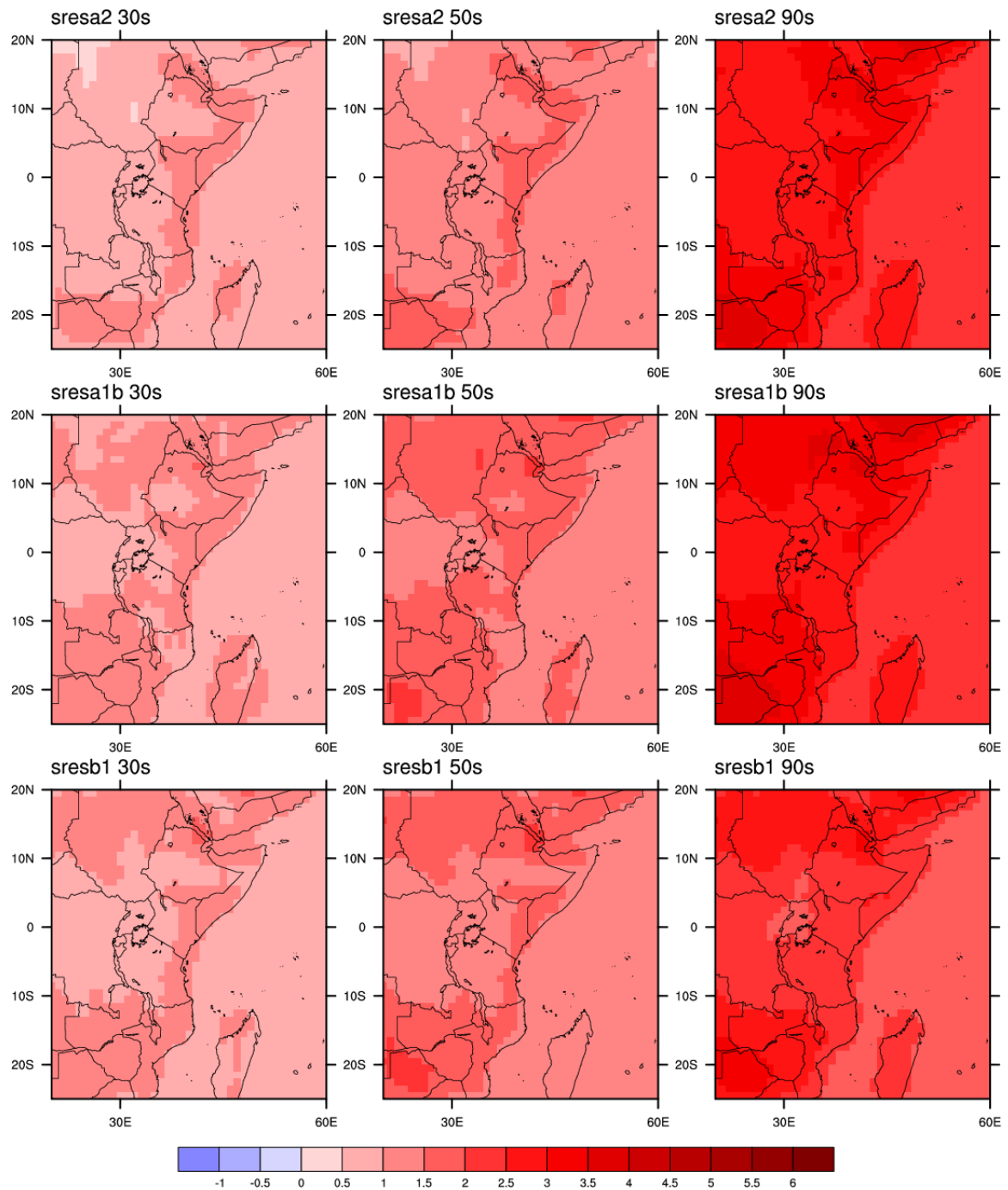


Figure 3.16 Model ensemble temperature response to 3 SRES scenarios, change in °C relative to 1971-2000. Left column shows response for 2030s, middle for 2050s and right for 2090s, the highest emissions scenario is in the top row and lowest in the bottom row.

The universal pattern is for warming with temperatures increasing throughout the twenty-first century, peaking at a positive anomaly of 4°C under the SRES A2 scenario by the 2090s for the ensemble over coastal East Africa. Whilst the individual models show

similar patterns of temperature increase over the twenty-first century, the magnitude of change and spatial pattern of this change shows considerable variation between models. In some models warming in excess of 5.5° is seen under the high emissions scenario by the 2090s, with Ethiopia, Sudan and South Sudan especially vulnerable to large projected temperature increases. With respect to a more mitigated low emissions scenario there is equally a spread of projections between models. The ensemble projection for the 2090s under the B1 scenario is for a positive anomaly of 2.5°C to 3.0°C but individual models show temperature increases from the climatology of 1.0°C to 3.0°C.

3.6.2.2 Rainfall

Figure 3.17 shows projected changes in rainfall for the ensemble relative to the climatology for both the three time slices and three emissions scenarios. Individual plots for each model were also computed but are not shown here. The ensemble of CMIP3 models shows a consistent spatial pattern in projected rainfall changes over the twenty-first century for the whole year over East Africa, for increased rainfall over the majority of countries of the domain. This change is centred over Uganda and the surrounding area, and measures 1.5-2.0mm/day at its peak. A projected decrease in rainfall is not a common feature for the focus countries in this region but is present over South Sudan and coastal Tanzania and Uganda during low emissions scenarios.

With the reaction of rainfall systems to increased anthropogenic greenhouse gases being more unpredictable in terms of spatial variation and amplitude than the expected increases in temperature it is no surprise that the individual models show much more variation in rainfall projections for the twenty-first century. For the short rains, the majority of the models follow the dominant annual projection of increased rainfall with GFDL the only one that departed from this projection. The model that projects the highest increases during

this season is CNRM, which projects rainfall increases of more than 1.5mm/day in December under the A2 scenario by the 2090s (and between 0.3 and 0.9mm/day during the rest of the OND season).

This agreement between models is less evident during the March to May long rains. Here, several models including CCCMA, CSIRO, GFDL and MPI project decreases in mean daily rainfall of up to 2mm over Ethiopia by the 2090s under the high A2 emissions scenario. Conversely other models (CNRM, MIUB and MRI) project increases in daily mean rainfall for these same months to the values of 0.6mm/day. This highlights the uncertainty in rainfall projections in AOGCMs for the long rains.

During the boreal summer of East Africa, a dominant trend in projections emerges once more. Here, the models converge on a drier pattern to the magnitude of 1.2mm/day over much of continental East Africa. In terms of links between projections of temperature and rainfall over the twenty-first century, locations where projected temperature increases are lowest match those where increased rainfall is projected.

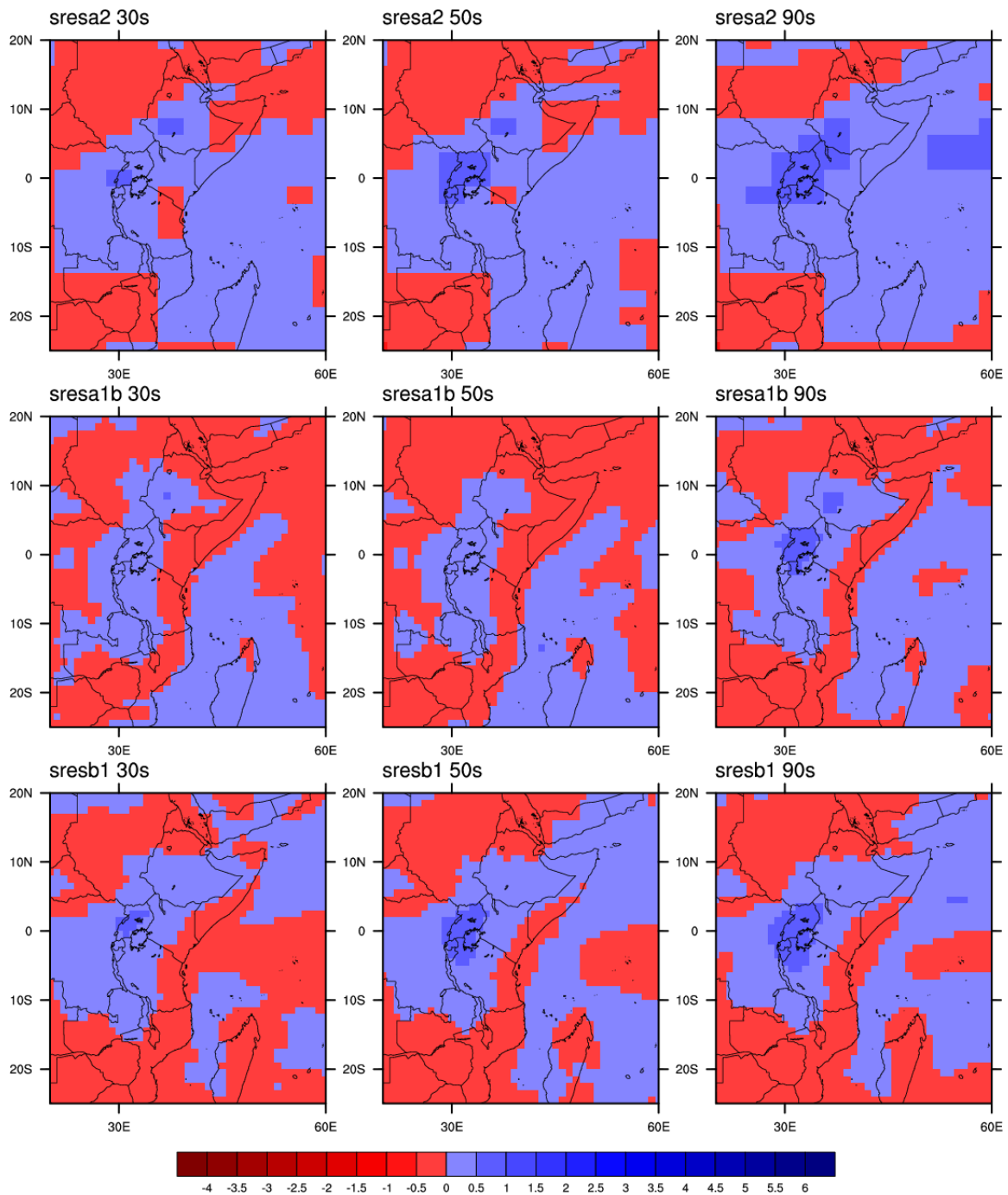


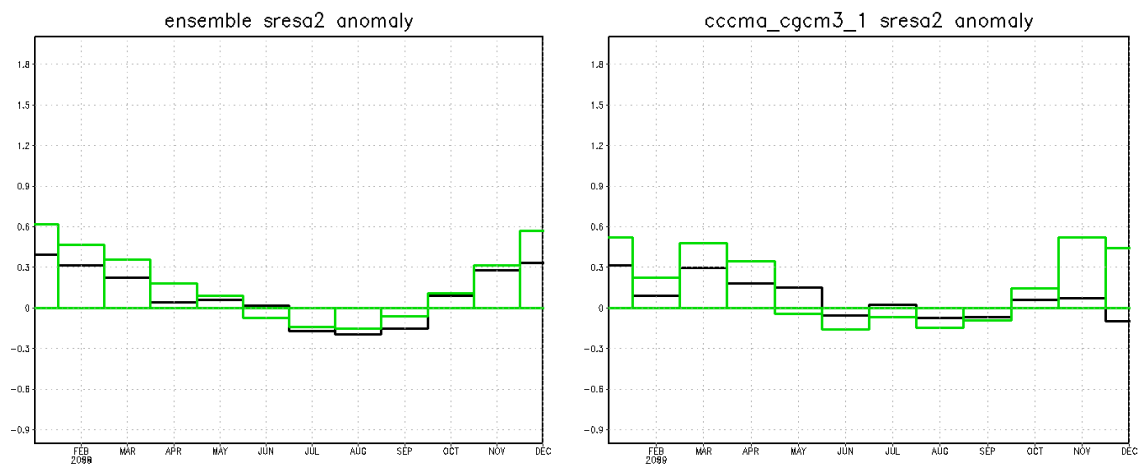
Figure 3.17 Model ensemble rainfall response to three SRES scenarios, change in mm/day relative to 1971-2000. Left column shows response for 2030s, middle for 2050s and right for 2090s, the highest emissions scenario is in the top row and lowest in the bottom row.

These patterns can be visualised for the East African domain as a whole, with rainfall anomalies from the baseline climatology for the A2 scenario for the ensemble and the individual models shown in Figure 3.18 for the 2050s (black) and 2090s (green). The 2030s are excluded from these plots due to the expectation that changes will be greater in the latter half of the twenty-first century. The GISS model is excluded due the gross

overestimation of rainfall over East Africa in this model. These plots help to reinforce the seasonal patterns described above. For the ensemble, the patterns are consistent between scenarios, with the highest emissions outcome underlying the projections that emerged earlier in the twenty-first century and under lower emission possibilities.

To summarise, the overriding ensemble projection is for increased rainfall from October to May, with maximum increases seen for November to May (of up to 0.6mm/day), but with a slight decrease (0.15mm/day) in rainfall projections for the boreal summer months.

Whilst many models concur with these patterns there is considerable variation in the magnitude of rainfall change projected. There are some exceptions within the individual models to this though. These include the CSIRO model which has inconsistency in the sign of projected rainfall changes with variation between individual months and the GFDL and MRI models which tend towards a drier future over the year as a whole.



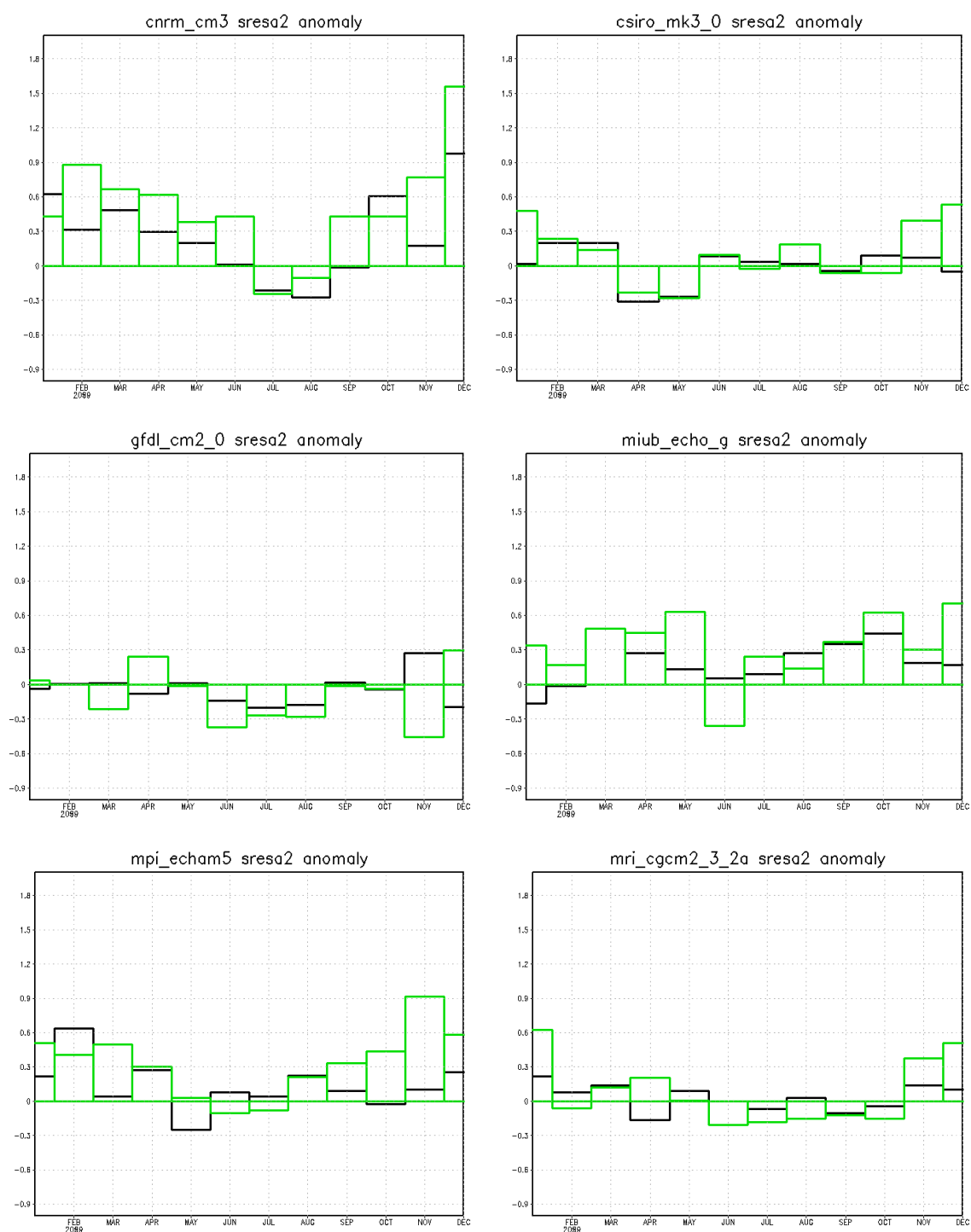


Figure 3.18 Monthly rainfall anomalies in mm/day relative to the climatology under the A2 scenario for the ensemble and individual models for the 2050s (black line) and 2090s (green line)

3.6.2.3 Extremes

The 95% rainfall event anomalies for the three emissions scenarios and time slices are shown in Figure 3.19 (a), (b) and (c) for the B1, A1B and A2 scenarios respectively. The key feature of these plots, especially by the end of the twenty-first century under the high emissions scenarios is the increased intensity of the most extreme rainfall events over the domain. This intensity is measured as an increase in daily rainfall during the heaviest rainfall events relative to the 1970-1999 climatology of between 1mm and 4mm a day. Centred over Kenya and Tanzania, this projected increased intensity is visible over the focus countries, although not Sudan. Here a slight decrease in intensity of up to 1mm/day during the heaviest 95% rainfall events is projected by the 2090s under all emissions scenarios.

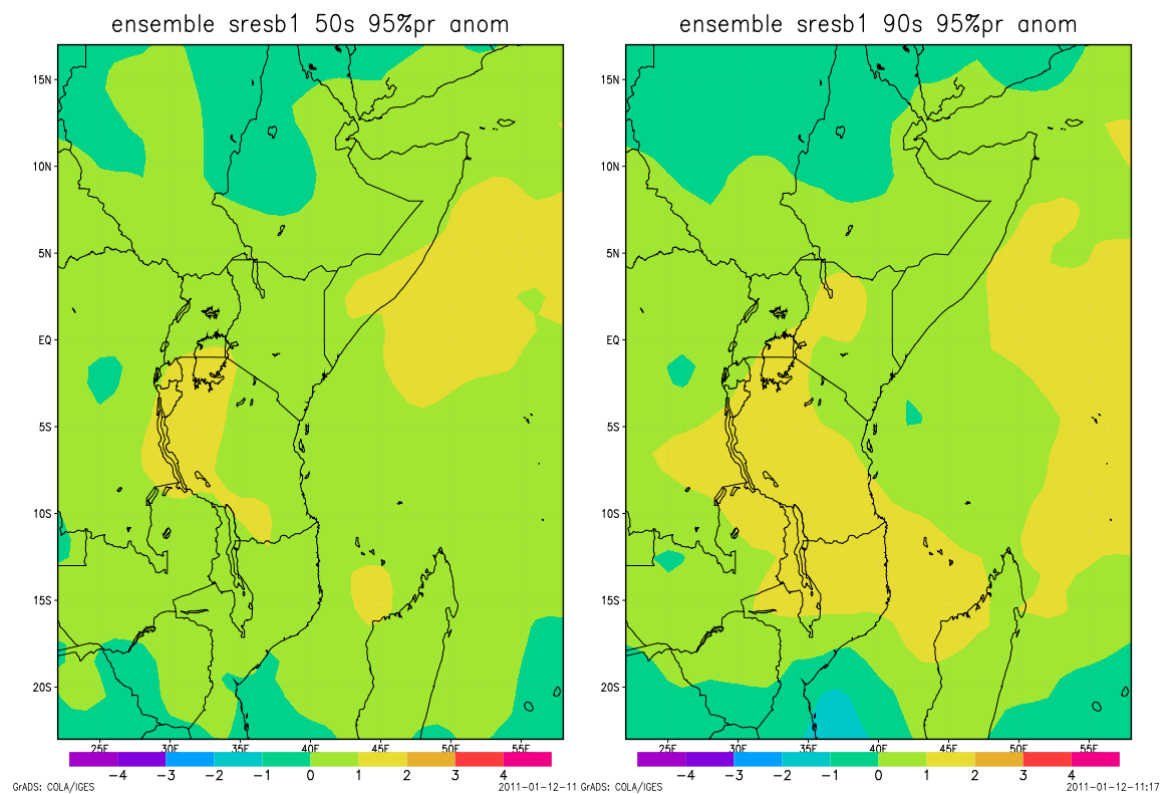


Figure 3.19 (a) 95% rainfall event anomalies under the B1 scenario in mm/day for the 2050s (left) and 2090s (right)

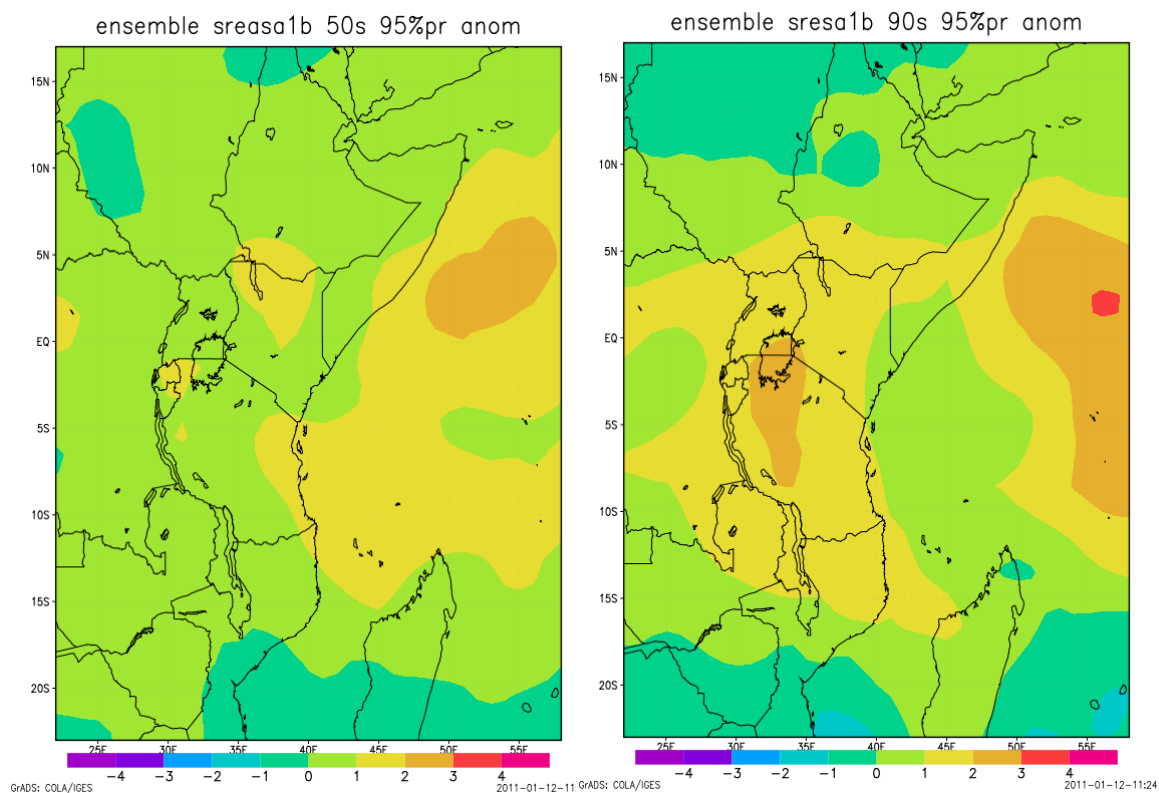


Figure 3.19 (b) 95% rainfall event anomalies under the A1B scenario in mm/day for the 2050s (left) and 2090s (right)

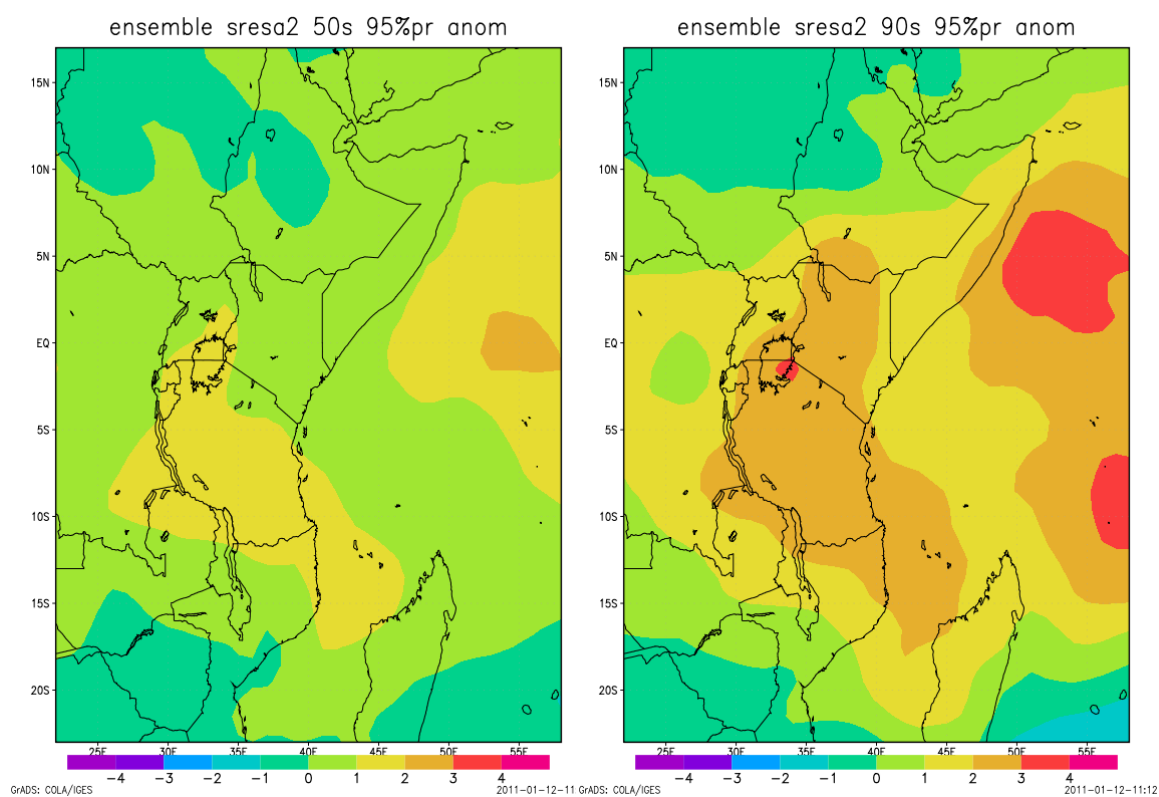


Figure 3.19 (c) 95% rainfall event anomalies under the A2 scenario in mm/day for the 2050s (left) and 2090s (right)

3.7 Chapter summary

Chapter 3 examines the mean annual background climate of East Africa, as portrayed in reanalysis, observed and statistical datasets and a subset of CMIP3 global climate models. This is followed by a preliminary evaluation of projections for East Africa under increasing greenhouse gas emissions throughout the twenty-first century. The purpose of this chapter was twofold. First, it was to provide background information to contextualise the contemporary mean climate region for the analysis that follows, on daily rainfall regimes during the contemporary short rains from Chapter 4 onwards. Second, this chapter answered the sub question of the second research aim; to evaluate the ability of a sample of coupled models from the CMIP3 dataset to simulate the mean rainfall climatology of East Africa.

Rainfall over East Africa is dominated by convective events co-located with the movement of the ITCZ in conjunction with solar insolation throughout the year. The two datasets studied in this chapter had different peak rainfall values for the short rains; 5-9mm/day in the reanalysis NCEP series and 11-12mm/day in the statistically obtained H08 series. These disparities highlight the difficulties in capturing true rainfall values in a region where observations are sparse and gauge series inadequate for representing even the annual cycle of rainfall (Sylla et al, 2013).

Nonetheless, the NCEP and H08 series provided a basis for examining the representation of the annual cycle of rainfall over East Africa in a subset of CMIP3 models to answer the sub question of research question two. The annual cycle was well represented in two of the models; MPI and MRI. These two models had realistic monthly rainfall climatology in terms of temporally and latitudinally mean daily rainfall amounts in the range seen in the NCEP and H08 datasets.

In relation to the models which show weaknesses in their replication of the mean East African rainfall, all of the AOGCMs replicated the ITCZ related movement of the rainfall band of the annual cycle throughout the year, as demonstrated through comparison of mean daily rainfall at each latitude throughout the year. However, in five of the eight AOGCMs the rainfall band moved too far north over East Africa during the boreal summer months of June to August, as seen in CCCMA, CNRM and MIUB to the greatest extent, with ITCZ related rainfall moving to a peak latitude of 16°N in comparison to the peak of 13°N seen in the observational datasets. The rainfall band was too latitudinally widespread (specifically in CCCMA and CNRM) and simulated peak rainfall which was muted in intensity in comparison to observed datasets but contributed to an overestimated annual rainfall total. GISS had daily mean rainfall of more than 8mm throughout the year at all latitudes from 20°S to 15°N, with no spatial variation or annual cycle visible, making it a poor simulation of the annual cycle. To summarise the conclusion to the sub question, MRI and MPI have the most realistic simulations of the annual rainfall cycle with GISS highly unrealistic in the rainfall field. Two weaknesses were found in the other models to some extent; the band of ITCZ related rainfall moving too far north and the region of rainfall being too widespread during the year.

For projections of future climate under the SRES emissions scenarios, the short rains of October to December have a coherent wetting projection under a climate change scenario of increasing emissions. Results for other seasons are disparate, with no clear model agreement as to the sign of potential changes to seasonal rainfall totals over East Africa. These results are considered in future chapters in the addressing of the thesis aims. Despite the coherent wetting sign for model rainfall projections over East Africa for the short rains, it is clear that difficulties remain in both the representation of the rainfall climatology in available datasets and the simulation of these features in climate models. From this point

onwards in this thesis, the focus is on an innovative approach in order to redress the previous neglect of the daily scale of rainfall. Daily rainfall regimes over East Africa during October to December are established in reanalysis and satellite datasets and the replication of these daily regimes are evaluated in climate models. Finally, projections from the perspective of daily rainfall regimes are examined, in order to achieve the overall aim of this thesis of assessing the reliability of model rainfall projections for East Africa.

Chapter 4: Contemporary characteristics of daily rainfall regimes during the short rains in the ERA-40 dataset

This chapter addresses the first research question of this thesis; to establish the contemporary structure of daily rainfall regimes of the short rains over East Africa from 1971-2000. The first focus is on the sub question of this first research question; is the self-organising map an appropriate tool for this investigation? To reiterate, the OND short rains have been chosen for investigation because this rainfall season, and associated teleconnections, are more comprehensively understood than the March to May long rains. Seasonally, they are more predictable (Hession and Moore, 2011) and have greater spatial coherence than the long rains (Manatsa et al, 2012). For these reasons, the short rains represent a better constrained test of AOGCM rainfall production. To establish and evaluate the structure of daily rainfall regimes a method which can classify these daily regimes objectively and in a reproducible way is required. The self-organising map has the potential to be an effective tool for this purpose, but before the structure of the daily rainfall field within the short rains can be established it is necessary to determine if the self-organising map is an appropriate tool for this investigation, hence the sub question to the first research aim. The SOM has not previously been applied to the convection dominated climate of East Africa, but has been selected as potentially useful to facilitate the analysis of daily events over East Africa (see Chapter 2, Section 2.9).

A series of stages are undertaken to address the sub aim of the first research question. First, the self-organising map is evaluated for its ability to cluster daily circulation parameters. The output here consists of both the generalised classification in the form of

maps and the days assigned to each of the maps. The reason that the classification does not begin with rainfall, despite daily rainfall events being the focus of this study, is that within the convective environment of the tropics, rainfall may be expected to be difficult to classify given that there is normally a strong stochastic element to its spatial distribution. Moreover, reanalysis rainfall tends to be less coherent than reanalysis circulation parameters. Accordingly, it is deemed safer to begin with classifying circulation variables as a precursor to deciding whether to use the SOM methodology to classify rainfall events. As a cross-check, the ERA-40 derived classification scheme is then imposed on the partially independent NCEP reanalysis data set. This step is undertaken to check whether the classification scheme is robust through testing both whether the two reanalysis products resolve the same circulation regimes and the fidelity of the classification scheme itself. It is therefore a stringent test. The next step is to classify daily rainfall itself and then to consider the circulation parameters which accompany the rainfall classification. The latter step goes on to form an important basis for the evaluation of coupled climate models in Chapter 6.

Within the classification of rainfall itself, the second stage of addressing the sub aim of the first research question, three configurations of the ERA-40 rainfall dataset are examined; an anomaly rainfall series (daily total minus daily mean), a normalised series (anomalies divided by the standard deviation) and a normalised series constrained over a smaller, land-based domain. The purpose of these configurations is to establish the most appropriate way to harness the occurrence of daily rainfall events within the four East African focus countries of Ethiopia, Kenya, Tanzania and Uganda. The first configuration does not cluster daily rainfall events over the four focus countries, with convective events over the Congo and Indian Ocean dominating. It is not until a normalised series is

constrained over a smaller domain that identifiable rainfall events over the focus countries of Tanzania, Kenya, Uganda and Ethiopia are classified using the SOM.

Following this, the normalised small domain SOM is evaluated in more detail, looking for physical reasons for the daily rainfall events by considering the structure of moisture flux, moisture flux divergence, zonal winds, geopotential height and specific humidity associated with the daily classification. Is there a reason for the rainfall being clustered spatially that could be explained in circulation data (moisture flux, moisture flux convergence, zonal winds, geopotential height and specific humidity)? A variable that shows physically consistent results with the presence of clustered rainfall events could prove useful as an evaluation tool for the parameterised rainfall events in AOGCMs as evaluated using the SOM in Chapter 6.

Following assessment of the SOM as a methodological tool and the establishment of the answer to the sub question of the first research question, the most and least closely matched dates to key nodes (by quantisation error) are examined. This is to ensure that the accuracy of the classification is not being adversely affected by the inclusion of days assigned to individual nodes that do not match the rainfall state of that node. The build-up of circulation patterns to rainfall events is investigated to see if rainfall events over East Africa occur in a stochastic manner or if there are indications in conditions over the domain on preceding days of certain rainfall events occurring.

4.1 Evaluation of the SOM classification method: A simple test case

Self-organising maps are used on several variables from the ERA-40 dataset in a simple test case to establish the capabilities of SOMs on daily circulation regimes. This work is carried out on a number of variables from the ERA-40 dataset, excluding rainfall. This is the first stage in testing the methodology for its appropriateness to the East African domain

as well as its usefulness in establishing contemporary conditions of the daily rainfall regimes of the short rains. Part of the test is to determine whether the classification system is objective and reproducible. The variables and levels for this simple test case are zonal and meridional winds, specific and relative humidity, geopotential height and moisture flux at the surface and 850hPa levels and temperature and sea level pressure at the surface level. The detailed stages of the SOM methodology were explained in Chapter 2, Section 2.10.3. In summary, these stages are:

1. Formatting the input data, an ASCII file with one line for each input (i.e. each day for daily data), with a header with information on the number of grid points for each input.
2. Initialisation of each node of the SOM array on a random basis, specifying the topology (hexagonal) and neighbourhood type (bubble).
3. Using the Euclidean distance to find the ‘winning’ node in the array, the one that is closest to the daily input.
4. The winning node is adjusted in the direction of the input data, and there is also modification of the surrounding nodes to a lesser degree, according to both a learning rate that decreases through iterations and a radius of surrounding nodes that also decreases through iterations.
5. Repetition of stages three and four through all days of the input data and through multiple iterations to converge on the most tightly classified realisation of the SOM, evaluated through the quantisation error (QE). Repeat the process multiple times to achieve the SOM for that input data with the lowest possible QE.

6. The matching of each input date with an individual SOM node and visualisation of the SOM.

4.1.1 Absolute data

The SOMs are created on the absolute input data, defined as the output from the reanalysis data interpolated at a resolution of 1° by 1° over the domain 20°E to 60°E , 15°S to 20°N , following the approach which was outlined in Chapter 2 as an initial application of the method over East Africa. With different units and scales for each variable, the weighting given to each variable is potentially dependent on the units rather than the contribution that the particular variable would make to the classification itself. Since the overriding aim is to evaluate climate model rainfall projections over East Africa and it is also important to be able to evaluate the circulation systems that relate to rainfall events, the effect of the variable units needed to be assessed. This is achieved by computing the classification on both absolute and normalised data. The first stage evaluation of the method for this region is carried out on the absolute input data so that the patterns and magnitude of patterns can be easily compared from one variable to the next and between the ERA-40 and NCEP datasets.

4.1.2 Evaluation of the classification scheme

The self-organising map methodology has never previously been applied to the East African region. Most previous applications of this method to climate science have been in regions where mid-latitude synoptic scale dynamics play an important role, for example work by Hewitson and Crane over South Africa (2006). In comparison, over East Africa, convection is a defining feature of the climate. The location astride the equator means that synoptic scale systems of midlatitude origin cannot play a dominant role, if any, in the organisation and production of rainfall. Scale analysis, as illustrated by Webster (1983),

explains why the synoptic scale is not prevalent in the tropics (rotational effects are minimal in the tropics). It follows that there has been considerable uncertainty over whether the method can be successfully applied to this region, necessitating the sub aim of the first research question to establish if the SOM is a suitable tool to facilitate analysis of daily rainfall events over East Africa.

The first stage in the evaluation process of the self-organising map is a visual inspection to determine if the SOMs are identifying coherent and distinct daily events in circulation variables. A 7 by 5 array (that is, 35 individual nodes or types in the classification scheme) has been chosen as giving an appropriate level of detail for each SOM. This is the same array size that is employed for all SOMs in Chapter 4. For each SOM and each of the 35 nodes within it, six dates are randomly selected that have been matched to that node. The data for each of these days is plotted using the partly independent NCEP reanalysis data. The use of two different datasets, which are expected to represent the East African climate to slightly varying degrees for the different variables (because of the different assimilation schemes and model physics in the two datasets- see Chapter 2, Section 2.1 and 2.2), provides an additional evaluation method. Each of these plots is qualitatively compared with the visualised ERA-40 node for similarities. This is to evaluate the SOM in two ways. First, to see if the ERA-40 self-organising map has clustered into states occurring in more than one dataset for each variable over the East African region and second, to examine the extent of the variability between days belonging to the same node and the similarity of each day to the node to which it was being compared.

To reiterate, the SOMs are produced using the ERA-40 reanalysis dataset and the plots for each randomly selected individual day are produced using the NCEP reanalysis dataset. If states emerge from the ERA-40 SOM which also appear in the NCEP data and are correctly matched to the equivalent SOM node it increases confidence in the suitability of

the SOM methodology to be applied to East Africa and facilitates further, more rigorous evaluation of the method.

Tests are done at the surface and 850hPa levels. The 850hPa level has also been used in many studies over East Africa (for example Mpeta and Jury (2001), Mapande and Reason (2005), Camberlin and Planchon (1997), Kijazi and Reason (2005), Camberlin et al (2010) and Philippon et al (2002)). Success is defined based on visual criteria. If the NCEP plot of the selected day shows the strongest qualitative resemblance to the node it has been assigned to it is deemed ‘successful’. If similarities between the plots and the node are not apparent, that assignment is considered ‘unsuccessful’. One example is shown here, that of the SOM for the zonal component of surface level winds and the six dates selected for two of the nodes; 0 0 (top left) and 6 4 (bottom right).

These two nodes are selected for this evaluation due to their location at opposite corners of the SOM, with nodes located far apart in the array representing the greatest difference within the data space. The SOM is shown in Figure 4.1 and the six dates for the selected nodes are shown in Figure 4.2

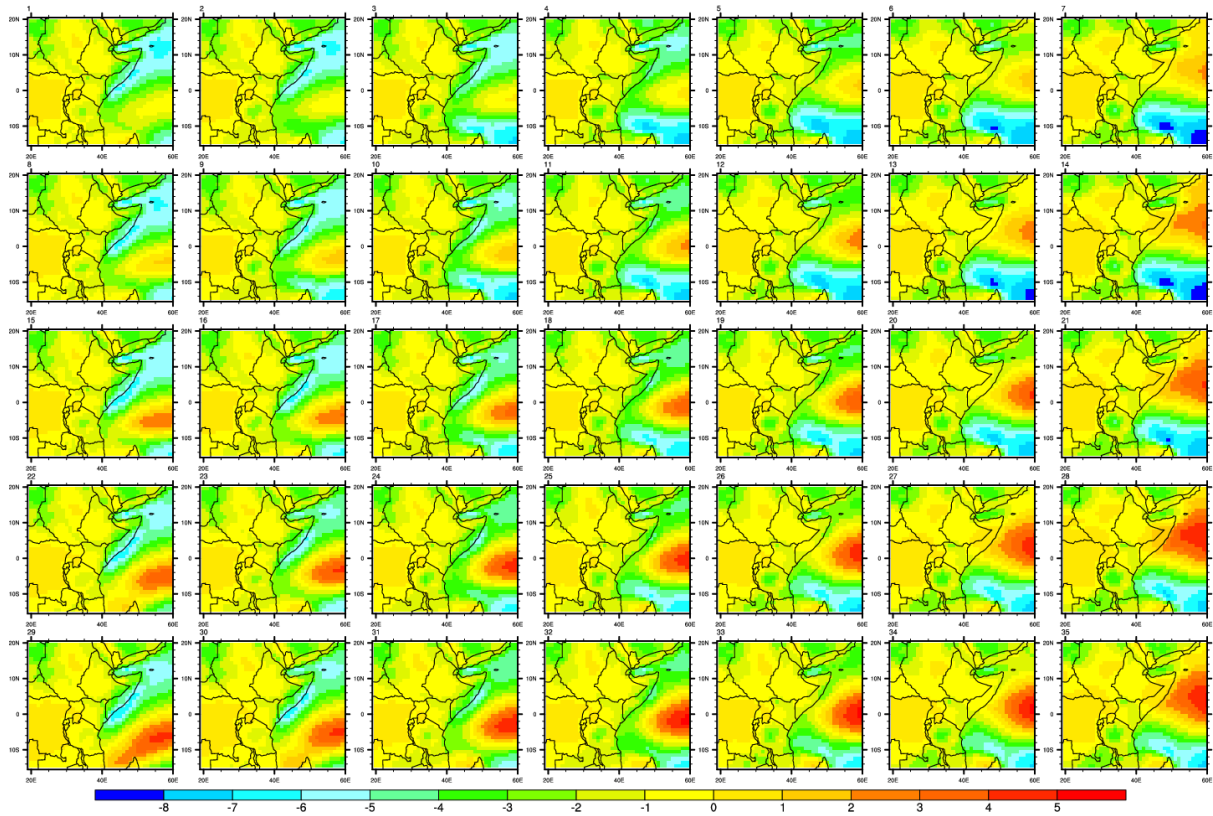


Figure 4.1 Scaled SOM of raw zonal wind (m/s) data at the surface level, associated with reference vector 6

The SOM clusters together easterly and westerly wind centres, with the dominant pattern being centres of wind located offshore of East Africa. The top left part of the array clusters easterly wind centres to the north and south of the ocean region, with a westerly centre of varying strengths and spatial positions in between. The right part of the domain has two wind centres (one westerly, one easterly), with the westerly centre located northwards of the easterly centre.

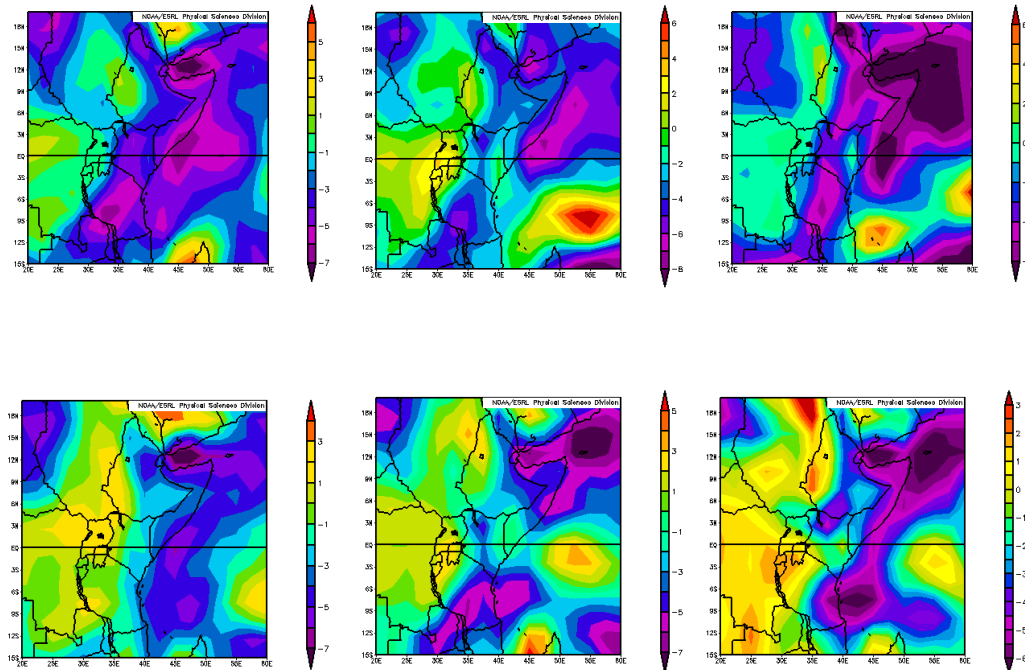
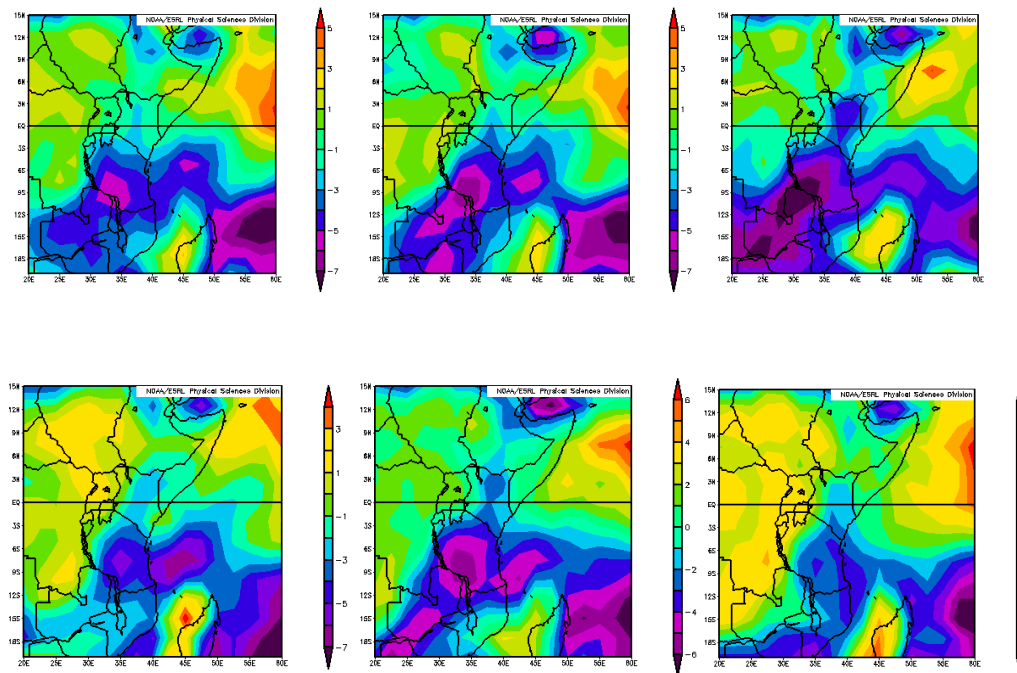


Figure 4.2 NCEP surface winds (m/s) associated with SOM node 0 0 (above) for the dates, in a clockwise direction from top left 10/11/1971, 19/12/1981, 21/12/1981, 18/11/1997, 30/11/1991 and 09/12/1984 and 6 4 (below) for the dates also in a clockwise direction from the top left 10/10/1974, 11/10/1974, 07/10/1979, 12/10/1992, 11/10/1992 and 13/10/1983.



For the two nodes examined here, 0 0 and 6 4 (see Chapter 2, Section 2.10.3 for the example SOM numbering), the pattern of easterly and westerly winds and the strength of the flow matches well with the NCEP plots. In all of the selected 0 0 node plots the key feature of the easterly centre located south east of the Arabian Peninsula and Somalia is well replicated, with maximum flow speeds of 7ms^{-1} . Node 6 4 picks up the contrasting wind centres in the Indian Ocean, with westerly winds of up to 5ms^{-1} off the coast of Somalia and easterly winds with speeds up to 7ms^{-1} further south. Here, the dominant pattern is the easterly centre and it is these winds which are emphasised in the NCEP plots. This relates well to the SOM; whilst this SOM visualisation has been scaled to cover the full range of input wind speeds, the automated self-organising map visualisation strongly emphasises the easterly winds with the scale bar covering easterly patterns in strong detail with westerly winds compromised in their colour representation.

This comparison method confirms that the SOM methodology is an appropriate base from which to proceed with the examination of rainfall events and associated circulation parameters over East Africa. SOMs created in other variables (meridional winds, geopotential height, sea level pressure and surface temperature), whilst not shown here, are also successfully evaluated using this same methodology. That is, NCEP plots of individual days selected as a match to each SOM node show the same patterns in terms of amplitude and spatial patterns as the node to which they are matched. This goes some way towards answering the sub question of the first research question, allowing more in-depth analysis to establish the best way to engage the SOM to establish contemporary conditions of the short rains in terms of daily rainfall events and their associated circulation patterns.

4.2 Application of the SOM classification to three realisations of the ERA-40 rainfall series

SOMS are capable of classifying circulation data over the East African domain and the classification scheme is reproducible in an independent data set. As the next step, SOMs are used to classify daily rainfall events (the focus of this study) in the same way. If successful, this will facilitate achieving the aims of this thesis through a methodology that allows daily rainfall events to be classified in an objective and reproducible way; the results of which can then be applied to the evaluation of AOGCM daily rainfall under contemporary conditions in Chapter 6 and under the high emissions SRES A2 scenario at the end of the twenty-first century from 2090-2099.

SOMs are applied to a variety of daily rainfall datasets. The SOM that best manages to harness the daily rainfall events over East Africa is then analysed to see which circulation patterns are responsible for creating different rainfall events over the East African domain. For each individual node, the days associated with it are isolated in the remaining variables, with the mean state calculated and plotted for each variable. For example, all the days in the OND 1971-2000 series associated with the node 0 0 are extracted from the data series for the variable being examined. The mean state associated with that node is then plotted. This is repeated for each of the thirty-five nodes, so as to establish the circulation signature associated with the rainfall SOM.

Before the investigation into the variables representing rainfall events during the East African short rains, three realisations of the ERA-40 rainfall SOM are evaluated. Given the lack of previous studies employing the self-organising map as an analysis tool over East Africa, this step is in place to ensure that the methodology capable of capturing the

contemporary daily characteristics of the short rains in the most appropriate way is selected.

4.2.1 ERA-40 rainfall anomaly SOM

First, SOMs are computed on the rainfall anomaly data. Here anomalies are the departure of each daily rainfall value for a grid box from its long term mean. Aside from creating the anomaly series (series minus mean), the only pre-processing before employing the SOM methodology is to interpolate the series to a 1° by 1° grid, from the original Gaussian N80 grid, and to select the time period and domain (20°E to 60°E, 15°S to 20°N). This leads to an input series of 2760 dates (92 days each year for a thirty year period) and a grid of 1435 values (1 value for each grid point in the gridded East African domain). The standard SOM methodology as described in Chapter 2 is then applied. The self-organising map with the lowest quantisation error is that associated with reference vector 5 (out of the fifteen realisations calculated) and this SOM is visualised. Each input date from the rainfall anomaly series is matched to an individual SOM node. The distribution of the occurrence of each node can be seen in Table 4.1.

Once each date is matched to a node for the control period data, it is then possible to look at other variables in relation to the reference vector 5 self-organising map. For this ERA-40 rainfall anomaly SOM the other variables examined are moisture flux, moisture flux divergence, specific humidity, zonal winds and geopotential height.

Table 4.1 Distribution of occurrence of each SOM node for the ERA-40 rainfall anomaly SOM, reference vector 5

Node No.	Node Ref	Total Mapped	% Total	No. Oct	% Oct	No. Nov	% Nov	No. Dec	% Dec
1	0 0	106	3.84	28	3.01	48	5.33	30	3.23
2	1 0	62	2.25	7	0.75	27	3.00	28	3.01
3	2 0	75	2.72	1	0.11	19	2.11	55	5.91
4	3 0	114	4.13	0	0.00	14	1.56	100	10.75
5	4 0	95	3.44	0	0.00	21	2.33	74	7.96
6	5 0	86	3.12	8	0.86	40	4.44	38	4.09
7	6 0	83	3.01	16	1.72	39	4.33	28	3.01
8	0 1	49	1.78	5	0.54	26	2.89	18	1.94
9	1 1	73	2.64	0	0.00	30	3.33	43	4.62
10	2 1	67	2.43	0	0.00	16	1.78	51	5.48
11	3 1	60	2.17	4	0.43	32	3.56	24	2.58
12	4 1	48	1.74	21	2.26	22	2.44	5	0.54
13	5 1	37	1.34	18	1.94	15	1.67	4	0.43
14	6 1	76	2.75	23	2.47	32	3.56	21	2.26
15	0 2	81	2.93	2	0.22	9	1.00	70	7.53
16	1 2	62	2.25	7	0.75	13	1.44	42	4.52
17	2 2	89	3.22	2	0.22	41	4.56	46	4.95
18	3 2	97	3.51	9	0.97	65	7.22	23	2.47
19	4 2	91	3.30	47	5.05	44	4.89	2	0.22
20	5 2	80	2.90	45	4.84	30	3.33	3	0.32
21	6 2	82	2.97	40	4.30	26	2.89	16	1.72
22	0 3	38	1.38	4	0.43	6	0.67	28	3.01
23	1 3	61	2.21	8	0.86	25	2.78	28	3.01
24	2 3	75	2.72	18	1.94	49	5.44	8	0.86
25	3 3	111	4.02	70	7.53	39	4.33	2	0.22
26	4 3	54	1.96	50	5.38	4	0.44	0	0.00
27	5 3	37	1.34	26	2.80	11	1.22	0	0.00
28	6 3	87	3.15	38	4.09	38	4.22	11	1.18
29	0 4	108	3.91	7	0.75	12	1.33	89	9.57
30	1 4	74	2.68	23	2.47	20	2.22	31	3.33
31	2 4	113	4.09	73	7.85	32	3.56	8	0.86
32	3 4	88	3.19	76	8.17	11	1.22	1	0.11
33	4 4	104	3.77	97	10.43	7	0.78	0	0.00
34	5 4	100	3.62	91	9.78	8	0.89	1	0.11
35	6 4	97	3.51	66	7.10	29	3.22	2	0.22
Total	-	2760	100.00	930	100.00	900	100.00	930	100.00

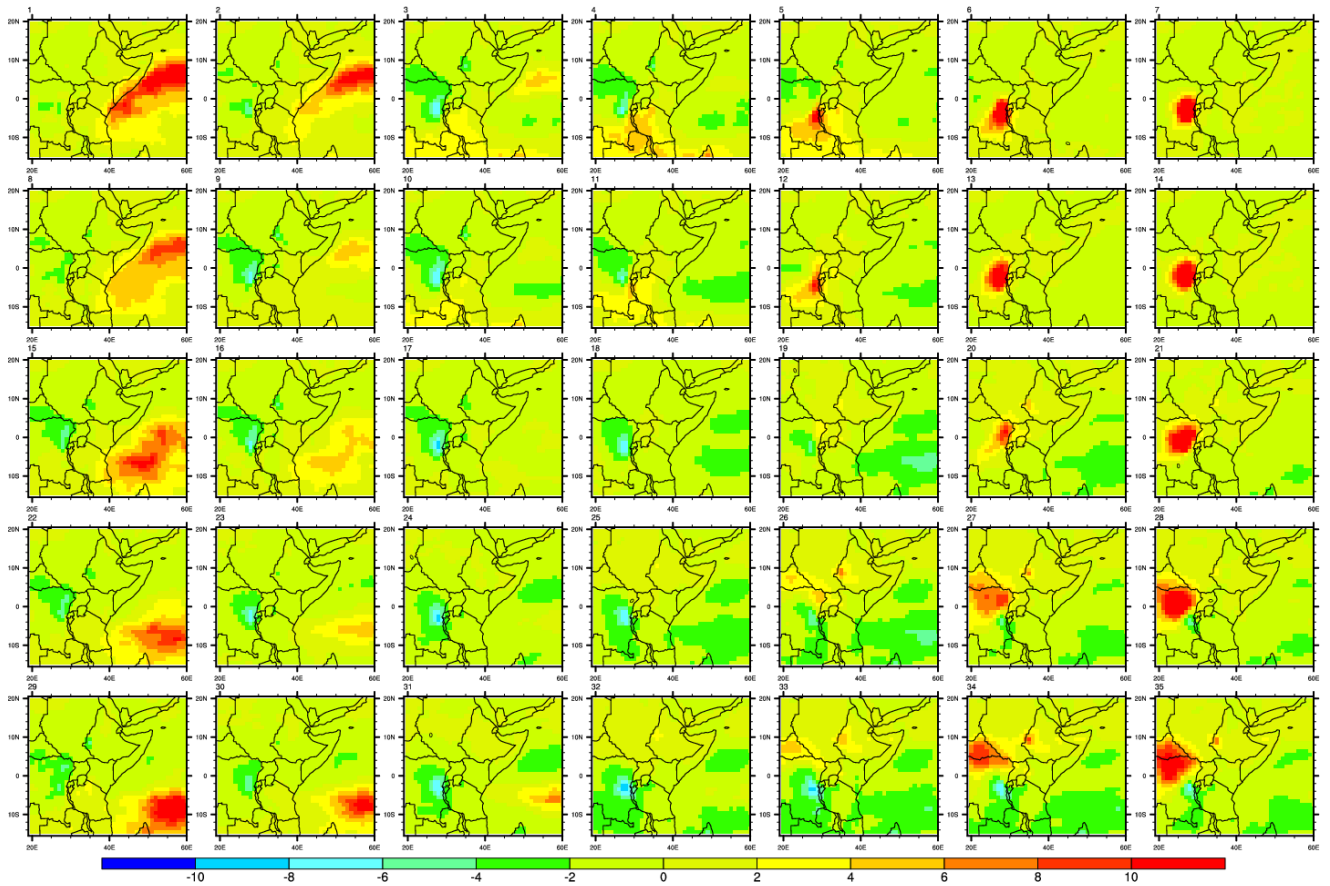


Figure 4.3 Optimal SOM for the ERA-40 rainfall anomaly (mm/day) series, reference vector 5, scaled.

Figure 4.3 shows the ERA-40 anomaly rainfall SOM with the lowest quantisation error, reference vector 5. Shown is the scaled version of the SOM which covers the full range of data entry. This SOM has classified the rainfall into nodes with differing rainfall events (both spatially and in terms of rainfall amounts), with nodes located furthest apart showing the biggest differences in state and nodes adjacent to each other being closer in the data space; it is important to note that the nodes are not evenly separated in phase-space. There are four interesting features of note in this SOM, which dominate the majority of the thirty-five nodes. Two of these features relate to daily rainfall events in the Indian Ocean and occupy the left hand side of the SOM array, the first is to the north of the domain with a centre of rainfall off Somalia with the second, a centre of oceanic rainfall located southwards off the coast of Tanzania to the north east of Madagascar. The third feature is intense rainfall over the western Congo, located on the right hand nodes of the 7 by 5

array. The fourth main feature of this SOM realisation is the suppression of convection over the land region in particular in the western Congo. These features, as seen in specific nodes, are described in more detail below.

Node 0 0 (top left) is associated with intense anomalous oceanic rainfall off the coast of Somalia extending southwards and slightly inland along the East African coast, of up to 10mm/day with slightly drier than normal conditions seen over Uganda (anomaly of up to 2mm/day). The node 0 4 (bottom left in the SOM array) is also characterised by anomalous oceanic rainfall events, this time centred off the coast of Tanzania with lesser anomalous rainfall events in northern Tanzania. Dry conditions are seen over the Congo. Node 6 0 (top right) classifies extreme heavy rainfall conditions over the Congo (more than 10mm/day), with dry conditions over much of the rest of the region. Node 6 4 (bottom right) has an anomalously wet centre to the west of the land domain on the border between the Congo and Central African Republic with dry conditions found further south in the Congo.

These nodes, whilst showing distinct patterns, are dominated by rainfall events in the Democratic Republic of the Congo and the Indian Ocean with patterns within the four East African focus countries (Tanzania, Kenya, Uganda and Ethiopia) muted, despite the use of rainfall anomalies. There are nodes that identify rainfall events over these countries. Node 3 0 (top row, middle node) identifies anomalous rainfall over western Tanzania, node 3 2 (central node of the array) has lighter rainfall events over Uganda, with dry conditions in the Congo and node 5 4 (bottom row, second from right) depicts anomalous rainfall events over the Ethiopian Highlands.

Circulation indicators at the surface and 850hPa level

Circulation parameters at the 850hPa level (moisture flux divergence, moisture flux, geopotential height, specific humidity and zonal winds) are assigned to the appropriate node for each date of the input data and the mean state and mean anomaly state for each node plotted. These are examined in turn to see if the patterns in each variable are physically consistent with the rainfall events seen in each node. The patterns are found to be physically consistent but these plots are not shown here. Despite variety of daily rainfall events evident over the focus countries, it is clear that when classifying the anomaly rainfall series, the SOM patterns are dominated by much larger convective events, namely Indian Ocean rainfall events and the heavy rainfall events in the Democratic Republic of the Congo where the convection is likely to be more spatially coherent and organised than over the drier region of East Africa. This, however, means that the detail in rainfall events over the area of interest is lost; individual rainfall events over Kenya, Tanzania, Ethiopia and Uganda are the focus of the study, not the large scale convection that is currently dominating each of the thirty-five nodes.

The lack of meaningful information on daily rainfall events in this SOM over the area of interest means that more investigation is needed into a suitable way to employ SOMs using daily rainfall data over the region to harness the detail of daily rainfall regimes over East Africa during the short rains.

4.2.2 ERA-40 rainfall normalised SOM

Given the dominance of regions of large-scale convection over the rainfall anomaly series classification scheme, the ERA-40 rainfall series is normalised to the climatology over the domain to attempt to reduce the influence of larger scale convective rainfall events over the Congo and the Indian Ocean that appear in the SOM calculated from the anomaly

series. The resulting SOM (Figure 4.4) is an improvement on the domination of large scale rainfall over the Congo with increased variation in the continental rainfall centres and a much reduced emphasis on the heavy rainfall located to the west of Burundi and Rwanda that dominated the anomaly series SOM.

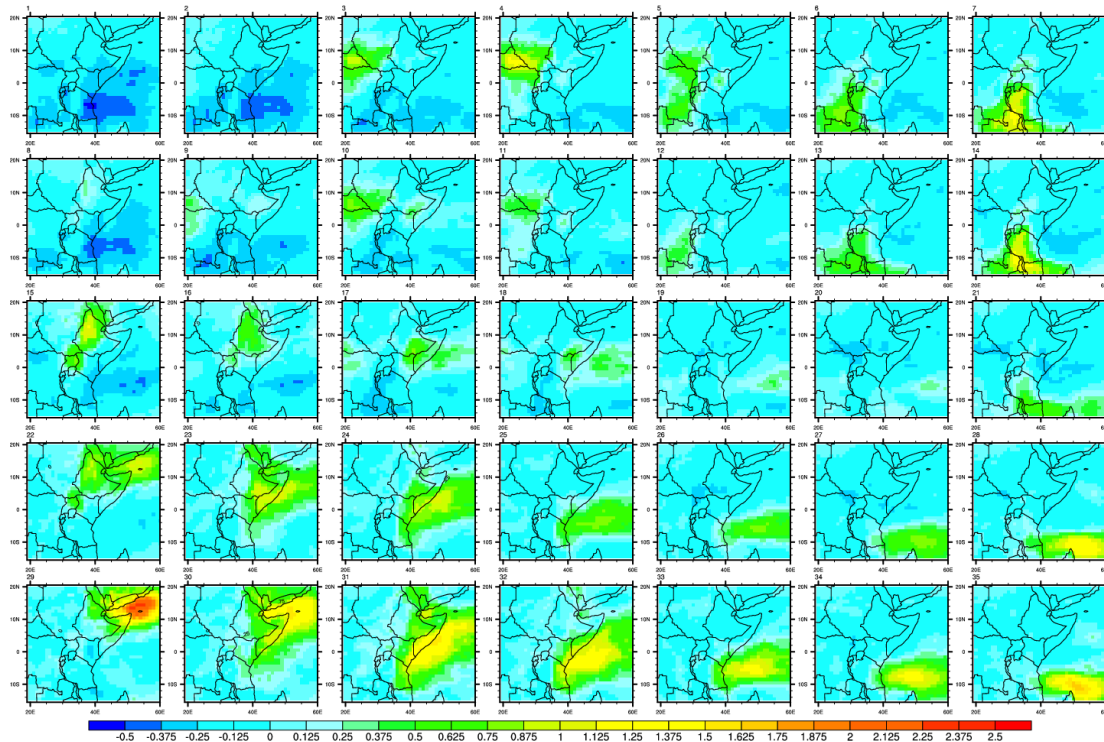


Figure 4.4 Reference vector 7 SOM for ERA-40 normalised rainfall, scaled to cover the whole data range

However, oceanic rainfall events become even more widely represented, featuring in nineteen of the thirty-five SOM nodes. This can be seen in Figure 4.4, the realisation of reference vector 7 scaled to cover the full range of input rainfall data. Suppressed convection over the Indian Ocean to the top left of the array features heavily. Distinct patterns over the four focus countries are still not clear, with the SOM methodology not achieving the aim of being a useful classification method for rainfall events in Kenya, Tanzania, Ethiopia and Uganda in this format. It has been decided to constrain the ERA-40 rainfall just to the region of Kenya, Tanzania, Ethiopia and Uganda to see if further improvement can be made to the SOM classification method.

4.2.3 ERA-40 rainfall SOM normalised reduced domain

Whilst the normalised version of the ERA-40 rainfall SOM is an improvement on the anomaly series SOM, there is still a tendency for oceanic rainfall to dominate many of the nodes. Therefore, it has been decided create a SOM over a smaller domain, with the aim of highlighting detail in the daily rainfall events over the focus countries. The resulting grid covers just the focus countries of Uganda, Kenya, Tanzania and southern Ethiopia; a grid size of 308 values (30°E to 43°E, 12°S to 10°N). The same methodology is followed as previously, using the self-organising map with the lowest quantisation error, in this case associated with reference vector 15. This map is then visualised and each input date matched to an individual node with the distribution of the occurrence of nodes shown in Table 4.2. Figure 4.5 shows the SOM for the ERA-40 normalised daily rainfall over the small East African domain focusing on Ethiopia, Kenya, Tanzania and Uganda.

Table 4.2 Distribution of SOM nodes for ERA-40 normalised data, small domain

Node No.	Node Coords	Total Mapped	% Total	No. Oct	% Oct	No. Nov	% Nov	No. Dec	% Dec
1	0 0	77	2.79	22	2.37	20	2.22	35	3.76
2	1 0	50	1.81	21	2.26	18	2.00	11	1.18
3	2 0	73	2.64	24	2.58	22	2.44	27	2.90
4	3 0	67	2.43	15	1.61	28	3.11	24	2.58
5	4 0	72	2.61	33	3.55	17	1.89	22	2.37
6	5 0	64	2.32	12	1.29	26	2.89	26	2.80
7	6 0	39	1.41	13	1.40	17	1.89	9	0.97
8	0 1	42	1.52	9	0.97	19	2.11	14	1.51
9	1 1	76	2.75	41	4.41	22	2.44	13	1.40
10	2 1	74	2.68	44	4.73	18	2.00	12	1.29
11	3 1	76	2.75	22	2.37	27	3.00	27	2.90
12	4 1	52	1.88	3	0.32	14	1.56	35	3.76
13	5 1	31	1.12	2	0.22	15	1.67	14	1.51
14	6 1	66	2.39	34	3.66	21	2.33	11	1.18
15	0 2	91	3.30	25	2.69	27	3.00	39	4.19
16	1 2	58	2.10	28	3.01	24	2.67	6	0.65
17	2 2	95	3.44	43	4.62	37	4.11	15	1.61
18	3 2	115	4.17	50	5.38	28	3.11	37	3.98
19	4 2	70	2.54	7	0.75	25	2.78	38	4.09
20	5 2	82	2.97	15	1.61	29	3.22	38	4.09
21	6 2	71	2.57	22	2.37	29	3.22	20	2.15
22	0 3	70	2.54	24	2.58	23	2.56	23	2.47
23	1 3	75	2.72	25	2.69	30	3.33	20	2.15
24	2 3	65	2.36	23	2.47	22	2.44	20	2.15
25	3 3	110	3.99	67	7.20	17	1.89	26	2.80
26	4 3	72	2.61	26	2.80	17	1.89	29	3.12
27	5 3	63	2.28	14	1.51	29	3.22	20	2.15
28	6 3	81	2.93	27	2.90	26	2.89	28	3.01
29	0 4	114	4.13	43	4.62	31	3.44	40	4.30
30	1 4	103	3.73	35	3.76	32	3.56	36	3.87
31	2 4	122	4.42	28	3.01	41	4.56	53	5.70
32	3 4	167	6.05	43	4.62	55	6.11	69	7.42
33	4 4	87	3.15	20	2.15	32	3.56	35	3.76
34	5 4	111	4.02	39	4.19	32	3.56	40	4.30
35	6 4	79	2.86	31	3.33	30	3.33	18	1.94
Total	-	2760	100.00	930	100.00	900	100.00	930	100.00

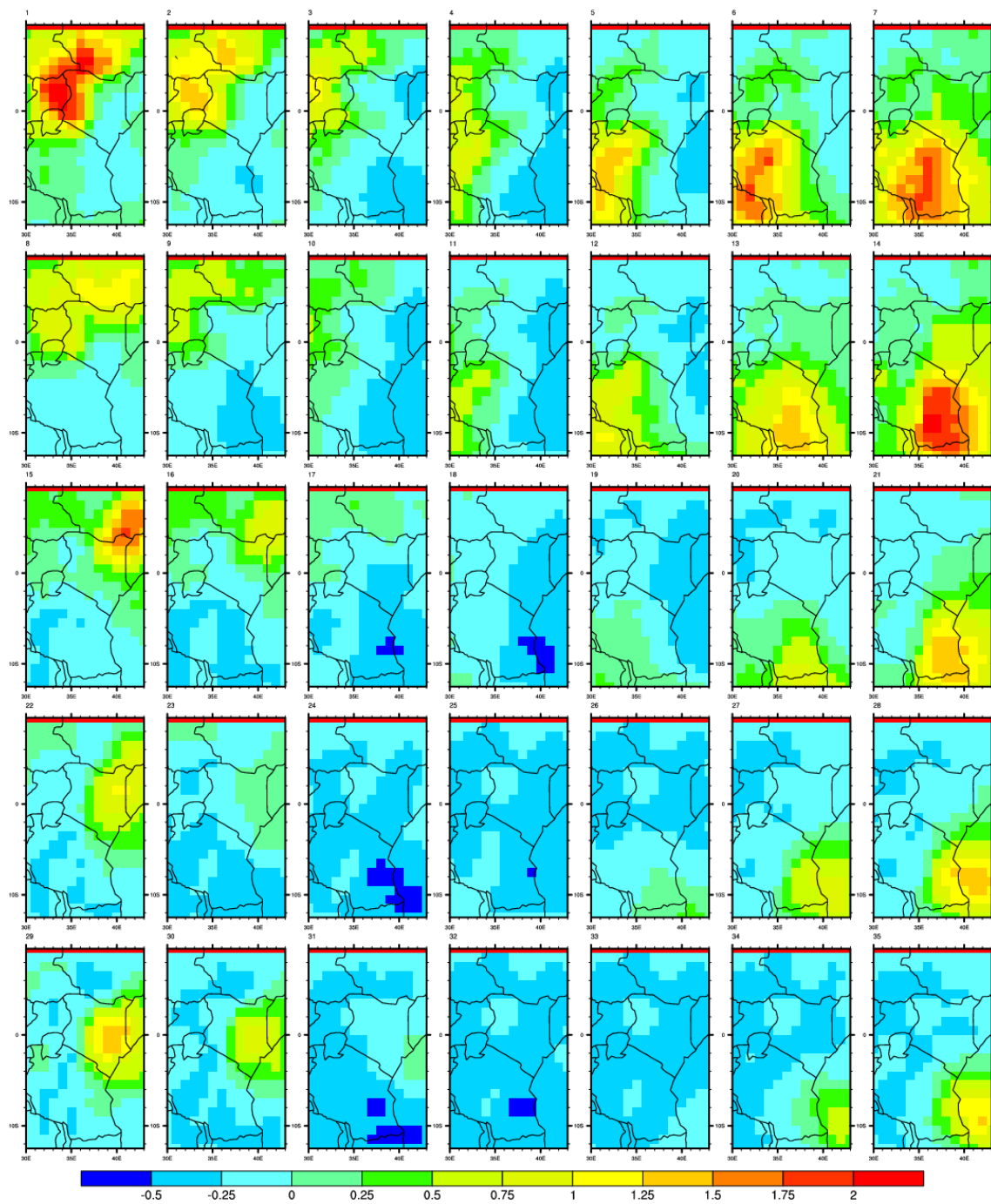


Figure 4.5 Reference vector 15 SOM associated with ERA-40 normalised daily rainfall, domain constrained to focus countries of Kenya, Tanzania, Uganda and Ethiopia

This SOM is qualitatively more successful than the previous ERA-40 SOMs in harnessing rainfall events over the focus countries of the domain, with distinct rainfall centres classified over Tanzania, Kenya, Uganda and Ethiopia as well as oceanic rainfall events and dry nodes showing the suppression of convection over the land region.

Over Tanzania, the SOM clusters four rainfall patterns. Nodes occur with rainfall events covering the west and east of the country, extensive rainfall and coastal rainfall events. Nodes covering Kenyan rainfall are less prevalent in this SOM being either completely dry or wet over the country with coastal rainfall centres occurring most frequently. Rainfall events centred elsewhere infringe on Kenya including those over the remaining three focus countries. Over Uganda and southern Ethiopia, rainfall of varying strengths is seen but there is rarely a rainfall centre clustered here when Tanzania has rainfall events.

Following the matching of each date of the input series to an individual node, investigation of the states in circulation data associated with each node is undertaken. The reason for inspection of the circulation data is to evaluate the extent to which identifiable states of atmospheric perturbations are associated with the rainfall events in the classification scheme. If spatially clustered daily rainfall events over East Africa can be linked to an independent circulatory parameter this could provide an extra layer of model evaluation. Rainfall is parameterised and thus a challenging variable to simulate in models. Coherence between rainfall simulation and an additional variable for the clustered rainfall states, both during the climatology period and under climate change conditions, could potentially reveal whether the rainfall is being simulated for the right reasons, allowing for the assessment of model reliability.

A suite of variables is examined for reasons of thoroughness, comprising moisture flux, moisture flux divergence, zonal winds, geopotential height and specific humidity fields at the 850hPa level.

After the automatic assignment of dates to each node has taken place based on the closeness of fit to the nodes (through quantisation error) the individual days for each node are combined to find the mean state associated with that node of rainfall. For example, for

the node 0 0, seventy-seven dates are assigned as being most closely matched to that node over the other thirty-four. These seventy-seven dates are isolated in the daily series, the mean calculated and this mean state plotted. This process is repeated for all nodes and all variables. In this section, associated circulation parameters are examined for seven individual nodes, selected as representative of each distinct rainfall ‘type’ being classified over this East African domain during the short rains.

Node 0 0 (top left node of Figure 4.5)

Node 0 0 is characterised by a bi-centred region of rainfall, with centres located over northern Uganda and to the north-east in southern Ethiopia. The region of rainfall expands into western Kenya and South Sudan. This node has seventy-seven days allocated to it over the control period, thirty-five of them occurring during the month of December with the remaining 42 split approximately evenly between October and November (22 and 20 occurrences respectively). For all plots discussed in this chapter moisture flux units are g kg m s^{-1} , moisture flux divergence are g kg m^{-1} , zonal winds in m/s, geopotential height is in m and specific humidity is in kg kg^{-1} .

Moisture flux

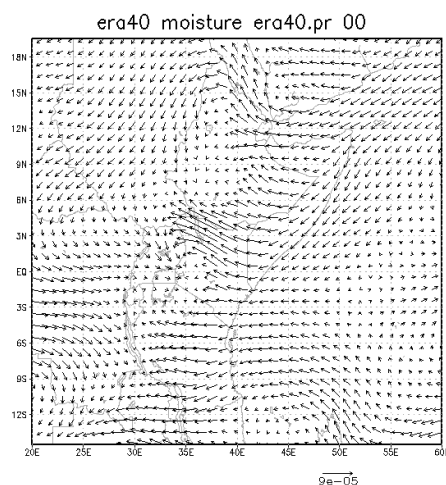


Figure 4.6 (a) Moisture flux mean for node 0 0

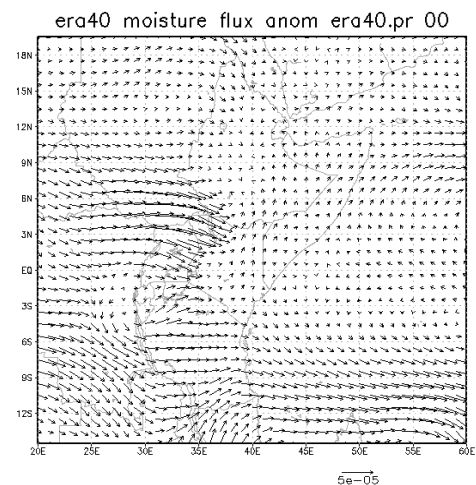


Figure 4.6 (b) Moisture flux anomaly for node 0 0

Figures 4.6 (a) and (b) show the moisture flux absolute and anomaly vectors respectively associated with the 0 0 node. Moisture flux associated with node 0 0 is characterised by strong onshore monsoon flow from the north east. This flow extends through the dry regions by the coast, converging with westerly moisture flow over Uganda. Convergence is increased by onshore flow at the south of the domain that joins the north easterly flow over the East African land domain. The moisture flux patterns are indicative of rainfall events over the Ugandan and Ethiopian borders and this pattern is reinforced by the anomaly moisture flux patterns. Figure 4.6 (b) illustrates strong westerly moisture flow anomalies originating from the Congo that causes convergence over the area of rainfall when combined with the normal onshore monsoon flow. The second feature of the anomaly moisture flux plot is the westerly flow in the south of the domain from 9°S southwards which travels offshore by Tanzania, preventing convergence in this region over the land domain.

Moisture flux divergence

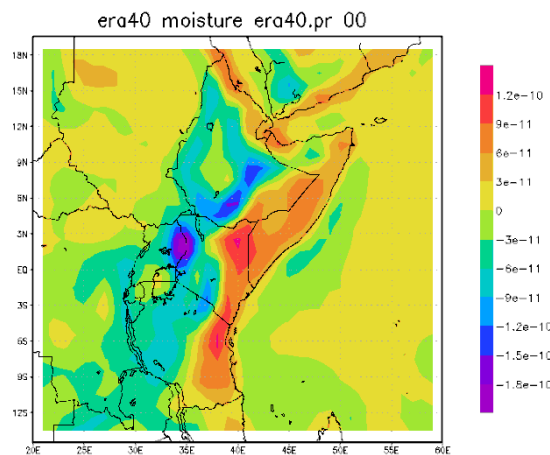


Figure 4.7 (a) Moisture flux divergence mean for node 0 0

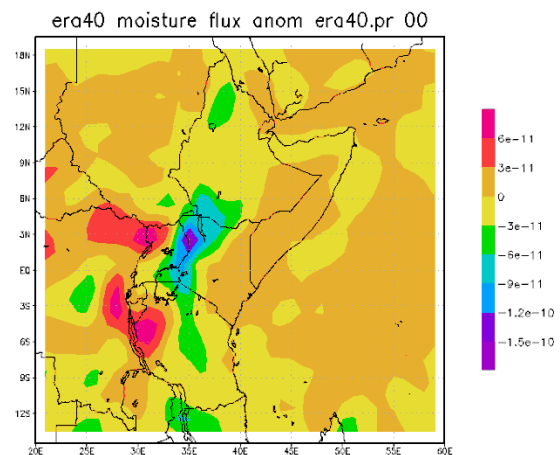


Figure 4.7 (b) Moisture flux divergence anomaly for node 0 0

Figures 4.7 (a) and (b) show the mean and mean anomaly 850hPa level moisture flux divergence over the domain associated with the 0 0 node. Strong convergence is co-located with the rainfall centre over Uganda, with a second maximum over southern Ethiopia. Lower levels of convergence continue down the domain through western Tanzania. Along the East African coast, where there is no rainfall associated with this node, divergence of moisture flux dominates, as expected. These features are physically consistent with the rainfall patterns associated with this node. The primary anomaly feature is convergence co-located with the rainfall centre with much of the rest of the domain close to climatology features.

Geopotential height

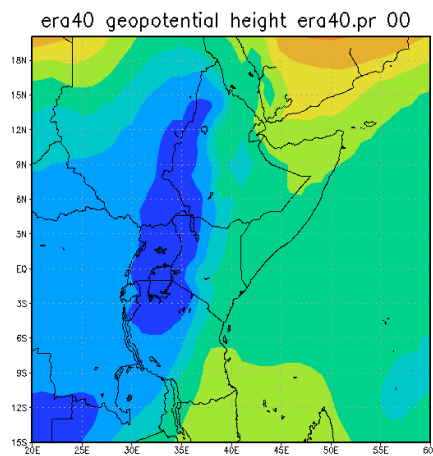


Figure 4.8 (a) Geopotential height mean for node 0 0

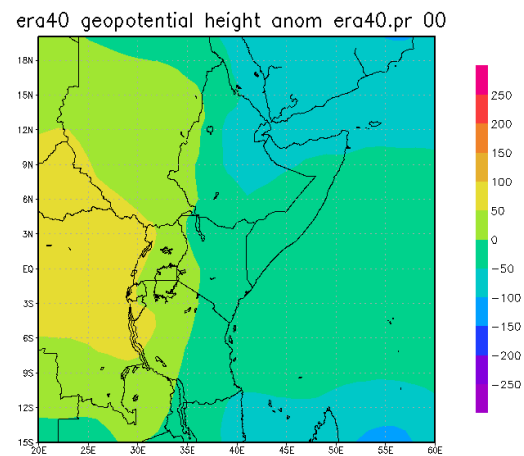


Figure 4.8 (b) Geopotential height anomaly for node 0 0

Figure 4.8 (a) and (b) show the mean and anomaly 850hPa geopotential height state for the 0 0 node. These patterns are less clear in explaining a rainfall centre over northern Uganda. As can be seen in the anomaly plot, geopotential height over the East African domain remains close to the mean, especially at the rainfall centre location. It is not surprising that

geopotential height is less successful than moisture flux in differentiating rainfall regimes in the tropics. The former is particularly useful in identifying midlatitude synoptic scale systems which do not occur over the domain in question.

Specific humidity

Figures 4.9 (a) and (b) shows the specific humidity and specific humidity anomaly mean states at the 850hPa level associated with the 0 0 rainfall node.

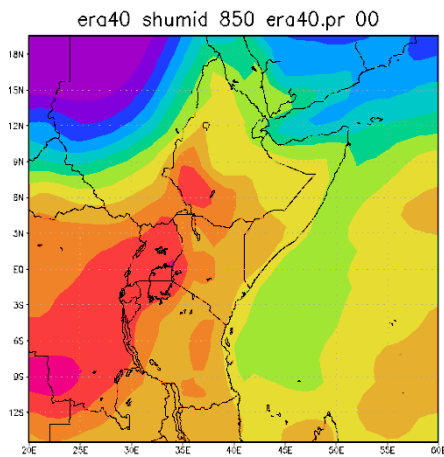


Figure 4.9 (a) Specific humidity mean for the node 0 0

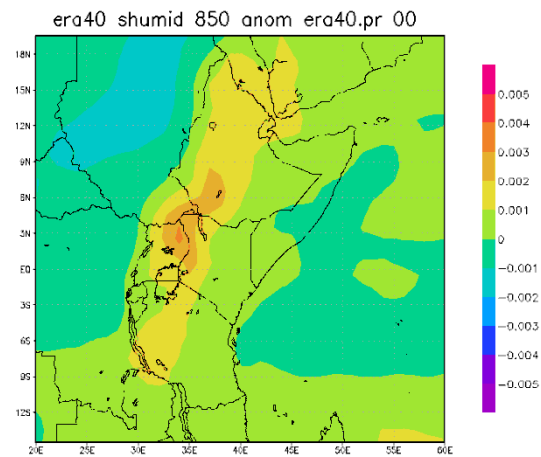


Figure 4.9 (b) Specific humidity anomaly for the node 0 0

A positive specific humidity anomaly is co-located with the region of rainfall. This is emphasised with the greatest anomaly over the strongest rainfall centre in northern Uganda, with the second greatest anomaly situated concurrent with the secondary rainfall maxima over southern Ethiopia. This anomaly additionally covers an area further south over Uganda and western Tanzania and is a pattern physically consistent with a rainfall centre over northern Uganda.

Zonal wind

Mean zonal winds at the 850hPa level and mean zonal wind anomalies at the same level associated with SOM node 0 0 are shown in Figures 4.10 (a) and (b).

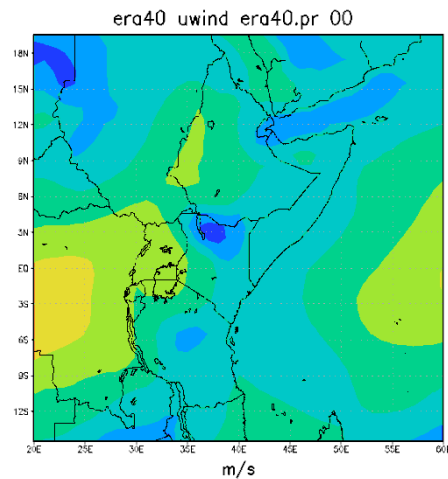


Figure 4.10 (a) Zonal wind mean for node 0 0

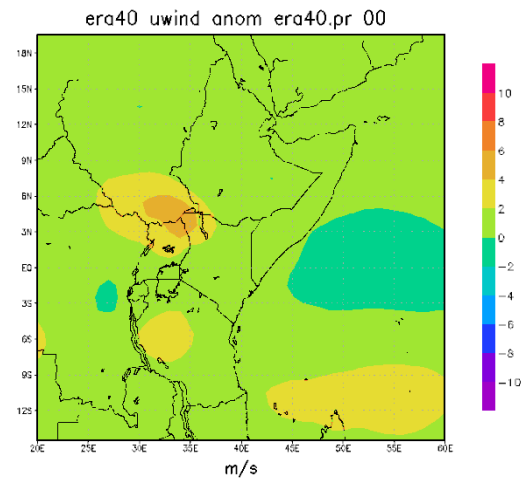


Figure 4.10 (b) Zonal wind anomaly for node 0 0

The strongest zonal wind anomaly is a westerly anomaly located on the borders of South Sudan, Uganda and Ethiopia. This is consistent with the anomaly westerly moisture flow and increased specific humidity over the same region. For node 0 0, with the exception of the geopotential height, the circulation indicators are consistent with one another and point to rainfall in the same location as the daily events characterising node 0 0.

This evaluation of circulation indicators is promising in its capacity to assess the rainfall SOMs; more nodes are now examined to see if this physical consistency is maintained through other key self-organising map nodes.

Node 0 2 (Left most column, third row down)

The rainfall centre for node 0 2 is near the East African coast, just to the north of the Kenyan and Ethiopian border. Daily rainfall events at this location are characterised by an

extension both south and north through Ethiopia and Kenya and additionally eastwards through Somalia and into the coastal Indian Ocean. Light rainfall also covers the land domain towards the west into South Sudan. Node 0 2 is assigned as the best matching node ninety-one times, thirty-nine of these dates occurred in December, with twenty-five and twenty-seven in October and November respectively.

Moisture flux

Figures 4.11 (a) and (b) show the mean and anomaly moisture flux vectors respectively that are associated with node 0 2.

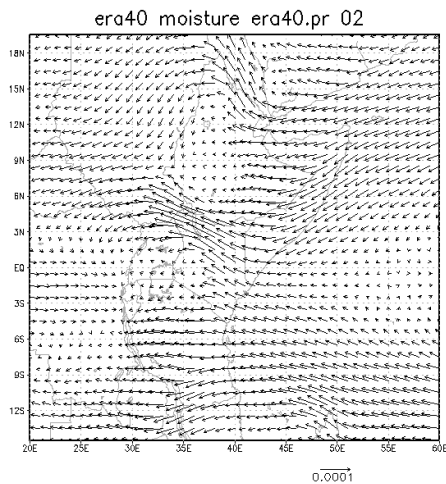


Figure 4.11 (a) Moisture flux mean for node 0 2

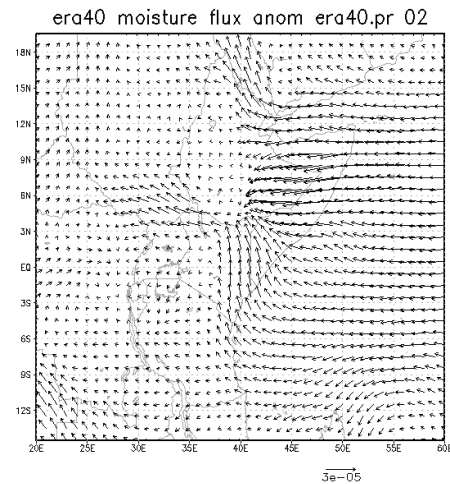


Figure 4.11 (b) Moisture flux anomaly for node 0 2

The strongest moisture flux anomaly is onshore flow into Ethiopia which originates from the Indian Ocean at all latitudes of the domain and converges over the rainfall centre. This pattern can also be seen in the raw moisture flux vectors, with flow originating from the north and south east of the domain which travels along the coastline and then inland adjacent to Ethiopia.

Moisture flux divergence

Figures 4.12 (a) and (b) show the mean and anomaly moisture flux divergence for all dates associated with the node 0 2. For dates associated with this node there is a large centre of convergence concentrated over south eastern Ethiopia, just to the north of the border with Kenya. This coincides with the rainfall centre. There is also weaker convergence over much of the northern part of the domain, including through South Sudan. However to the south of the domain, from approximately 3°N southwards, there is a mixture of convergent and divergent regions over the majority of the land area, with neither state dominating.

As with node 0 0, the overriding feature of the anomaly moisture flux is anomalous convergence consistent with the rainfall location with the rest of the domain close to the mean.

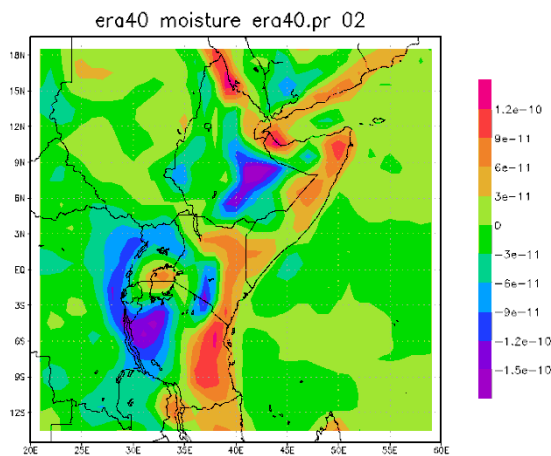


Figure 4.12 (a) Moisture flux divergence mean for node 0 2

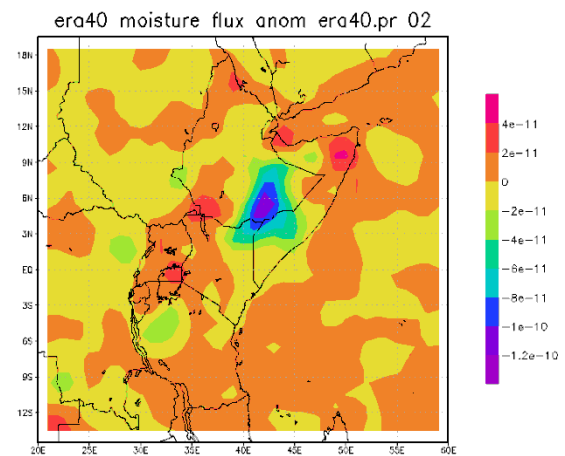


Figure 4.12 (b) Moisture flux divergence anomaly for node 0 2

Geopotential height

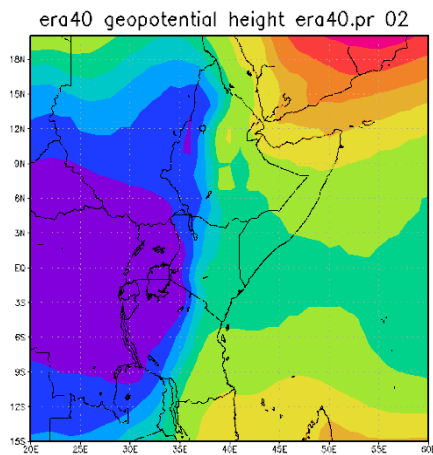


Figure 4.13 (a) Geopotential height mean for node 0 2

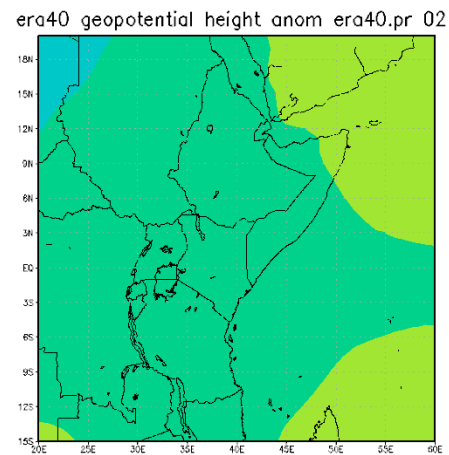


Figure 4.13 (b) Geopotential height anomaly for node 0 2

Figures 4.13 (a) and (b) show the 850hPa geopotential height and geopotential height anomalies linked to the 0 2 ERA-40 normalised rainfall SOM node. As with the first node examined, 0 0, geopotential height is inconclusive in providing physical evidence as to the cause of the rainfall centre on these days. The geopotential height shows no significant anomalies from the mean state for the October to December period.

Specific humidity

The specific humidity and specific humidity anomalies associated with node 0 2 are shown in Figures 4.14 (a) and (b) respectively. A specific humidity pattern linked with this node is a positive anomaly over south east Ethiopia, co-located with the rainfall centre. Whilst this is not the highest specific humidity over the larger domain (figures peak over much of the Congo), it is the greatest over the smaller domain where the ERA-40 rainfall is being constrained and combined with the convergence, is a physically consistent factor in the creation of rainfall over south eastern Ethiopia during these dates. The specific humidity

anomaly shows positive values over the rainfall centre, reinforcing the physical link with daily rainfall events in the region.

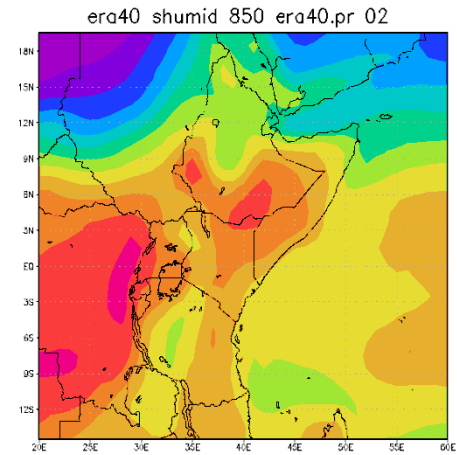


Figure 4.14 (a) Specific humidity mean for node 0 2

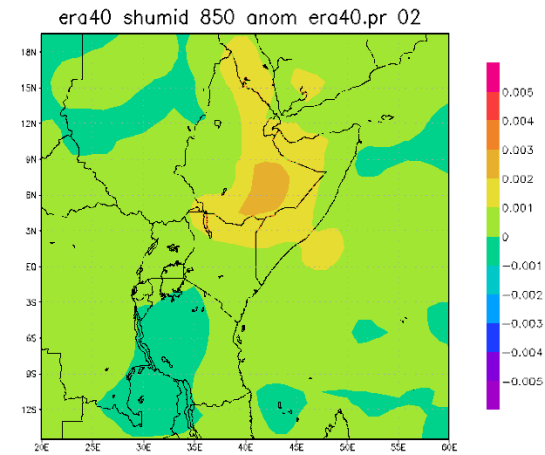


Figure 4.14 (b) Specific humidity anomaly for node 0 2

Zonal winds

Figure 4.15 illustrates the zonal wind anomaly associated with the 0 2 node.

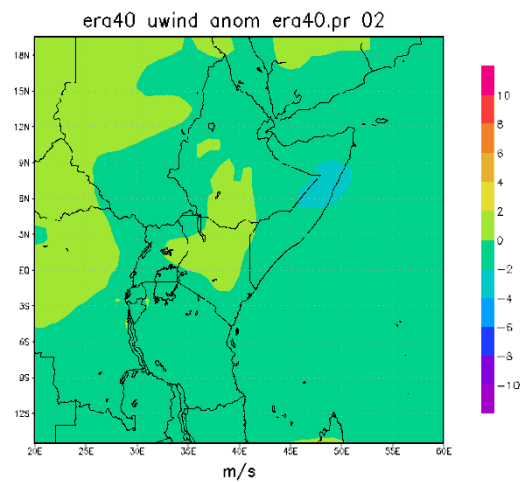


Figure 4.15 Zonal wind anomaly for node 0 2

Whilst node 0 0 has a large zonal wind anomaly co-located with the region of rainfall, this is not the case for this node where the rainfall is located over south eastern Ethiopia. Zonal winds associated with the ninety-one dates allocated to this node show conditions very close to the mean across the domain. The mean plot for this node is not shown, because it does not provide any insight into the rainfall centre for this node.

Node 6 0 (top right of the SOM array)

SOM node 6 0 is exemplified by a large area of rainfall centred over central Tanzania, with a heavy concentration of rainfall northwards up to the border with Kenya. Lighter rainfall covers much of the land domain expanding from this centre through Kenya and Uganda. There is also some rainfall eastwards of the Tanzanian centre over the coastal Indian Ocean. The 6 0 node is allocated as the node with the lowest quantisation error in the matching process on only thirty-nine dates, thirteen in October, seventeen in November and nine in December.

Moisture flux

Figures 4.16 (a) and (b) show the moisture flux mean and anomaly associated with node 6 0. There are two primary features to the anomalous moisture flux flow associated with this node. First, there is westerly flow originating from the Congo. Second, northerly flow originating from the north east of the domain is present through Somalia and Kenya along to the Tanzanian coastline. These two anomalous flows converge over Tanzania, accounting for the rainfall-causing convergence seen in Figure 4.16 (b). The plot for the mean moisture flow conditions shows onshore flow directly onto Tanzania which contributes to this convergence.

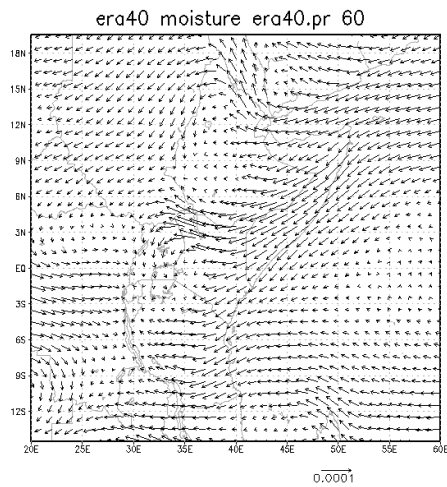


Figure 4.16 (a) Moisture flux mean for node 6 0

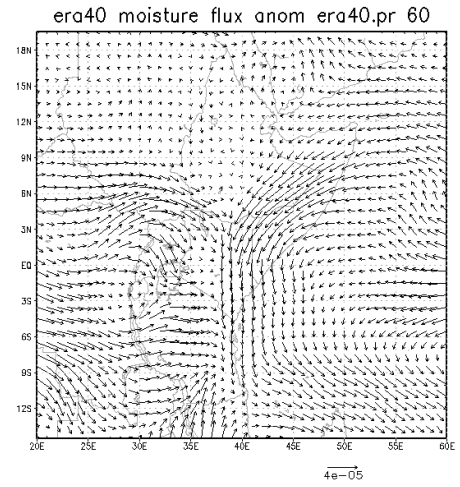


Figure 4.16 (b) Moisture flux anomaly for node 6 0

Moisture flux divergence

The mean and anomaly state for node 6 0 moisture flux divergence is shown in Figures 4.17 (a) and (b) respectively. Convergence is widespread, specifically over the southern half of the domain. In comparison to the mean OND plot, the region of convergence is much greater in width and is physically consistent with the large centre of rainfall in Tanzania. The anomaly plot shows convergence along the East African coast in particular over Tanzania.

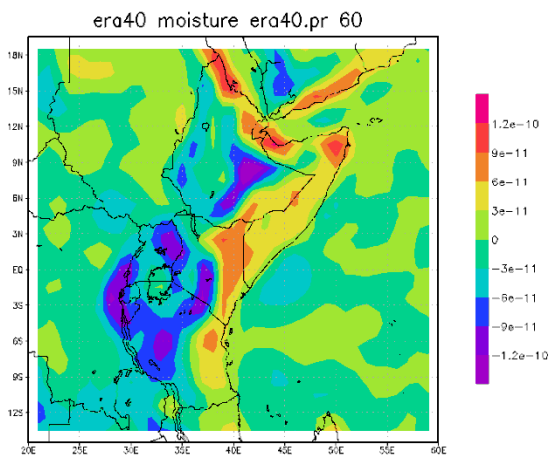


Figure 4.17 (a) Moisture flux divergence mean for node 6 0

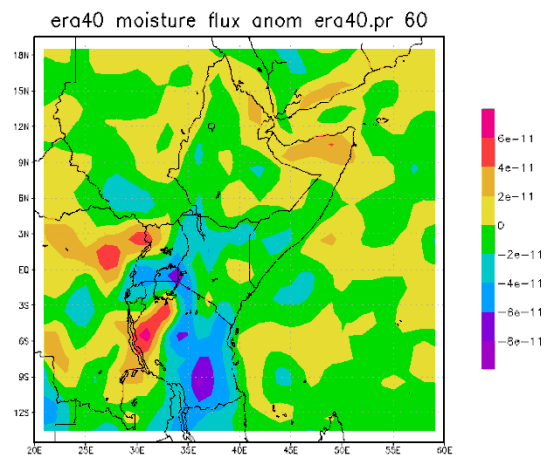


Figure 4.17 (b) Moisture flux divergence anomaly for node 6 0

Geopotential height

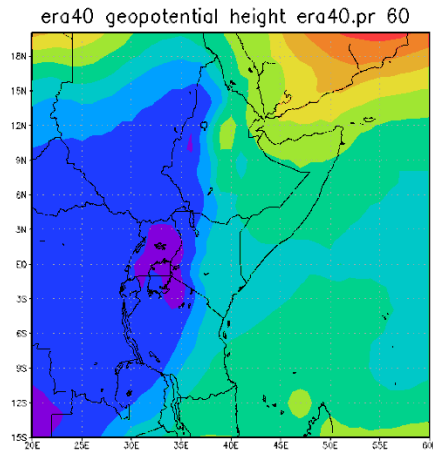


Figure 4.18 (a) Geopotential height mean for node 6 0

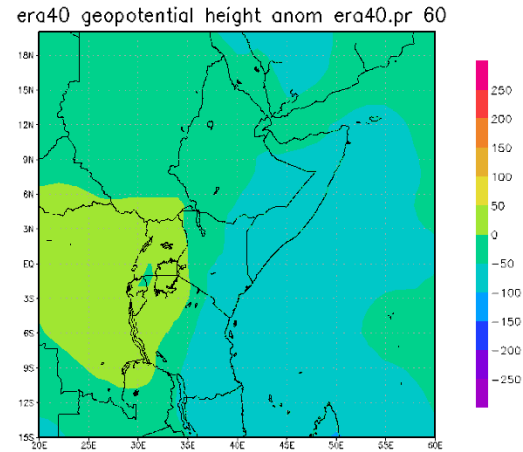


Figure 4.18 (b) Geopotential height anomaly for node 6 0

Figures 4.18 (a) and (b) show the geopotential height conditions associated with the 6 0 node. As with previous geopotential height plots they are inconclusive in physically supporting the characteristics of rainfall in this node, suggesting that this variable in the ERA-40 dataset at this level is not appropriate to use to investigate mechanisms leading to rainfall over the region.

Specific humidity

Figures 4.19 (a) and (b) show the mean and anomaly specific humidity 850hPa fields that are associated with the 6 0 node over East Africa. Of the individual nodes examined so far, this one has the greatest positive specific humidity anomaly associated with a rainfall centre and this supports the extensive Tanzanian rainfall that is characteristic of dates assigned to this node. Although the anomaly is centred over Tanzania, there is increased specific humidity over much of the region of the focus countries.

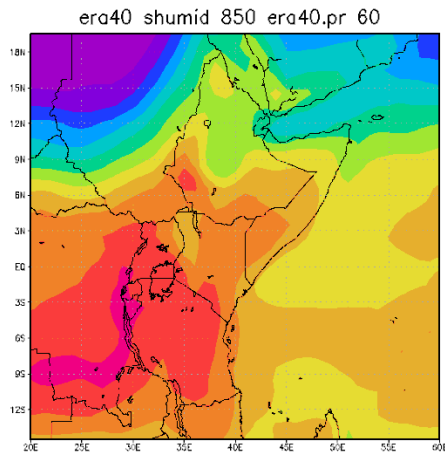


Figure 4.19 (a) Specific humidity mean for node 6 0

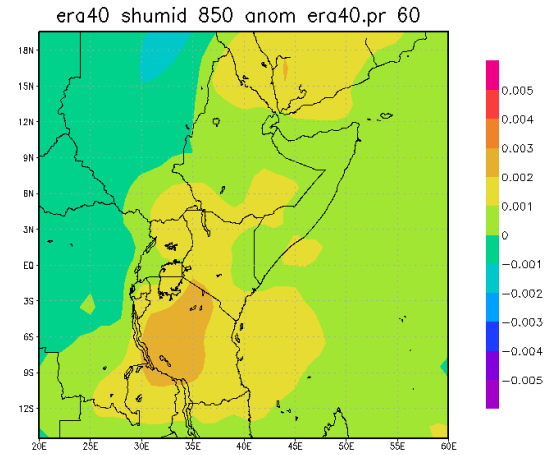


Figure 4.19 (b) Specific humidity anomaly for node 6 0

Zonal winds

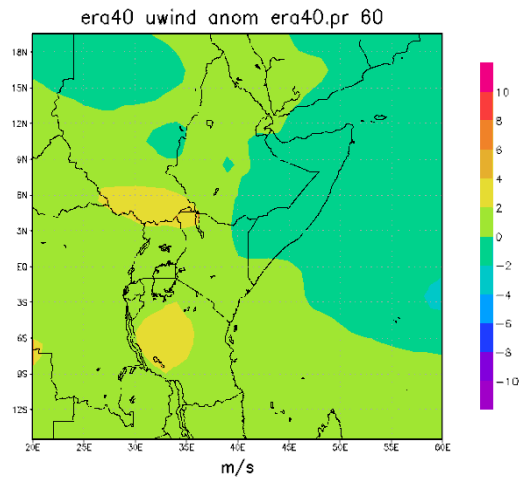


Figure 4.20 Zonal wind anomaly for node 6 0

Figure 4.20 shows zonal wind anomalies at the 850hPa level associated with the 6 0 node.

Whilst the magnitude of anomalies is small, with this node having no anomalous wind speeds of greater than 4ms^{-1} , there are two regions of anomalous westerly winds within the

domain over Tanzania and South Sudan. The mean zonal wind plot for this node does not show conclusive supporting evidence for the rainfall centre and so is not shown here.

Node 6 4 (bottom right)

Node 6 4 occurs as the best matched node on 79 occasions during the thirty year control period. These dates are distributed through October (31 dates) and November (30 dates) evenly, with a lower number of occurrences during December (18 dates). The rainfall centre occurring during this node is centred offshore in the Indian Ocean adjacent to Tanzania. Rainfall events associated with this node have a tendency to encroach onshore into coastal regions of Tanzania, but no further.

Moisture flux

Figures 4.21 (a) and (b) show moisture flux (mean and anomaly respectively) over East Africa associated with node 6 4 during the climatology period.

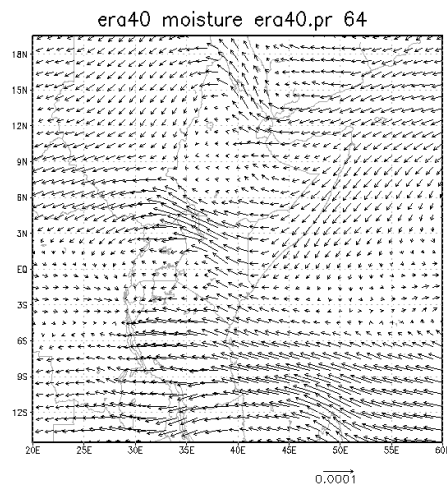


Figure 4.21 (a) Moisture flux mean for node 6 4

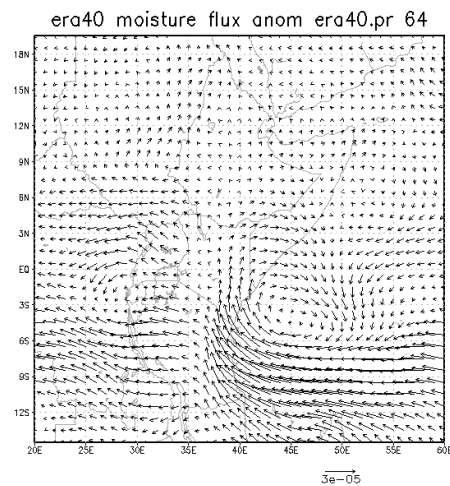


Figure 4.21 (b) Moisture flux anomaly for node 6 4

The strongest anomalous moisture flux associated with this node is located to the south of the domain (at 6°S to 12°S). This flow tracks onshore into Tanzania and then curves

northwards along the East African coast, curving offshore into the Indian Ocean at 3°N. This concurs with the raw data mean plots showing onshore moisture flux in the same direction. The pattern continues along the whole East African coastline, with the main pattern being easterly flow which travels onshore and across the domain with no convergence with flow from the west.

Moisture flux divergence

Figures 4.22 (a) and (b) show the mean and anomaly moisture flux divergence respectively associated with the 6 4 node.

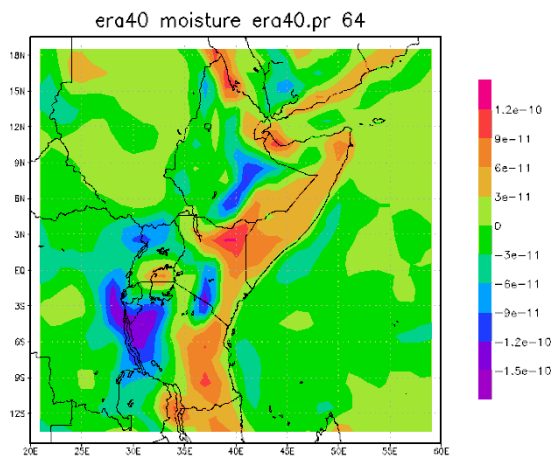


Figure 4.22 (a) Moisture flux divergence mean for node 6 4

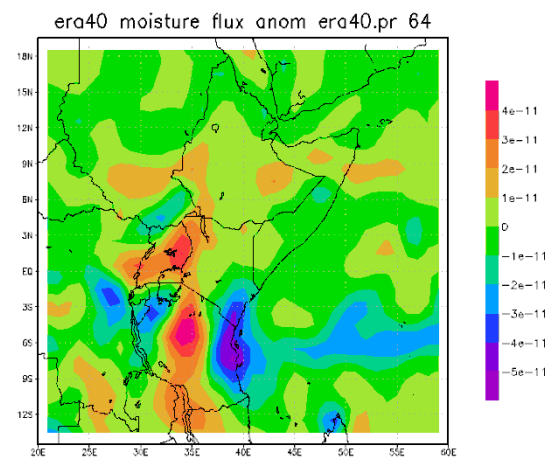


Figure 4.22 (b) Moisture flux divergence anomaly for node 6 4

There is a region of anomalous convergence offshore in the Indian Ocean adjacent to Tanzania, which strengthens over the Tanzanian coastline, where the fringes of the rainfall region are located. The weakening of the divergence often seen over Tanzania is what allows the anomalous convergence to encroach from the Indian Ocean over the land area, bringing rain to this region. Elsewhere over the focus countries there is anomalous divergence, physically consistent with the dry East Africa seen in this node.

Geopotential height

The 850hPa geopotential height for this node shows a small positive anomaly over the far south of the domain at 37°E and is shown in Figures 4.23 (a) and (b), the mean and anomaly respectively. However, as with previous nodes, this variable does not highlight a conclusive physical state attributable to the presence of the Indian Ocean/Tanzanian rainfall centre.

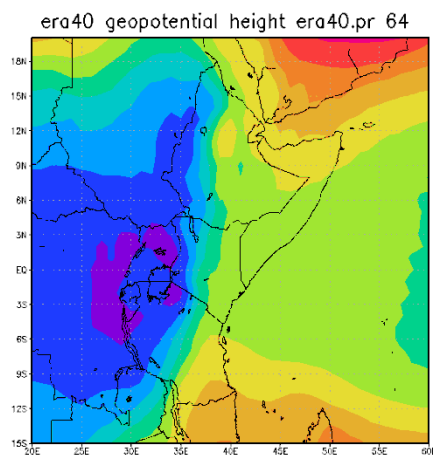


Figure 4.23 (a) Geopotential height mean for node 6 4

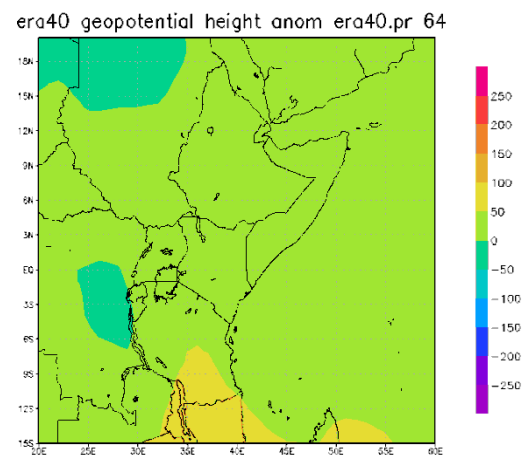


Figure 4.23 (b) Geopotential height anomaly for node 6 4

Specific humidity

Figures 4.24 (a) and (b) show the specific humidity mean and anomaly respectively for days linked to the 6 4 SOM node. In line with patterns identified in other nodes, there is a positive specific humidity anomaly co-located with the region of rainfall in the Indian Ocean and coastal Tanzania. This is physically consistent with rainfall in this region. Additionally there is a slight negative specific humidity anomaly over the dry focus countries. The mean state does show very high specific humidity over the eastern Congo. This is one of the reasons why the larger East African domain SOM is dominated by

rainfall in this region. Excluding this area from the ERA-40 rainfall series allows varying rainfall states over the four focus countries to be identified.

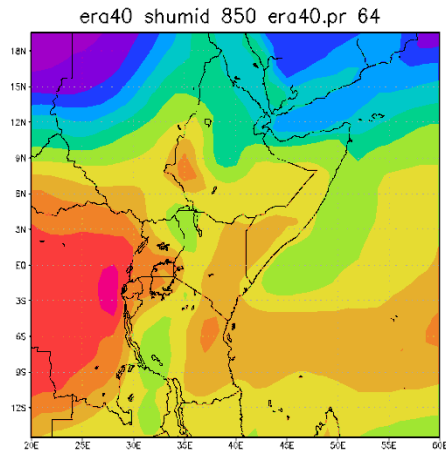


Figure 4.24 (a) Specific humidity mean for node 6 4

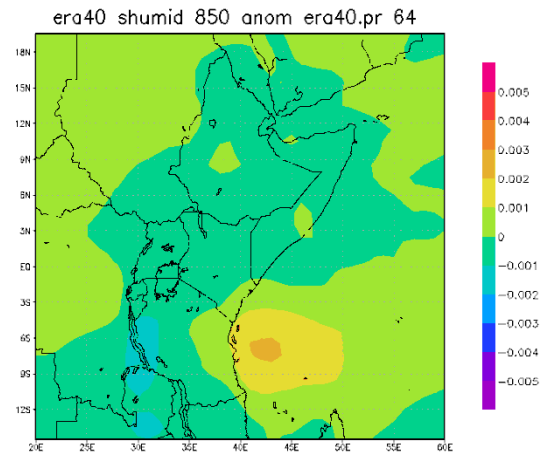


Figure 4.24 (b) Specific humidity anomaly for node 6 4

Zonal winds

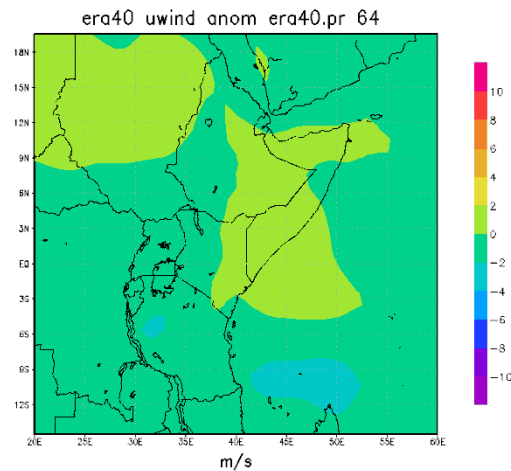


Figure 4.25 Zonal wind anomaly for node 6 4

Zonal wind anomalies, Figure 4.25, are close to the mean state over much of the domain. There is just one small region where stronger anomalies are seen, which is in the Indian Ocean north of Madagascar where anomalous winds reach speeds of up to 4ms^{-1} . This is in line with patterns seen in the moisture flux variable. Again, the mean state for this node has been omitted due to the lack of additional information it provides.

Node 6 1 (far right column, second row from top)

Node 6 1 features the most intense rainfall of all of the thirty-five reference nodes. The heavy rainfall, which is the main feature here, is centred over eastern Tanzania with some extension up into eastern Kenya and additionally into the Indian Ocean. The node is linked to daily rainfall in the region 66 times during the control period. These days are distributed between the three focus months with occurrences 34 times in October, 21 in November and 11 in December.

Moisture flux

Figures 4.26 (a) and (b) show the moisture flux mean and anomaly respectively, for node 6 1. The moisture flux anomaly for days associated with this node is dominated by one feature, onshore easterly flow onto Tanzania which then divides into three strands and travels either further eastwards at a latitude of approximately 10°S or further southwards or northwards along the East African coastline. The north of the plot is characterised by the mean state for this time of year; monsoon flow originating from the north east.

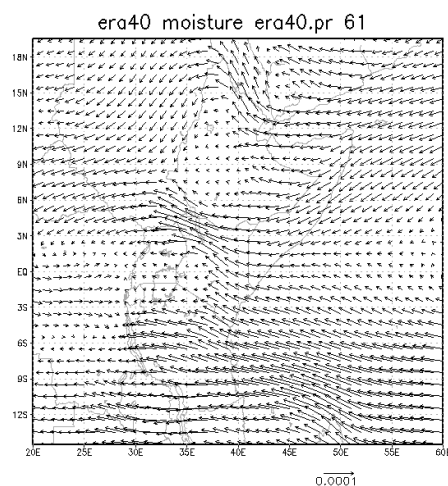


Figure 4.26 (a) Moisture flux mean for node 6 1

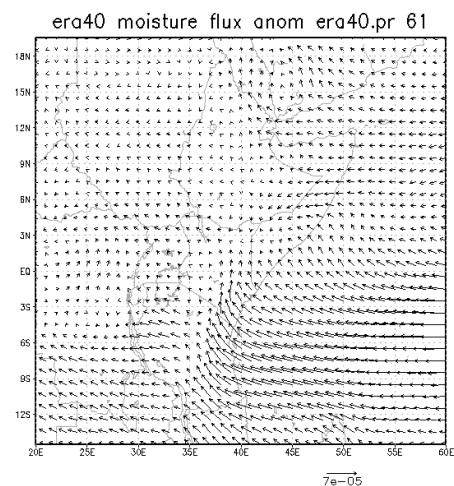


Figure 4.26 (b) Moisture flux anomaly for node 6 1

Moisture flux divergence

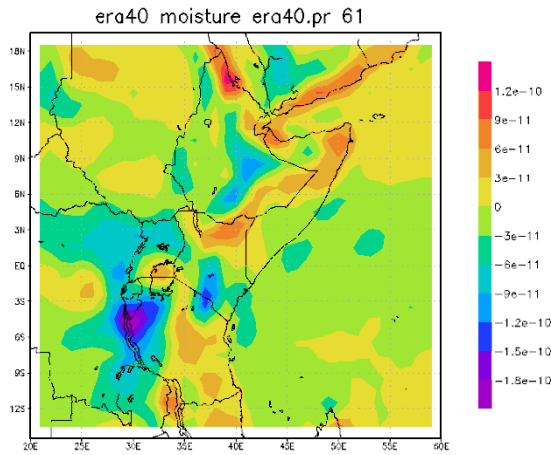


Figure 4.27 (a) Moisture flux divergence mean for node 6 1

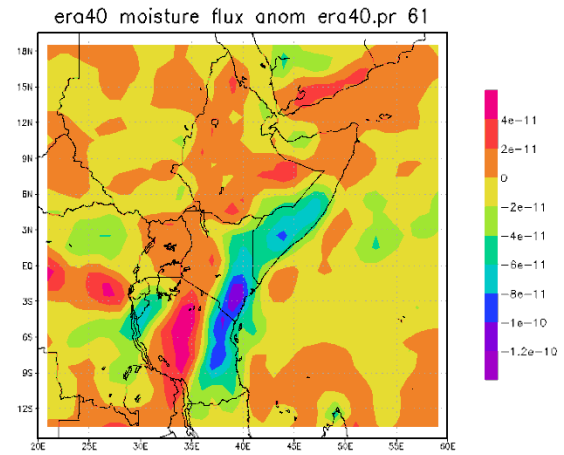


Figure 4.27 (b) Moisture flux divergence anomaly for node 6 1

Figure 4.27 (a) and (b) show the moisture flux divergence mean and anomaly for node 6 1.

The mean moisture flux for this node shows a complex pattern of a mix of convergence and divergence over the land regions, although the regions of convergence are stronger, in particular further inland. Interestingly, the main region of convergence is to the left of the rainfall centre although there is also convergence on the Kenya and Tanzania border.

When considering this in conjunction with the anomaly plot a pattern of moisture flux convergence more physically consistent with the location of heavy rainfall emerges. Here, there is anomalous convergence all up the east of the land area of the domain, with the strongest convergence seen on northern and eastern Tanzania, concurrent with the rainfall pattern. It is important to note that this node is slightly unusual in that such heavy rainfall is triggered by a relatively small region of convergence in comparison to other rainfall states over the domain.

Geopotential height

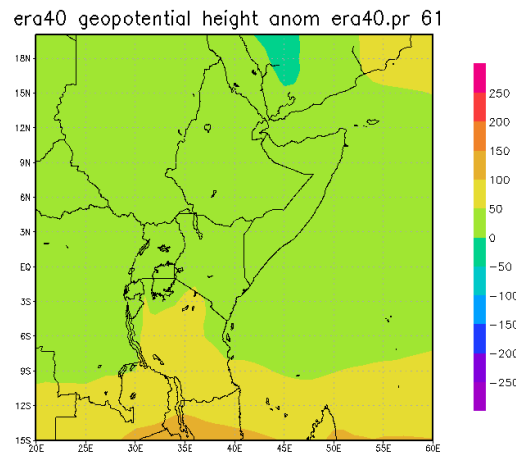


Figure 4.28 Geopotential height anomaly for node 6 1

As with the other selected nodes, the 850hPa geopotential height plots gives little information. The anomaly plot is shown in Figure 4.28, the mean plot for these days has been omitted through lack of insight it provides. The only distinguishing feature from the mean state for this time of year is a small positive anomaly in the south of the domain, as present in the mean anomaly states for many nodes for this variable.

Specific humidity

Figures 4.29 (a) and (b) show the specific humidity patterns associated with the 6 1 node of the ERA-40 normalised to climatology values small domain rainfall SOM.

This node is characterised by widespread regions of high specific humidity over the land domain and the western Indian Ocean, something which is emphasised by the positive specific humidity anomalies seen in Figure 4.29 (b). Unusually in comparison to other nodes examined, the greatest specific humidity anomalies are not co-located with the region of most intense rainfall but are instead offshore, with high humidity relative to the mean also evident over northern Ethiopia and the Arabian Peninsula.

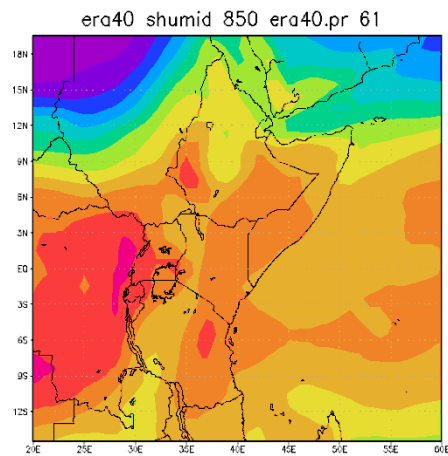


Figure 4.29 (a) Specific humidity mean for node 6 1

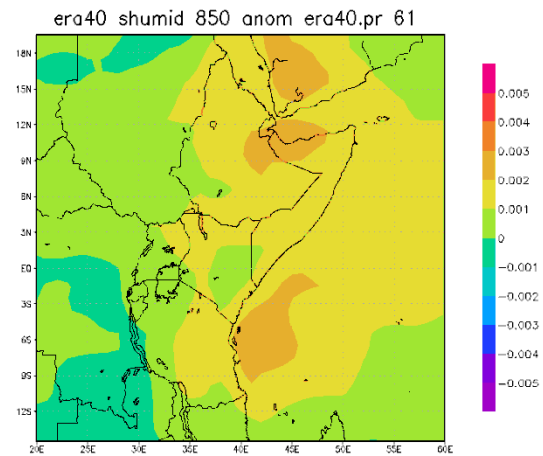


Figure 4.29 (b) Specific humidity anomaly for node 6 1

Zonal winds

Concurring with the moisture flux plots, the 850hPa zonal winds shown in Figure 4.30 (anomaly plot) show a strong easterly anomaly offshore of Tanzania. This emphasises the transport of moisture advected from the Indian Ocean over the land region.

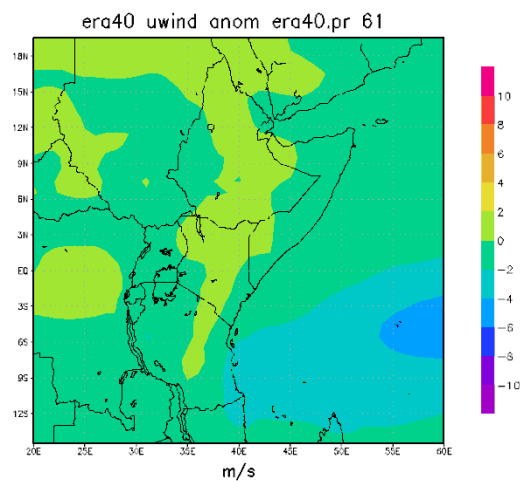


Figure 4.30 Zonal wind anomaly for node 6 1

The plots shown in association with the node 6 1, linked with the heaviest rainfall seen over the land domain do not show associations as clearly, with the location of rainfall as

the others examined. The physical characteristics of the circulation parameters still point to the generation of rainfall, but it is clear that more investigation is required to understand how certain patterns arise and the link to rainfall production over East Africa.

Two more nodes are analysed to demonstrate further the patterns linked with production of rainfall, and also with dry conditions. The same plots as the other nodes are analysed but have not been shown here.

Node 0 4 (left most column, bottom row)

The 0 4 node is characterised by a rainfall event in eastern Kenya, centred midway between the Kenyan coast and the border with Somalia. This region of rainfall extends both westwards and eastwards into central Kenya and the Indian Ocean respectively and occurs 114 times during the control period. These occurrences are split over all three months, with 43 dates in October allocated to the 0 4 node, 31 dates in November and 40 dates in December.

Patterns associated with this node are as follows:

Moisture flux

There are two main anomaly moisture flows over the domain. These are flows originating in the Indian Ocean in the north east and south east of the domain flowing inland and converging over eastern Kenya. For this node there is very little westerly flow coming into the domain, explaining the presence of rainfall only in the east of the land area where the monsoon and south east flows converge.

Moisture flux divergence

Convergence occurs in Kenya co-located with the rainfall centre along the East African coast. There are additional regions of convergence further inland, although the presence of divergence along the rest of the East African coastline outside of Kenya blocks moisture flow inland from the Indian Ocean. This is reinforced by the anomaly field which shows anomalous convergence located where the rainfall is and light divergence elsewhere.

Geopotential height

As with other nodes, the geopotential height plots at the 850hPa level do not show useful information in relation to the production of rainfall over East Africa.

Specific humidity

There are two positive specific humidity anomalies associated with the centre of rainfall for node 0 4. However, neither is perfectly matched with one slightly further north west and the other to the east offshore, although rainfall does extend to this region. In general specific humidity is high over all the East African countries although especially so over Kenya.

Zonal winds

Much of the domain has slight easterly wind anomalies, which is the expected pattern given the dominance of easterly moisture flux moving onshore. However the zonal wind fields at the 850hPa level do not give us any extra information in this case.

Node 3 4 (middle column, bottom row)

Given that so far the examination of individual nodes has covered rainfall events over East Africa it is important to also look at a dry node, especially given that much of East Africa

is relatively arid for the latitude (Nicholson, 1996). The node 3 4 is dry over the whole of the land region. Node 3 4 is mapped as the best matched node on 167 occasions, 43 in October, 55 in November and 69 in December.

Moisture flux

The mean moisture flows associated with this node show similarities to the climatology moisture flux for the region. However, the anomaly plot highlights the widespread divergence over the region, creating the dry conditions seen in the node. There is considerable anomalous moisture flux flowing away from the focus countries.

Moisture flux divergence

This node has divergence much more widespread than the rainfall producing nodes, with a very similar appearance to the climatology plot. The divergence that dominates along the East African coastline ‘squeezes’ the regions of convergence inland with less moisture flux coming in from the Indian Ocean, blocking rainfall systems to the region. The anomaly moisture divergence plot shows little deviation from the mean, emphasising that the region tends to less rainfall than would be expected for the latitude.

Geopotential height

This variable at the 850hPa level is very close to the climatology state for the OND period; reinforcing that this variable does not provide useful information on the daily level as to the presence of rainfall centres over an East African domain.

Specific humidity

In contrast with rain producing nodes, this node is notable for the widespread negative specific humidity anomaly over the focus countries, a key characteristic of dry conditions.

Zonal winds

Similarly to other variables, the zonal winds associated with this variable are close to the mean state for the climatology period. This pattern is of easterly winds over much of the domain, including the monsoon flow originating in the north east of the domain with two westerly wind centres, one offshore in the Indian Ocean over the equator with wind strengths of up to 4ms^{-1} and the second with lower winds speeds located west of Rwanda and Burundi.

Conclusions

It is concluded that the ERA-40 normalised rainfall series produces a sensible classification of different rainfall centres over East Africa when the SOM domain is constrained to exclude the more dominant rainfall events over the Congo and Indian Ocean. Whilst here the analysis focuses on the nodes classifying rainfall events at different locations within the domain, from looking at the overall SOM array the predominant mode is for dry conditions and the suppression of rain over the domain, despite the latitude and co-location with the ITCZ during the October to December months.

The best indicator of the presence of rainfall is the moisture flux and moisture flux divergence field at the 850hPa level. This close relationship can be explained through the strong influence the provision of water vapour (through moisture flux) has on the generation of rainfall, a relationship found to be true over the nearby Congo Basin (Washington et al, 2013). This strong relationship could prove useful in later chapters; the zonal and meridional wind fields exert an important control on the provision of water vapour and, given that winds are better represented in climate models than rainfall, this could provide an additional method of evaluation of differing daily rainfall fields in the

CMIP3 AOGCM subset under both climatology and climate change conditions (Washington et al, 2013).

Geopotential height is the least useful indicator in the ERA-40 dataset over the smaller domain; deviations from the mean state on a daily basis are minimal and not indicative of the presence of isolated rainfall events over any of the four focus countries. This result is not unexpected for the tropics. Geopotential height would be better suited to midlatitude regions, where it indicates the configuration of rainfall producing synoptic systems; such a scale does not exist to the same extent in the tropics where convection dominates the production of rainfall. Scale analysis, as illustrated by Webster (1983), states that thermodynamic perturbations in the tropics are theoretically an order of magnitude smaller than in the midlatitudes. This is due to rapid adjustment of pressure gradient imbalances in the tropics, owing to this adjustment being less constrained by rotational effects (Webster, 1983).

Typically, the circulation plots associated with the rainfall nodes do present a physically consistent pattern for the occurrence of daily rainfall events at that particular location but there are occasions when not all variables are consistent. In order to check the consistency of the match between the circulation fields and the rainfall nodes, the five most closely matched individual dates and the five least closely matched individual dates, selected by the size of the quantisation error, are examined. This will allow an evaluation of whether it would be appropriate to deselect the least closely matched dates as being representative of the node in question, whilst leaving the node as still representative of some days over the East African domain. The case for deselecting specific days is simply that inclusion of those days into the circulation composite may unnecessarily complicate the composite by adding avoidable noise to it.

4.3 Most closely matched nodes

Each individual day is assigned to the SOM node that best describes the spatial pattern of rainfall for that date, that is, the node with the lowest error when the individual date is tested against each node. This is done through the quantisation error (QE); a measure of how well each grid point of the domain for the input date matches to the most similar node. Consequently the date with the lowest quantisation error is the one which is most closely matched to the node to which it is assigned. Table 4.3 shows the five best matched days (date and associated quantisation error) for each node evaluated in Section 4.2.3.

Table 4.3 the 5 best matched dates to selected SOM nodes and their quantisation error for the ERA-40 normalised SOM, small domain

Node	1st	2nd	3rd	5th	5th
0 0	02/10/1988	20/12/1985	06/12/1986	03/10/1988	06/12/1989
	14.6009	17.6173	18.6225	18.8845	19.2384
0 2	05/10/1973	03/11/1992	10/11/1980	11/12/1995	26/11/1980
	9.02664	9.63305	10.0128	10.9909	11.5336
6 0	20/11/1973	16/11/1980	16/11/1978	22/12/1991	04/12/1982
	15.1408	15.1929	15.8696	16.3013	16.5337
6 4	13/12/1991	08/12/1995	04/10/1977	18/11/1990	03/10/1972
	6.32319	6.90335	6.90419	6.99628	7.1712
6 1	22/10/1995	21/10/1995	14/11/1982	14/10/1994	27/11/1977
	13.5688	14.5185	14.6715	15.1647	15.54
0 4	12/11/1979	18/12/1988	08/12/1973	11/10/1976	12/11/1977
	7.11417	8.15634	8.50445	8.69654	8.92596
3 4	12/10/1996	09/12/1993	12/10/1990	03/12/1996	10/12/1993
	3.86535	3.96621	3.99494	3.99621	4.00457

The selected dry SOM state, node 3 4, has the lowest quantisation errors of all featured nodes. This is the expected result; a dry node with very little rainfall present over the

domain is simpler to classify than rainfall events over different regions given the great spatial and amplitudinal variation in individual rainfall events and the limited number of nodes available for classification. The selected days, shown in the table above, whilst not shown here, are plotted alongside the mean and mean anomaly states for the circulation parameters and examined. In general, the best matched dates are those where rainfall centres are more spatially widespread. This is the expected result, given that the SOM classification is, by its nature, a generalisation which will tend towards representation of more widespread rainfall occurring on a greater number of days than localised events occurring only occasionally.

4.4 Least closely matched nodes

The next step is to examine the nodes that are least closely matched to the SOM node to which they are assigned, that is the five worst matched nodes (also judged through quantisation error). Table 4.4 shows the five worse matched nodes for each of the featured nodes for the ERA-40 small domain normalised rainfall SOM, alongside the date they occurred and the quantisation error.

There is considerable variation in how closely matched the ‘worst matched’ nodes are actually matched to the node to which they are assigned. Node 3 4, a dry state over the land region, has quantisation errors for the least well matched dates that are smaller than the best matched dates for nodes 0 0, 6 0 and 6 1 where there are rainfall centres over the land domain. In contrast node 6 0, representing heavy rainfall over Tanzania, has a very high quantisation error. This shows the difficulties in classifying closely all rainfall days with a limited number of states in the SOM.

Table 4.4 the 5 worst matched dates to selected SOM nodes and their quantisation error for the ERA-40 normalised SOM, small domain

Node	1st	2nd	3rd	4th	5th
0 0	28/12/1983	24/12/1983	26/10/1977	13/11/1987	21/11/1997
	57.6271	53.2227	48.3649	47.5091	47.1506
0 2	04/12/1983	12/12/1990	14/11/1997	13/11/1997	01/11/1977
	63.4433	42.4443	40.6723	40.5533	39.6703
6 0	19/10/1971	27/11/1982	29/10/1973	28/11/1982	20/11/1997
	77.764	52.481	48.7624	46.8859	39.3849
6 4	21/12/1994	07/10/1994	31/10/1995	28/11/1994	22/11/1993
	20.9276	18.5444	17.2692	16.9506	16.7637
6 1	15/12/1997	06/12/1997	19/10/1997	29/11/1982	18/10/1997
	56.1479	51.8733	50.9698	48.3992	46.1046
0 4	14/10/1997	10/12/1992	26/11/1997	02/12/1997	11/12/1992
	40.4725	37.0298	36.2936	36.203	35.8368
3 4	10/10/2000	14/12/1985	18/10/1999	12/10/1987	08/11/1981
	13.5442	10.8911	9.9483	7.99203	7.70001

Nevertheless, examination of the individual worst matched days shows that in many cases they are still matched to an appropriate node. Whilst not all cases are shown here, Figures 4.31 (a) and (b) show the individual dates for both the ERA-40 rainfall and the associated 850hPa circulation parameters for the worst matched dates for the nodes 0 0 and 6 1.

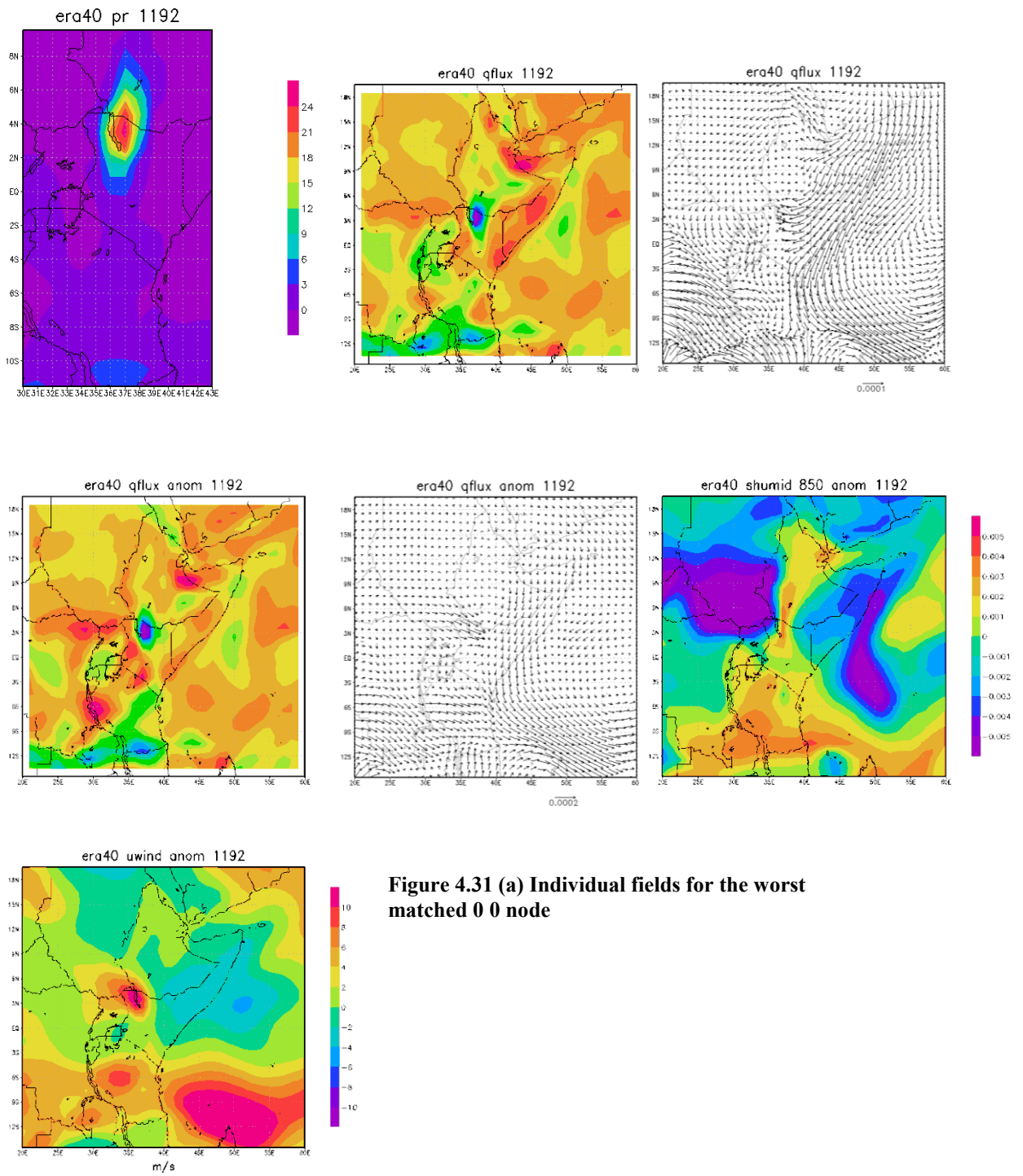


Figure 4.31 (a) Individual fields for the worst matched 0 0 node

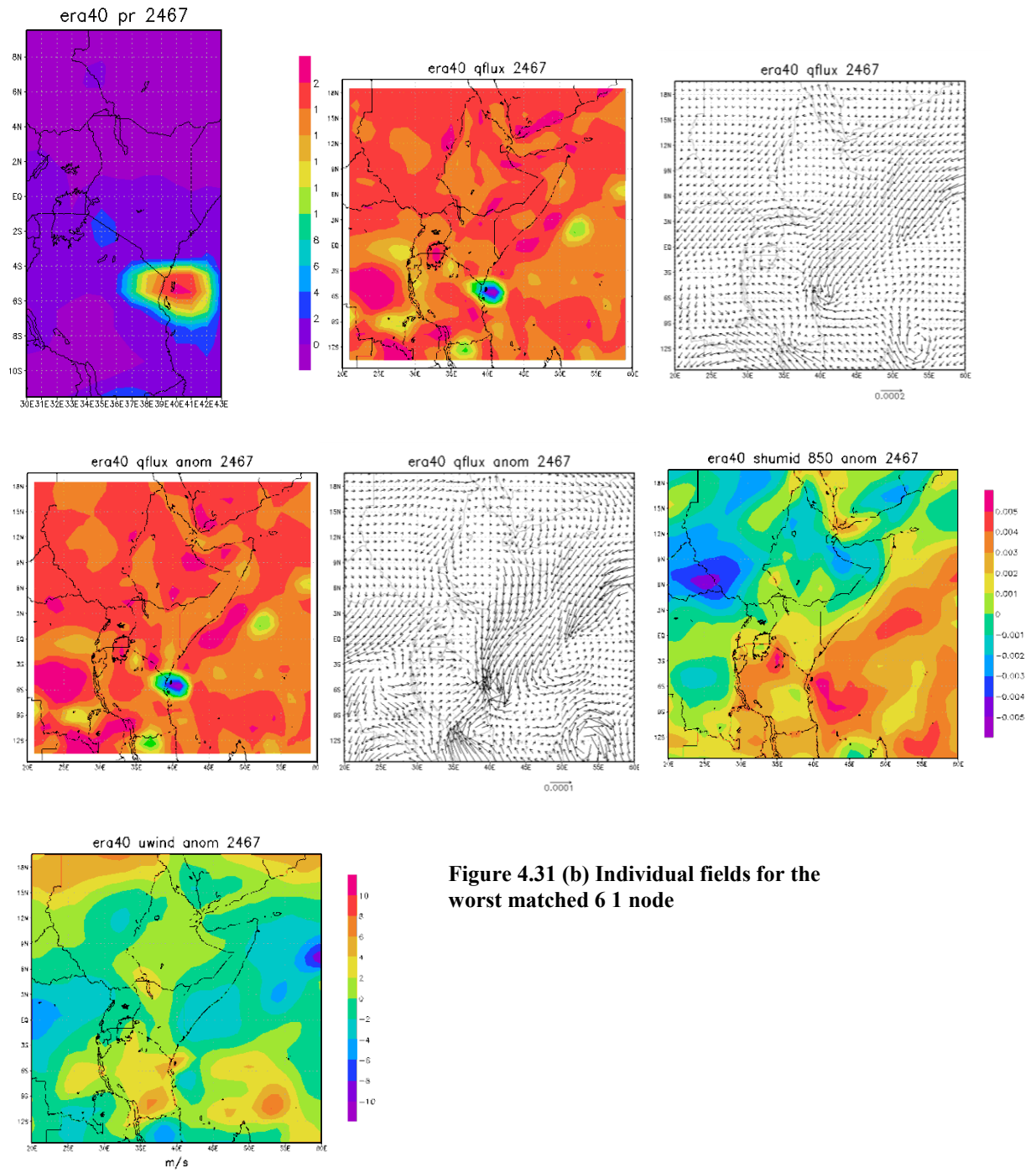


Figure 4.31 (b) Individual fields for the worst matched 6 1 node

The highest quantisation error for node 0 0 is 57.6271. Despite this, Figure 4.31 (a) shows that this rainfall event is appropriately allocated to this node, with the rainfall event centred on the border of Kenya and Ethiopia, a region which is within the spatial pattern of rainfall linked to this node. Moreover, the remaining 850hPa circulation parameters reinforce this with the physical patterns pointing to convergence in the region, a positive specific humidity anomaly and anomalous westerly and easterly zonal winds meeting in this region. Similar patterns are seen in the other least matched 0 0 dates.

The node 6 1 shows a slightly different picture. Here, for the least matched node with a QE of 56.1479, the location of rainfall does not completely overlap with the spatial pattern classified for this node, being located eastwards of it in the Indian Ocean. The reason this is the best classification for this day is the strength of the rainfall event with the combination of heavy rainfall and nearby location making this the best suited node. The other ERA-40 variables show patterns physically consistent with this rainfall event.

Although the other selected worst matched dates are not shown, similar reasoning applies to all of the dates; the classification is sensible in allocating rainfall events to the correct node from a limited number of states. Often the larger quantisation error arises from being a day with more localised rainfall that is not explicitly represented within the SOM array. Therefore, the weaker matches for each node are not having a negative impact on the mean field of the circulation parameters responsible for bringing rainfall to East Africa.

Therefore, all dates assigned to each node are kept within the classification system with none discarded for the remainder of this thesis. This also increases confidence that a 7 by 5 self-organising map with 35 states is a suitable array size to examine as being representative of the recurring East African short rains daily rainfall events without being overly complex.

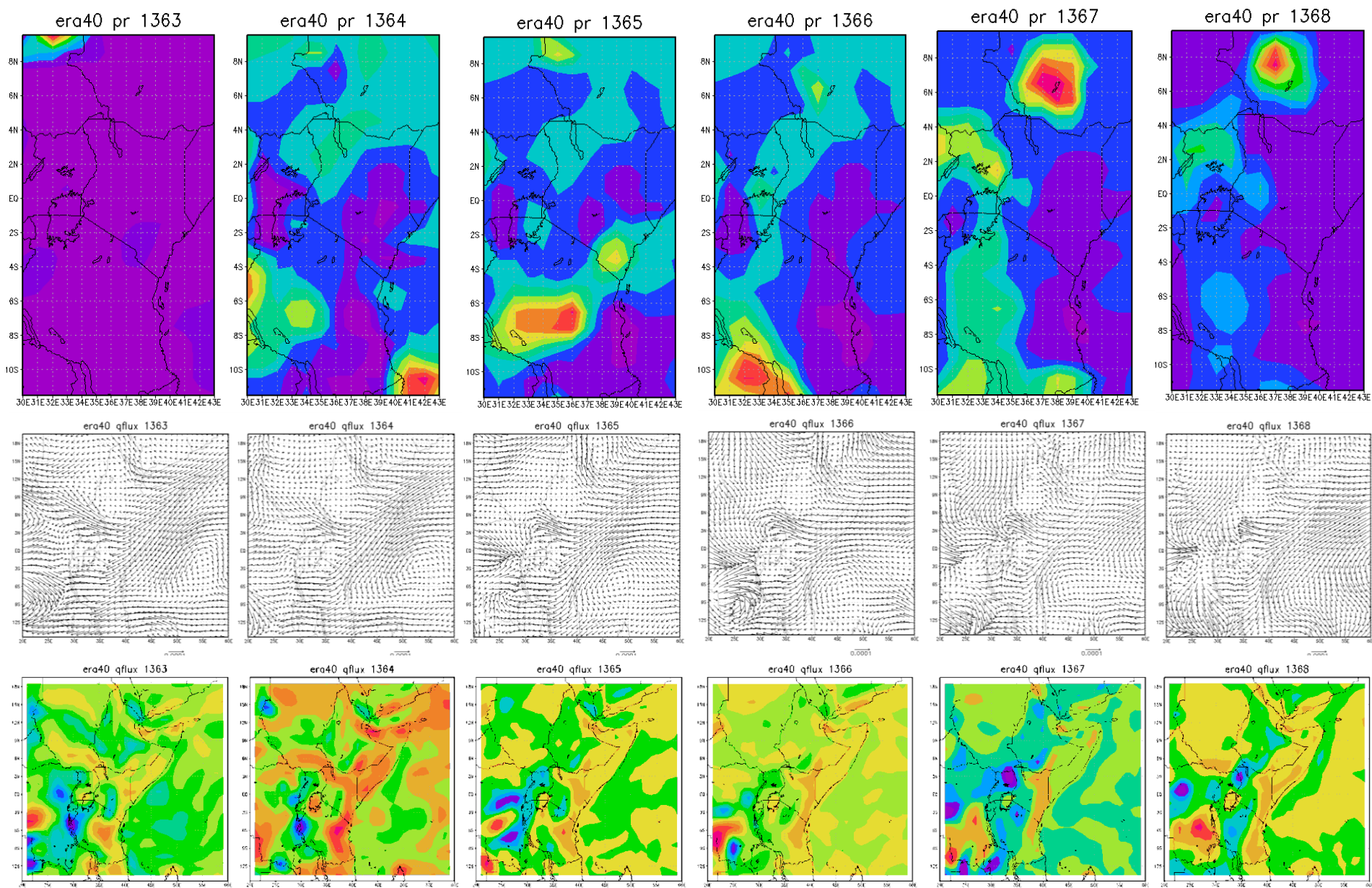
4.5 Progression of circulation fields over the days preceding a rainfall event

The day to day progression of the circulation patterns within the East African climate that relate to the production of rainfall over the focus countries is studied next. The particular questions in focus are: Do rainfall events occur on single, isolated days? If not, are the circulation patterns sustained over a number of days?

This is done through examination of the five day build up to each of the best matched dates for node 0 0. This node is selected due to the coherent picture that emerged in the mean state for days associated with this node in all circulation parameters. Figures 4.32 to 4.34 show the build up to the node 0 0 for three of the dates assigned to it; 20/12/1985, 06/12/1986 and 06/12/1989.

Figure 4.32 Node 0 0, date 20/12/1985 (no. 1368), run in from 15/12/1985 (no. 1363)

Node assignment from 15/12/1985 is 2 3, 4 4, 2 3, 3 2, 3 1



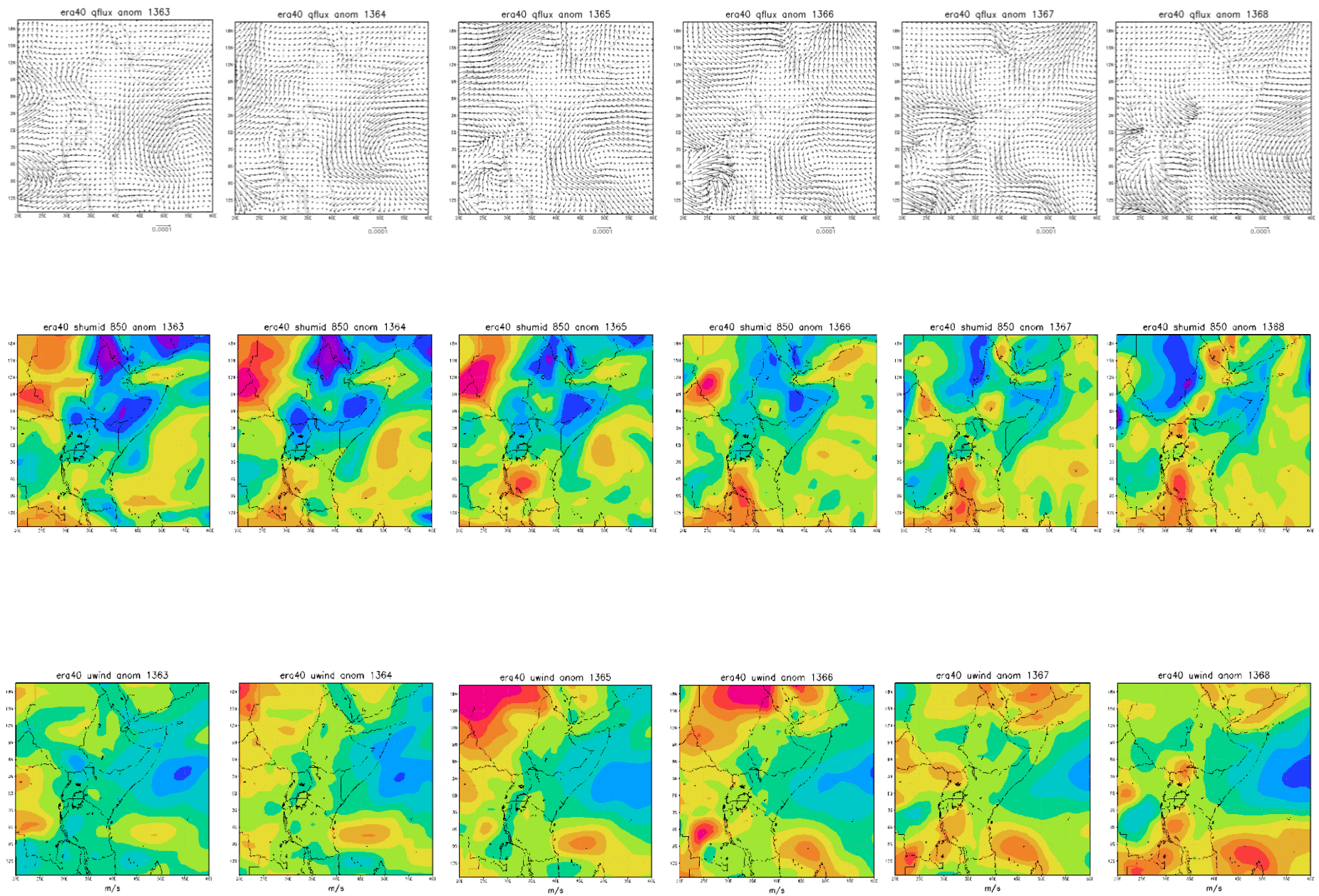
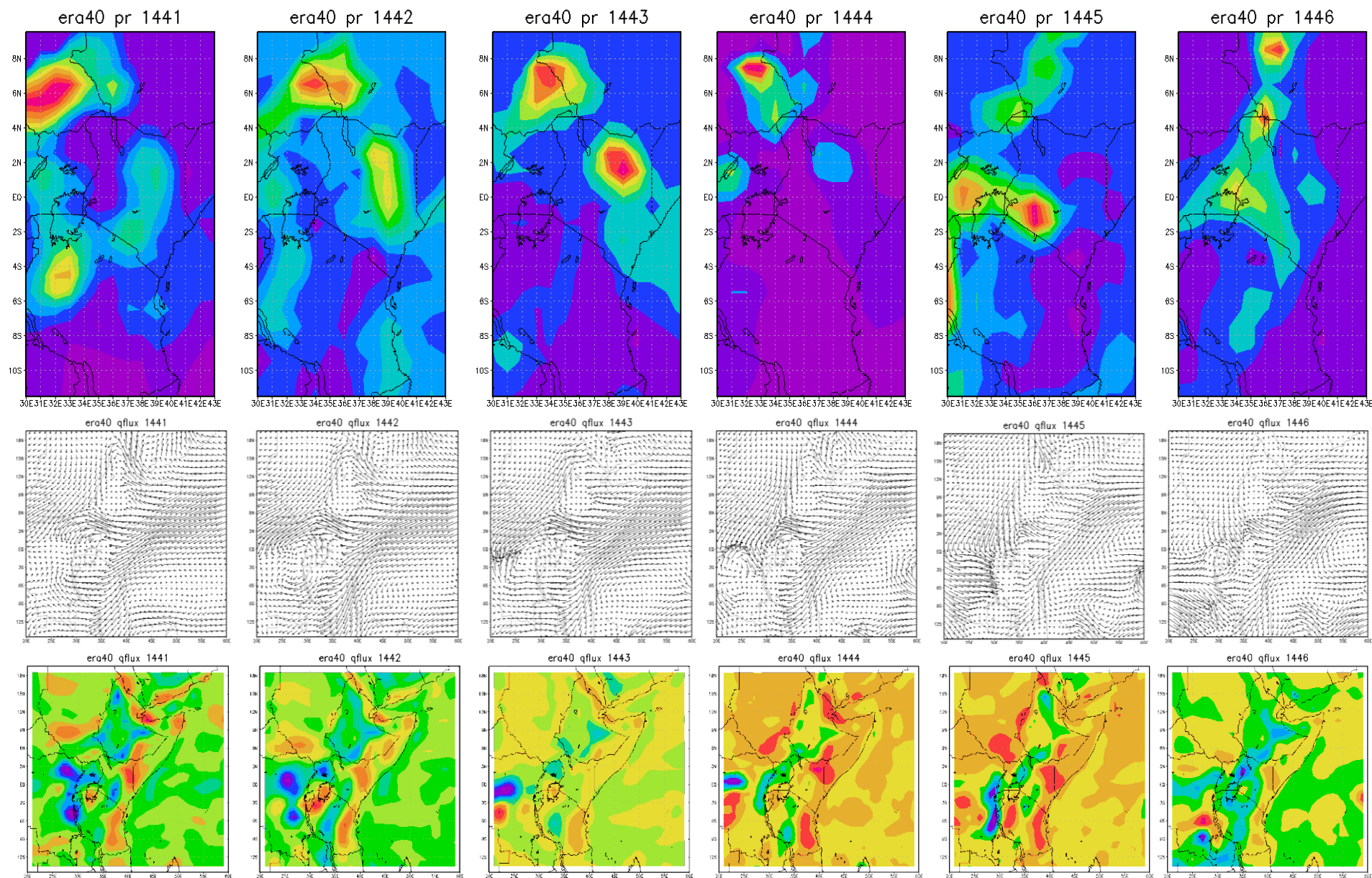


Figure 4.33 Node 0 0, date 06/12/1986 (no. 1446), run in from 01/12/1986 (no. 1441)

Node assignment from 01/12/1986 is 2 0, 1 1, 1 3, 2 0, 3 0



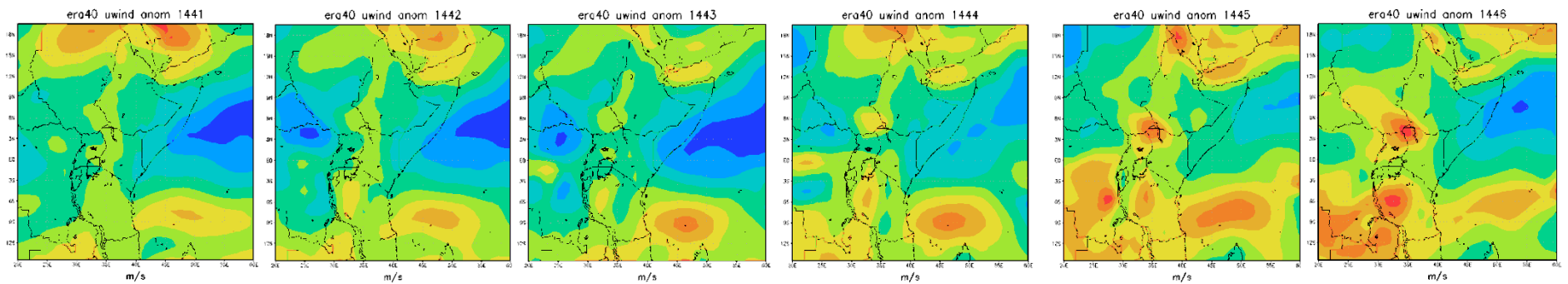
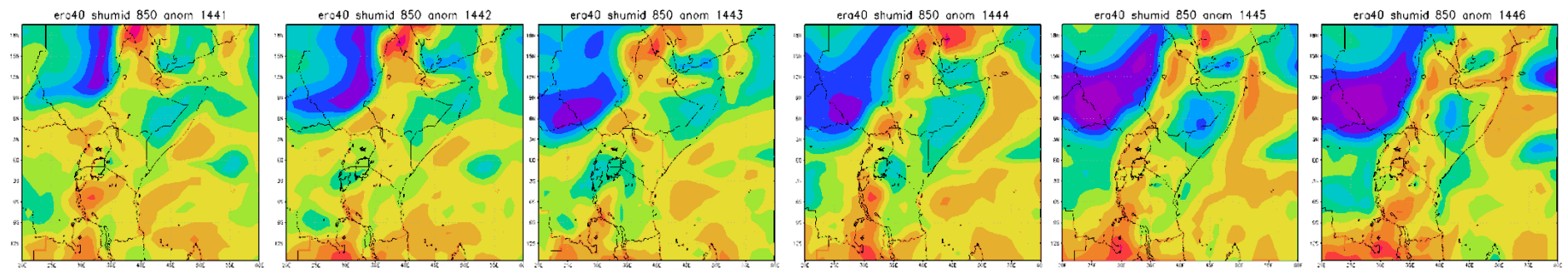
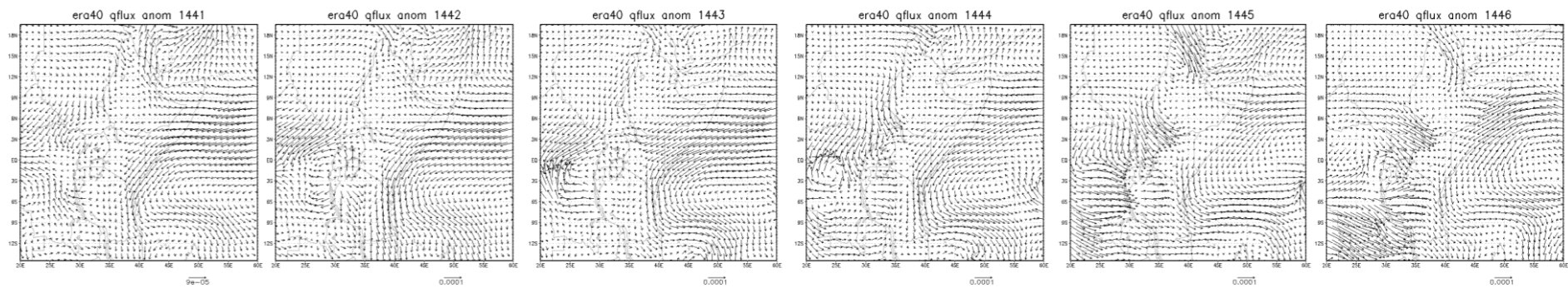
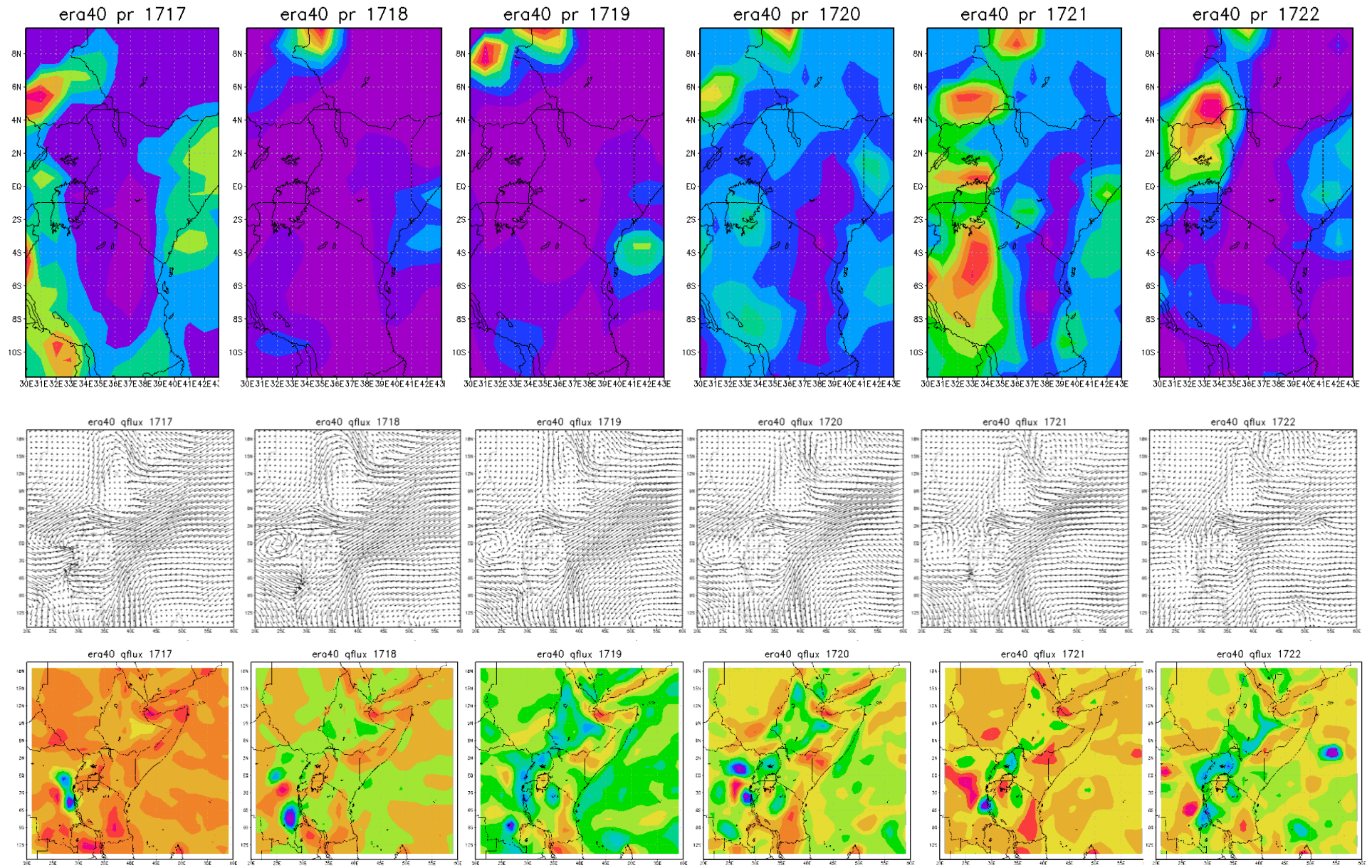
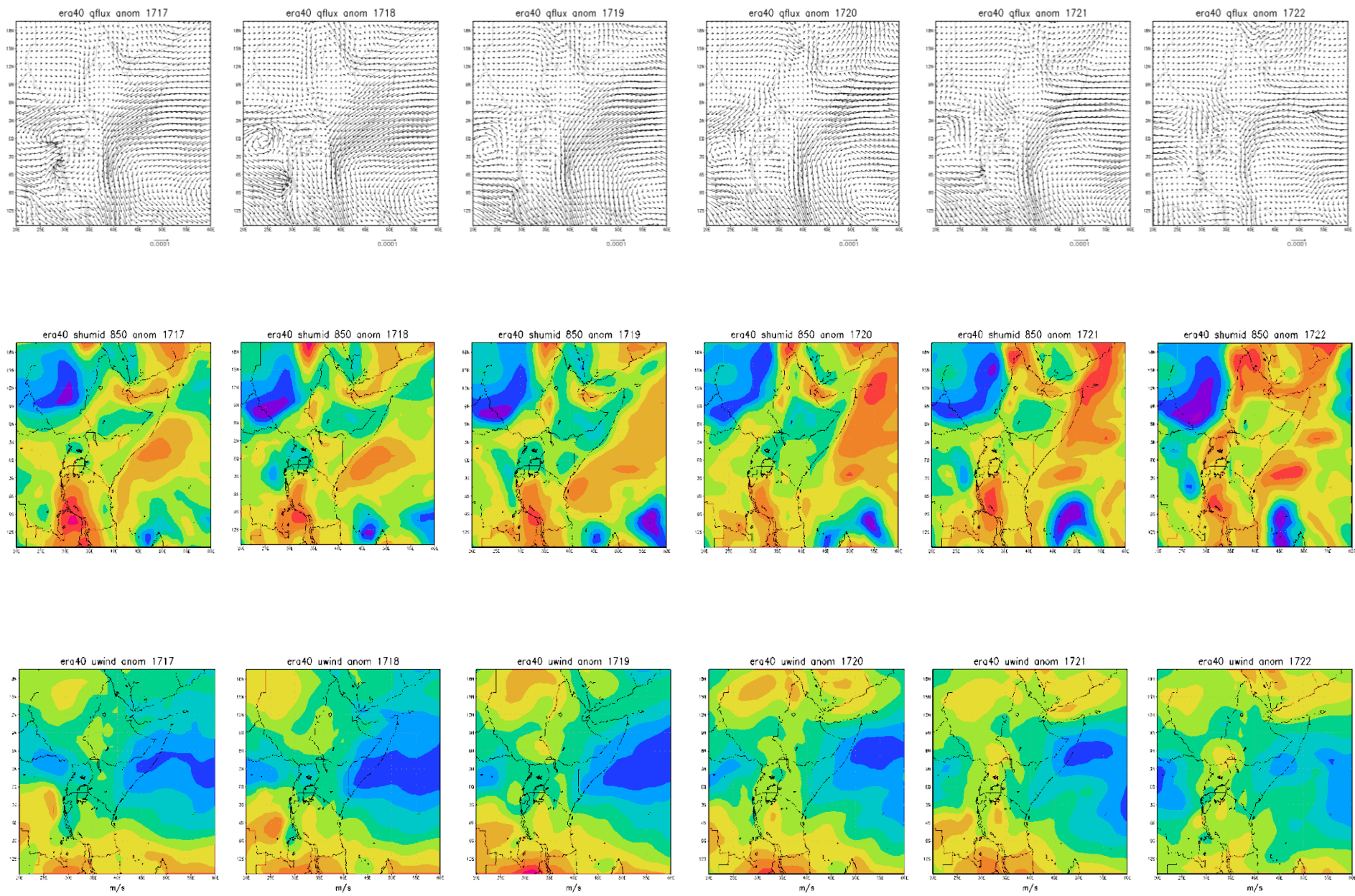


Figure 4.34 Node 0 0, date 06/12/1989 (no. 1722), run in from 01/12/1989 (no. 1717)
Node assignment from 01/12/1986 is 3 0, 1 1, 1 1, 2 2, 3 0





For each of the selected dates, 20/12/1985, 06/12/1986 and 06/12/1989, with related quantisation errors of 17.6173, 18.6225 and 19.2384 respectively, the structural features of circulation patterns for the preceding five days are examined.

Through examining the circulation patterns in December 1985, the development of the structure through to the rainfall event classified as 0 0 on 20/12/1985 and featuring a rainfall centre on the borders of South Sudan, Ethiopia, Kenya and Uganda can be seen. It is evident that the occurrence of this event is not independent with the build-up of several variables showing clear signs of the development of patterns that result in the final 0 0 classified rainfall event.

The moisture flux patterns show a state of flow which evolves gradually over the five day period leading to the region of convergence co-located with the rainfall event. Particularly evident is the moisture flow into the region which emerges from the north west, replacing a moisture flow out of the region towards the north west, this appears to be a key feature of rainfall creation at this latitude with the monsoon from the north east showing smaller adjustments during the five day period.

The development of structure leading to this rainfall event is also visible in other circulation parameters. Specific humidity anomalies show initially that specific humidity was low over the Uganda-South Sudan border, with very high specific humidity over South Sudan and Sudan. This region of high specific humidity moves eastwards through the five days, aided by westerly wind anomalies building in the region, until it settles at the location of the rainfall event by the fifth day.

The second example shown here, for 06/12/1986, culminates in a similar, although more widespread, rainfall event. Again, development of patterns in circulation variables is evident through the five preceding days. The moisture flux build up shows similar patterns to the

preceding example; north easterly monsoon flow is evident throughout with the trigger for rainfall appearing to be a change in flow to the west of the rainfall event from the moisture flowing away from this region to that flowing into the region from the north west leading to convergence, coupled with high specific humidity anomalies, that indicate the presence of the rainfall event seen in the 0 0 classified node. Equally, the third example follows the same patterns as described for the above two examples, with a clear build up in moisture flow leading to the rainfall event on 06/12/1989.

Therefore, the notable feature of the circulation lead in to these rainfall events is the 850hPa north-westerly flow leading to convergence over the Ugandan-South Sudan-Ethiopian border during the rainfall event. This feature is evident in all three build ups to a 0 0 node classification shown here and is a key component of rainfall events in this location. The similarities of the three examples shown here indicate that the ERA-40 dataset does not just have a rainfall ‘switch’ but that the creation of rainfall is bound up in circulation patterns over East Africa. To reiterate, for daily rainfall events over the Uganda-South Sudan-Ethiopia border, key for rainfall is anomalous north westerly moisture flow leading to convergence, a feature not previously identified in the literature.

4.6 Chapter summary

Chapter 4 has addressed the first research question of this thesis and the associated sub question. To summarise, the sub question of the first research question was to establish if the self-organising map was an appropriate tool to facilitate the establishment of contemporary conditions of daily rainfall regimes over East Africa during the short rains.

The self-organising map was proven to be an effective method of clustering daily rainfall events over East Africa during the short rains, facilitating further examination of these events. In order to obtain useful results it became clear that it was necessary to constrain the domain

to just the countries of focus; allowing a larger domain only served to enable the domination of larger scale events over the Congo and Indian Ocean, obscuring detail in the East African region. Nonetheless, once the domain had been limited to 30°E to 43°E, 12°S to 10°N, the methodology successfully clustered the daily rainfall field into thirty-five representative states characterising the continuum of daily rainfall patterns seen during the short rains control periods.

A further test of the ability of the method to classify daily rainfall regimes in an independent and reproducible manner was to investigate atmospheric circulation parameters associated with occurrences of each node of the SOM array. This was to establish if the clustering of daily rainfall events was associated with other, physically connected, circulation variables. Whilst a suite of variables were examined, moisture flux divergence and convergence was found to be the most reliable indicator of the presence of either dry or wet conditions at any location within the East African domain, with geopotential height the least useful at the daily level, as explained in Section 4.2. The characterisations of nodes as represented by days with the ‘best’ and ‘worst’ fit through quantisation error was examined, and it was concluded that even the ‘worst’ fitted days were assigned to an appropriate node and were not adversely affecting results for these nodes. Finally, the lead-in of circulation indicators to different rainfall events was analysed, with all evaluations contributing to creating a coherent and comprehensive characterisation of daily rainfall events during the short rains in the ERA-40 dataset, highlighting that during the short rains there was a development of circulation patterns over several days leading to rainfall events.

There are known weaknesses of reanalysis, or indeed any rainfall datasets over Africa. Therefore, the TRMM satellite rainfall series is used in Chapter 5 to introduce a further perspective on the East African short rains. This facilitates addressing the first research question in a truly comprehensive manner, allowing for confidence moving forwards in the

evaluation in later chapters of the daily rainfall field in AOGCMs over East Africa. The final conclusions to the first research question of this thesis, to establish the contemporary structure of daily rainfall regimes over East Africa during the October to December Short Rains season will be answered at the end of Chapter 5.

Chapter 5: Contemporary characteristics of daily rainfall regimes during the short rains – an ERA-40 and TRMM comparison

Chapter 5 introduces the TRMM satellite daily rainfall series as a second dataset with which to establish characteristics of daily rainfall regimes during the short rains events over East Africa. There are between 800 and 1200 gauge rainfall series providing daily rainfall measurements over the African continent, equivalent to one for every 23,300km² (Love et al, 2004). Given the difficulties in both utilising the station data available for this work and known weaknesses of reanalysis data, the TRMM dataset serves as a valuable alternative to explore the contemporary structure of daily rainfall events in the short rains during the twentieth century control period, under the first research question of this thesis.

It has been suggested that satellite derived rainfall estimates are the most appropriate option for overcoming the scarcity of station observations over Africa (Adeyewa and Nakamura, 2003). Furthermore, gridded satellite rainfall databases have so far been neglected as a tool for characterising seasonal rainfall features in Africa (Herrmann and Mohr, 2011). This work on the TRMM daily rainfall characteristics has the potential to contribute to these research areas for daily rainfall regimes of the short rains for an East African domain.

Chapter 5 provides the conclusions to the first research question of this work; to establish the contemporary structure of daily rainfall regimes over East Africa during the October to December short rains season. This is done through a comparison of characteristics of the ERA-40 dataset in a SOM rainfall array with the TRMM rainfall characteristics. First, the

TRMM series from 1998-2010, October to December, is clustered into a thirty-five node 7 by 5 self-organising map and compared to the SOM trained on ERA-40 data from 1971-2000. SOMs are trained on both datasets for the short overlap period of the two daily series of 1998-2000 and the resulting arrays evaluated, including their ability to assign the same days to the same nodes. The SOM array is reduced in size from thirty-five to twenty nodes as characteristic of daily rainfall events over East Africa during the short rains and the ability of the two datasets to assign the same types of events to each day is evaluated. The 1998-2010 TRMM rainfall series is then applied to the 1971-2000 ERA-40 trained SOM and the proportional occurrence of each type of key rainfall event during the short rains is compared to proportional occurrences of the same rainfall regimes in the ERA-40 series applied to the same ERA-40 trained SOM.

5.1 Classification of TRMM satellite rainfall

The TRMM rainfall series is a satellite derived product (see Chapter 2, Section 2.5). The data series employed here (and the total yearly availability for this study) covers the period 1998-2010, October to December with an input series of 1196 dates over thirteen years. The time period available has only a three year overlap with the 1971-2000 climatology period for which the ERA-40 data is utilised, so these SOMs are linked to the ERA-Interim reanalysis datasets for the circulation parameters following the methods of Chapter 4. For the TRMM SOM the smaller domain (30°E to 43°E, 12°S to 10°N) is used from the outset to allow for direct comparison with the ERA-40 dataset for the capture of daily rainfall events, resulting in an input field of 4576 values over the domain due to the higher 0.25° by 0.25° spatial resolution.

5.1.1 Daily rainfall states in the TRMM SOM

Figure 5.1 shows the best matched 7 by 5 TRMM rainfall SOM, associated with reference vector 7, for the normalised to climatology TRMM daily rainfall series.

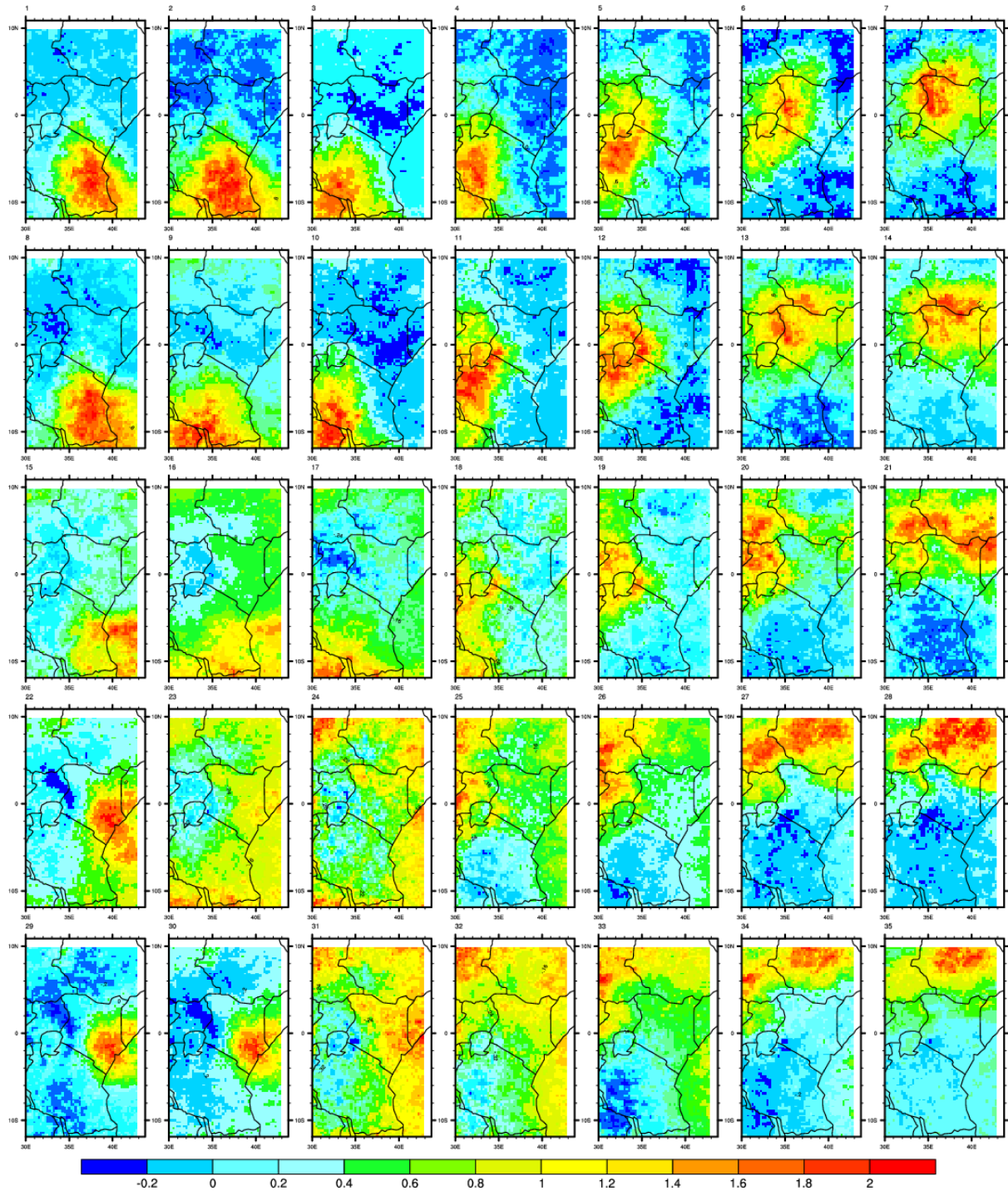


Figure 5.1 TRMM normalised rainfall SOM associated with reference vector 7

The TRMM SOM highlights similar daily rainfall states to the ERA-40 SOM array. There are daily events located over all four focus countries, Tanzania, Kenya, Uganda and southern Ethiopia, and also oceanic rainfall events as in the ERA-40 SOM. However, widespread suppression of convection leading to dry conditions over the whole of the East African domain, as commonly found in several states in the ERA-40 SOM, is not seen within the states classified in the TRMM rainfall SOM.

Daily rainfall events over Tanzania show considerable variation in the TRMM SOM. The most distinct daily events are found to the top left corner of the array with the centre of rainfall located to the south of the country at several different longitudes, from the southernmost point of the domain at 12°S to northern Tanzania at approximately 3°S. Rainfall over Tanzania, as part of more widespread rainfall events, is found in the central-left area of the bottom half of the array with this region covering events originating in the Indian Ocean, encroaching from Kenya and also from further south of Tanzania. Daily events over Kenya are located in two regions of the SOM array. Localised rainfall events centred over coastal Kenya are found in the bottom left corner of the array. There are also rainfall events over Kenya to the right of the array, which represent more widespread rainfall over Kenya. The top right hand corner of the array is the location for daily rainfall events located on the border of Uganda, Ethiopia and Kenya. Rainfall events over southern Ethiopia are located in the bottom right corner of the array.

5.1.2 Comparison of daily rainfall states in the TRMM and ERA-40 SOMs

In order to examine whether circulation parameters of daily rainfall events in the ERA-Interim dataset for TRMM rainfall are the same as for the ERA-40 dataset as discussed in Chapter 4, SOM nodes similar to those selected as key for ERA-40 rainfall are identified and the accompanying circulation patterns examined.

TRMM node 5 1

The first node to be examined for the ERA-40 constrained domain rainfall SOM was the node 0 0 (top left corner). The equivalent node in the TRMM constrained domain normalised SOM is node 5 1 (sixth column across, second row). This node is characterised by a bi-centred region of rainfall with the two centres located over northern Uganda and to the north-east in southern Ethiopia. The strength of rainfall is reduced in the TRMM SOM in comparison with the ERA-40 node, but the spatial distribution of rainfall for daily events assigned to this node is very similar, although extended slightly further into Kenya in the TRMM SOM.

TRMM node 6 2

The second node to be examined in the ERA-40 SOM was node 0 2. A similar rainfall pattern is found in the TRMM SOM for node 6 2 (rightmost column, third row down). The extension of rainfall associated with this node is much greater in the TRMM dataset than in the ERA-40 dataset, but the two nodes share the same location for most intense rainfall. This centre is located at the border of Kenya and Ethiopia. Both nodes show an extension of rainfall across southern Ethiopia associated with this pattern, but the node is characterised by much heavier rainfall in the TRMM SOM.

TRMM node 1 0

Node 1 0 (top row, second from left) in the TRMM SOM is equivalent to node 6 0 in the ERA-40 SOM. This node is characterised by extensive Tanzanian rainfall. The heaviest rainfall is centred over central south Tanzania, but rainfall can be found over the whole of the country, also infringing into the Indian Ocean and southern Kenya. Node 6 1 in the ERA-40 rainfall SOM array had been selected due to it representing the most intense daily rainfall events found in the ERA-40 data series. In the TRMM dataset, the most intense

rainfall events are again found over Tanzania in node 1 0. The region of heaviest rainfall in this node was found to the south of the country.

TRMM node 2 3

Node 6 4 in the ERA-40 SOM was linked to an oceanic rainfall centre offshore of Tanzania. There is no direct equivalent in the TRMM SOM array, where oceanic rainfall is not neatly compartmentalised with a clear centre. The node 2 3 (third column from left, second row from bottom) has widespread oceanic rainfall in conjunction with the driest conditions over land for this array and so this node is examined to look at circulation patterns associated with oceanic rainfall in the TRMM dataset.

TRMM node 1 4

Node 0 4 in the ERA-40 dataset was characterised by rainfall events centred over the eastern Kenyan border with Somalia. The equivalent node in the TRMM SOM is node 1 4 (bottom row, second from left). Days associated with this node are characterised by an intense centre of rainfall co-located with the rainfall in the 0 4 ERA-40 node, with a very similar spread of rainfall over the eastern Kenya, Somalia and coastal Indian Ocean regions.

TRMM node 6 3

The final node to be examined from the TRMM dataset is a dry node (equivalent to node 3 4 in the ERA-40 SOM). As previously mentioned, the suppression of convection leading to dry conditions is prevalent in the ERA-40 SOM with many of the central lower nodes of the array featuring this. However, in the TRMM SOM widespread dry conditions over the East African domain do not feature to the same extent and there are no nodes that are dry over the whole land area. For this node, the node 6 3 (right most column, second row from

bottom) is examined. This node has been selected due to the suppression of convection over three of the four focus countries, namely Tanzania, Kenya and Uganda and it is the conditions of circulation indicators of these three countries which are evaluated.

To summarise, six nodes from the TRMM SOM have been selected for their similarities to the examined nodes from the ERA-40 rainfall SOM and the accompanying circulation patterns for these selected nodes are now analysed. The six nodes chosen from the TRMM SOM array, to reiterate, are node 5 1 (bi-centred rainfall in Uganda and Ethiopia), node 6 2 (rainfall centred at the border of Kenya and Ethiopia), 1 0 (extensive and intense Tanzanian rainfall), 2 3 (Indian Ocean rainfall), node 1 4 (rainfall on the Kenyan border with Somalia) and finally node 6 3 with dry conditions over the southern half of the continental domain.

5.1.3 Atmospheric circulation parameters associated with TRMM daily rainfall types

As in Chapter 4, Section 4.2.3 it is important to examine circulation parameters in other variables for different types of rainfall events over East Africa during the short rains. This provides a test of the classification methodology in a second dataset, namely to establish whether the spatially clustered rainfall events appear to have a physical basis in an independent dataset.

The atmospheric state is averaged over the days assigned to each node of the thirty-five state rainfall array. For the TRMM series, the circulation parameters are calculated from the moisture flux, zonal winds, specific humidity and geopotential height fields at the 850hPa level of the ERA-Interim reanalysis dataset. This is necessary to have concurrent date series between the TRMM and ERA-Interim series, which is not possible using the ERA-40 series. By using a dataset independent from TRMM to examine the state of the circulation for each clustered rainfall event, a more stringent test to establish the

parameters associated with rainfall or dry conditions over East Africa during the short rains season is introduced. Analysis is focused on the nodes outlined in Section 5.1.2 as representing the key event types over East Africa during the short rains. Figures 5.2 to 5.7 show the mean state for the ERA-Interim variables for each node representative of a rainfall event over the domain.

Node 5 1

Mean and mean anomaly states for moisture flux, moisture flux divergence and mean anomaly fields for specific humidity, zonal wind and geopotential height at the 850hPa level for node 5 1, characterised by a bi-centred rainfall event across Uganda and southern Ethiopia are shown in Figure 5.2 (a) to (g). The primary feature of the ERA-Interim plots for this node is the anomalous moisture flux convergence over East Africa, north of 3°S, with the peak convergence co-located with the rainfall centre present in this node. The moisture flow plots highlight anomalous moisture flow originating in the north-west and converging with southerly flow over the rainfall centre. The moisture flux fields indicative of the rainfall centre are supported through the anomaly fields for specific humidity, zonal winds, and to some extent geopotential height. Notable, for the consistency between the ERA-Interim fields, is the slight displacement of the circulation parameters of rainfall further east of the centre depicted in the TRMM dataset. Slight differences are to be expected, nonetheless it is reassuring that the ERA-Interim fields support the rainfall location identified by the SOM clustering method for the 5 1 node. As is Chapter 4, units for the accompanying variables are g kg m s^{-1} (moisture flux), g kg m^{-1} (moisture flux divergence), m/s (zonal winds), m (geopotential height) and kg kg^{-1} (specific humidity). These units remain the same for all analysis of circulation parameters throughout the thesis.

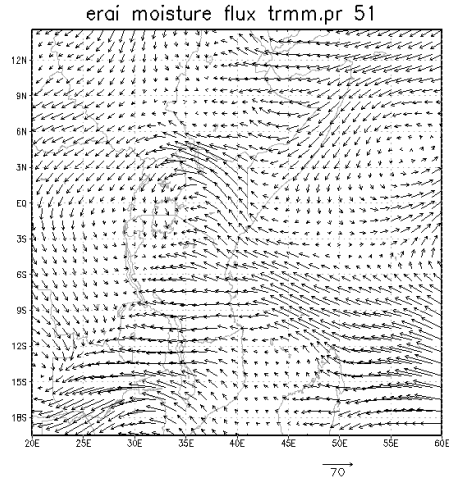


Figure 5.2 (a) Moisture flux mean for node 5 1

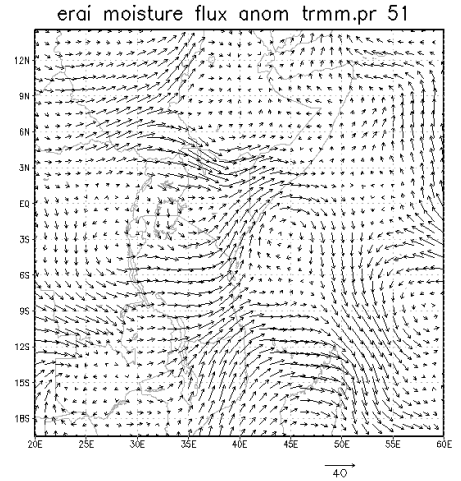


Figure 5.2 (b) Moisture flux anomaly for node 5 1

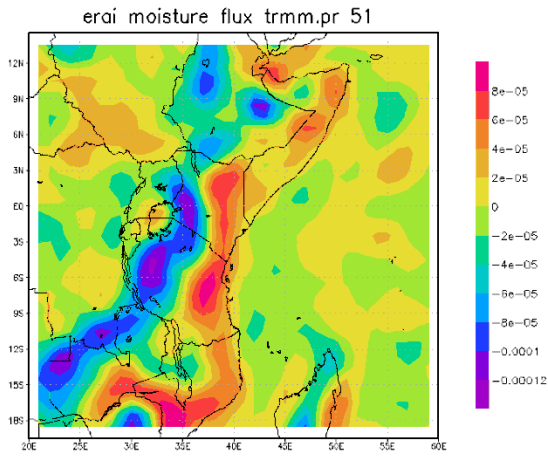


Figure 5.2 (c) Moisture flux divergence mean for node 5 1

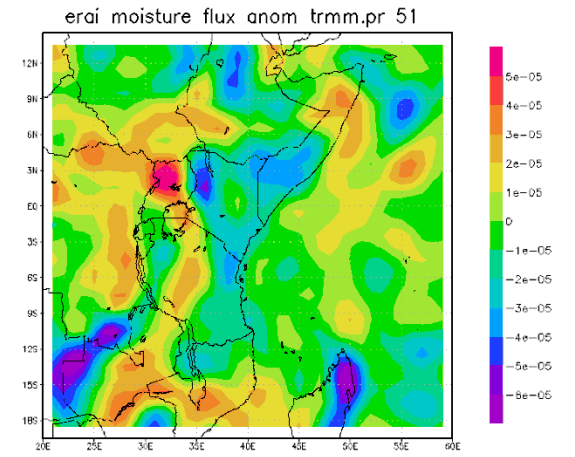


Figure 5.2 (d) Moisture flux divergence anomaly for node 5 1

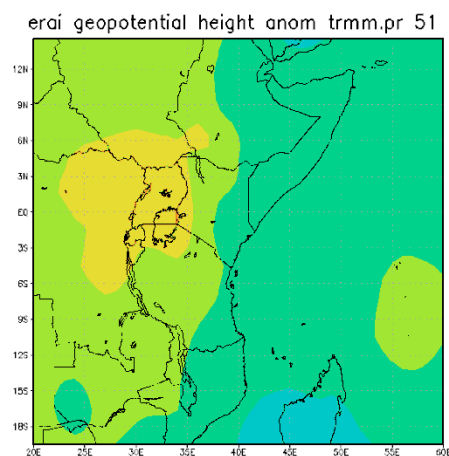


Figure 5.2 (e) Geopotential height anomaly for node 5 1

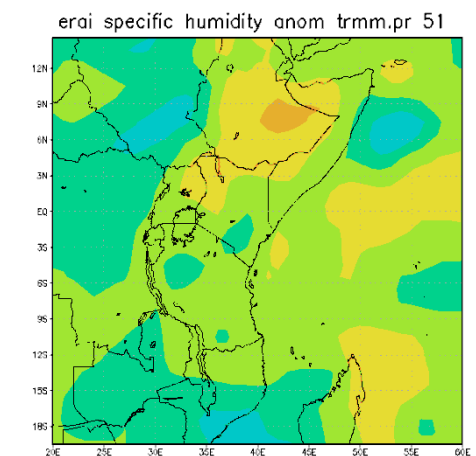


Figure 5.2 (f) Specific humidity anomaly for node 5 1

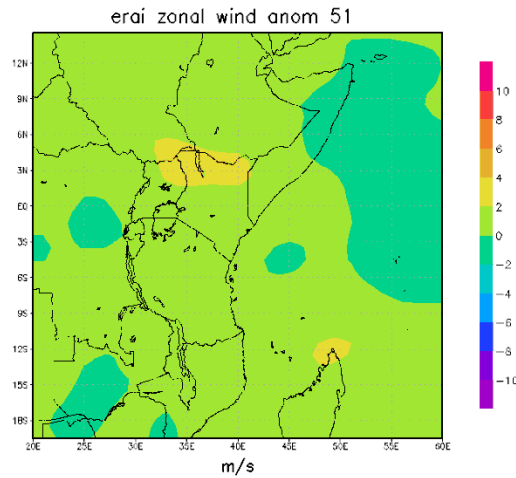


Figure 5.2 (g) Zonal wind anomaly for node 5 1

Node 6 2

Figures 5.3 (a) to (g) show the mean and anomaly states for days associated with the node 6 2, representing daily rainfall events centred across the eastern borders of Kenya and Ethiopia but also covering the border further inland as well. The ERA-Interim fields for this node are characterised by anomalous moisture flux convergence maxima near the rainfall centre, with anomalous divergence further south in East Africa, specifically over eastern Tanzania. Anomalous moisture flux from the Indian Ocean onshore through the rainfall location is present and the anomalous zonal wind and specific humidity fields also support the rainfall centre over the Kenyan-Ethiopian border. Again, the ERA-Interim fields are slightly displaced from the TRMM rainfall centre whilst still clearly supporting the presence of rainfall events at the location depicted by the clustering of the TRMM rainfall.

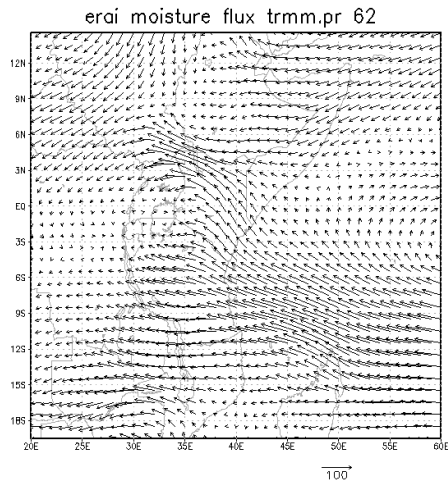


Figure 5.3 (a) Moisture flux mean for node 6 2

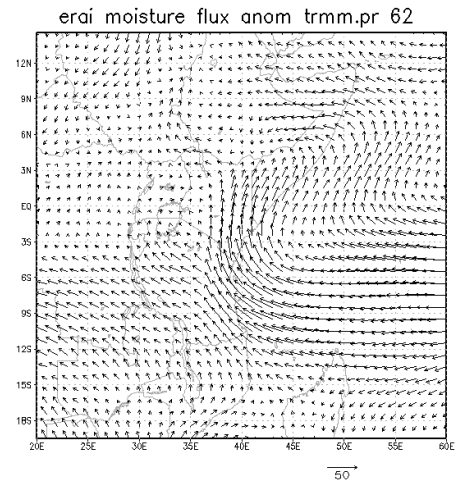


Figure 5.3 (b) Moisture flux anomaly for node 6 2

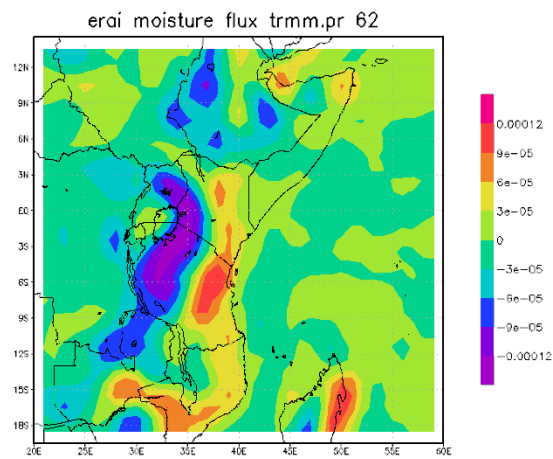


Figure 5.3 (c) Moisture flux divergence mean for node 6 2

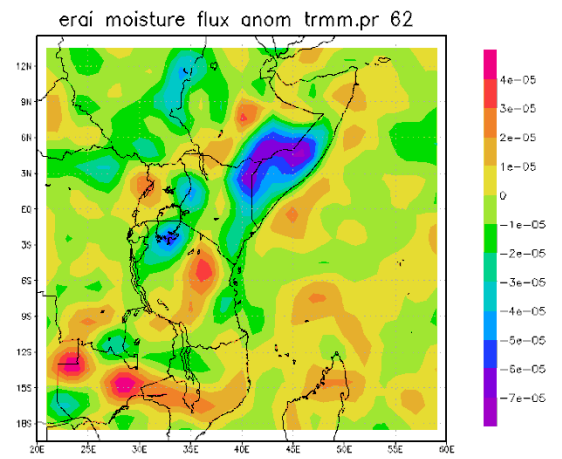


Figure 5.3 (d) Moisture flux divergence anomaly for node 6 2

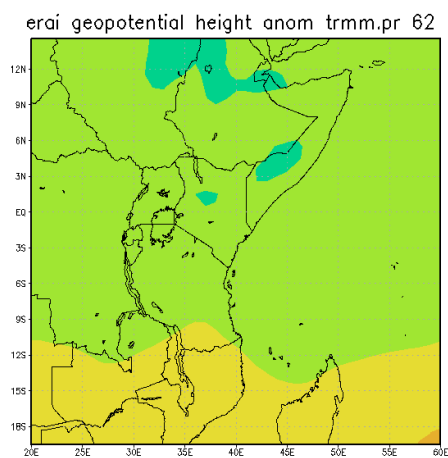


Figure 5.3 (e) Geopotential height anomaly for node 6 2

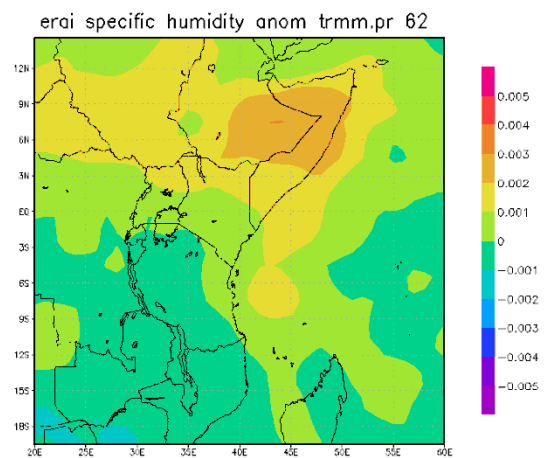


Figure 5.3 (f) Specific humidity anomaly for node 6 2

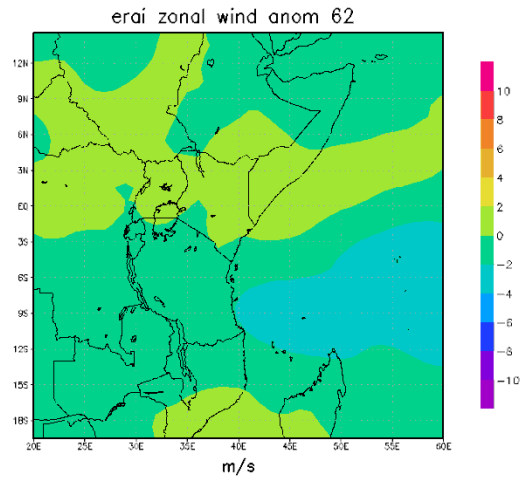


Figure 5.3 (g) Zonal wind anomaly for node 6 2

Node 1 0

Extensive and intense Tanzanian rainfall is located in node 1 0 of the TRMM SOM array, with the accompanying mean and mean anomaly ERA-Interim fields shown in Figures 5.4 (a) to (g). Characteristic of this node is the convergence of two anomalous moisture flows, originating in the north-east and from the Congo respectively, over eastern Tanzania, co-located with the most intense rainfall. This pattern is reinforced through the presence of anomalous moisture flux convergence over eastern Tanzania. Westerly wind anomalies in this region and a positive specific humidity anomaly reinforce the patterns highlighted in the moisture flux field for the occurrence of Tanzanian rainfall centres for days assigned to this node.

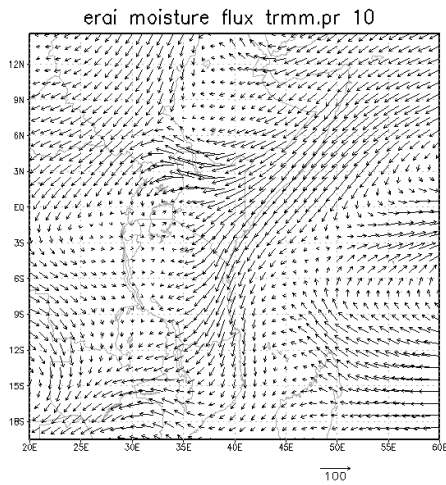


Figure 5.4 (a) Moisture flux mean for node 1 0

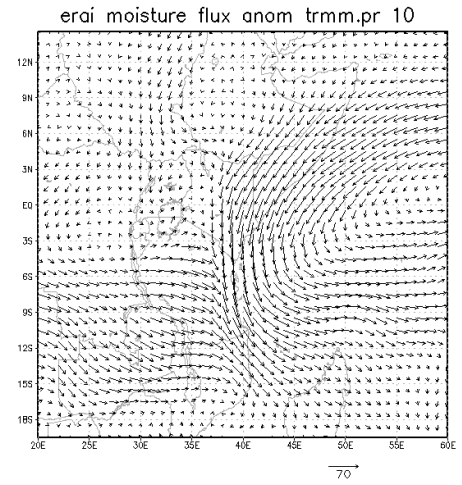


Figure 5.4 (b) Moisture flux anomaly for node 1 0

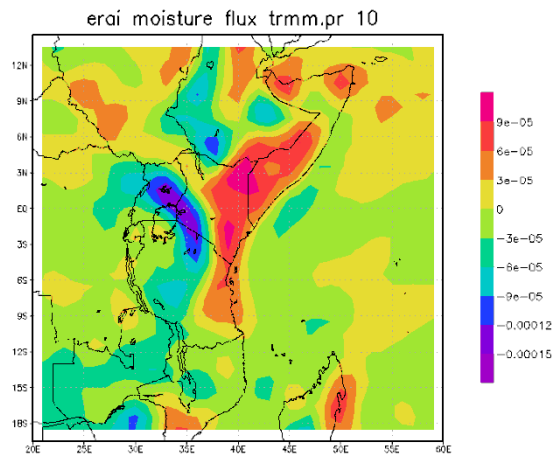


Figure 5.4 (c) Moisture flux divergence mean for node 1 0

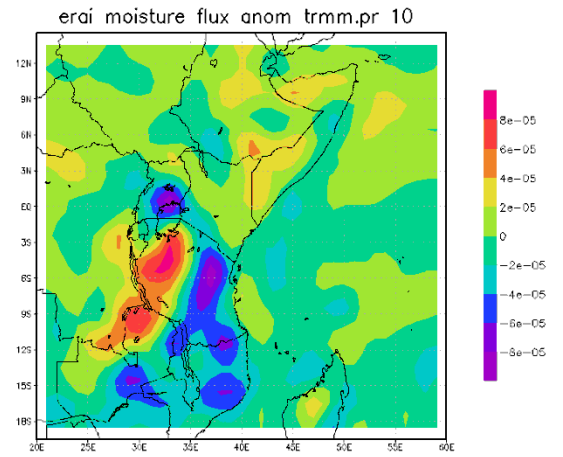


Figure 5.4 (d) Moisture flux divergence anomaly for node 1 0

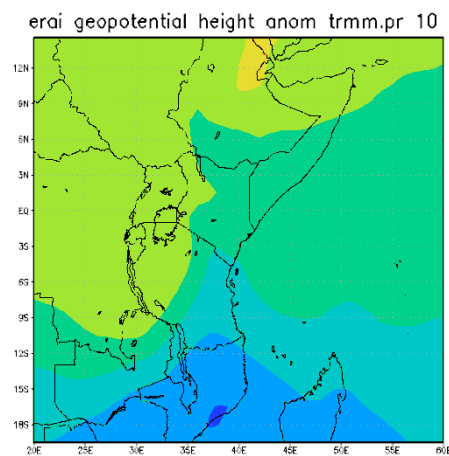


Figure 5.4 (e) Geopotential height anomaly for node 1 0

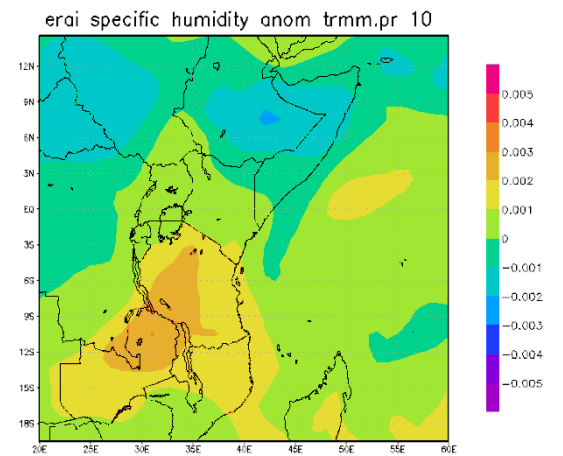


Figure 5.4 (f) Specific humidity anomaly for node 1 0

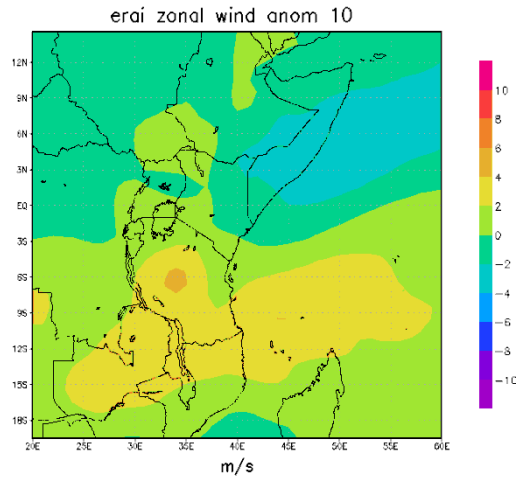


Figure 5.4 (g) Zonal wind anomaly for node 1 0

Node 2 3

Figures 5.5 (a) to (g) shows ERA-Interim fields associated with days assigned to node 2 3 of the TRMM normalised rainfall SOM, characterised by widespread Indian Ocean rainfall, with dry conditions over land, specifically for the west of the East African domain. For this node, the key moisture flux divergence feature is not the presence of a strong moisture flux convergence anomaly over the Indian Ocean but instead the presence of strong moisture flux divergence along the entirety of coastal East Africa. This ‘blocking’ of moisture laden flow onshore can be highlighted in the moisture flux plots; the anomaly plot clearly identifies anomalous moisture occurring along the East African coastline and across to the Arabian Peninsula. There is also moisture flow towards the Congo. For days assigned to this node, there are no clear patterns identifiable in the other variables of the ERA-40 dataset. This reinforces earlier conclusions that it is the specific combination of wind and humidity fields contained in the moisture flux divergence field that is the most useful indicator of rainfall centres over East Africa during the short rains.

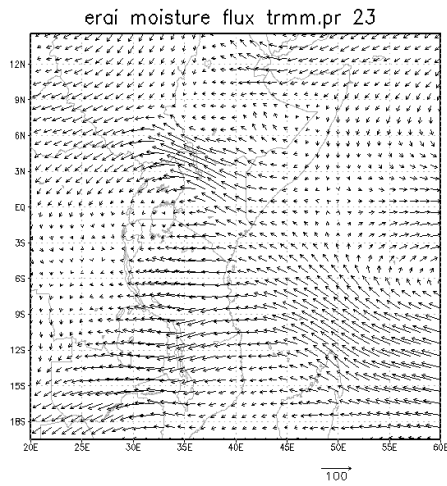


Figure 5.5 (a) Moisture flux mean for node 2 3

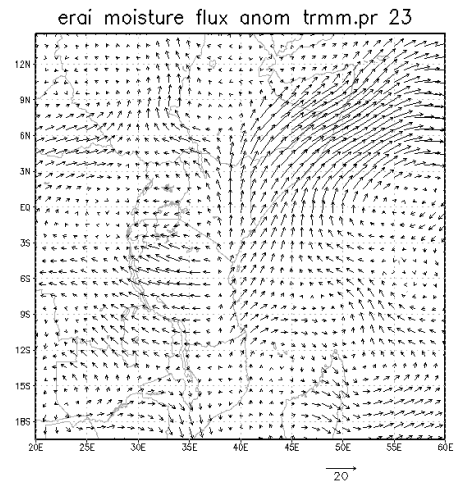


Figure 5.5 (b) Moisture flux anomaly for node 2 3

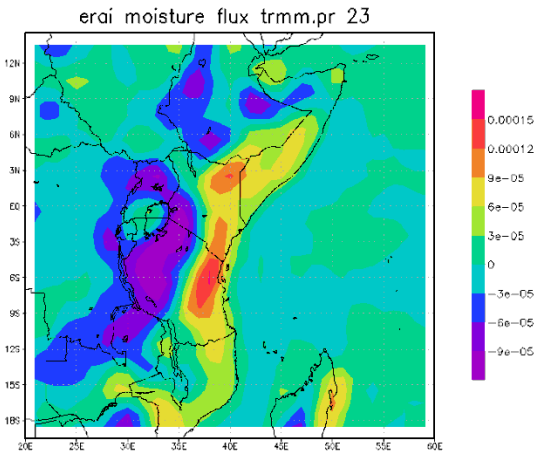


Figure 5.5 (c) Moisture flux divergence mean for node 2 3

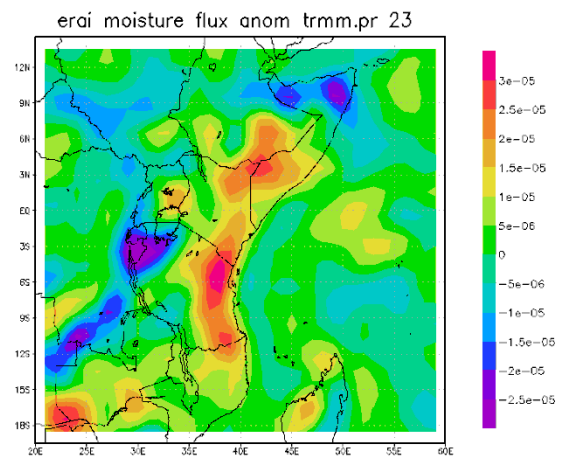


Figure 5.5 (d) Moisture flux divergence anomaly for node 2 3

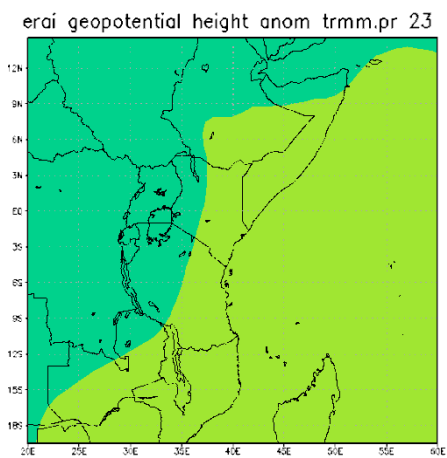


Figure 5.5 (e) Geopotential height anomaly for node 2 3

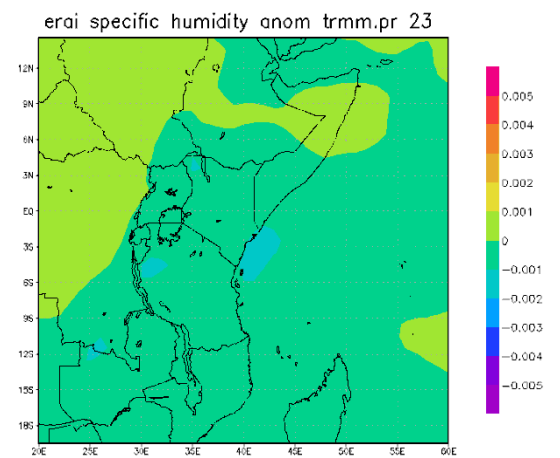


Figure 5.5 (f) Specific humidity anomaly for node 2 3

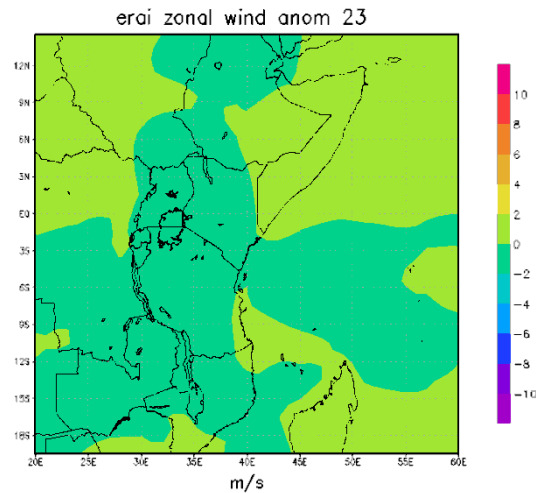


Figure 5.5 (g) Zonal wind anomaly for node 2 3

Node 1 4

Node 1 4 is characterised by a coastal Kenyan rainfall centre. Figures 5.6 (a) to (g) show the mean and mean anomaly ERA-Interim fields for days assigned to this node. There is a moisture flux convergence anomaly co-located with the region of rainfall, reinforced by the convergence of onshore moisture flow, originating from both the north and south, at the Kenyan coastline. There is evidence of a small positive specific humidity anomaly over Kenya for these days, but again it is the moisture flux field that is most representative of the TRMM rainfall events.

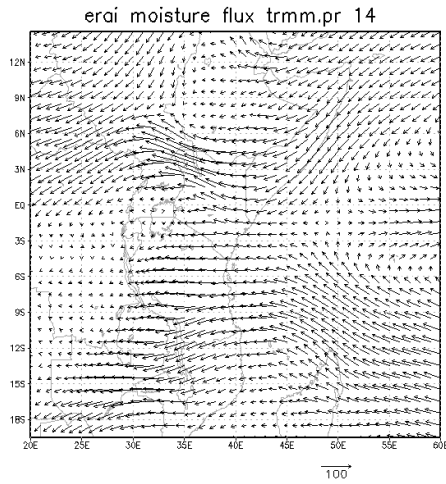


Figure 5.6 (a) Moisture flux mean for node 1 4

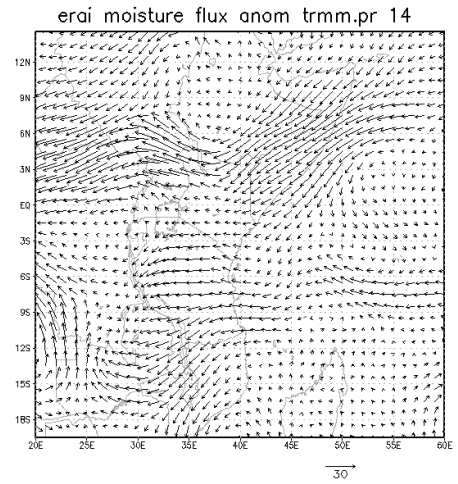


Figure 5.6 (b) Moisture flux anomaly for node 1 4

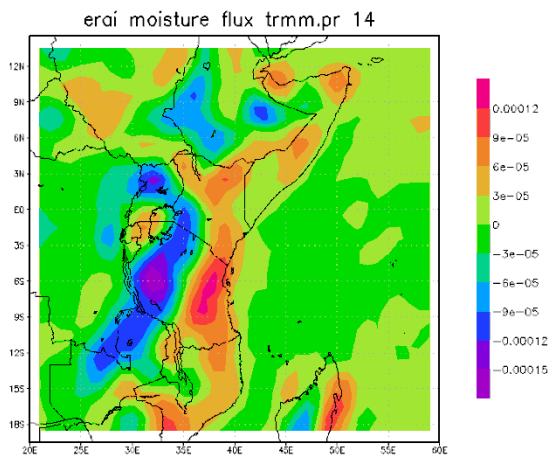


Figure 5.6 (c) Moisture flux divergence mean for node 1 4

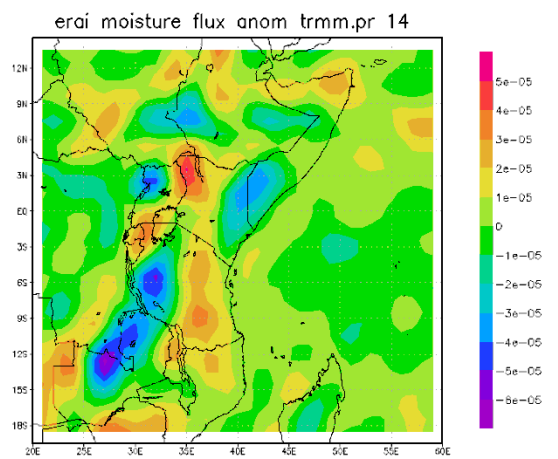


Figure 5.6 (d) Moisture flux divergence anomaly for node 1 4

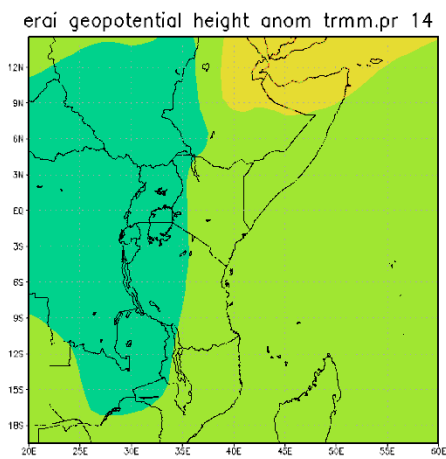


Figure 5.6 (e) Geopotential height anomaly for node 1 4

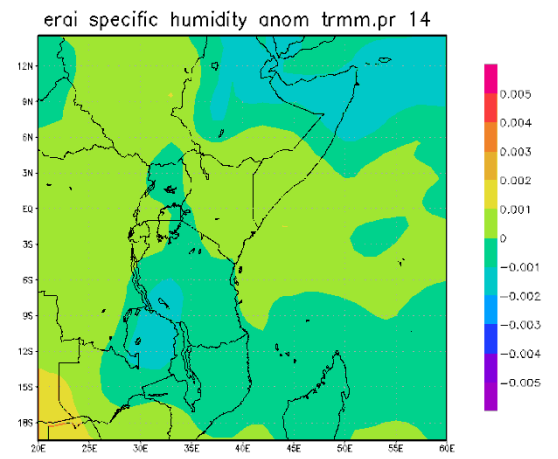


Figure 5.6 (f) Specific humidity divergence anomaly for node 1 4

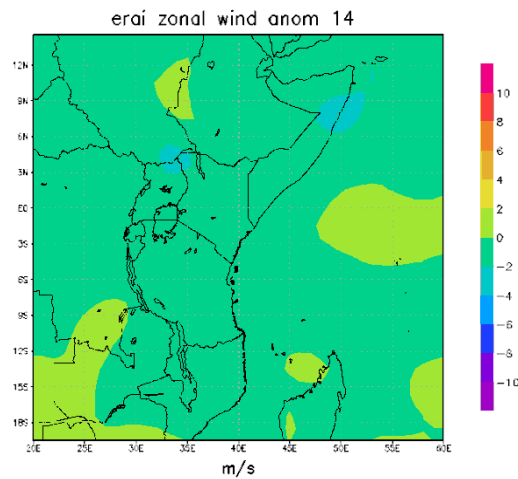


Figure 5.6 (g) Zonal wind anomaly for node 1 4

Node 6 3

Whilst there are no overwhelmingly dry nodes in the TRMM daily series as in the ERA-40 dataset, node 6 3 does see suppressed convection over three of the four focus countries of the East African domain; Uganda, Kenya and Tanzania. The associated fields for the ERA-Interim dataset are shown in Figures 5.7 (a) to (g). In this instance, the ERA-Interim fields are not as indicative of the TRMM rainfall conditions as for other nodes. There is evidence that the ERA-Interim fields are more indicative of the rainfall centre over Ethiopia, for example through the anomalous convergence in this region, than they are of the dry conditions further south in the domain.

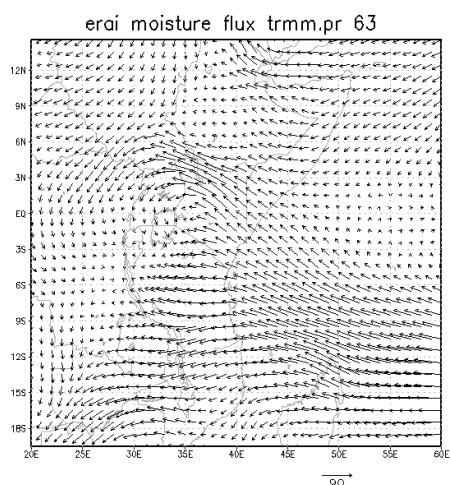


Figure 5.7 (a) Moisture flux mean for node 6 3

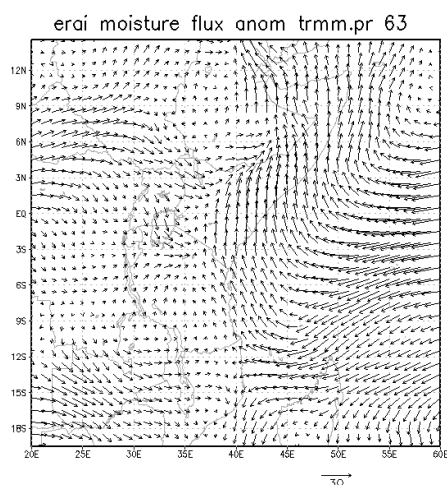


Figure 5.7 (b) Moisture flux anomaly for node 6 3

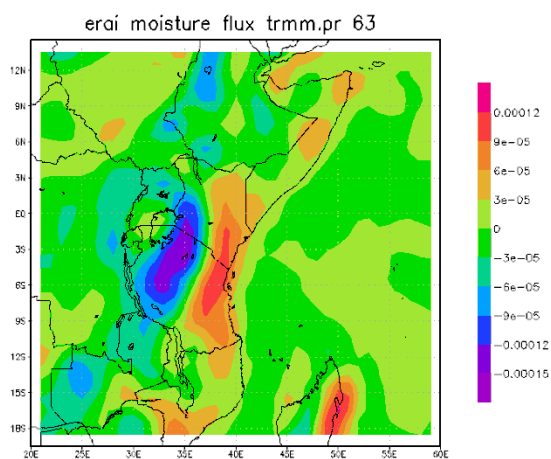


Figure 5.7 (c) Moisture flux divergence mean for node 6 3

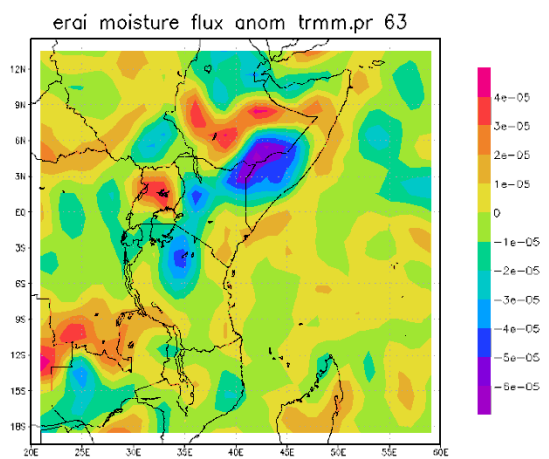


Figure 5.7 (d) Moisture flux divergence anomaly for node 6 3

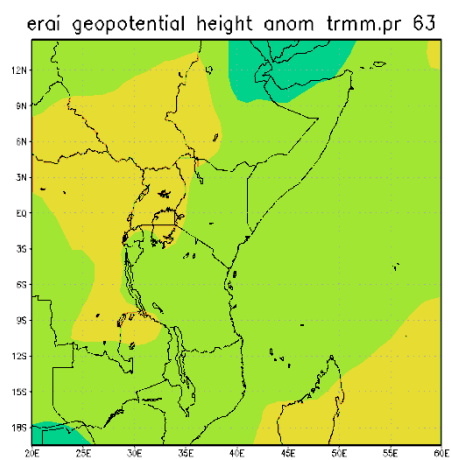


Figure 5.7 (e) Geopotential height anomaly for node 6 3

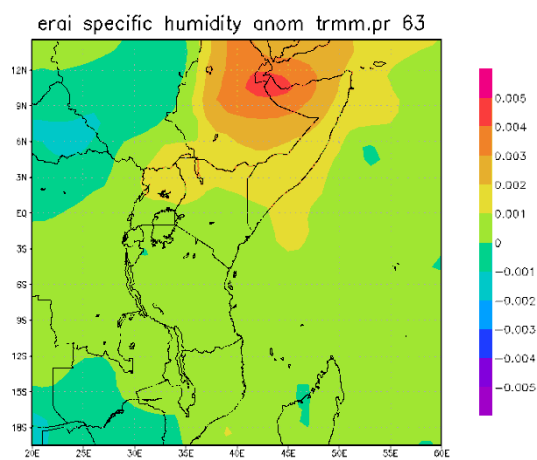


Figure 5.7 (f) Specific humidity anomaly for node 6 3

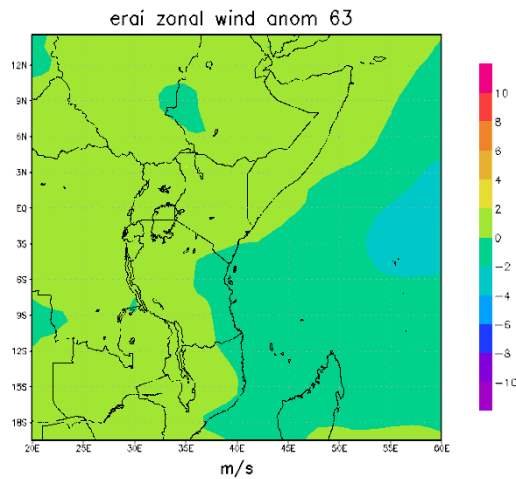


Figure 5.7 (g) Zonal wind anomaly for node 6 3

In conclusion, there is good spatial agreement between the TRMM daily rainfall clusters and the ERA-Interim daily circulation fields during the short rains from 1998-2010. The use of two independent datasets for analysis reinforces that the moisture flux convergence field is in strong agreement with TRMM rainfall; without exception anomalous moisture flux convergence indicates the presence of a co-located rainfall centre. In the ERA-40 dataset, anomalous moisture flux divergence is a very strong feature of dry nodes of suppressed convection. Such nodes have not been clustered as a feature of the TRMM SOM and so it has been difficult to assess this feature here. Whilst specific humidity, geopotential height and zonal wind anomalies provide extra support for the clustering of rainfall events to specific nodes, the results are not as clear as for moisture flux. The focus of analysis of daily rainfall moving forwards in the research will therefore be on moisture flux, specifically for the analysis of daily AOGCM data in Chapter 6. It is the water vapour provided by the moisture flux field that is a key stage in the generation of rainfall events as asserted in Chapter 4, Section 4.2. This result is in agreement with other studies over differing domains such as the Congo (Washington et al, 2013) and the Amazon (Yin et al 2013).

5.2 Comparison of ERA-40 and TRMM daily rainfall events

The overall objective of this thesis is to investigate the fidelity of rain producing systems in coupled climate models. Furthermore, the ERA-40 dataset, whilst aiming to reproduce historical climate, is also model generated, albeit with the input of observations every six hours. Therefore, a key point of investigation is to look at how well the ERA40 dataset reproduces daily rainfall events seen in a dataset from a different source, that is, the TRMM satellite rainfall being evaluated here. The extent to which a model experiment designed to reproduce the observed climate, as in the case of the reanalysis experiments, can accurately achieve the spatial characteristics and magnitude of daily rainfall events in East Africa provides insight into the potential ability of AOGCMs to simulate daily rainfall events and how they may change under increased greenhouse gas conditions. This evaluation of the ERA-40 rainfall series is done by comparing the clustering of daily rainfall events in the SOMs for the reanalysis dataset, ERA-40, and the TRMM satellite rainfall dataset. The two datasets cover different time periods; the ERA-40 dataset covers the period 1971-2000 and the TRMM rainfall dataset covers the period 1998-2010 as previously explained. Whilst the overlap three years (1998-2000) are examined for each dataset, the types produced over the whole climatology period and their similarity to each other are also examined. Both datasets have had strengths and weaknesses examined in the scientific literature, with these previous findings summarised in Chapter 2, Sections 2.1 and 2.5.

5.2.1 Daily rainfall events; correlation between ERA-40 and TRMM

The first comparative evaluation is undertaken on the climatology period for each of the datasets; 1971-2000 for ERA-40 and 1998-2010 for TRMM. As a first measure of the statistical similarity between daily rainfall occurrences in the two datasets, the Pearson's

product moment correlation coefficient is calculated for each node in the ERA-40 SOM to each node in the TRMM SOM. Each correlation calculation consists of the pair of SOM values making up the node map. The total number of correlations computed is 1225. Each node in the TRMM dataset is statistically significantly correlated to at least one node in the ERA-40 dataset when the confidence level set at 99%. This indicates a high degree of similarity in spatial patterns and normalised amplitude of the daily rainfall events in each dataset, although it does not provide any information about the frequency of occurrence of different types of daily rainfall events over East Africa in the ERA-40 and TRMM datasets.

For each of the key ERA-40 nodes identified in Chapter 4, Section 4.2.3, a similar node in the TRMM dataset was selected as showing a visually comparable daily rainfall characterisation in terms of spatial pattern and amplitude. In order to examine if these visually similar nodes are also statistically similar, the most highly correlated node to each of the key ERA-40 nodes is established, with the results shown in Table 5.1.

Table 5.1 Highest correlation of key ERA-40 SOM nodes with a node in the TRMM SOM, normalised rainfall series

ERA-40 Key Node	Linked TRMM Node	Correlation
0 0	5 2	0.709
0 2	6 2	0.730
6 0	1 0	0.878
6 4	0 2	0.873
0 4	1 4	0.763
3 4	2 4	0.400

For four of the six nodes, the most closely related TRMM node is the same as that selected as being most visually similar during the initial evaluation of the TRMM SOM in comparison to the ERA-40 rainfall SOM. The two exceptions are ERA-40 nodes 0 0 and 3

4. Node 0 0 in the ERA-40 SOM is most highly correlated with TRMM node 5 2. From a visual inspection node 5 1 was selected, this equally has a statistically significant correlation of $r = 0.615$ but is not as highly correlated as node 5 2 ($r = 0.709$). The second exception is node 3 4 of the ERA-40 dataset. This represents the suppression of convection over the domain, with such widespread dry conditions not present in the TRMM SOM to the same extent. Through a visual examination this had been countered through selecting a node with suppression of convection over the focus countries of Kenya, Tanzania and Uganda but it transpires that this node has a correlation of only $r = 0.250$ with the dry ERA-40 SOM node. The node with the highest correlation to the dry ERA-40 node is node 2 4, although the correlation measures only $r = 0.400$.

Key similarities and differences in the clustering of daily rainfall events between the two datasets can be identified, increasing understanding in the different representations of daily rainfall in a model and satellite rainfall dataset. In terms of differences, the ERA-40 rainfall SOM array is characterised by more clearly defined daily rainfall events, whereas TRMM sees daily rainfall events represented over more of a continuum. Rainfall in the TRMM SOM is more expansive in a greater number of nodes and spatial configurations with the widespread suppression of convection seen in the ERA-40 SOM not present. There is a relatively low correlation between the very dry 3 4 ERA-40 SOM node and the most closely related TRMM node over the same time period, node 2 4. A further difference between the two SOM arrays is the presence of daily rainfall events in the Indian Ocean off coastal Tanzania in the ERA-40 SOM which cannot be seen in the TRMM rainfall SOM. The TRMM SOM has daily rainfall events centred over Uganda and also oceanic rainfall centres in the north of the domain encroaching significantly over the land region.

There are similarities in how the two datasets represent daily East African rainfall during their respective climatology period with regards to the location within the domain of distinct daily rainfall events. Daily events represented in both SOM arrays include rainfall occurrences over the Ugandan-Ethiopian border, in both western and eastern Tanzania, over coastal Kenya, over southern Ethiopia and into Uganda and down the western side of the domain.

5.3 Daily rainfall events 1998-2000; comparison between ERA-40 and TRMM

5.3.1 Daily states in the ERA-40 and TRMM rainfall datasets

In order to undertake a more comparable analysis of the two datasets a common period of 1998-2000, October to December, is established. This 276 day period allows direct comparison of the daily rainfall events in each of the datasets. Figure 5.8 shows the best representation of the ERA-40 7 by 5 rainfall SOM associated with reference vector 13. As with the longer climatology period SOM, some key nodes and characteristic daily rainfall events over the domain can be identified.

Node 3 4 – This node is characterised by bi-centred rainfall over Uganda and southern Ethiopia. The pattern is similar to the node 0 0 of the thirty year climatology SOM, although over the shorter three year period the rainfall centre over Uganda is of noticeably greater amplitude than over Ethiopia.

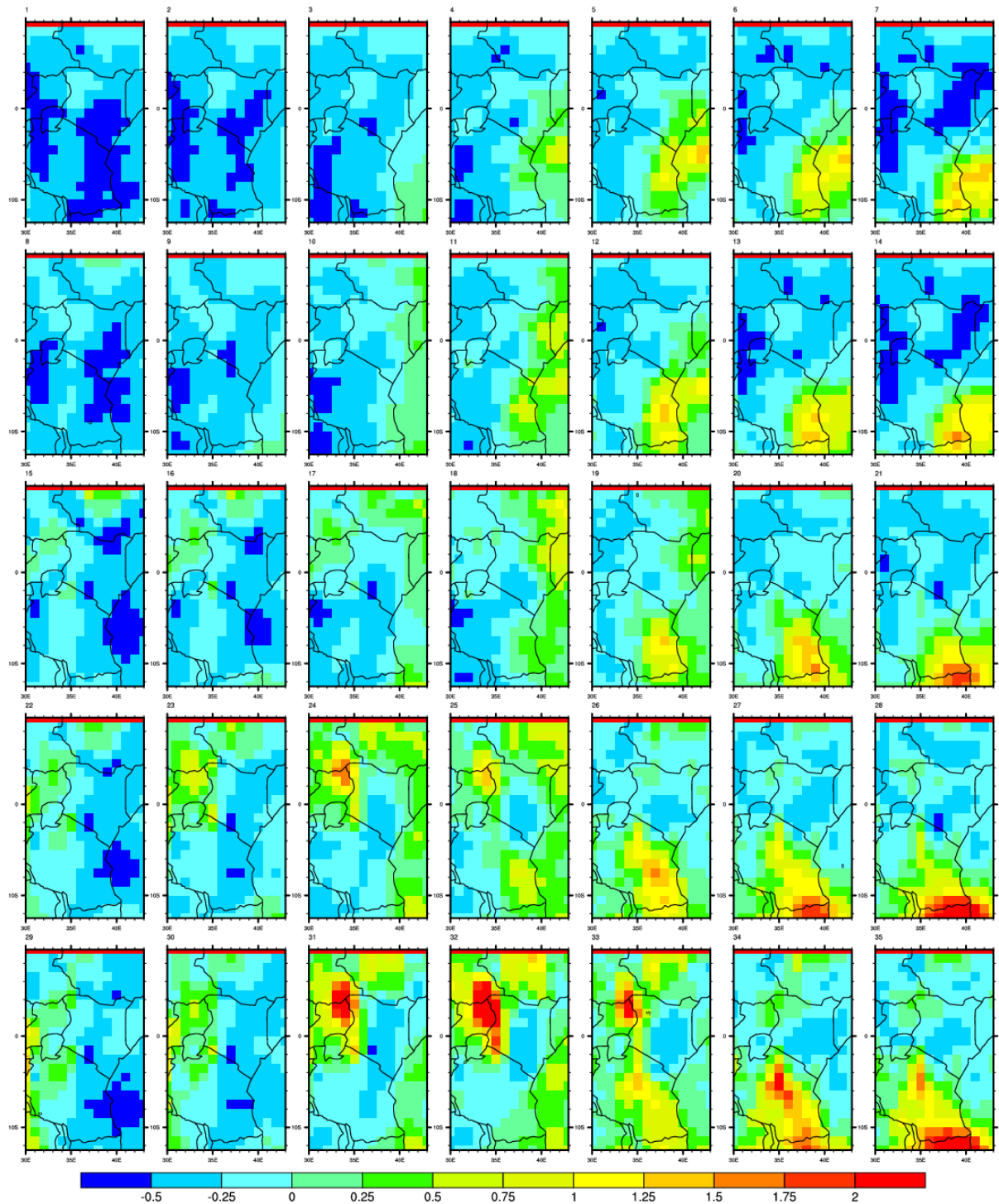


Figure 5.8 ERA-40 normalised rainfall SOM, 1998-2000, reference vector 13.

Node 3 2 – This node has rainfall centred over Ethiopia and along the East African coastline. The three year SOM has no node where daily rainfall was confined to Ethiopia with this the closest pattern.

Node 5 4 – Node 5 4 is characterised by rainfall over much of Tanzania, although it is less widespread than the extensive Tanzanian rainfall seen in the thirty year climatology.

Node 6 0 – Much of the land domain is characterised by suppressed convection in this node. There is a region of oceanic rainfall encroaching on coastal regions of Tanzania.

Node 6 4 – This is the node associated with the most intense daily rainfall events seen within the SOM array. These events are located over Tanzania as in the thirty year climatology SOM but here the most intense rainfall is located over a more localised region.

Node 0 0 – This is a dry node with widespread suppression of convection over the land domain to a greater extent than seen in the climatology SOM.

The highlighted nodes illustrate similar key patterns to those found during the thirty year climatology period, although with a difference in amplitude or spread of rainfall than during the thirty year period. However, the three year SOM has no states representing daily rainfall events in Kenya. This is not unexpected. Drought conditions in Kenya from 1998 to 2000 have been described as some of the worst that have been recorded (Mworia et al, 2008). Costs incurred during this prolonged drought are estimated at US\$2.8 billion, with livestock depletion measuring up to a third in some regions of Kenya (Kenya Ministry of Environment and Mineral Resources, 2013). Therefore, given the training algorithm of the SOM, which clusters states where there is most data occurring within the data space, it is unsurprising that when training data is limited to this three year period of the short rains season, no distinct rainfall events over Kenya are represented in the nodes of the SOM array.

Figure 5.9 shows the 7 by 5 TRMM SOM for the period 1998-2000 OND associated with reference vector 3. This array does cluster daily rainfall events in Kenya during this period,

although it is often as part of a widespread pattern of rainfall over much of the domain.

Daily rainfall in isolation over Kenya is not seen during the 1998-2000 period; this again reflects the interannual characteristics of rainfall over Kenya for this three year period.

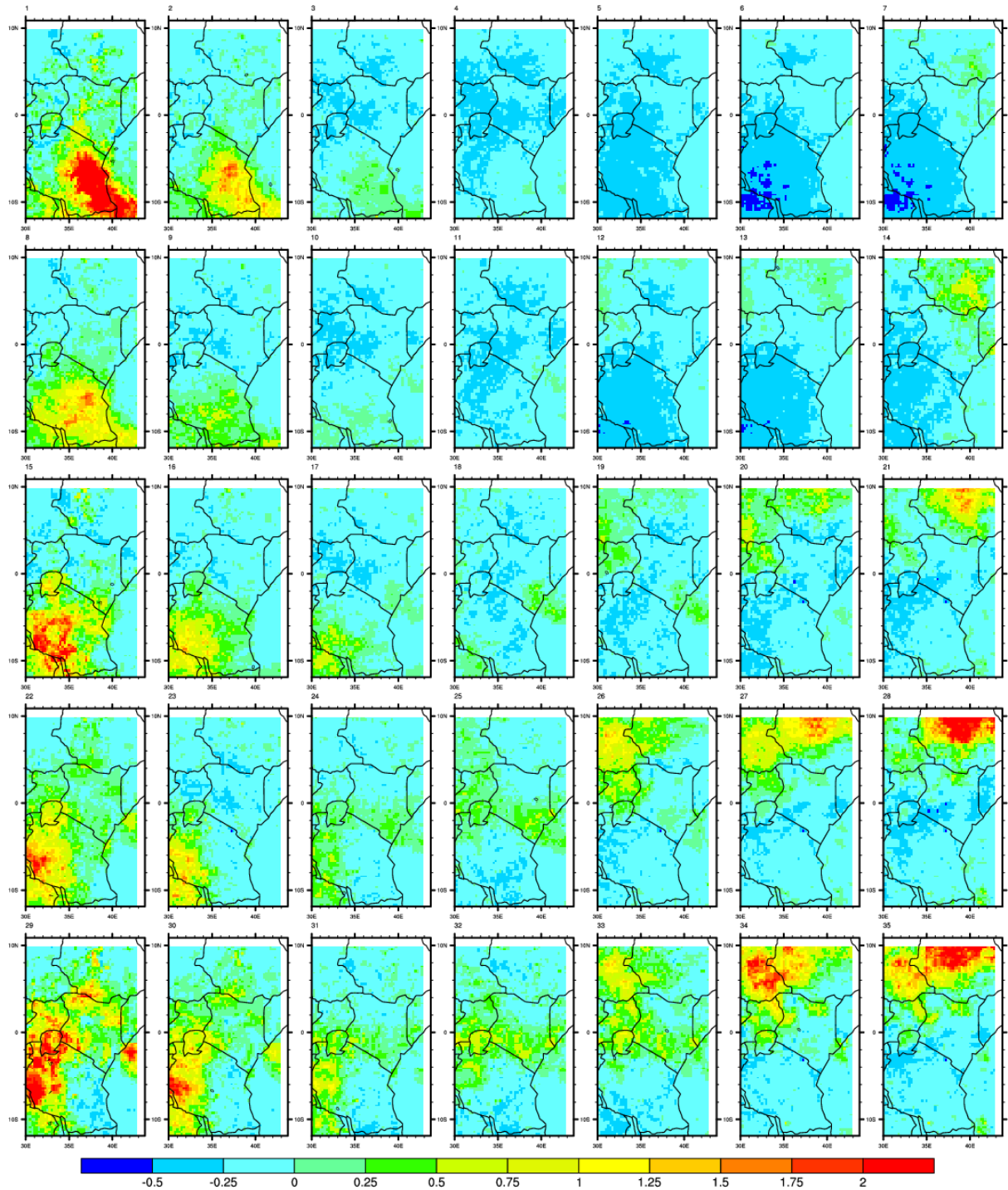


Figure 5.9 TRMM normalised rainfall SOM, 1998-2000, reference vector 3

As in Section 5.2, each node in the ERA-40 SOM array is correlated to each node in the TRMM SOM array, with the results shown in Table 5.2.

Table 5.2 Most closely correlated TRMM nodes to the ERA-40 key states, 1998-200.

ERA-40 Key Node	Linked TRMM Node	Correlation
3 4	4 3	0.663
3 2	6 0	0.708
5 4	0 1	0.708
6 0	3 0	0.633
6 4	0 1	0.714
0 0	6 4	0.439

Table 5.2 illustrates the close relationship between daily rainfall events in the TRMM and ERA-40 datasets during the 1998-2000 period. Given that these SOMs represent a common period, if the ERA-40 model produces rainfall that is a good approximation for the observed daily rainfall over East Africa then it might be expected to retain similarities with the TRMM rainfall over a common period. The SOM arrays for both datasets contain nodes representing daily rainfall events over Tanzania, in particular in a NW to SE diagonal leading out into the Indian Ocean. There are also nodes with daily rainfall events over the Uganda, South Sudan and Ethiopia region, and along the western sector of the domain in conjunction with suppression of convection over eastern Tanzania. As previously explained, neither dataset sees a node in the three year SOM array representative of daily rainfall events over Kenya, attributable to the drought conditions during this period in Kenya.

There are also differences in the two rainfall classifications despite the complete overlap in time period. The ERA-40 SOM array has stronger suppression of convection than TRMM, which generally sees heavier daily rainfall events recorded. In conjunction with this more

intense rainfall, the TRMM SOM array has nodes representing different daily rainfall events including more events in western Tanzania, heavier events in southern Ethiopia and more rainfall recorded on the fringes of Ethiopia. One region where ERA-40 sees a greater representation of daily rainfall is the Indian Ocean.

5.3.2 Date matching of ERA-40 and TRMM daily rainfall states

The ERA-40 and TRMM datasets have been sub-sampled to cover the same dates in time, resulting in the selection of three years, 1998-2000 October to December, thus allowing a direct comparison of the dates. In order to compare the daily rainfall states directly, the ERA-40 data is forced on to the TRMM SOM array. Each input date of the ERA-40 data is compared to, and allocated, to the most closely matching TRMM SOM node. The allocations are compared to evaluate how many of the ERA-40 days are allocated to the same node within the SOM array as the equivalent TRMM day.

Over the three month period for three years, only 47 of the 276 days (17.03%) in the ERA-40 rainfall dataset are correctly matched to the TRMM SOM array. The distribution of these correct matches is shown in Figure 5.10. For many nodes there are no correctly matched dates; in particular for the lower half of the array where detailed, less spatially widespread events are located in all regions of the domain. The highest proportion of matches occurs for nodes 40 and 50 (both have more than a 40% success rate at matching), these are nodes representing widespread oceanic rainfall that spreads across much of the land domain of East Africa.

0 0 (2) [25.00]	1 0 (0) [0.00]	2 0 (2) [28.57]	3 0 (7) [25.00]	4 0 (6) [42.86]	5 0 (15) [41.67]	6 0 (2) [28.50]
0 1 (2) [100.00]	1 1 (2) [15.38]	2 1 (0) [0.00]	3 1 (1) [14.29]	4 1 (1) [16.67]	5 1 (0) [0.00]	6 1 (1) [9.09]
0 2 (0) [0.00]	1 2 (0) [0.00]	2 2 (0) [0.00]	3 2 (0) [0.00]	4 2 (1) [20.00]	5 2 (1) [10.00]	6 2 (0) [0.00]
0 3 (0) [0.00]	1 3 (0) [0.00]	2 3 (0) [0.00]	3 3 (0) [0.00]	4 3 (0) [0.00]	5 3 (0) [0.00]	6 3 (0) [0.00]
0 4 (1) [11.11]	1 4 (0) [0.00]	2 4 (0) [0.00]	3 4 (1) [7.69]	4 4 (0) [0.00]	5 4 (2) [28.57]	6 4 (0) [0.00]

Figure 5.10 Distribution of the correct matches when ERA-40 rainfall series allocation is compared to the TRMM allocation to the TRMM normalised rainfall SOM 1998-2000 OND. Red is for no correct matches, orange for up to 10%, yellow up to 40% and green for more than 40% of correct matches.

Although it is possible the low match rate is due to fundamental differences between the representation of East African rainfall in the ERA-40 and TRMM datasets, there are also other possible reasons. The thirty five states of the 7 by 5 SOM cover a number of key patterns with the large number of states allowing detailed depiction of the small, but important, variations in spatial patterns and amplitude of the daily rainfall events. However, this raises the possibility that the ERA-40 dataset could be simulating daily rainfall events of largely the correct amplitude or spatial location but that subtle differences are leading them to be linked to similar, but adjacent nodes and therefore not being marked as a correct match. This could lead to a much lower percentage of correct

matches. In order to investigate this, closely related nodes are examined to allow a wider range of matches to be considered ‘correct’. This will allow a fuller examination of which observed rainfall states the ERA-40 can replicate.

5.3.3 Closely matched nodes in the TRMM and ERA-40 datasets

There are at least two alternative options for deciding how to define ‘closely related’ when selecting nodes classified as a successful match. The SOM algorithm allows for the arrangement of the most similar data types next to each other within the data array, so one method to select closely related nodes is to simply use the adjacent nodes. However, there are two disadvantages to this method. First, using each adjacent node will result in a different and arbitrary number of nodes being considered as correct matches depending on where in the array a particular node was located. For example, for the node 0 0, located in the top left corner of the array an additional three nodes will be considered as closely related, whereas for the more central 1 2 node an additional eight nodes will be considered closely related with there being no scientific reason to explain the disparity in extra matches considered correct. Furthermore, the nodes are not distributed evenly within the data space, with more nodes located in areas of dense data and less where data is sparse. This means that each daily rainfall type will not necessarily have the same number of nodes representing it and that adjacent nodes will not all represent the same event type.

Therefore, it has been decided to undertake the second option for defining closely matched nodes, which is to base the classification on correlation, with each node correlated with every other node in the 7 by 5 array. The level of correlation for statistical significance is relatively low and so it has been decided to set a more stringent test to examine the ability of the ERA-40 dataset to produce the same rainfall types as the TRMM rainfall dataset. Two levels of correlation are examined, $r = 0.75$ and $r = 0.5$. For both levels of correlation

nodes are deemed closely matched if they meet or exceed the value of correlation in question and there is no maximum or minimum set for the number of nodes that can be allocated the closely matched label.

Each node has at least one other node correlated to it when setting the correlation at 0.75 and the greatest number of additional nodes being considered as closely related is five when the correlation was set at 0.75 and nine when it is decreased to 0.5. Results of correct matches for the 7 by 5 SOM over the period OND 1998-2000 are shown in Table 5.3.

Table 5.3 Correct matches for ERA-40 data to TRMM SOM when expanding ‘correct’ nodes through correlation

	Exact Match	Node Correlation $r = 0.75$	Node Correlation $r = 0.5$
No. of Correct Matches	47	64	108
% of Correct Matches	17.02	23.19	39.13

Table 5.3 illustrates that expanding the definition of what is considered a success in the ERA-40 data in reproducing the same rainfall event type as the TRMM dataset does still not result in a particularly high success rate. Indeed with there still being a less than four in ten chance of each individual day in the ERA-40 dataset reproducing the same rainfall state as the TRMM dataset for that day there are clearly considerable differences in the model representation of rainfall over East Africa. Therefore, it has been decided that reducing the number of states in the SOM array could be useful in identifying the key differences in rainfall types represented in the two datasets more efficiently.

5.4 A 5 by 4 rainfall self-organising map over the East African array

A 5 by 4 SOM array has been selected as a more appropriate size to constrain further daily rainfall over East Africa whilst not compromising the emergence of key daily rainfall states within the array when compared with the 7 by 5 array. The SOM process is followed to establish four ‘best fit’ realisations over the two currently employed datasets; the ERA-40 and TRMM rainfall series. The ‘best fit’ realisation for each dataset for the overlap period of 1998-2000 is reference vector 8 for the ERA-40 dataset and reference vector 1 for the TRMM satellite dataset, when the realisation with the lowest quantisation error is selected for each daily series. Additionally the ‘best fit’ realisation of reference vector 1 is established for the 1971-2000 period for the ERA-40 dataset to be employed in the evaluation of AOGCM daily rainfall. The final realisation is reference vector 6 for the complete TRMM rainfall series covering 1998-2010.

5.4.1 A common period: 1998-2000 5 by 4 SOM realisations

Figure 5.11 shows the realisation of the 1998-2000 5 by 4 TRMM daily rainfall SOM associated with reference vector 1, the reference vector with the lowest quantisation error and so the best fit to the data. Figure 5.12 shows the equivalent SOM for the ERA-40 daily rainfall data, associated with reference vector 8. As with the larger SOM, key rainfall events can be identified within the two arrays and are summarised below.

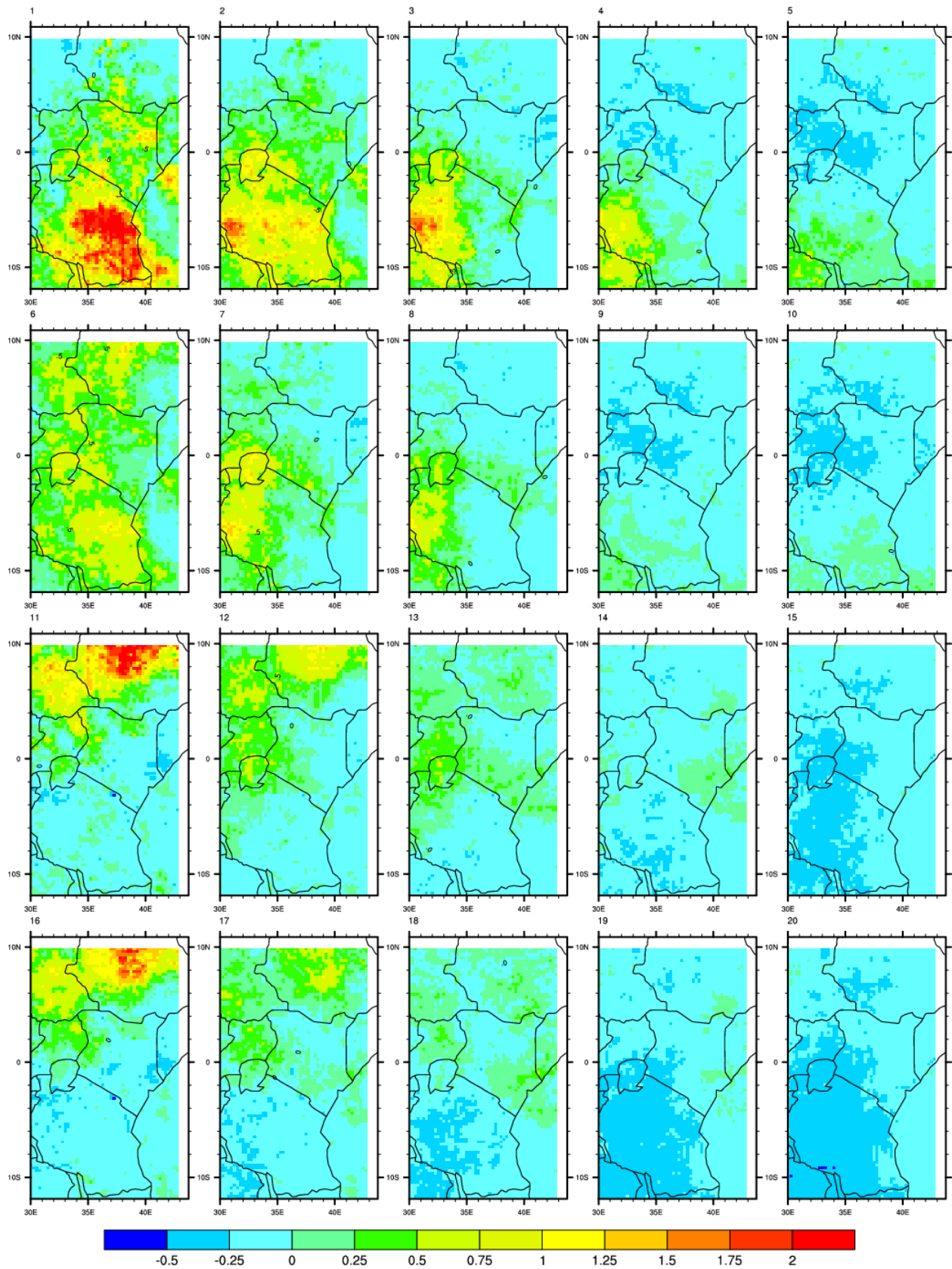


Figure 5.11 5 by 4 TRMM normalised rainfall SOM 1998-2000 OND, reference vector 1

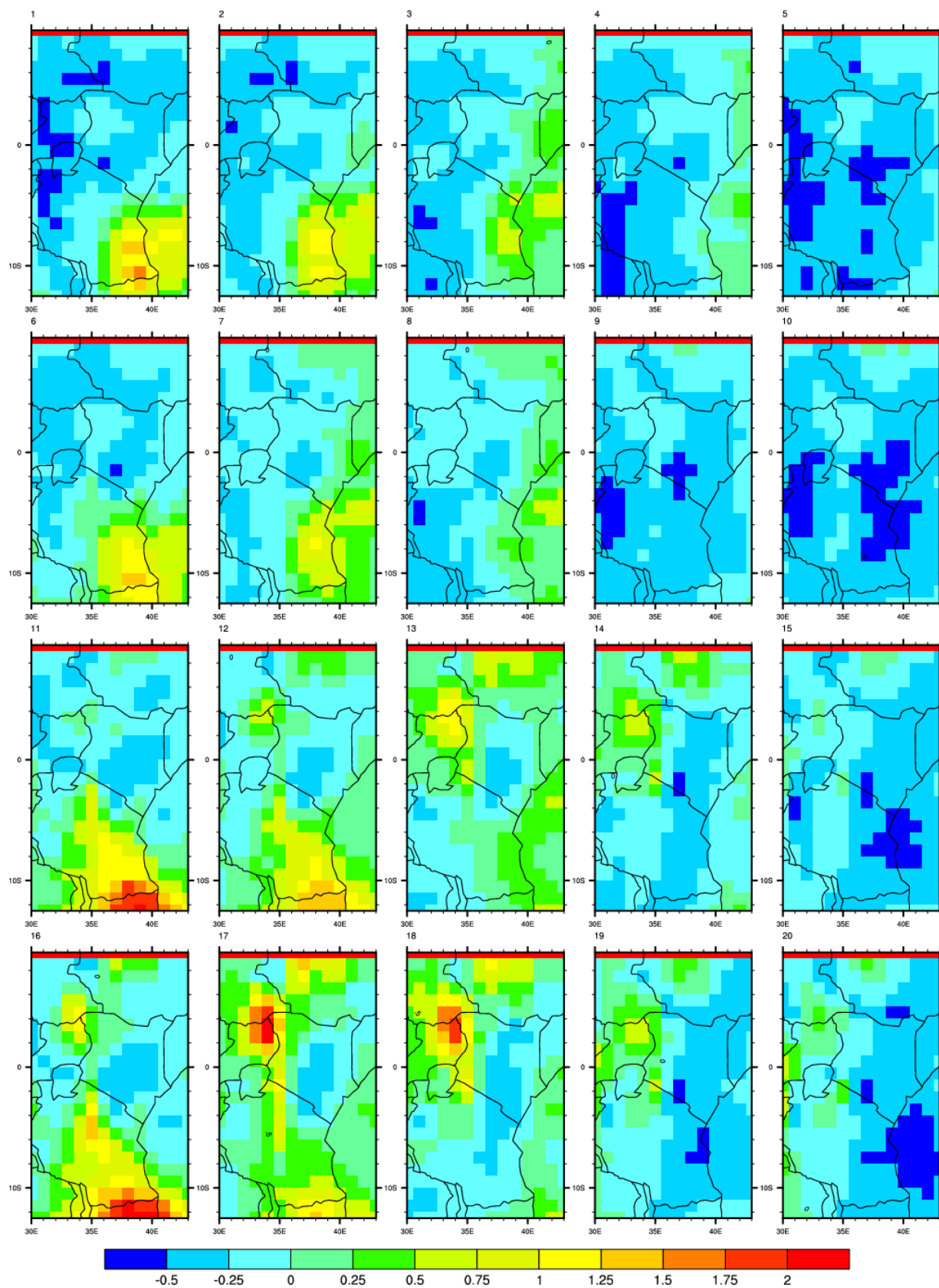


Figure 5.12 5 by 4 ERA-40 normalised rainfall SOM 1998-2000 OND, reference vector 8

TRMM SOM

As with the larger 7 by 5 SOM array, the smaller array also has no states representing daily rainfall events that are centred over Kenya for the 1998-2000 period. The lower right side of the array contains events of suppressed convection over East Africa, although again this feature is not as pronounced as in the ERA-40 dataset. The top left corner of the array has events representing heavy rainfall centres over Tanzania; node 0 0 has the strongest and most widespread rainfall centre of the array, with a further very heavy rainfall centre located in node 0 2 over southern Ethiopia and a second rainfall centre of lower amplitude located to the south east over South Sudan.

ERA-40 SOM

In this SOM suppressed convection dominates the right side of the array although on balance there is a similar proportion of dry to wet states as contained in the TRMM SOM. Heavy rainfall centres over Tanzania are located in the bottom left of the array (node 0 3) with lighter Tanzanian rainfall and Indian Ocean rainfall centres located in the same column. The ERA-40 dataset also has nodes representative of the bi-centred rainfall pattern over southern Ethiopia and South Sudan/Uganda although it differs to the TRMM array in that this dataset clusters the weaker centre of the pattern over southern Ethiopia instead of the strongest. When the number of nodes is constrained to 20, it is the Indian Ocean and coastal East African daily rainfall patterns that receive the greatest numerical representation of nodes within the array containing rainfall. Dry nodes are equally prominent in the array but the number of detailed patterns over the focus countries decreases although it is important to note that the key states from the larger array remain. It is only rainfall centres over Kenya which are missing, which is also the case in the large

array for this short three year time period, due to severe drought conditions over Kenya during these years.

5.4.2 Matching the ERA-40 rainfall data to the 5 by 4 TRMM rainfall SOM

In this section, the same methodology as in Section 5.3 is followed, with each input date in the ERA-40 dataset allocated to the TRMM node it most closely resembles. These allocations are compared to the allocations of the TRMM dataset, on which the SOM was also trained. By reducing the number of nodes by more than 40%, the states characterising the daily rainfall characteristics are necessarily more generalised. This generalisation increases the correct matches when only exact node matches are considered, to 84 out of 276 days (30.43%). However, this success rate is still very low and the raw figures suggest that there are significant differences between the two datasets.

0 0 (5) [41.67]	1 0 (1) [25.00]	2 0 (0) [0.00]	3 0 (0) [0.00]	4 0 (6) [33.33]
0 1 (1) [50.00]	1 1 (1) [10.00]	2 1 (0) [0.00]	3 1 (1) [10.00]	4 1 (11) [50.00]
0 2 (2) [20.00]	1 2 (2) [40.00]	2 2 (2) [15.38]	3 2 (3) [23.08]	4 2 (3) [12.00]
0 3 (0) [0.00]	1 3 (2) [11.76]	2 3 (0) [0.00]	3 3 (7) [53.85]	4 3 (37) [72.55]

Figure 5.13 Distribution of the correct matches when ERA-40 rainfall series allocation is compared to the TRMM allocation to the TRMM normalised rainfall SOM 1998-2000 OND for 5 by 4 SOM, red is for no correct matches, orange for up to 10%, yellow for up to 40% and green for more than 40% correct matches.

Figure 5.13 illustrates how well the ERA-40 dataset manages to reproduce the different daily rainfall states found in the TRMM SOM for the OND season for 1998-2000. For each node, the figure in parentheses states the absolute number of dates in the ERA-40 dataset that are correctly matched to that node. The figure in square brackets is the percentage of the total dates that should be allocated to that node that are indeed correctly allocated in the ERA-40 dataset. Whilst there are still nodes in this smaller 20 node array that the ERA-40 dataset never manages to replicate on the correct date, these are far fewer than in the larger 35 node array, totalling 20% of nodes in the smaller array in comparison to 54% in thirty-five state array. With the patterns in a smaller SOM array more generalised, this is the expected result. The nodes with the most successful hits, both in terms of absolute matched days and percentage of the possible total matches, are found where the most number of days are matched; nodes representing rainfall over much of the East African domain.

5.5 Proportional occurrence of rainfall events in TRMM and ERA-40

Whilst the two rainfall series cluster many qualitatively similar spatial patterns of daily rainfall over East Africa during both the common period of 1998-2000 and the climatology periods of 1971-2000 and 1998-2010 for ERA-40 and TRMM respectively, in terms of the date matching, even with a smaller twenty state SOM array there are difficulties in the two datasets allocating the same calendar dates to each individual node.

However, the models on which the analysis for the second research question, the replication of daily regimes of the short rains in AOGCMs, will be carried out are not constrained to the calendar dates as occurred through the twentieth century. Therefore, the matching of individual dates to specific nodes based on either the TRMM or ERA-40 series is not possible. This leaves two primary markers for the evaluation of daily rainfall

regimes within the short rains in models; the structure of daily rainfall regimes in models, in particular in the clustering of spatial rainfall patterns and the proportional occurrence of these patterns. Sections 5.1 to 5.4 addressed the qualitative similarity between the two ERA-40 and TRMM series. This section will evaluate the proportional occurrence of the main daily rainfall states during the short rains in the two daily series, when the daily rainfall is assigned to one of the twenty states in the 5 by 4 SOM array trained on the ERA-40 rainfall data. Figure 5.14 shows the ERA-40 normalised rainfall SOM associated with reference vector 1.

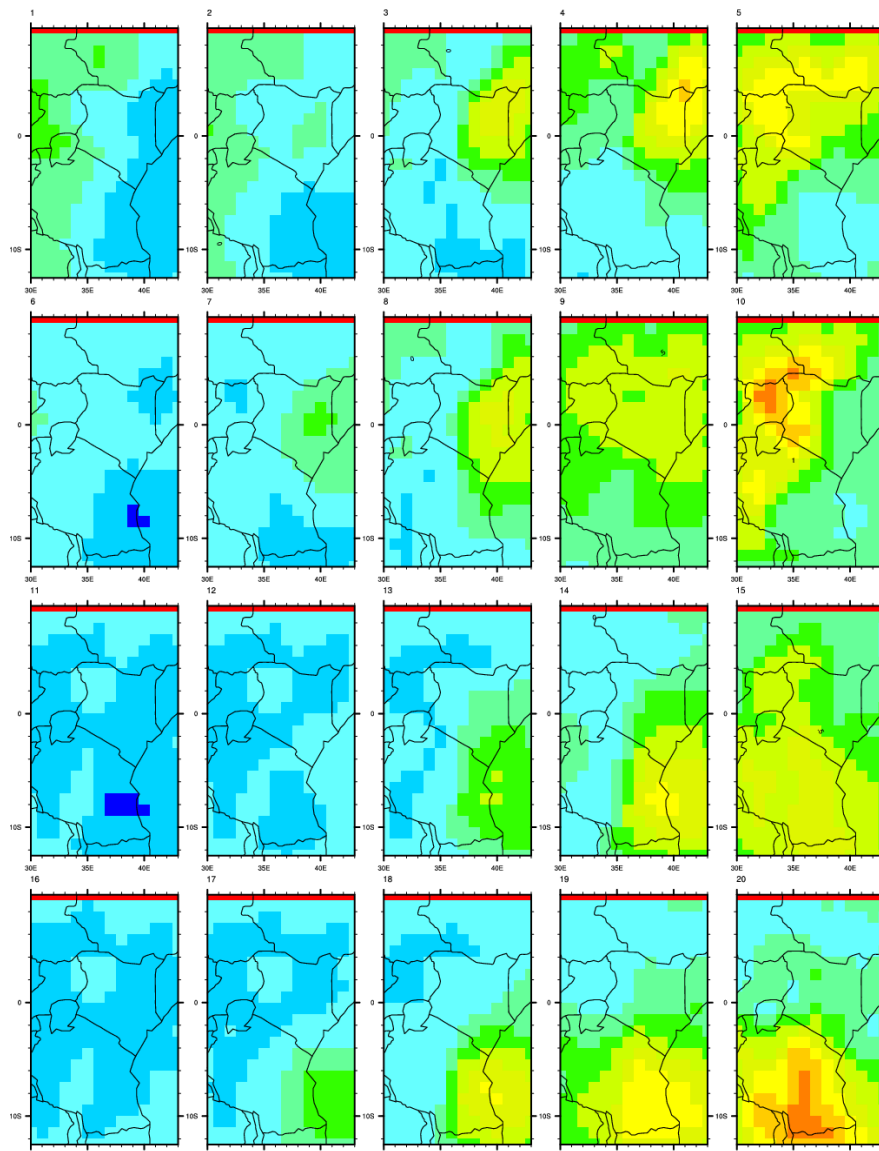


Figure 5.14 ERA-40 normalised 5 by 4 rainfall SOM 1971-2000, reference vector 1

Six main rainfall states were identified earlier through the qualitative analysis of clustering of daily rainfall events in the two series. These are rainfall events over Uganda-southern Ethiopia, the eastern border of Kenya-Ethiopia, Tanzania, the Indian Ocean adjacent to the Tanzanian coast, the western length of the domain originating in the Congo and dry events of suppressed convection across the four focus countries. These events are grouped together within the SOM array for analysis. They are shown in Figure 5.15 and described below.

Group A – These nodes cover states depicting rainfall down the western part of the domain, originating in the Congo and covering the western parts of Ethiopia, Uganda and Tanzania along with South Sudan, Rwanda and Burundi.

Group B – This group contains dry nodes, that is, those nodes where suppressed convection covers the whole of the East African land region within the domain. Node 13 has been allocated to this grouping as opposed to group D, even though it has a light Indian Ocean rainfall centre. This is because in this node conditions over land are very dry (less than 2mm/day mean rainfall in both ERA-40 and TRMM series), following the same spatial patterns of suppressed convection as other nodes in group B.

Group C – Nodes clustered in group C depict a rainfall event centred over the Kenyan-Ethiopian border, with rainfall focused along the east side of these two countries. Sometimes, light rainfall also occurs over South Sudan.

Group D – This group contains Indian Ocean rainfall events. The rainfall centre is located off the East African coast, adjacent with Tanzania. Some rainfall encroaches on coastal East Africa mainly through Tanzania but also through coastal Kenya.

Group E – Group E contains rainfall events characterised by a bi-centred rainfall pattern centred over northern Uganda and south-western Ethiopia with widespread rainfall over much of the north of the domain and north-west Tanzania.

Group F – These nodes are characterised by widespread rainfall over Tanzania, with the rainfall centred over this country. There is some encroachment of rainfall northwards through the domain over Kenya and Uganda.

0 0	1 0	2 0	3 0	4 0
0 1	1 1	2 1	3 1	4 1
0 2	1 2	2 2	3 2	4 2
0 3	1 3	2 3	3 3	4 3

Figure 5.15 Sub groups of daily rainfall types for the quantification of projected rainfall projections at the end of the twenty-first century under the SRES A2 scenario

The TRMM daily rainfall series from 1998-2010 and the ERA-40 series from 1971-2000 are both applied to the ERA-40 normalised rainfall SOM; each calendar day is assigned to the node most closely depicting the rainfall field over East Africa for that day. The percentage occurrences of each of the six groupings for the two series are calculated with results shown in Table 5.4.

Table 5.4 Percentage occurrence of each rainfall state grouping when the ERA-40 and TRMM rainfall series applied to the ERA-40 trained normalised rainfall SOM

Group	ERA-40 % occurrence 1971-2000	TRMM % occurrence 1998-2010	Difference
A	15.54	15.47	0.07
B	37.42	45.74	-8.32
C	16.93	14.38	2.55
D	12.71	5.59	7.12
E	8.41	9.27	-0.86
F	8.99	9.53	-0.54

Three of the groupings, A, E and F, representing light rainfall along the western side of the domain, Uganda-southern Ethiopian bi-centred rainfall and Tanzanian rainfall centres respectively, show almost identical occurrences in both datasets, with differences of less than 1% in occurrence over the two different control periods. Rainfall events over eastern Kenya and Uganda, as depicted in group C, also show very similar occurrences with only a 2.55% difference in allocations to these nodes. The main origins of difference are in the occurrence of dry conditions over the land domain and rainfall centres in the Indian Ocean, adjacent to Tanzania. Dry conditions characterising group B nodes occur 8.32% more frequently in TRMM than ERA-40, with Indian Ocean rainfall events occurring 7.12% more frequently in ERA-40. This is an interesting result given that the TRMM series does not cluster as many nodes of suppressed convection independently. Section 6.4 in Chapter 6 investigates the mean daily rainfall associated with each node, which will be important for establishing if these two datasets agree on mean rainfall amounts for each type of daily event. Given the differences in time period of the two datasets (1971-2000 for ERA-40 and 1998-2010 for TRMM) and in the different methodologies of calculating daily rainfall values over East Africa, the similarities in the occurrence of each type of rainfall event is good.

5.6 Chapter summary

From the analysis undertaken throughout Chapter 5, in addition to earlier research in Chapter 4, it can be said that the first research question of this thesis has been answered, to establish contemporary short rains conditions over East Africa on the daily timescale. Two datasets (satellite and reanalysis) were employed due to the difficulties in a classification of gauge based rainfall data.

Two main markers of similarity were employed to establish the characteristics of daily rainfall regimes within the contemporary short rains. These two markers were, first, the examination of daily rainfall regimes qualitatively, to see if the two datasets produced the same daily rainfall regimes, in particular in terms of the clustering of spatial rainfall patterns and second, the proportional occurrence of these patterns.

It was established that when the two datasets were independently trained into thirty-five node SOM arrays there were considerable visual similarities in the type of events that were clustered for the short rains period, despite the difference in the length and time period of data availability. Moreover, these similarities were confirmed through correlating the nodes in the two arrays with each other; each node in both the TRMM and ERA-40 arrays showed a statistically significant degree of similarity with at least one node from the opposite array. Moreover, the key indicator of the presence of rainfall centres in the ERA-40 SOM also held true in the TRMM SOM; anomalous moisture flux convergence was always found to be present and co-located with a rainfall centre.

Conversely, further investigation revealed differences between the two datasets, in particular when a comparison of which nodes each individual date in the three year overlapping period was allocated to was undertaken. It is accepted that different datasets will represent the African climate to varying degrees, as will the AOGCMs utilised in later

chapters of this thesis (Gianinni et al, 2008). This does not mean that these differences compromise the evaluation of the daily rainfall field in AOGCMs to be evaluated in Chapter 6. Analysis of specific calendar days cannot occur in the twentieth century model runs, which are not forced to simulate exact days during this period and so this marker of similarity between the ERA-40 and TRMM series is less important in this work. It is however, an important result which needs careful consideration in the evaluation of AOGCM data in later chapters. The reanalysis run is heavily constrained to behave as observed and yet there are still clear differences with the satellite rainfall series.

An awareness of the differences between the TRMM and ERA-40 depictions of daily rainfall events during the short rains over East Africa provides signals as to which characteristics of the daily field need particular attention in the evaluation of the daily rainfall field in the AOGCM simulations. Specifically, the prolific occurrence of days of suppressed convection over the East African land region in ERA-40 and scarcity of such events in the TRMM rainfall series when it was independently trained highlighted a state which needs particular attention in AOCGM representations of East African rainfall.

To answer the first research question of this thesis, the contemporary structure of daily rainfall regimes of the short rains of East Africa, six main states were clustered, with high qualitative similarity between the TRMM and ERA-40 series. These were rainfall events centred over Uganda-southern Ethiopia, the eastern border of Kenya-Ethiopia, Tanzania, the Indian Ocean adjacent to the Tanzanian coast, the western length of the domain originating in the Congo and dry events of suppressed convection across the four focus countries. Both TRMM and ERA-40 datasets clustered similar spatial patterns, reinforced by statistically significant correlations between the two arrays at the 99% level. The most frequently occurring states, accounting for 37.42% of days in the ERA-40 series were those associated with suppressed convection. States representing rainfall along the western

side of the array accounted for approximately 15.5% of days in both series, with Indian Ocean rainfall centres having a similar frequency of occurrence in the ERA-40 series. Rainfall centres over Tanzania, Uganda-southern Ethiopia and eastern Kenya-Ethiopia accounted for a lesser proportion of days, measuring 8.99%, 8.41% and 12.71% of occurrences in ERA-40. All states had a physically consistent co-located moisture flux anomaly (divergent or convergent, dependent on dry or wet conditions respectively) in both datasets, providing an atmospheric reason for the clustering of these six key types of events to have occurred.

This work has provided valuable insight into the East African short rains season, which will be considered in the AOGCM evaluation, in particular in relation to research question 2; to what extent can coupled climate models replicate the contemporary structure of the East African daily rainfall regimes of the short rains season. Additionally, having identified moisture flux divergence as a key indicator of rainfall events in two independent datasets, supported by theory as to why this variable is particularly useful for the presence of rainfall events, it is this circulation indicator which becomes the focus in later chapters. The suite of variables (moisture flux, moisture flux convergence, specific humidity, geopotential height and zonal winds) analysed in Chapters 4 and 5, whilst supporting the moisture flux divergence results and providing thoroughness of analysis, is omitted for the remainder of this thesis since they provided less detailed and insightful evidence of daily rainfall events during the short rains.

Chapter 6: East African daily rainfall regimes: a comparison of CMIP3 models with ERA-40 and TRMM

The previous chapter analysed daily rainfall regimes during the late twentieth century in reanalysis and satellite based datasets. This chapter considers daily rainfall regimes in coupled climate models and compares these regimes with those established in the reanalysis and satellite rainfall data sets in order to address the second research question.

Climate models are the main tool available to produce climate change projections and for them to be employed effectively in research there needs to be confidence, at least, in their ability to simulate the contemporary climate system. Such confidence is often not as high as users would like in the case for parameterised variables such as rainfall, with Giannini et al (2008), for example, highlighting divergent rainfall projections emerging from various AOGCM modelling groups globally. Added to this problem, there is no widely accepted way to weight the models according to their fidelity. To further reiterate, the premise of this study is based on the East African region having one of the most coherent climate change rainfall projections for the end of the twenty-first century, that of increased rainfall during the short rains. However, the majority of studies have focused on the use of monthly rainfall means; this chapter redresses this imbalance through the study of daily rainfall regimes in a CMIP3 model subset, since daily events sum to make the much studied monthly means. If these monthly means come about because of daily events which are unrealistic, then the model may not be as reliable as it seems from examination of the monthly totals alone.

This chapter addresses the second research question of this thesis; to establish to what extent coupled climate models can replicate the contemporary structure of the East African

daily rainfall regimes of the short rains season. This research question is answered through a number of measures. First, in Section 6.1 SOMs are trained independently on the rainfall data for each model and the ensemble mean, that is, the normalised rainfall series for each model is classified into an independent SOM. These SOMs are examined and the type and distribution of rainfall events classified are compared to those that emerged in the ERA-40 SOM.

Following this, a more constrained test of the individual model and ensemble mean data is undertaken, with the normalised model rainfall series applied to the ERA-40 trained SOM. In this case, the 5 by 4 best realisation SOM from 1971-2000 that was trained on ERA-40 daily rainfall is used as the base SOM array and each date in the model daily rainfall data over the East African study domain is assigned to one of these states, in other words the model daily rainfall data is 'forced' on to the ERA-40 SOM. The reason for forcing the model data onto the ERA-40 SOM is that (as will be shown) the models produce unique SOMs. Direct comparison between the SOMs is therefore difficult to make. However, the distribution of the coupled models when forced onto the ERA-40 SOM does allow for direct comparison. It is these realisations that form the basis of the work for much of the remainder of this chapter. Goodness of fit of the model data to the ERA-40 trained SOM is evaluated through the quantisation error. The distribution of the daily rainfall data between the nodes of the ERA-40 SOM array are appraised for their qualitative likeness to the TRMM and ERA-40 data over East Africa. Trends in daily short rains rainfall over the region are examined in model and reanalysis datasets. Finally, the daily rainfall means in millimetres associated with each SOM node in the ERA-40 and TRMM datasets and the individual models are evaluated and the associated atmospheric dynamics in the 850hPa moisture flux divergence field are compared.

Through these investigative measures, insight into the daily rainfall field in the CMIP3 model subset, strengths and weaknesses, is gained. This facilitates examination of the daily rainfall field under climate change conditions in order to address the final two research questions of the thesis in Chapters 7 and 8.

6.1 Daily rainfall in AOGCMs during the twentieth century

This chapter employs data from the eight CMIP3 AOGCMs introduced in Chapter 2, Section 2.8. These models are CCCMA, CNRM, CSIRO, GFDL, GISS, MIUB, MPI and MRI. An ensemble mean of these eight models is also employed. Shongwe et al (2011) asserts that research on the most appropriate way to weight climate models objectively is still ongoing and so the ensemble employed here gives each model equal weighting. Mean climate characteristics of these models were evaluated in Chapter 3 and so this chapter immediately focuses on the daily rainfall field during the OND period, 1971-2000, in order to address the second research question of this thesis. The majority of studies over the East African domain rely heavily on monthly data, for example those undertaken by Shongwe et al, 2011, Anyah and Semazzi, 2007, Black et al, 2003 and Mutai and Ward, 2000, and so focus on mean rainfall change during individual seasons. There is a significant impact on East Africa as a result of the daily distribution and intensity of rainfall over the domain. For example, intense rainfall days interlaced with long dry periods can adversely affect crop yields in East Africa (Camberlin et al, 2009). Thus, the use of daily data in this instance adds considerable knowledge and value.

6.1.1 Daily rainfall characteristics of eight CMIP3 AOGCMs

The first stage of investigating daily rainfall in AOGCMs is to examine the daily rainfall characteristics in each model during their twentieth century climatology runs. This is done

through using the 20C3M run which has been designed to respond to the atmospheric forcing of the twentieth century. The same time period as in previous chapters, 1971-2000, October to December, is used and a 5 by 4 array, twenty node, self-organising map is created for each of the eight models and also for an ensemble mean of these models. Examination of each of the SOMs for the individual models and a comparison of the type of states which emerge from the clustering of the model data in relation to the observed datasets forms the basis for a preliminary qualitative evaluation of the ability of the AOGCMs to accurately simulate the observed East African climate.

Figure 6.1 shows the ERA-40 5 by 4 daily normalised rainfall SOM against which the AOGCMs are evaluated. Several key states are identified from this array. The two leftmost columns of the array are dominated by the suppression of convection over much of the East African land area, in particular nodes 0 1, 0 2 and 0 3. Node 0 0 represents light rainfall along the far western part of the domain with the centre of this rainfall pattern located in western Uganda. The right half of this array is where the majority of daily rainfall centres over East Africa are located in this SOM realisation. Node 4 1 (right most column, second row down) represents a bi-centred rainfall pattern, with the two centres located over northern Uganda and southwest Ethiopia. This bi-centred rainfall location is accompanied by dry conditions on the East African coast. Also to the north of the domain, node 3 0 (top row, second from right) depicts a daily rainfall centre on the border of eastern Kenya and eastern Ethiopia, with Tanzania largely dry. A daily rainfall centre over Tanzania is seen in node 4 3 (bottom right corner). Of the nodes representing Tanzanian rainfall, the rainfall is of greatest amplitude in this node with the strongest rainfall centre located in southern Tanzania extending northwards through the centre of the country. A final key pattern identifiable in this twenty node array is of Indian Ocean rainfall centres as depicted in node 2 3 (bottom row, centre column). Rainfall in this node extends inland

through eastern Tanzania and the coast of Kenya with dry conditions elsewhere in East Africa.

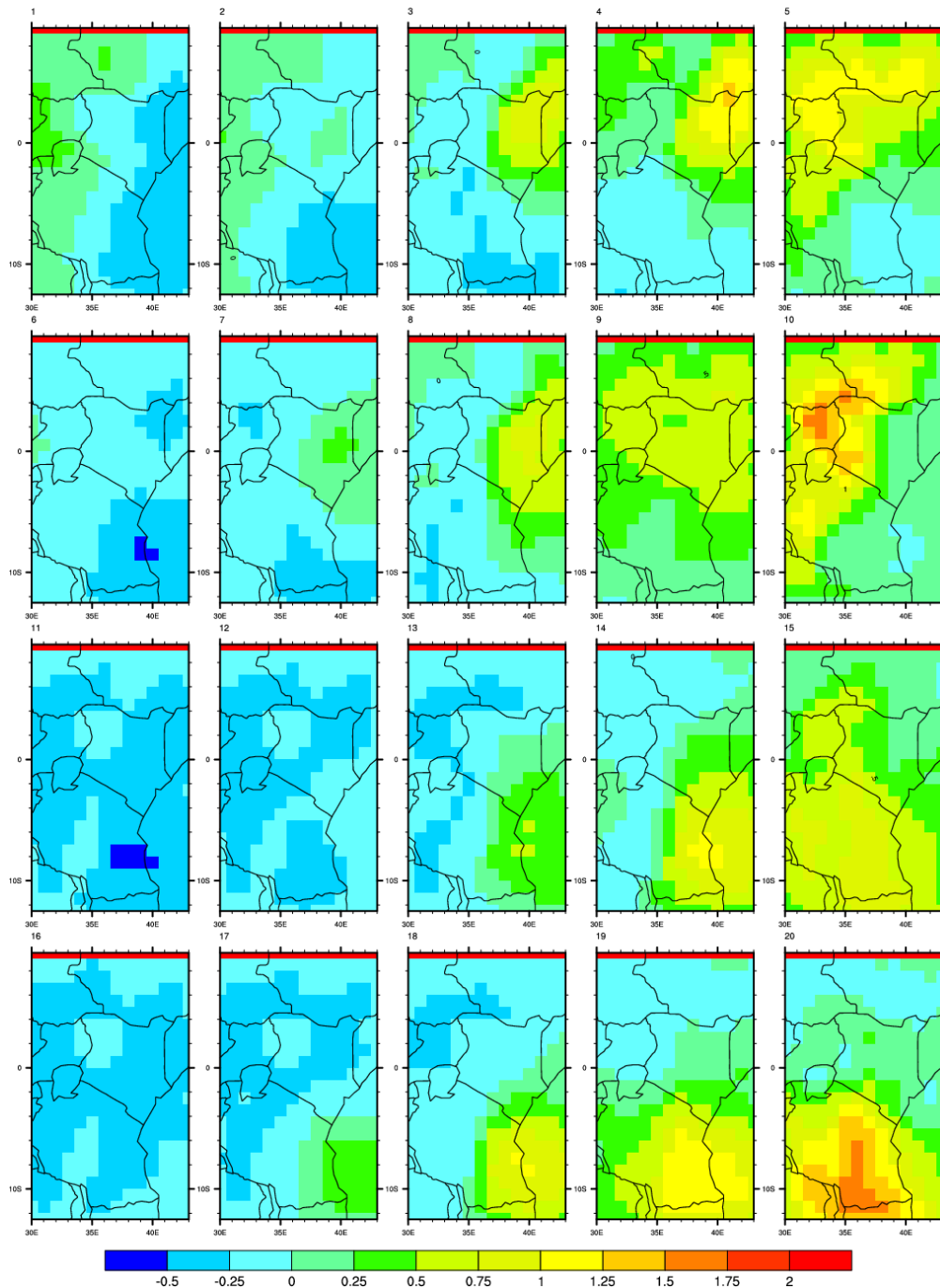


Figure 6.1 ERA-40 normalised 5 by 4 rainfall SOM 1971-2000, reference vector 1

For the eight AOGCMs analysed in Chapter 3 and an ensemble mean comprising these eight individual models, SOMs of daily rainfall are calculated. The same methodology as used previously is followed. To summarise the methodology, this involves interpolating all

the models to the same 1° by 1° grid, selecting the domain of 30°E to 43°E, 12°S to 10°N and converting the data to text format with one line of text for each date in the series to input into the SOM algorithm. This is done for the months of October to December and time period of 1971-2000. Fifteen realisations of each SOM are then created using the SOM_PAK software, with the SOM with the lowest quantisation error associated with its reference vector selected as the best fit SOM for that model. Table 6.1 shows the reference vector number of the selected SOM realisation and the associated quantisation error for each model and the ensemble.

Table 6.1 Reference vector of the SOM realisation with the lowest quantisation error for the ensemble and each individual model for control period daily rainfall.

Model	‘Best Fit’ SOM Reference Vector	Associated Quantisation Error
Ensemble	Reference 14	12.898
CCCMA	Reference 1	10.964
CNRM	Reference 4	11.610
CSIRO	Reference 1	11.867
GFDL	Reference 1	11.643
GISS	Reference 10	12.262
MIUB	Reference 2	11.774
MPI	Reference 5	11.714
MRI	Reference 1	12.077

As with the ERA-40 dataset, key rainfall patterns are identified in the AOGCM normalised rainfall SOMs. Figure 6.2 shows the SOM array for the ensemble mean rainfall. Arrays for the individual AOGCMs are described below, but not shown here. Key rainfall patterns are defined as the distinct major rainfall centres that are evident within the automated clustering of daily patterns that have emerged in the visualised array, alongside states where suppressed convection is widespread across the majority of the four focus countries.

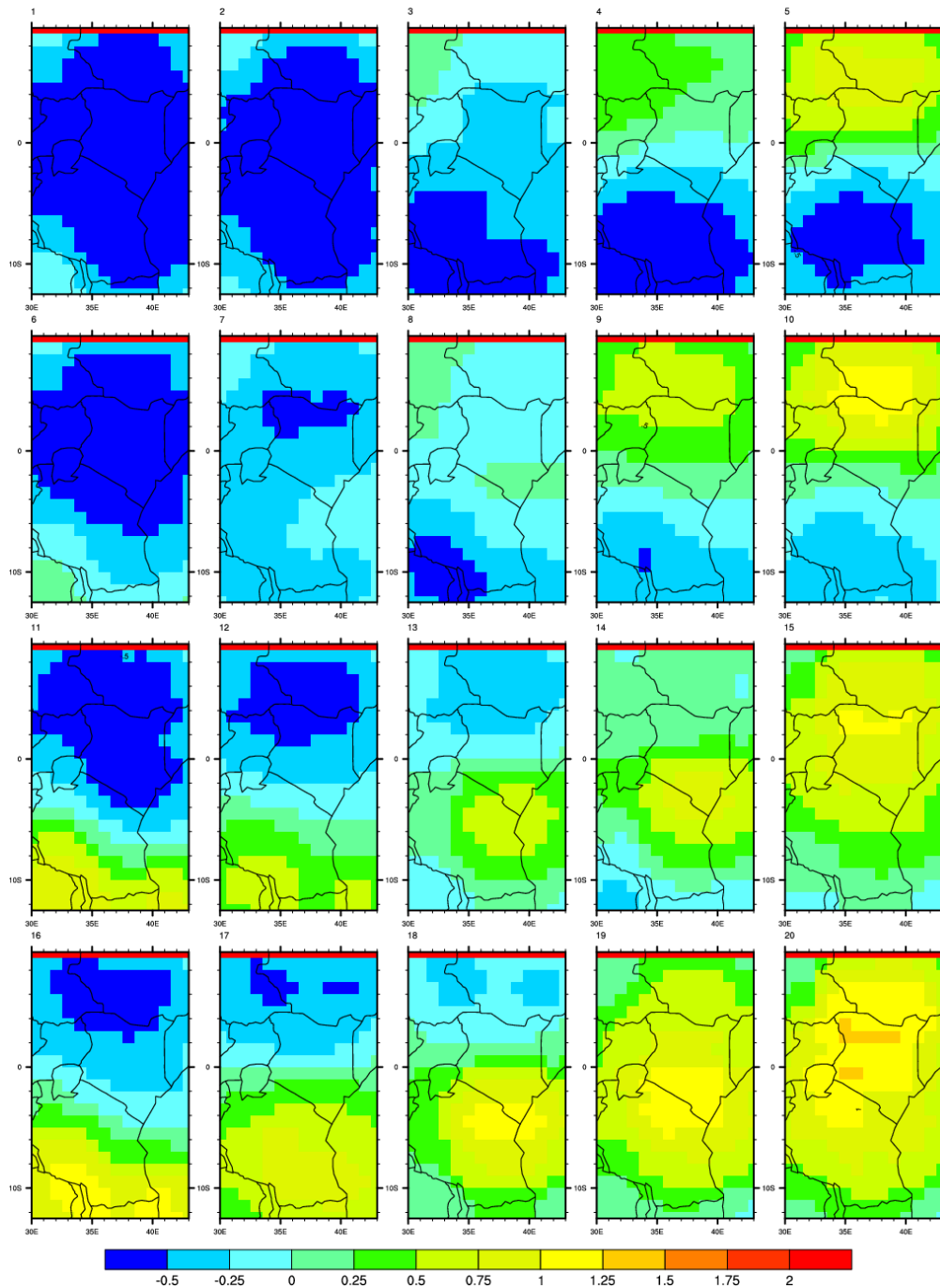


Figure 6.2 Ensemble mean normalised 5 by 4 rainfall SOM 1971-2000, reference vector 14

Ensemble – The ensemble mean rainfall SOM array over East Africa shows much stronger suppressed convection than in the ERA-40 SOM. In addition, strong convection is widespread over much of the domain. Rainfall patterns in this array generally show less spatial distinction than the satellite and reanalysis datasets, with rainfall often present

across the width of East Africa. However, there are some similarities in rainfall patterns present in the ensemble SOM. For example, node 43 (bottom right corner) shows a pattern equivalent to the bi-centred rainfall seen over the Ugandan- Ethiopian border.

Additionally, Tanzanian rainfall centres are well represented in the ensemble SOM.

CCCMA – In this SOM array, the major rainfall centres tend to be stronger than in the ERA-40 SOM. A second differing feature is that rainfall regions characteristic of particular SOM nodes are larger, more widespread, and less detailed spatially than the reanalysis and satellite rainfall patterns. Additionally, there are fewer nodes within the array showing suppressed convection than in the ERA-40 dataset. Several observed rainfall patterns are absent from the array altogether. These include more compact rainfall centres along the eastern side of the domain, for example over eastern Kenya and Ethiopia. The CCCMA model does cluster smaller scale events down the western side of the domain into a node within the array and also successfully simulates the bi-centred pattern over the Ugandan – Ethiopian border, although with less clarity than has been seen in observed datasets.

CNRM – This model produces a good simulation of the amplitude of rainfall centres over East Africa. However, the array does not contain centres over the Indian Ocean which is a key component of the reanalysis and satellite datasets. This absence of some rainfall patterns is continued into the land on the eastern part of the domain. There is some evidence of the bi-centred Ugandan – Ethiopian rainfall, although a less clear spatial pattern is seen. The key feature of rainfall nodes in this model SOM array is that they tend to spread across the whole East African domain from East to West at each latitude. There is a rainfall centre over Tanzania in the SOM array for this model but this extends too far north in comparison to the satellite and reanalysis datasets.

CSIRO – A key feature of this model is very strong suppressed convection over a large part of the domain, which appears to a greater extent than in the ERA-40 dataset. This model has nodes within the SOM array representing Indian Ocean rainfall to the south of the domain, the bi-centred Ugandan – Ethiopian rainfall centre and some evidence of coastal Kenyan rainfall, with the caveat that all these features occur with less clarity than in the ERA-40 rainfall climatology. A general observation, as with many of the AOGCMs, is that when rainfall occurs in a node of the SOM array it has a tendency to be too widespread, reflecting a shortcoming in the model.

GFDL – A weakness in this model in the simulation of the daily rainfall states in comparison to the satellite and reanalysis datasets is a clear tendency towards latitudinal rainfall, with a large number of nodes where rainfall covers the whole width of the domain. This is especially the case through Kenya and the central part of the domain, a daily rainfall state which is not seen in the ERA-40 or TRMM datasets. As with several of the AOGCMs, very strong suppressed convection is found over the East African domain to a greater extent than in the ERA-40 dataset. This model also misses nodes with a strong rainfall centre over Tanzania. However similarities between the GFDL model and the ERA-40 dataset are seen in nodes with Indian Ocean rainfall centres encroaching on to the land domain.

GISS – Rainfall in the GISS model is zonally oriented and far too prevalent in too many nodes in comparison to the satellite and reanalysis SOM arrays. Additionally, suppression of convection is also very strong and widespread. There are no clear matches between daily rainfall patterns in the SOM arrays of the GISS model and ERA-40 dataset. As is the case in the model climatology examined in Chapter 3, daily rainfall characteristics in GISS are particularly poorly represented in comparison to the remainder of the selected AOGCMs.

MIUB – The MIUB model places the strongest daily rainfall centre just north of the Kenyan border with Tanzania and in comparison to the strongest observed rainfall centre in Tanzania this is too far north. A further feature of this SOM array that is not seen in the ERA-40 dataset is the presence of several nodes with daily rainfall in south-west Tanzania expanding north-east through the country. The bi-centred Ugandan-Ethiopian rainfall pattern that is a key feature of the observed climate is not seen in the MIUB SOM array; there is a rainfall centre displaced too far east into Ethiopia. Also absent from the array is Kenyan and Ethiopian coastal rainfall. Similarities between nodes in this AOGCM and the ERA-40 array occur in the Indian Ocean and Tanzanian daily rainfall patterns.

MPI – In this SOM array there is a good representation of the satellite and reanalysis climate array in some aspects. In particular, this model successfully simulates the heavy rainfall centre in southern Tanzania alongside a good approximation of the nodes representing Tanzanian rainfall in the ERA-40 array. A feature with a heavy presence in this SOM array that is not seen in the ERA-40 dataset to the same extent is the presence of several nodes with a rainfall centre in the western part of the domain and into Uganda. The bi-centred rainfall centre over Uganda and Ethiopia is located too far to the east in this dataset with the rainfall also too extensive spatially. This AOGCM shows strong suppression of convection but successfully replicates some of the patterns seen in the ERA-40 SOM array.

MRI – A key feature of this SOM array is very strong suppressed convection, which is not seen in the ERA-40 array. The model also misses the majority of the Indian Ocean and coastal rainfall centres seen in the ERA-40 array. In common with several of the AOGCMs, the rainfall over Ethiopia (bi-centred over Uganda-Ethiopia in the ERA-40 array) is too far east. The light rainfall along the western side of the domain in the ERA-40 dataset is much stronger and does not extend the length of the domain in the MRI model

but is focused to the north. However, Tanzanian rainfall centres are well replicated in this array.

To summarise, the AOGCMs do not independently replicate the same daily rainfall states that are seen in the SOM array for the ERA-40 dataset. Some common issues between the models are apparent. AOGCMs have difficulty clustering the rainfall into detailed rainfall centres, with a strong tendency for the individual rainfall centres to be combined in the array to show largely latitudinal rainfall regions and much more widespread rainfall over many nodes in comparison to the ERA-40 and TRMM SOM arrays. Additionally the AOGCMs have a propensity to stronger suppression of convection over the land area of the East African domain.

However, on the other hand, one region where the AOGCMs have most success in simulating daily rainfall states is over Tanzania, with the majority of the models having nodes with rainfall centres in this part of the domain. This matches characteristics of the satellite and reanalysis SOM arrays.

6.1.2 Proportion of occurrence of different daily rainfall states during the control period

Whilst it was necessary to look at the spatial type of rainfall events that occurred, it is also important to consider the occurrences of individual rainfall events as a proportion of total days during the October to December short rains. Do AOGCMs during the 20C3M climatology period reproduce the same type of rainfall events in the same proportions as the ERA-40 dataset during the same timescale? This initial investigation is carried out on the SOMs created on each of the model datasets independently as described above (that is, ERA-40, the ensemble mean and the eight AOGCMs), to see if the models are able to

replicate the proportional occurrence of different daily rainfall types within the control period.

Figures 6.3 (a) and (b) show the percentage occurrence of each node within the SOM array for the ERA-40 dataset and the model ensemble for the October to December season, 1971-2000 respectively. The TRMM array and individual models are not shown. Two states in the ERA-40 SOM array occur more frequently than any of the others. These are the suppression of convection seen in node 0 2 and light rainfall along the western side of the domain (node 0 0) occurring on 13.51% and 12.03% of days within the OND rains respectively. Although occurring far less frequently, the most often seen widespread daily rainfall events in this dataset are daily rainfall centres over Tanzania (4.46% of days, node 4 3) and bi-centred rainfall over Uganda and Ethiopia (4.64% of days, node 4 1). For the TRMM dataset, the 0 0 node, representing suppressed convection over East Africa is the most frequently occurring node within the array occurring on 16.14% of days. Although this represents a different time period (1998-2010 instead of 1971-2000) it is a useful comparison as a second dataset attempting to represent observed conditions over the region.

For the model ensemble, all of the nodes of the array occur to a very similar degree. The most frequently occurring nodes, representing 7.43% and 7.25% of days within the control period are nodes 0 0 and 4 3 respectively. These nodes show a pattern of either widespread strong suppression of convection over the land region of the East African domain or widespread rainfall over the domain. This indicates that the ensemble is unable to combine the daily characteristics of each individual AOGCM to a replicate a good approximation of the ERA-40 daily rainfall; when the ensemble is allowed to create a SOM rainfall array for daily data independently it fails to classify detailed daily rainfall states over the region and has a tendency to default to either dry or wet conditions over much of East Africa.

12.03	3.51	4.28	4.17	1.96
6.74	5.69	2.79	1.81	4.64
13.51	3.95	3.62	3.80	1.85
6.05	7.17	5.29	2.68	4.46

Figure 6.3 (a) Percentage occurrence of nodes within the ERA-40 normalised rainfall SOM array for the period October to December, 1971-2000.

7.43	4.06	5.40	5.76	5.07
5.36	5.33	4.60	3.84	5.26
5.33	3.30	4.02	4.46	4.24
5.76	5.29	4.86	3.37	7.25

Figure 6.3 (b) Percentage occurrence of nodes within the ensemble mean normalised rainfall SOM array for the period October to December, 1971-2000

Patterns (not shown) are also identifiable in each individual AOGCM. A high occurrence of nodes within the SOM array representing suppressed convection over East Africa is a recurring theme common across all of the AOGCMs. Indeed, the node of strongest widespread convection is the most frequently occurring node for each of the models, with the percentage occurrence ranging from 8.30% for node 4 3 in the MRI AOGCM to 15.35% for the same node in CCCMA.

Conversely, some of the models show significant occurrences of other rainfall states which differ from those seen in the observed rainfall SOM arrays. The CCCMA model has a 9.42% occurrence of the node 3 0. This node has clustered light rainfall in the north-west of the domain, centred over South Sudan which is a pattern that does not appear in the same form in either the TRMM or ERA-40 SOM arrays. A second AOGCM which has significant occurrences of rainfall in the north-west of the domain centred over South Sudan and not seen in the satellite and reanalysis SOM arrays is the MIUB AOGCM. This AOGCM has two nodes (2 0 and 3 0) representing this pattern being assigned in total 12.87% of the total days during the control period. A further AOGCM, the GFDL model, has considerable occurrences of a state not seen in the satellite and reanalysis datasets. During the thirty year control period, 6.25% days of this model climatology are assigned to the 0 0 node, showing a rainfall centred astride the Tanzania – Kenyan border in a state not clustered into the SOM elsewhere. One model which from all aspects has an unrealistic representation of East African rainfall is the GISS model. This is reinforced by the examination of the percentage distribution of the individual days onto the SOM array. The two most commonly occurring states are nodes 0 0 and 4 3, characterised by very widespread rainfall over the East African domain and very strong suppression of convection, respectively.

Despite these weaknesses across the subset of AOGCMs as a whole, there are two models which do replicate frequently occurring reanalysis rainfall states; the MPI and MRI AOGCMs. For both of these models the most frequently occurring rainfall type (not node) is a centre in the north of the domain. Whilst neither model maps exactly on to the bi-centred Ugandan – Ethiopian rainfall centre that occurs frequently in the ERA-40 dataset, they are the closest approximation of all the states within the array allowing the conclusion

that the AOGCMs can replicate on a daily basis some of the individual rainfall patterns seen within East Africa.

6.2 Application of AOGCM data on to the ERA-40 rainfall self-organising map

Whilst examining the SOM of each AOGCM independently was a useful tool in obtaining an understanding for the types of daily rainfall events occurring in the model, it did not directly evaluate how well the models can replicate the daily types observed during the 1971-2000 climatology period when constrained to the clustered daily rainfall states of the ERA-40 rainfall series. Moreover, whilst the similarities are limited when the daily AOGCM data is independently trained into a self-organising map, it is feasible that restricting the model data to a SOM trained on a more spatially complex representation of East African daily short rains rainfall will be more fruitful. This hypothesis is evaluated here through applying the AOGCM data to the ERA-40 daily rainfall SOM to evaluate how well the data corresponds.

Normalised model data is applied to the best fit ERA-40 5 by 4 SOM for the 1971-2000 OND period. Each individual date in the input file is linked to the node to which it is most closely matched and an output for the 2760 days with a different node assignment for each day is created.

The AOGCMs do not replicate the weather of the calendar date for each of the days of the twentieth century and so it is not possible to examine individual days within the thirty year climatology as to whether they are linked to the same node as the ERA-40 data on the same day. However, over the thirty year climatology period the percentage occurrence of each node is evaluated, and therefore each daily rainfall event type. It is expected that the

ability of the model to reproduce the same distribution of types of rainfall events over the domain during the twentieth century short rains will be a sign that the model climatology is skilful in its representation of East African rainfall.

6.2.1 Quantisation Error Evaluation

An initial evaluation of the AOGCM data is done through examination of the quantisation error. The quantisation error is a measure of how closely each input line of data fits to one of the SOM array nodes, averaged over the input data as a whole. Each input line is then allocated to the node with the lowest quantisation error with the overall quantisation error defined as the mean error of all the input lines. Table 6.2 shows the mean quantisation error for the ERA-40 dataset and each selected AOGCM over the climatology period when applied to the ‘best fit’ realisation SOM of the ERA-40 dataset relating to reference vector 1. This means that for this evaluation, all data, irrespective of the dataset which it originated from, is applied to the same SOM trained on the ERA-40 daily rainfall dataset.

The results for this are initially somewhat surprising. When the climatology period daily rainfall for each model is forced on to the ERA-40 rainfall SOM, six of the eight AOGCMs have a lower quantisation error than the ERA-40 SOM with the other two AOGCMs and the ensemble mean within 0.75 of the ERA-40 value. This means that the majority of the models align more closely to the twenty state SOM array than the ERA-40 dataset on which the SOM was trained. One explanation for this is that it could have been that the twenty generalised states to which the 5 by 4 SOM is confined are not adequate to fully represent the continuum of the natural variability of daily rainfall states in the ERA-40 dataset. If the models contain less natural variability and fewer detailed spatial patterns in the daily rainfall field then they may appear to fit more closely to an array of daily rainfall states. This result is encouraging for further analysis of the model datasets.

Table 6.2 Mean quantisation error for the ERA-40 SOM and each AOGCM applied to this same SOM for the period OND 1971-2000.

Model	Quantisation Error
ERA-40	13.239
ENSEMBLE	13.991
CCCMA	11.603
CNRM	11.989
CSIRO	12.553
GFDL	12.545
GISS	13.410
MIUB	12.865
MPI	12.315
MRI	13.303

6.2.2 Distribution of AOGCM nodes to the ERA-40 SOM

The next stage in evaluating the model daily rainfall states is to examine in more detail the representation of the ERA-40 rainfall climatology SOM in the ensemble and the individual models. Figure 6.4 shows the percentage distribution of occurrence of each rainfall state for data applied to the ERA-40 SOM array during the period October-December, 1971-2000 for (a) the ERA-40 data and (b) the rainfall ensemble of the eight AOGCMs.

12.03	3.51	4.28	4.17	1.96
6.74	5.69	2.79	1.81	4.64
13.51	3.95	3.62	3.80	1.85
6.05	7.17	5.29	2.68	4.46

Figure 6.4 (a) Percentage occurrence of nodes within the ERA-40 normalised rainfall SOM array for the period October to December 1971-2000

6.60	2.46	2.79	6.02	6.74
2.83	2.83	2.43	5.47	3.73
16.31	1.81	2.54	2.90	5.84
8.05	6.49	2.68	4.49	7.00

Figure 6.4 (b) Percentage occurrence of nodes within the ensemble rainfall series when applied to the ERA-40 normalised rainfall SOM array, OND 1971-2000

To recap, for the ERA-40 dataset, during the months of the short rains, the occurrence of different spatial patterns of daily rainfall occurrences is not divided evenly over the SOM array; some daily rainfall states are much more frequently occurring than others over the domain. Dry states, of a pattern of suppressed convection, occur frequently during the short rains over East Africa. The most frequently occurring state, accounting for 13.51% of days during the OND period from 1971-2000, is of suppressed convection over the whole land region. In terms of daily rainfall occurrences over the domain, 12.03% of days see

rainfall along the western part of the domain, with rainfall occurring through western Ethiopia, Uganda and western Tanzania of the focus countries, with dry conditions seen through Kenya. Other regions of the East African domain often have more than one SOM node defining the daily rainfall events. For example, nodes representing daily rainfall over Tanzania account for approximately 9% of days (nodes 4 2, 4 3 and 3 3), one in five days is assigned to a node representing coastal East African rainfall at all latitudes (nodes 2 0, 2 1, 2 2, 2 3 and 3 0) and bi-centred rainfall over Uganda and southern Ethiopia account for approximately 8.4% of days (nodes 4 0, 4 1 and 3 1). The remaining days are allocated to nodes covering various states of suppressed convection and coastal Kenyan and Tanzanian rainfall.

For each of the eight AOCGMs and the ensemble series a chi-squared test is undertaken to establish if the occurrence of daily rainfall types in the models belongs to the same distribution as the observed ERA-40 dataset. The result is that no model has the same distribution of rainfall types over the climatology period as a whole over East Africa as the ERA-40 rainfall dataset; each model and the ensemble have a statistically significant distribution at the 99% confidence level as shown in Table 6.3. This result points to a clear weakness within the models.

Given that each model shows a different distribution of daily rainfall types to the ERA-40 dataset it is important to understand where these differences lie. This is done through examination of the percentage occurrence of each node within the SOM array during the 1971-2000 climatology period. The differences in distribution of daily rainfall types when all data is applied to the ERA-40 rainfall SOM are described below. Figure 6.4 (b) shows the distribution for the ensemble, Figure 6.5 (a) to (h) show the distributions for the eight individual models.

Table 6.3 Results of chi-squared test on distribution of node occurrence when AOGCM data was applied to the ERA-40 normalised rainfall SOM 1971-2000, OND

Model	P Value (probability that node distribution is the same as ERA-40)
Ensemble	7.51E-97
CCCMA	5.71E-81
CNRM	3.50E-82
CSIRO	1.40E-142
GFDL	2.29E-221
GISS	5.97E-223
MIUB	1.45E-161
MPI	2.96E-176
MRI	6.11E-182

Ensemble – The most frequently occurring day in the ensemble rainfall is the same as in the ERA-40 dataset; node 0 2 represents a state of suppressed convection across East Africa and is associated with 13.51% and 16.31% of the short rains days in the ERA-40 and ensemble series respectively. However the second most common state in the observed data (node 0 0, 12.03%, rainfall along the western part of the domain) occurs only half as frequently in the ensemble rainfall. There are generally fewer dry days in the ensemble rainfall series. Instead, the ensemble sees a greater number of days assigned to nodes representing rainfall centres over Tanzania (nodes 3 3, 4 2 and 4 3 and accounting for 17.33% of days in the ensemble but less than 9% in ERA-40) and additionally to the north of the domain over Kenya, Ethiopia, and to a lesser extent, Uganda.

CCCMA – Visually, the distribution of nodes in the CCCMA model is similar to the ERA-40 dataset, despite it being statistically different. This model has a large number of days of suppressed convection in common with the ERA-40 SOM. Differences include less oceanic rainfall in CCCMA and a greater number of days with widespread rainfall

over the north of the domain that are assigned to node 4 0, accounting for 5.62% of days in CCCMA versus only 1.96% of days in ERA-40.

CNRM – This AOGCM has a heavy emphasis on suppressed convection during the short rains, in particular the strong suppressed convection nodes of 0 2 and 0 3, which occur with greater frequency than in the ERA-40 dataset. These nodes account for 26.26% of short rains days in this model, an extra 6% of days in comparison to ERA-40. A further difference is a greater number of rainfall days over Tanzania, accounting for 12.32% of days in CNRM, an occurrence almost 50% greater than in ERA-40.

CSIRO – CSIRO has a strong tendency towards nodes of suppressed convection with these nodes accounting for an extra 10% of days in relation to the observed ERA-40 SOM. Oceanic rainfall is less common in this model, but when rainfall occurs it tends to be assigned to a node with widespread rainfall rather than the spatially focused nodes.

GFDL – In common with many of the AOGCMs examined above, and with a particular similarity to the CSIRO model, distribution of days for the GFDL AOGCM tend towards the nodes representing suppressed convection and dry days by approximately 10%. Additionally, there is a tendency for a greater number of days with widespread rainfall over the north of the domain, with node 4 0 accounting for 8.55% of days in GFDL in comparison to 1.96% of days in the ERA-40 dataset.

GISS – This model is closer in the proportion of days representing a state of suppressed convection though the distribution of the individual nodes is different over the five nodes associated with suppressed convection (0 1, 0 2, 0 3, 1 2 and 1 3). However, there are too many days in the series with widespread heavy rainfall; approximately double the number of days are assigned to the right hand column of the SOM array where this heavy rain is clustered to in GISS than in the ERA-40 series.

MIUB – Light rainfall occurs less frequently in this model in comparison to the ERA-40 series, with nodes located within the centre of the array occurring less frequently than expected. The difference is made up through almost twice as many days being assigned to nodes representing the heaviest rainfall of the short rains. These nodes comprise the 4 nodes of the rightmost column of the SOM array and account for 22.25% of days in MIUB in comparison to only 12.91% of days in the ERA-40 rainfall series.

MPI – This model has a very strong bias to suppressed convection. The five nodes of suppressed convection (0 1, 0 2, 0 3, 1 2 and 1 3) account for 44.27% of days during the control period. There are also a great number of days with widespread rainfall. Fewer days are assigned to nodes with rainfall over the Indian Ocean (2 2, 2 3 and 3 2), recorded as only 6.05% of the short rains total. These three nodes account for twice as many days (12.71%) in the ERA-40 rainfall series.

MRI – As with many of the other models, MRI shows fewer representations of lighter continental and also oceanic rainfall and more occurrences of heavy or widespread rainfall over East Africa. The rightmost column accounts for 24.24% of days in this model in comparison to only 12.91% of days in the ERA-40 SOM.

9.6	4.46	2.21	3.80	5.62
6.23	3.41	2.21	1.96	3.73
14.96	4.67	2.75	2.72	2.97
10.62	6.85	3.12	2.93	5.18

Figure 6.5 (a) CCCMA node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

7.79	3.66	3.12	4.24	3.99
5.18	5.58	2.93	2.21	4.53
13.26	4.31	2.75	2.61	2.86
13.01	6.05	2.46	3.19	6.27

Figure 6.5 (b) CNRM node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

10.51	2.79	2.97	3.12	5.62
4.49	3.04	4.06	2.57	6.05
11.96	2.79	3.26	2.10	3.15
13.91	7.21	2.21	1.85	6.34

Figure 6.5 (c) CSIRO node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

9.46	3.62	2.43	4.31	8.55
4.71	3.80	2.17	2.50	3.99
9.35	3.44	2.64	2.97	2.93
14.42	7.28	3.70	2.79	4.93

Figure 6.5 (d) GFDL node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

6.70	2.97	3.01	5.14	6.67
2.03	3.12	2.25	5.33	3.33
17.32	1.49	2.86	3.04	5.18
9.02	6.67	2.86	4.93	6.09

Figure 6.5 (e) GISS node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

10.18	3.15	2.17	5.15	6.23
4.42	2.39	1.92	3.30	4.42
17.18	2.03	1.74	2.25	3.37
11.27	4.42	3.04	3.12	8.23

Figure 6.5 (f) MIUB node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

11.34	3.12	2.50	3.95	5.76
6.59	3.33	1.96	2.28	4.09
12.03	3.80	2.32	1.99	2.32
16.34	5.51	1.74	2.86	6.16

Figure 6.5 (g) MPI node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

11.52	3.08	2.39	5.00	6.81
3.22	3.01	3.08	3.95	5.58
13.84	1.30	1.92	2.90	4.57
10.40	4.46	2.57	3.12	7.28

Figure 6.5 (h) MRI node distribution when applied to the ERA-40 normalised SOM 1971-2000 OND

To conclude, the differences in the distribution of daily rainfall types between the different models and the observed SOM follow the same pattern for all of the models and the ensemble. The AOGCMS have a greater number of days with widespread suppressed convection (assigned to the dry nodes of the SOM) and also a greater number of days that are very wet with widespread rainfall over a large proportion of the domain. In summary, the AOGCMS fail to accurately simulate the correct occurrences of light rainfall and more detailed spatial patterns of rainfall over East Africa.

6.3 Trends in East African rainfall during the 1971-2000 period

Whilst it is important to understand the daily rainfall characteristics of AOGCMS in order to interpret their climate change projections with confidence, this work needs to be carried out in the context of the evaluation of the mean climatology of each model and trends over the climatology period. With the ERA-40 dataset being in comparison with the AOGCMS in relation to the rainfall SOMs it is imperative to also evaluate this series over the climatology for its likeness to contemporary conditions in East Africa as evident in independent datasets.

The mean rainfall over the small East African domain (30°E-43°E, 12°S-10°N) is calculated on a yearly basis for the climatology period, the trend over this period calculated and the significance of the trend tested for the ERA-40 series and the eight selected AOGCMS. This is compared with the trend in mean rainfall in the CMAP rainfall dataset.

Figure 6.6 (a) and (b) show the mean short rains (October to December) rainfall for the period 1979-2000. Plots for each individual model are available but not shown here.

Both graphs show high interannual variability in the short rains over East Africa, but also a slight decreasing visual trend in mean rainfall over the climatology period. Only the CMAP satellite rainfall has a statistically significant trend. The trend in CMAP rainfall is for a significant decrease in mean short rains rainfall over the period 1979-2003 at the 99% confidence level. However, the available CMAP series is relatively short (24 years). Some studies argue that at least forty years of data are necessary to detect an anthropogenic trend from natural variability (Henson et al, 2010). The only AOGCM which shows a statistically significant trend in rainfall during the climatology period is the GISS model; which shows a highly unrealistic rainfall climatology over East Africa. Given the assertion of Henson et al (2010), of the length of series needed to detect a trend, it cannot be said that there is any significant trend in the short rains rainfall totals over the last decades of the twentieth century and so it is assumed that short rains rainfall remains unchanged during the control period.

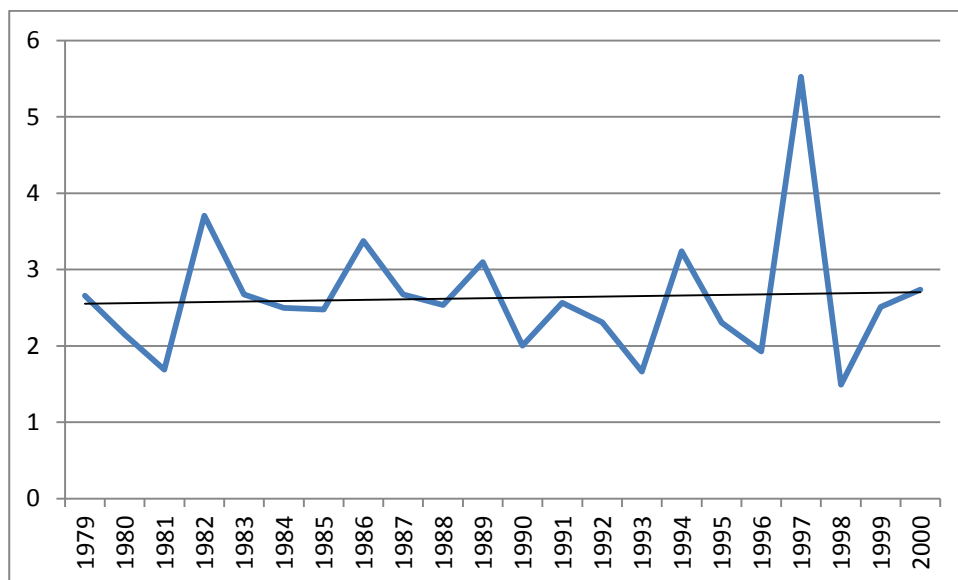


Figure 6.6 (a) ERA-40 short rains (OND) rainfall daily mean (mm) 1979-2000

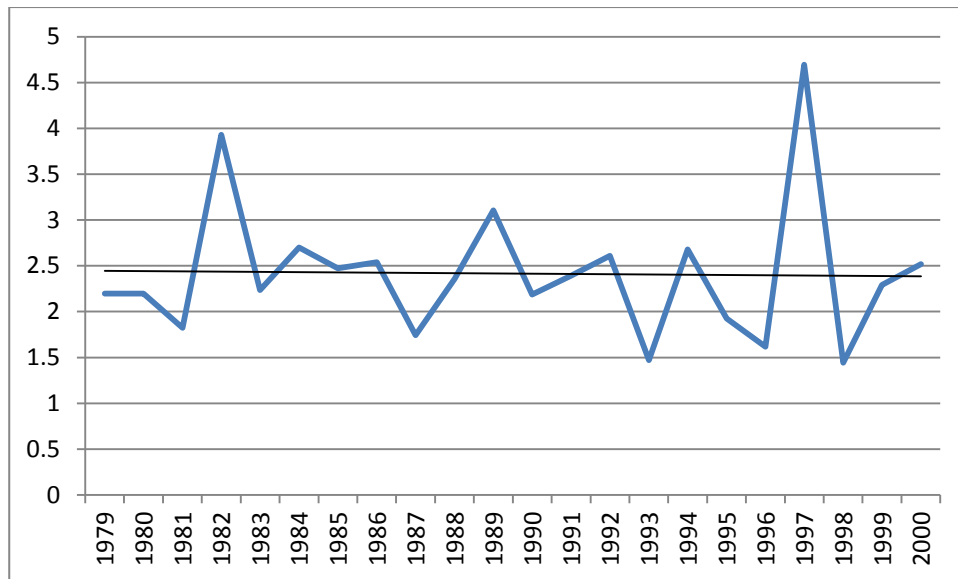


Figure 6.6 (b) CMAP short rains (OND) rainfall daily mean (mm) 1979-2000

6.4 Daily rainfall per day (mm) during the East African short rains in AOGCMs in relation to ERA-40

A further evaluation of the AOGCMs in their simulation of daily rainfall is undertaken through the examination of mean rainfall per day (in mm) produced through days allocated to each node (for the normalised rainfall ERA-40 SOM evaluated in Section 6.2). This provides answers to the question as to whether the AOGCMs are capable of simulating realistic daily rainfall patterns in terms of the amplitude of rainfall falling, independent of the occurrence of the different spatial patterns across the domain and the capability of the model to reproduce realistic short rains. For example, do the limitations in the model simulations highlighted above, namely a lack of focused rainfall and too many very wet or dry days, also impede the ability of the models to simulate the correct amounts of daily rainfall on the occasions they are matched to each node? Which models have the most realistic quantities of rainfall on a daily level in comparison to the short rains as a whole?

6.4.1 Mean daily node rainfall in reanalysis and satellite rainfall datasets over East Africa

This section examines the mean daily precipitation in millimetres associated with each node in the ERA-40 and TRMM rainfall datasets. This provides a basis for comparison for the mean daily AOGCM rainfall associated with each node when the data is applied to the ERA-40 normalised rainfall SOM.

6.4.1.1 Mean daily rainfall associated with each SOM node in the ERA-40 dataset

Figure 6.7 shows the mean daily rainfall in millimetres per day associated with each node within the 5 by 4 SOM array for the best-fit normalised ERA-40 rainfall SOM for the period 1971-2000, October to December. Key nodes representing the range of daily rainfall states described previously are re-examined here to look at associated mean rainfall amounts over all days allocated to that node.

States representing suppressed convection located to the left side of the array have very little rainfall associated with them. Over the four focus countries there is mean rainfall of less than 2mm/day. Mean daily rainfall associated with these nodes is a little higher over the Indian Ocean and to the far west of the domain with mean rainfall values of up to 6mm/day, but these regions are outside of the focus countries and the main area of suppressed convection. Node 0 0 is associated with light rainfall of up to 8mm through western Tanzania, Uganda and South Sudan, with indications that this is originating from a rainfall centre located outside of the small East African domain in the Democratic Republic of the Congo; there is a very small region with mean daily rainfall of 14mm evident in the far west of the domain.

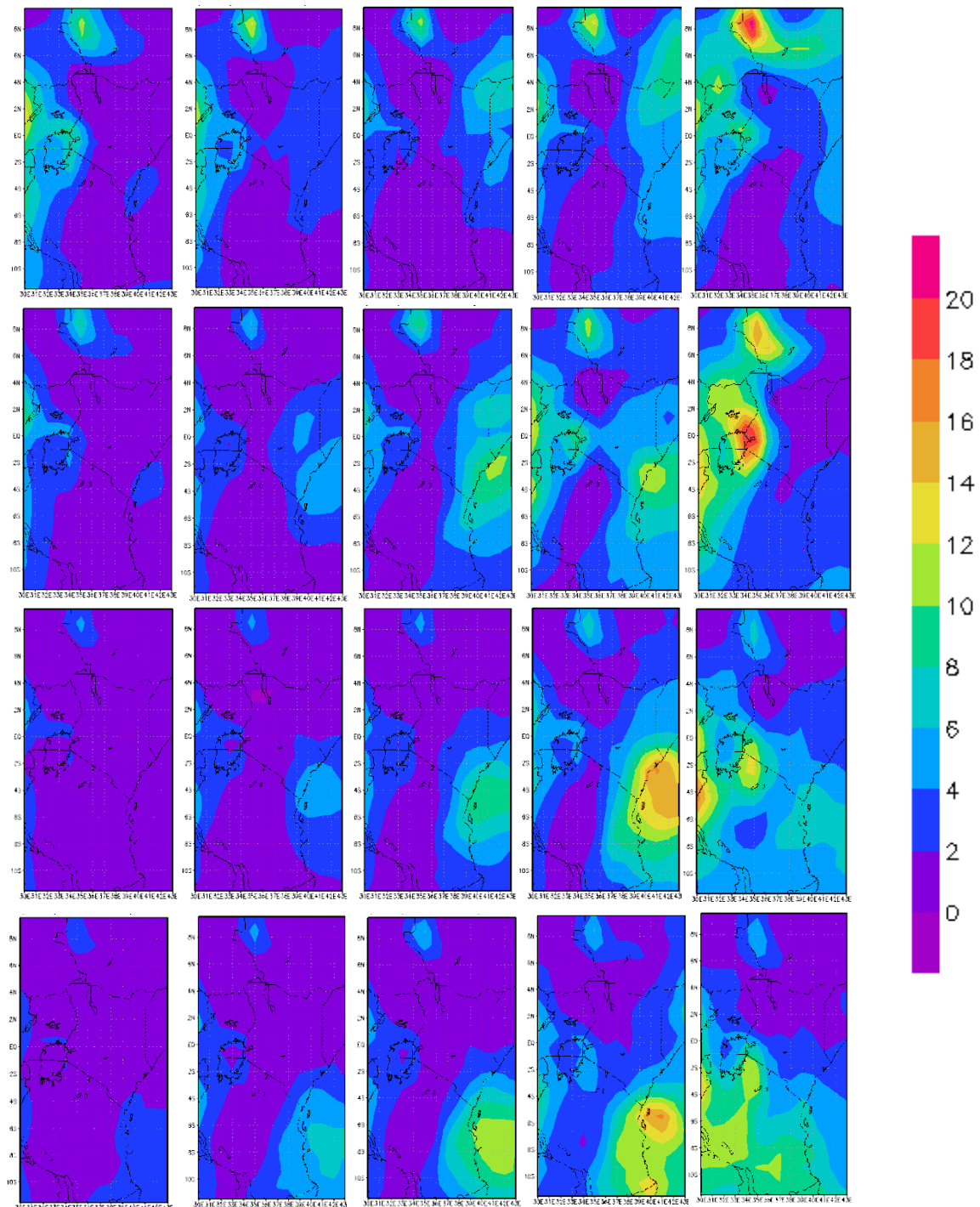


Figure 6.7 Mean ERA-40 daily rainfall (mm) associated with each SOM node in the 5 by 4 ERA-40 normalised SOM array, 1971-2000, OND.

The other key rainfall centres have the strongest rainfall centres located with the East African domain. Mean daily rainfall associated with node 4 1, the bi-centred rainfall state across the Uganda and Ethiopia reaches a maximum of 20mm just north of Lake Victoria

on the border of Uganda and Kenya. A daily maximum of 16mm is reached in southern Ethiopia. Light rainfall of 4mm/day is found to be associated with this node over eastern Tanzania, but eastern Kenya and Ethiopia remain dry. The node representing rainfall over Tanzania (4 3) has a slightly different mean pattern of daily rainfall to the depiction within the array where the heaviest rainfall is located to the south of the country. As in the SOM array, there is a 'tail' of heavier rainfall through the centre of Tanzania, here measuring up to 12mm/day. However, where the SOM array has a heavy rainfall centre located in southern Tanzania this is less evident in the mean rainfall plot with just a small centre of 12mm/day rainfall and lower mean rainfall rates surrounding it. The final key pattern identified, of an Indian Ocean rainfall centre (node 2 3) has a centre of 12mm/day mean rainfall in the ERA-40 dataset, located in the Indian Ocean off the coast of northern Tanzania. Light rainfall falls through the coastal regions of Tanzania and Kenya of up to 8mm/day.

It should be noted through looking at Figure 6.7, the mean rainfall reinforces an earlier evaluation of atmospheric dynamics over the region. It was decided in Chapter 4, Section 4.2.3 to further constrain the domain over East Africa to the current size of 30°E to 43°E, 12°S to 10°N. This was due to a tendency for the rainfall SOMs to be dominated by large scale convection over the Democratic Republic of the Congo and Indian Ocean at the expense of extracting any daily rainfall patterns over the focus countries of Ethiopia, Kenya, Tanzania and Uganda. Whilst constraining the domain removed the tendency for the SOM algorithm to be drawn to these centres and produced a SOM array with a wide variety of daily rainfall states over East Africa, the domination of the convective centres of the Indian Ocean and Democratic Republic of the Congo re-emerges in the plot of the mean daily rainfall associated with occurrences of each node. Nodes with suppressed convection over much of the domain show the fringes of rainfall centres to both the west

and east of the focus region. Furthermore, daily rainfall centres over Tanzania appear to be accompanied by evidence of significant rainfall originating to the west of the small domain in the Democratic Republic of the Congo. Nodes 2 2, 3 3, 2 3 and 3 3 all show a mean daily rainfall of up to 8mm to the far west of the domain in addition to the Tanzanian and Indian Ocean rainfall centres extracted for these nodes in the normalised rainfall SOM. This highlights the dominance of the large convective systems of the Congo in this region of Africa and could be one reason why East Africa receives less rainfall than expected for its latitude.

6.4.1.2 Mean daily rainfall associated with each SOM node in the TRMM dataset

The second dataset examined as an approximation of observed daily rainfall over the domain is the TRMM satellite rainfall series. Figure 6.8 shows the mean rainfall in millimetres per day associated with each node in the five by four SOM array for the ERA-40 trained SOM. That is, each day of the TRMM daily rainfall series for the time period 1998-2010, October to December, has been attributed to one of the ERA-40 1971-2000 SOM nodes. The mean rainfall in millimetres per day associated with each of the twenty nodes of the array is then calculated. Whilst the two time periods do not run concurrently, indeed the overlap is only three years, this exercise adds value for two reasons. First, evaluating the likeness between the mean rainfall per day associated with each node is a further way of understanding East African rainfall. Second, this provides an additional base from which to examine daily rainfall in the ensemble and individual AOGCMs.

There are numerous, strong similarities in the mean rainfall associated with each node of the SOM array. As in the ERA-40 array, nodes linked with the suppression of convection have very little rainfall associated with them. The four focus countries are associated with mean rainfall of less than 2mm/day, again with evidence of encroaching rainfall centres

from both the east and west fringes of the domain outside of the main region of suppressed convection. Node 0 0 shows both the same spatial pattern and the same mean daily rainfall range in the two datasets. The daily rainfall values seen in the rainfall centre down the left side of the domain are in the range of 4-10mm/day, with both the TRMM and ERA-40 mean rainfall plots showing evidence of the rainfall originating further west in the Congo.

Moving to the rainfall centres on the right hand side of the SOM array, some small differences begin to emerge between the mean daily rainfall patterns for each node between the two observed datasets. For node 4 1, the bi-centred pattern over Uganda and southern Ethiopia, the region of maximum rainfall, measuring 20mm/day, is transposed slightly to the west over Lake Victoria in the TRMM array. This is in comparison to a centring of maximum rainfall further east in ERA-40. A further difference in this node is the lack of a second maximum in the TRMM array and the more spatially widespread rainfall region. For the Tanzanian rainfall centre of node 4 3, TRMM has the same spatial pattern of mean rainfall but with increased amplitude of 16mm/day in comparison to 12mm/day in the same node in ERA-40. The TRMM dataset also has increased rainfall for days assigned to Indian Ocean nodes of 2 3 and 3 2. For example, in node 3 2 in the TRMM dataset the daily rainfall centre peaks at 20mm/day, 4mm higher than the same node in ERA-40.

To conclude, the two datasets show very similar mean daily rainfall patterns associated with each node of the ERA-40 rainfall SOM, in particular the spatial patterns show a very good likeness which is highly encouraging. There are some small differences in the amplitude of the mean daily rainfall between the two datasets, with the dominant tendency being for the TRMM dataset to have slightly higher rainfall maximums. The similarities between the two datasets add extra weight to the strength of the daily mean rainfall patterns associated with each of the characteristic daily rainfall states over East Africa that

are clustered using the ERA-40 rainfall series. These results are even more robust considering the differences in the rainfall series. These mean daily rainfall patterns are considered in the next section, when the ability of the ensemble and individual AOGCMs to reproduce such daily rainfall patterns is examined.

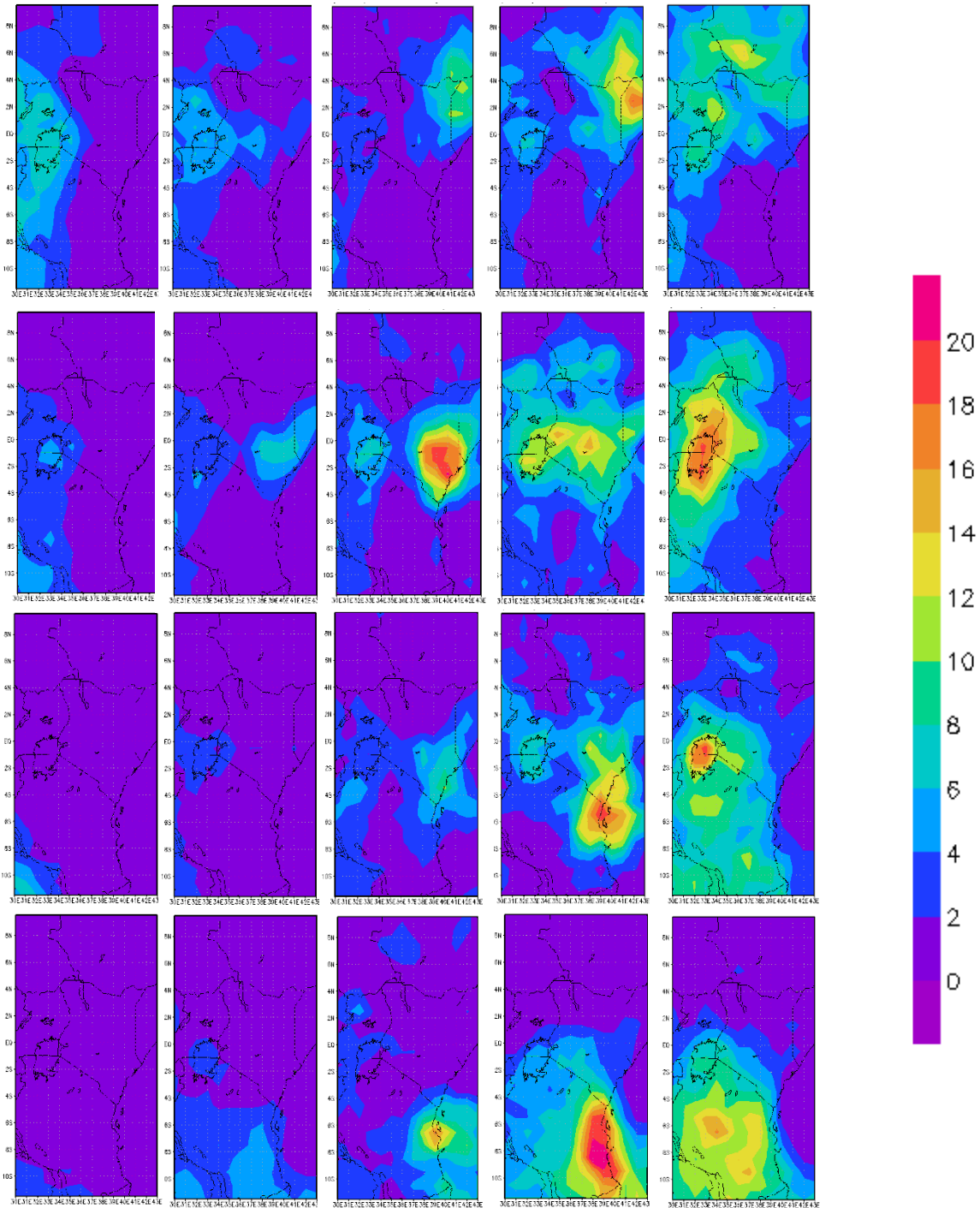


Figure 6.8 Mean daily TRMM rainfall (mm) associated with each node when applied to the ERA-40 normalised rainfall SOM

6.4.2 Mean daily node rainfall in model datasets over East Africa

In light of the observations of Section 6.4.1, namely the similarities in the reanalysis and satellite datasets, ERA-40 and TRMM, in the mean daily rainfall patterns in millimetres associated with each node of the same ERA-40 SOM array, these mean rainfall patterns are examined in the ensemble and individual AOGCMS. Again, the data for each rainfall series is applied to the same realisation of the ERA-40 normalised rainfall SOM, trained on the control period data of 1971-2000, October to December.

6.4.2.1 The Ensemble

Figure 6.9 (a) and (b) show the mean daily rainfall associated with the days assigned to each node in the ERA-40 5 by 4 SOM array for the ensemble rainfall series with and without the GISS model included for (a) and (b) respectively.

Immediately, differences are apparent between the mean daily rainfall patterns in the observed datasets and the ensemble of eight AOGCMS. Rainfall is present over the majority of the East African domain through all twenty SOM nodes in the ensemble array with no nodes showing any regions with mean rainfall of less than 2mm. This is in direct contrast to the ERA-40 and TRMM mean daily rainfall patterns where the nodes representing suppressed convection showed dry conditions over the whole of the region and it was additionally not uncommon to see dry conditions over part of the domain removed from the represented rainfall centre. This major difference between the two arrays indicates that there is a fundamental difference in the way that the AOGCMS simulate daily rainfall patterns from their occurrence in reality. Patterns in other atmospheric circulation indicators, such as moisture flux divergence, on days which are dry in the observational datasets, are studied in addition. This will indicate if the model's weaknesses

at simulating these dry days are pervasive through related variables, or if they are confined to the rainfall field.

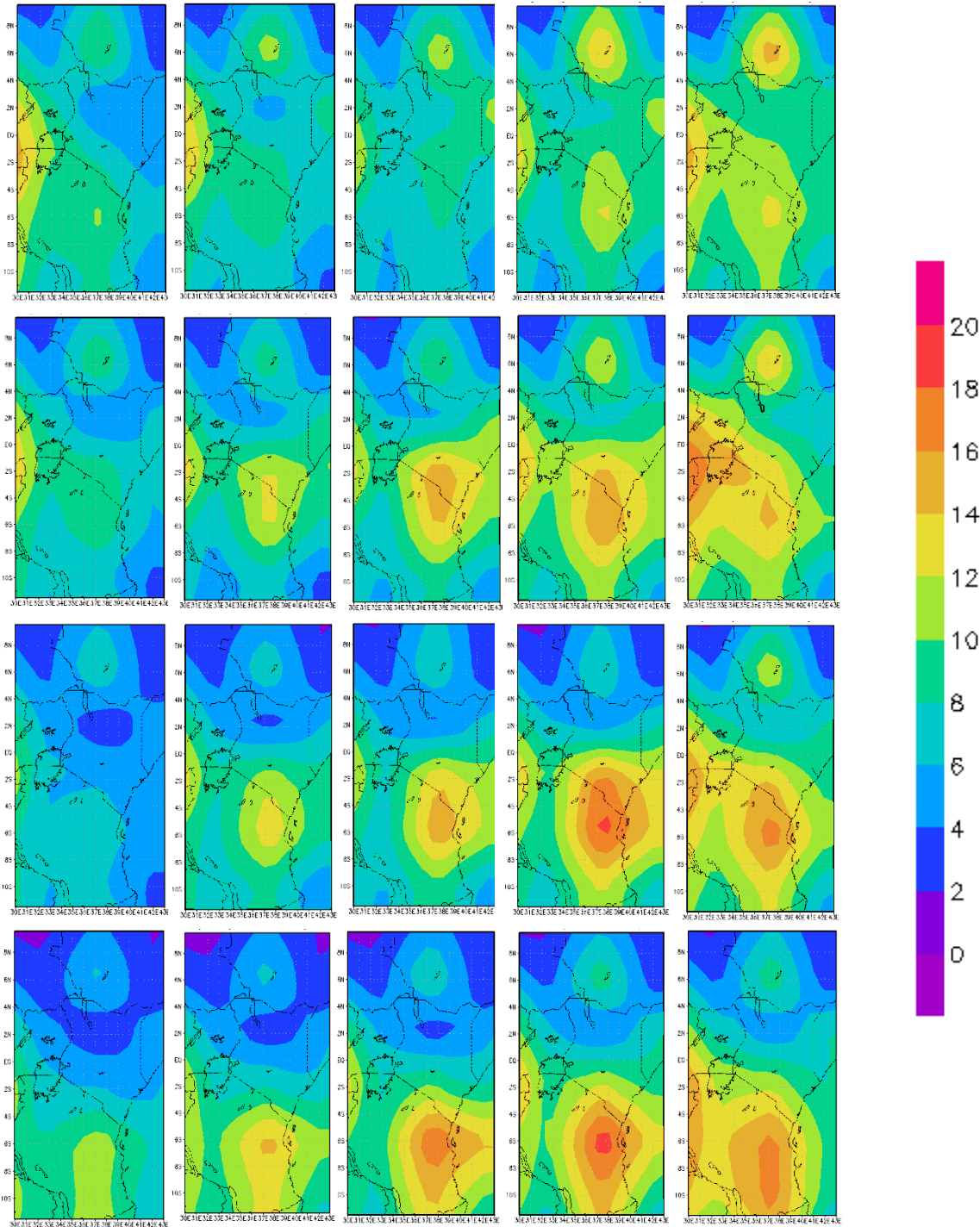


Figure 6.9 (a) Mean daily ensemble (with GISS) rainfall (mm) associated with each node when applied to the ERA-40 normalised rainfall SOM

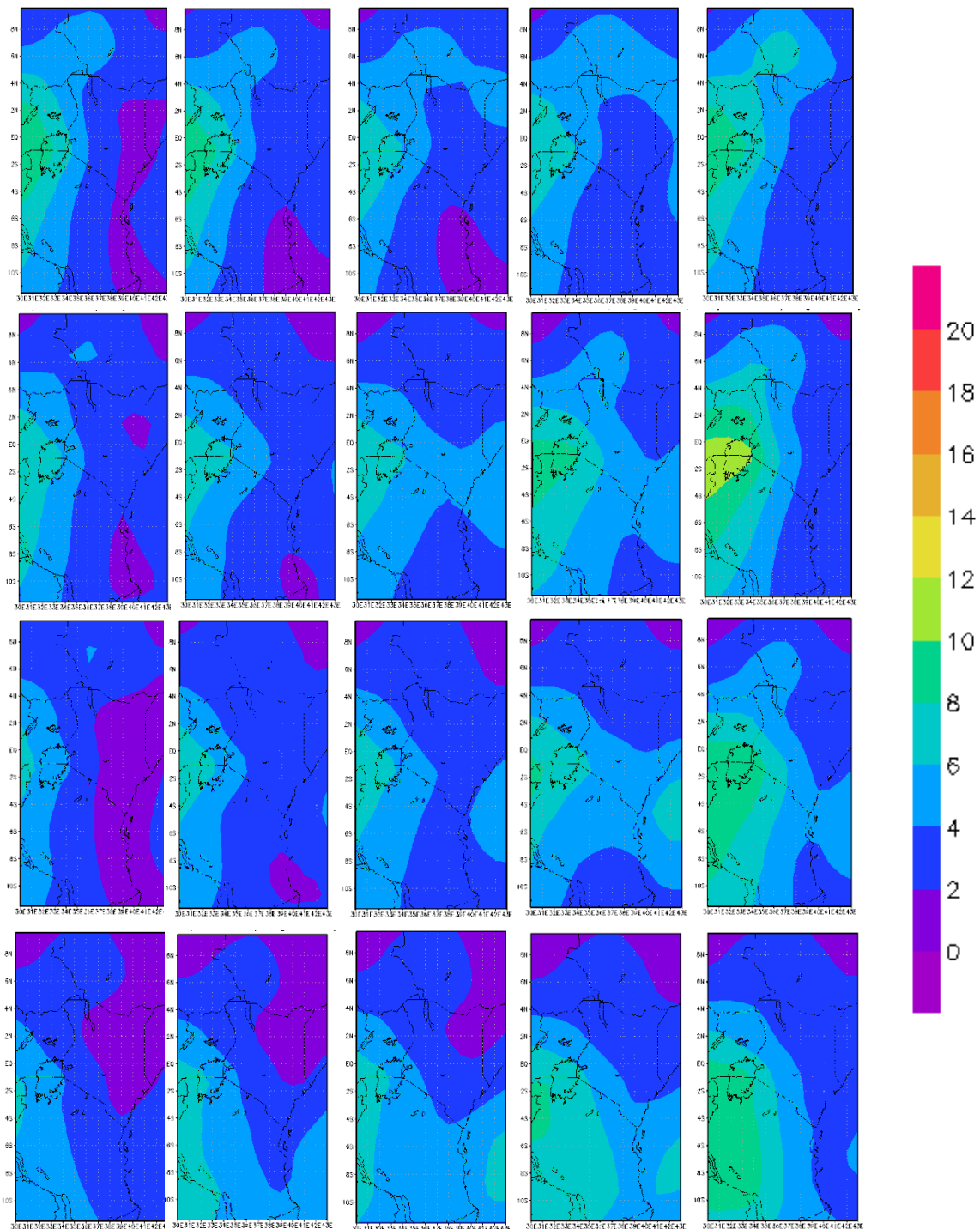


Figure 6.9 (b) Mean daily ensemble (without GISS) rainfall (mm) associated with each node when applied to the ERA-40 normalised rainfall SOM

Taking individual nodes, the rainfall ensemble is largely successful in placing the maximum rainfall centres in the correct part of the domain for each node. However, the region of maximum rainfall is frequently much more widespread in the model ensemble than is seen in the ERA-40 or TRMM rainfall datasets. For example, node 4 0 (top right), representing the bi-centred rainfall pattern has a very small maximum rainfall centre located in southern Ethiopia in the ERA-40 dataset and a maximum mean daily rainfall of 10mm/day elsewhere, whereas in the ensemble rainfall there is a large region of mean rainfall of up 14mm/day and the bi-centred pattern becomes diluted. Another example elsewhere in the domain can be seen in node 4 3 (bottom right) with a rainfall centre over southern Tanzania. In the ERA-40 dataset, mean maximum rainfall reaches 12mm/day over a narrow central corridor of Tanzania. In contrast, rainfall in the ensemble dataset reaches a daily mean of at least 14mm/day over the whole of the country, with a maximum daily mean of 18mm/day. This is higher even than the TRMM dataset where rainfall peaks at a mean of 16mm/day. These differences continue over the majority of the nodes with rainfall registering too high a value over a vast region of the domain. In addition, the ensemble cannot replicate the small but intense rainfall maximums. One such example is node 4 0, where the maximum daily rainfall associated with the rainfall centre over southern Ethiopia is greater than 20mm. In contrast the ensemble projects a maximum of 16mm/day in the same region.

In summary, the daily ensemble rainfall series produces variable results in its attempt to replicate the daily characteristics of rainfall across the East African region. Whilst the ensemble succeeds in simulating the spatial distribution of rainfall across the domain, rainfall is always too widespread and generally too heavy although equally the maximum spikes in daily rainfall when concentrated over a small and specific location are reduced within the ensemble.

It is feasible that some of the overestimation of mean daily rainfall is due to the inclusion of the GISS model in the ensemble rainfall, which is known to contain unrealistically high rainfall values. Figure 6.9 (b) shows the mean rainfall per node for the ensemble series excluding GISS. This does reduce the maximums seen in Figure 6.9 (a). However, the fundamental issue of the presence of rainfall across the domain during the days which should represent suppressed convection over East Africa remain. However, this analysis provides a basis on which to examine if any of the models individually are more successful in replicating mean daily rainfall amounts associated with any of the nodes within the ERA-40 rainfall SOM array.

6.4.2.2 Summary of daily rainfall mean in the individual AOGCMs

Each individual AOGCM has different strengths and weaknesses in their representation of the daily mean rainfall associated with each node in the SOM array. Many of the AOGCMs have success in replicating some of the daily rainfall features associated with the short rains in East Africa. Whilst the ensemble does not reproduce the mean dryness associated with states of suppressed convection and represented along the left side of the SOM array, this is not the case with the individual models. In particular, the MPI model has a good representation of states of suppressed convection with dry conditions over the East African land domain. The CCCMA and MIUB models also see dry conditions over the region when convection is suppressed. Many of the individual models have daily mean rainfall linked with each node of the same amplitude as in the observed datasets. For example, GFDL, MIUB and MPI all have a reasonable approximation of the amounts of rainfall in East Africa. However, these three models also have rainfall that is spatially too widespread, a common pattern that is seen consistently through the individual models. Some of the models, in contrast, have rainfall that is too strong for certain daily patterns.

The CNRM AOGCM has mean rainfall centres that are too strong as does the CCCMA model where there are large regions of rainfall with a daily mean of more than 20mm, figures that are not observed in the ERA-40 model.

One AOGCM which performs with a consistent weakness through all evaluations is the GISS model. Rainfall associated with this model is both much too heavy and too widespread through all parts of the domain and as such is likely to have skewed rainfall in the ensemble which may account for some of the overestimation of daily mean rainfall associated with certain nodes.

6.5 Atmospheric conditions associated with rainfall nodes within the ERA-40 self-organising map

The application of the daily AOGCM rainfall field to the clustering of rainfall states as quantified through the ERA-40 rainfall SOM highlighted some weaknesses that are known to be present in the parameterised rainfall field. The mean rainfall per day in millimetres associated with each node in the ensemble and each AOGCM was shown to have considerable disparities in comparison to the reanalysis and satellite datasets. In particular, for the nodes of suppressed convection located in the left hand side of the array the ensemble and five of the eight individual models (CNRM, CSIRO, GFDL, GISS and MRI) were unsuccessful in recreating the dryness associated with these days over the four focus countries. Given that the occurrence of dry days in the region could have a significant impact on agricultural livelihoods; it would be problematic to accept these rainfall overestimations in this region in the context of examining model projections under climate change scenarios without first investigating if this issue is embedded in other circulation parameters of rainfall within the model climatology.

In Chapter 4, the moisture flux and moisture flux divergence fields were found to be the most useful indicators of the presence, or not, of rainfall associated with each node in the observed datasets. This held true for both ERA-40 fields examined in conjunction with the ERA-40 rainfall and also, a more stringent test, for ERA Interim fields evaluated in conjunction with the independent TRMM satellite rainfall. The conclusion of this work was that rainfall centres were very strongly co-located with the presence of anomalous moisture flux convergence; with dry days having strong anomalous divergence in the same field. Examining the mean moisture flux divergence anomalies associated with the allocation of the model daily rainfall field to the ERA-40 SOM can provide insight into whether the daily rainfall field in the models is weak relative to alternative daily fields which could also indicate the presence of rainfall or if the model daily fields are consistent amongst themselves.

Given the clear overestimation of the rainfall field, leading to daily rainfall characteristics in the case of the GISS model, that are highly unrealistic, the GISS model is excluded from analysis both within the ensemble and as an individual AOGCM. Figure 6.10 shows the mean daily moisture flux divergence anomaly for the ensemble, excluding GISS, for days assigned to each SOM node of the ERA-40 rainfall SOM for the climatology period. The moisture flux divergence anomalies associated with each rainfall pattern are examined for indicators of the presence of rainfall or not. It is expected that anomalous moisture flux convergence is an indicator of a rainfall centre, whilst anomalous divergence is a signal for dry conditions. Particular focus is given to nodes 0 1, 0 2, 0 3, 1 2 and 1 3; the dry nodes characterised by suppressed convection across the domain and poorly represented in the daily rainfall field.

The three models with the best representation of the daily mean rainfall field associated with each node of the ERA-40 SOM array, CCCMA, MIUB and MPI, are evaluated first.

All three show a good representation of the expected moisture flux characteristics for each node. On days related to dry nodes, CCCMA shows large regions of moisture flux divergence across the focus countries of the domain. Further to the west, outside of the constrained domain used for rainfall events, there are regions of convergence, creating the rainfall centres that sometimes encroach on the smaller domain focusing on Tanzania, Kenya, Uganda and Ethiopia. MIUB has very clear anomaly patterns in the moisture flux divergence field, with the mean state for each node clearly indicating the presence of rainfall centres and dry days where appropriate. The MPI model also has appropriate moisture flux anomalies for each node. In this case, the convergent anomalies are stronger than the divergent ones, but the representation of dry days is still good.

The moisture flux field is a good indication of a characterisation of the nodes of the ERA-40 SOM for the above models where rainfall is relatively well simulated. Dry days located on the left hand side of the array are represented through divergent anomalies over the focus countries of East Africa, with other nodes having a consistent presence of regions of convergence co-located with the rainfall centre for that node. This is also the case for the models which have a poorer representation of daily rainfall characteristics. For example, the ensemble has realistic characteristics in the mean moisture flux divergence state for each of the twenty nodes. States of suppressed convection see anomalous divergence through the East African land domain, and although there is some convergence to the west, this is outside of the smaller rainfall constrained domain. All nodes with rainfall centres over the land domain have co-located regions of anomalous moisture flux convergence. The same picture emerges in the four individual models (CNRM, CSIRO, GFDL and MRI) that are weak in the daily mean rainfall figures once GISS has been excluded. All of these models show sensible and realistic moisture flux divergence patterns for each node.

Whilst anomalous convergence is stronger than the anomalous divergence values, the dry states still have clear mean anomalous divergence.

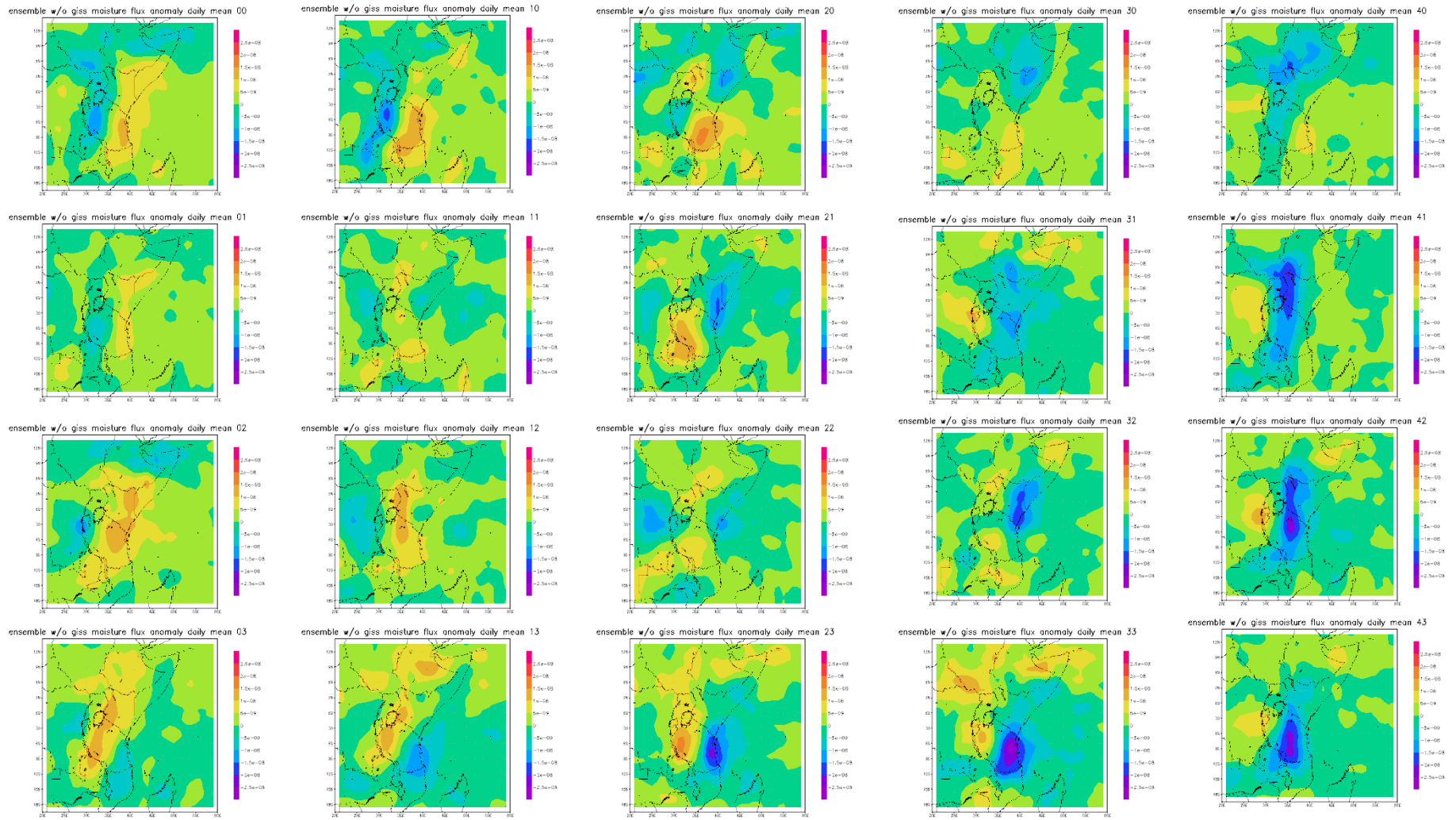


Figure 6.10 Mean moisture flux divergence anomaly (ensemble without GISS) associated with each node when the data is applied to the ERA-40 normalised rainfall SOM 1971-2000, OND.

6.6 Chapter summary

This chapter has addressed the second research question of this thesis; to establish to what extent coupled climate models could replicate the contemporary structure of the East African daily rainfall regimes of the Short rains season. In order to answer the overall aim of the thesis, to evaluate the fidelity of rainfall projections over East Africa during the short rains it was decided to first establish how well AOGCMs can simulate contemporary conditions in the daily rainfall field during the short rains.

When self-organising maps were trained independently on the daily rainfall field of the models, they largely failed to reproduce a SOM array akin to those from either the ERA-40 or TRMM datasets. Rainfall centres were too widespread and latitudinal. Centres over Tanzania were most effective. MPI and MRI had the most realistic distributions of daily rainfall types during this test.

It was evident that there was disparity in the effectiveness of different model fields in representing daily rainfall states over East Africa and so the test was further constrained through applying the model data to the ERA-40 5 by 4 SOM. All of the models and the ensemble had a different distribution of node occurrence that was statistically significant, with AOGCMS showing a tendency to a greater occurrence of days of suppressed convection.

In Section 6.4 it was demonstrated that both TRMM and ERA-40 produced very similar means of rainfall per day in millimetres when the daily rainfall series was applied to the normalised ERA-40 SOM trained during the 1971-2000 period. However, AOGCMs had limited success in replicating these values; in particular for the nodes of suppressed convection seen in the left hand side of the domain, where models could not replicate the

low rainfall values of less than 2mm/day across the domain, instead simulating greater rainfall amounts for these days.

It was decided to examine the mean daily states in the moisture flux divergence field for each AOGCM and the ensemble. This had earlier been proved to be an effective indicator of rainfall centres of dryness across an East African domain when related to the SOM in the reanalysis datasets. This was again the case here. All of the models and the ensemble showed a good characterisation of the expected moisture flux divergence anomalies associated with each node. Namely, dry states had a mean state of strong divergent anomalies associated with them and rainfall centres were always co-located with convergence.

Therefore, it was concluded that whilst there were weaknesses in the model representation of daily rainfall during the control period the moisture flux field showed considerable promise as a reliable indicator of the presence of dry or wet conditions as expected for days assigned to each SOM node. Therefore, the SOM continued to be used as a method to examine daily rainfall projections under a climate change scenario in following chapters; with particular care taken to conclusions drawn from the parameterised rainfall field unless reinforced through coherent moisture flux projections.

Chapter 7: Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario - Part I – Rainfall projections

Key to the climate change community is the ability to analyse climate change projections under increasing greenhouse gas emissions scenarios with a view to influencing policy, for example through adaptation to potential new climate conditions. For this it is necessary to establish what degree of confidence can be had in AOGCM projections, or, at a minimum, to consider projections with an understanding of the magnitude and nature of their weaknesses. Chapter 7, Section 7.1 answers the third research question, how do the East African short rains change in the daily field under the high emissions SRES A2 scenario at the end of the twenty-first century in coupled climate models. The remainder of this chapter addresses the fourth research question of this thesis; what patterns of daily rainfall are responsible for the change in future modelled rainfall? Work undertaken from Section 7.2 onwards is directed at attaining this result.

It is expected that greenhouse gases will increase in concentrations over the twenty-first century as depicted in the SRES emissions scenarios, for example in the A2 scenario (Nakicenovic and Swart, eds, 2000). Therefore, the period of the theoretically greatest projected changes in climate available from CMIP3 is examined here to evaluate the most extreme projections of potential changes to daily rainfall characteristics over East Africa, covering the period October to December 2090 to 2099.

The daily rainfall data for the 2090s for each model and the ensemble is applied to the ERA-40 normalised climatology SOM and the fit of the data and changes to the distribution of rainfall types during the short rains is analysed. Finally the projected rainfall changes are quantified in terms of the percentage of increases assigned to the different types of daily rainfall events over the domain.

7.1 Mean changes to East African rainfall under the SRES A2 scenario

Mean projected changes under the SRES A2 scenario are evaluated in order to provide the context within which to examine potential changes to the more complex daily rainfall characteristics of the region. This is initially a partial summation from Chapter 3, which examined the mean climate of the East African region, but here there is a focus on the short rains season, defined as the months of October to December for this project. For the duration of the chapter, as with earlier chapters, the model titled ‘Ensemble’ refers to the ensemble mean of seven of the eight models employed in this research. The GISS model, following the decision made in Section 6.5 of Chapter 6 is excluded from the ensemble. The annual cycle for this model shows in excess of 8mm/day rainfall at all latitudes and seasons from 20°S to 15°N over East Africa, with no spatial or seasonal variation. This makes the anomaly projections of the ensemble immediately unrealistic due to these known weaknesses of the climatology of this model, and so including it would distort the model ensemble anomaly results for the twenty-first century. The GISS model continues to also be excluded from analysis under a climate change scenario as an individual model for the same reasons.

Chapter 3 examined rainfall projections under three SRES emissions scenarios (A2, A1B and B1), with a comparative analysis of projected changes over the year as a whole and across individual seasons undertaken. For the year as a whole and the short rains seasons a

projected increase in monthly precipitation was found. Results for other seasons were disparate in nature. This chapter builds upon the results of Chapter 3, examining the projected rainfall increase for the short rains under increasing emissions in more detail, using daily data for the months of October to December.

7.1.1 Projected mean change to East African rainfall during the short rains under the SRES A2 scenario

Table 7.1 shows the mean daily rainfall during the short rains over the small East African domain (30°E-43°E, 12°S-10°N), for the climatology period (1971-2000) for the 20C3M run and the 2090s under the high emissions A2 scenario. All the models and the ensemble show a projected increase in daily mean rainfall during the months of October to December over the domain as a whole, with the increases in rainfall ranging between 0.4 mm/day and 1.2mm/day in the CSIRO and CNRM models respectively. The ensemble sees a mean daily increase of 0.806mm/day mean rainfall in the short rains period. One of the models with the most realistic annual cycle in terms of spatial patterns and amplitude is MRI. This model sees a projected increase of 0.7mm/day, in line with the projections for the ensemble and in the middle of the range of projections for the selected models.

The significance of the projected change in daily mean rainfall across the smaller East African domain is tested using a Z-Test. Each model has a statistically significant projected change in rainfall at the 99% level as demonstrated through the Z values shown in the final column of Table 7.1, all of which exceed the critical value of 2.326 necessary for significance at this level.

At the most rudimentary level, therefore, there is remarkable consistency in the projections for rainfall over the four focus countries as a whole, reinforcing the IPCC statement (Solomon et al, eds, 2007) for the East African short rains showing one of the most

consistent AOGCM rainfall projections globally. This statement requires further detailed evaluation, beginning with the examination of the spatial patterns of these projected changes to mean rainfall under a high emissions scenario.

Table 7.1 Daily mean short rains rainfall over East Africa (30°E-43°E, 12°S-10°N) for the climatology period, 2090s period and anomaly relative to the climatology

Model	1971-2000 Mean Daily Rainfall	2090s Mean Daily Rainfall	Change in Mean Daily Rainfall	Z Value
CCCMA	3.9	4.7	+0.8	8.866
CNRM	4.8	6.0	+1.2	12.389
CSIRO	2.9	3.3	+0.4	7.815
Ensemble	3.9	4.7	+0.8	20.586
GFDL	3.7	4.7	+1.1	14.137
MIUB	4.0	4.7	+0.7	7.420
MPI	3.0	3.7	+0.7	8.358
MRI	5.0	5.6	+0.7	11.190

7.1.2 Projected spatial changes to East African rainfall during the short rains under the SRES A2 scenario

Figure 7.1 shows the mean projected change in daily rainfall during the short rains season over East Africa in the 2090s from the climatology period (1971-2000) under the high emissions A2 scenario for the ensemble and each individual model. A common theme across all models is the projected increase in daily rainfall during the October to December period by the end of the twenty-first century over the geographical region of Tanzania, Kenya, Uganda and Ethiopia of between 0.3 and 3.0mm/day.

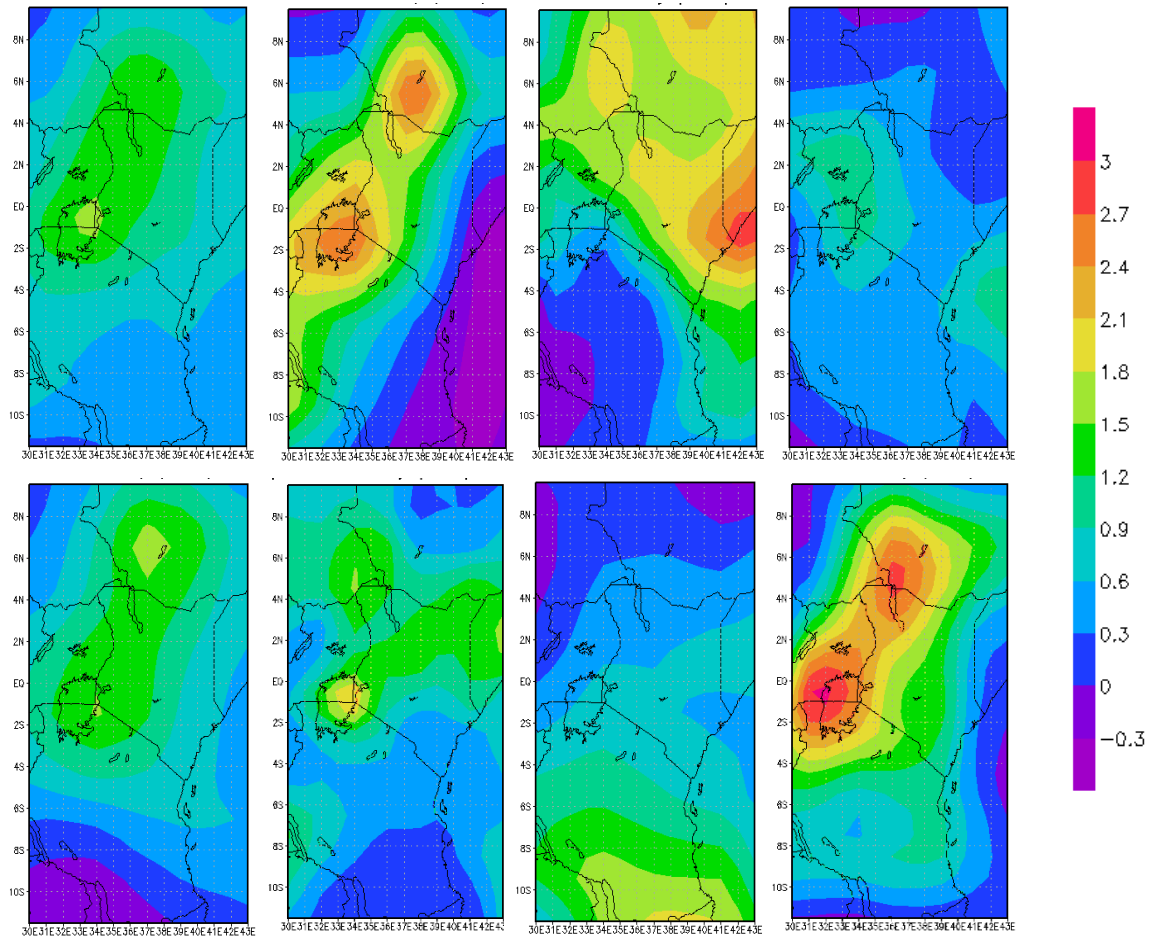


Figure 7.1 Mean projected rainfall change (mm/day for the short rains) for the 2090s from 1971-2000 (clockwise from top left) ensemble, CCCMA, CNRM, CSIRO, GFDL, MIUB, MPI and MRI

However there is considerable variation in the spatial patterns and magnitude of mean increase projected in the individual models. The ensemble sees a projected peak rainfall increase during the short rains of 1.8mm/day centred over Lake Victoria with increases of 1.5mm/day in the surrounding regions of eastern Uganda, west and north west Kenya and the south east of Ethiopia. The remainder of the East Africa domain also sees a projected increase in rainfall by the end of the twenty-first century of 0.3-0.9mm/day with the smallest projected changes in Tanzania measuring only 0.3-0.6mm/day.

Three of the individual models show this same spatial pattern of projected short rains rainfall increase; CCCMA, GFDL and MRI. The greatest increases of these three

AOGCMs is seen in GFDL, where peak projected rainfall anomalies of greater than 2.7mm/day relative to the 1971-2000 climatology are seen in a bi-centred pattern with the two centres located over southern Uganda (32°E, 1°S) and south-east Ethiopia (36°E, 5°N). In this model, the land regions of all four focus countries see projected increases upwards of 0.6mm/day to a maximum of 3mm/day, again with the smallest projected increases seen over Tanzania of 0.6-1.2mm/day. CCCMA also has some of the largest projected increases of the individual models following the same bi-centred pattern as GFDL with a peak increase of 2.7mm/day. Once more the lowest projected increases are over Tanzania, measuring 0-1.8mm/day; furthermore the east coast of Tanzania sees a slight projected decrease in rainfall by the end of the twenty-first century of 0.3mm/day which originates in a decreasing centre in the Indian Ocean. The MRI AOGCM is the final model of the seven to present the same spatial pattern as the ensemble, with this model also presenting a very similar magnitude of rainfall increase as the ensemble. The only difference is found over Tanzania where in MRI there is a gradient from north to south across Tanzania from a projected increase along the latitude of 2°S of 1.5mm/day to a slight projected decrease of -0.3mm/day along the south-west border of Tanzania with Malawi.

Of the four remaining models, three of them (CNRM, CSIRO and MPI) have very similar spatial patterns of projected rainfall increase during the short rains under a high emissions scenario with the greatest projected increases seen to the north of the domain and only very small projected increases of 0.3-0.9mm/day, if any, in the daily mean short rains rainfall over Tanzania.

The one model that dissents from this pattern of mean daily rainfall change by the end of the twenty-first century is the MIUB AOGCM. This model, whilst still seeing projected rainfall increases over the East African domain, has a different centre of maximum projected rainfall increase, here located in the far south east of the domain over southern

Tanzania (38°E, 12°S). Projected daily mean rainfall increases over Tanzania reach 1.8mm/day in this AOGCM, the largest increases projected over this country of any of the models. Elsewhere in the domain (northern Tanzania from 6°S northwards and the geographical areas of Kenya and Uganda) projected increases in rainfall are less than 0.8mm/day and over Ethiopia projected changes are minimal, with changes from the climatology period within the range of -0.3-0.3mm/day.

To summarise, it is over Tanzania where the models show the greatest differentiation in the projected short rains rainfall increases (ranging from decreases of -0.3mm/day to increases of +1.8mm/day), with the common factor the lesser magnitude of these projections in comparison to the projected daily rainfall over the north of the domain, specifically over Uganda and southern Ethiopia where the increase peaks at 3.0mm/day.

7.2 Application of 2090s data to the climatology ERA-40 SOM, fit to the SOM

Model projections of the mean change in daily short rains rainfall over East Africa at the end of the twenty-first century all show an increase in rainfall over the four focus countries of this region. There are, however, significant variations within these projections, with daily rainfall anomaly projections of between 0.3 and 3.0mm per day. This variation is evident in both the magnitude of the projected increase and the spatial patterns of the projections with centres of greatest increase located over Uganda, Tanzania, Ethiopia and the Kenyan-Somalia border across the model subset. With these differences evident in a mean plot of model projected rainfall changes for the months of October to December as a whole at the end of the twenty-first century, there is a high likelihood that a complex pattern of potential changes on the daily level could also emerge, which is now investigated. This follows the evaluation in Chapter 6 of how AOGCMs simulated daily

rainfall characteristics of the short rains over East Africa. It was concluded in Chapter 6 that the AOGCMS and model ensemble have two main weaknesses in their simulation of climatology conditions. First, rainfall has a tendency to be too longitudinally widespread covering the region from 30°E to 43°E, especially for those days associated with conditions of suppressed convection across the focus countries of East Africa where there are daily rainfall means of up to 8mm in comparison to less than 2mm/day in the ERA-40 and TRMM datasets. Second, rainfall centres have a tendency to be both too heavy in places (as demonstrated through days which should be dry) but also too dry particularly in the case of small intense rainfall centres which are not well simulated in the models. For example, node 4 1 has a maximum rainfall centre of 20mm in satellite and reanalysis datasets and only 12mm in the ensemble series. However, conversely the daily moisture flux divergence field in all of the models is a good approximation to the location of both rainfall centres and dry regions when applied to the daily rainfall matching of the model to the ERA-40 daily SOM. This suggests that it is weaknesses in the parameterised convection and related rainfall field that is the source of the difficulty in the replication of the rainfall daily fields of the models with the physically based moisture divergence field derived from specific humidity and wind fields being more successful at representing the daily characteristics of the East African short rains in a realistic manner.

As shown in Section 7.1, the mean projections for the OND season over East Africa is for increased rainfall over all of the focus countries. The SOM has been employed in Chapters 4 to 6 of this thesis, producing insight in to the characteristics of daily rainfall over East Africa under control conditions and here the aim is to further utilise this methodology to provide new insight into rainfall projections over the same region under a high emissions scenario.

As for the climatology period, data used in the analysis of the projected climate changes is normalised to the monthly climatology before being inputted into the SOM algorithm. The SOM size employed is the smaller of the two arrays used in this study; a 5 by 4 array covering twenty rainfall states in total.

7.2.1 Quantisation error of 2090s daily data when applied to the control period self-organising map

An initial evaluation of the daily rainfall data is undertaken by applying the 2090s timeslice to the same ERA-40 normalised rainfall SOM which formed the basis of the evaluation of the model climatology in Chapter 6. Whilst it is possible that different, and previously unseen, daily rainfall modes might emerge in East Africa under a climate change scenario, it is instructive to constrain the evaluation to changes in distribution of the same daily types (clustered in the ERA-40 5 by 4 SOM) that are present in the control period. This will allow for a more directed examination of projected changes to daily East African rainfall under a climate change scenario and provide the dominant direction for more thorough analysis if such a direction emerges.

Table 7.2 shows the mean quantisation error for each AOGCM and the ensemble when the SRES A2 2090s daily rainfall over East Africa was applied to the ERA-40 daily rainfall SOM. In a similar manner to the ensemble period quantisation error calculations (Chapter 6, Section 6.1.1), for many of the AOGCMs during the 2090s period the fit of the rainfall data has a lower mean error than for the dataset on which the SOM was trained and all models have a quantisation error of at least comparable size, covering the range of 12.317 to 13.914 in relation to the ERA-40 error of 13.239. It does not necessarily follow that the daily rainfall characteristics of the models over East Africa are similar to the ERA-40 dataset, as shown in Chapter 6 where the distribution of types of daily rainfall are

significantly different for each model in comparison to the ERA-40 dataset at the 99% significance level. Rather, it is likely that the models have a lower variability of daily rainfall types than seen in real world datasets and so can be assigned closely to a limited number of nodes without amassing an error too great over the time series. However, the relatively good fit (as defined by comparable quantisation errors) of the 2090s AOGCM data to the constrained rainfall patterns seen under control conditions do suggest that it is unlikely that any drastically different rainfall states to those set in the ERA-40 SOM will emerge by the end of the twenty-first century. This implies that changes projected are likely to be, whilst of no less importance, within the subset of daily rainfall states seen in the region during the control period with an increased likelihood of potential changes being related to the strength or number of occurrence of particular patterns within the short rains.

Table 7.2 Mean quantisation error when SRES A2 2090s data is applied to the ERA-40 1971-2000 normalised self-organising map

AOGCM	Quantisation Error
ERA-40 (1971-2000, for reference)	13.239
ENSEMBLE	13.914
CCCMA	12.687
CNRM	12.317
CSIRO	13.178
GFDL	13.242
GISS	13.609
MIUB	13.293
MPI	12.974
MRI	13.677

7.2.2 Distribution of daily states in the 2090s under the SRESA2 scenario

The distribution of daily rainfall types for each AOGCM and the ensemble is examined as an important stage in the evaluation of a twenty-first century climate change projection over East Africa under the third research question of this thesis. Changes in the distribution and relative occurrence of different daily types of rainfall could have a significant impact on the overall rainfall characteristics of the East African short rains which in turn could have consequences for agriculture and pastoral farming in the region. For example, Blackwell (2010) suggests that it is not changes to mean conditions which are the greatest risk to pastoralists but increasing extreme events, characterised by an expanding of the standard deviation of weather conditions in East Africa.

After the rainfall data for each model and the ensemble is applied to the ERA-40 climatology rainfall SOM, the number of occurrences of each node over the 920 (10 years of OND, 2090-2099) day test period is calculated and the distribution of the different rainfall types over the ten year period at the end of the twenty-first century is established. It is important to understand whether the distribution of different types of daily rainfall in each AOGCM and the ensemble are the same at the end of the twenty-first century under a high emissions scenario as they had been a century earlier. This is done through a chi-squared statistical test. Each model and the ensemble is considered as a separate entity, for example, the distribution of daily rainfall states in the ensemble in the 2090s is compared to the distribution of states between 1971 and 2000 in the ensemble. The occurrence of each node between 1971 and 2000 is treated as the expected population distribution with the 2090s occurrence of each node tested to see if it belongs to the same distribution. The absolute values of occurrence of each node (not the percentage) are inputted to the chi-squared test with the climatology values divided by three to scale from a thirty year period to a ten year period to align with the available daily timeslice of ten years of data for the

end of the twenty-first century. As a recap, Figure 7.2 shows the ERA-40 5 by 4 rainfall SOM to which the model data was constrained for this analysis.

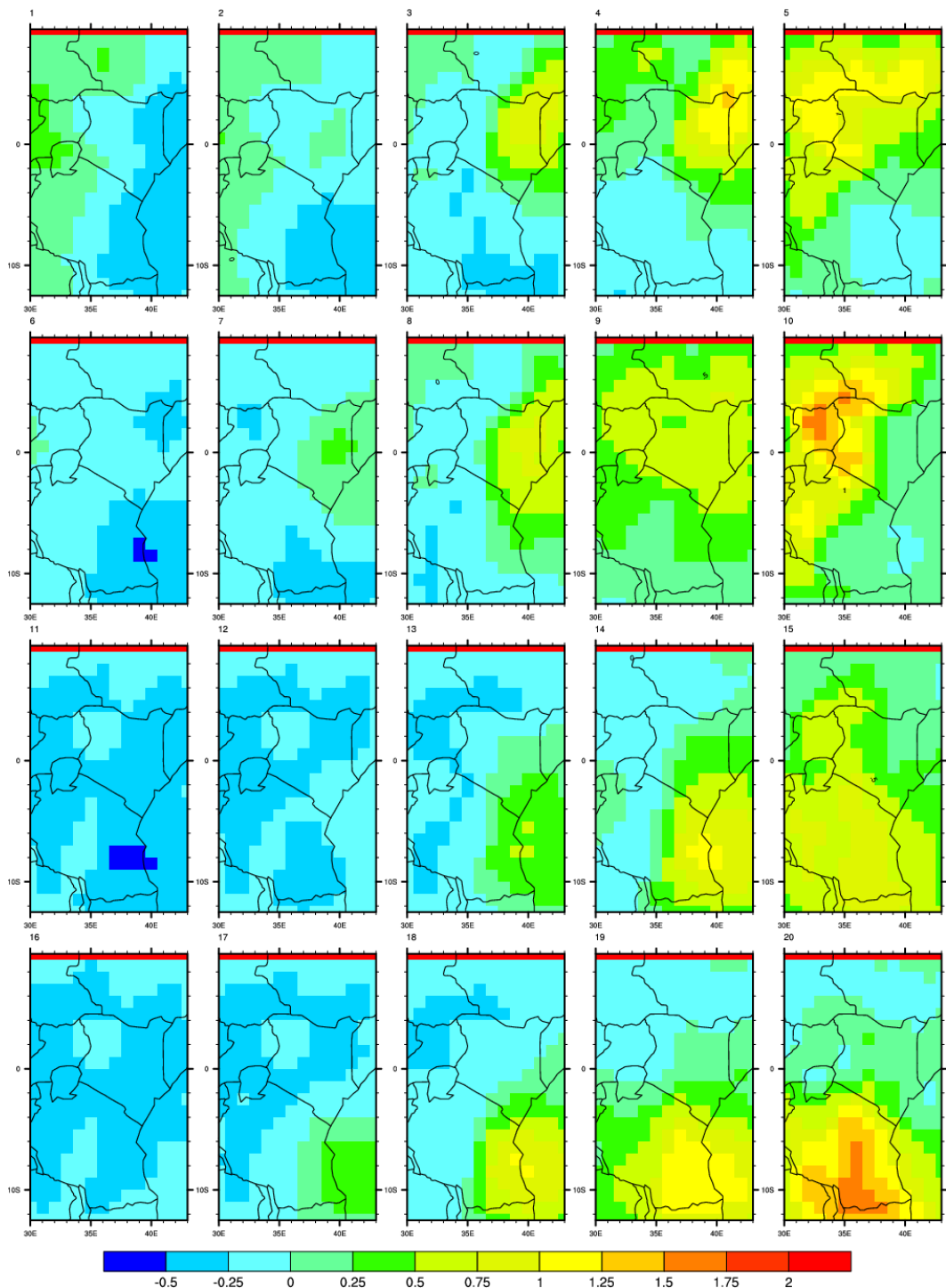


Figure 7.2 ERA-40 normalised rainfall SOM 1971-2000 OND, reference vector 1

Table 7.3 shows which of the AOGCMS have a statistically significant change in occurrence of daily rainfall nodes at the 95% level from the end of the twentieth century to the end of the twenty-first century.

Table 7.3 Models with a statistically significant (95% level) change in node distribution between 1971-2000 and 2090-2099 under the SRES A2 scenario

Model	Significant Change in Node Occurrence?	Chi-squared <i>p</i> Value
Ensemble	Yes	5.293E-05
CCCMA	Yes	3.933E-06
CNRM	Yes	0.0349
CSIRO	Yes	4.218E-06
GFDL	No	0.112
MIUB	No	0.069
MPI	Yes	0.001
MRI	No	0.219

As Table 7.3 shows, the ensemble and four of the seven individual CMIP3 AOGCMS show a statistically significant change in the occurrence of each node between 1971-2000 and 2090-2099 at the 95% level. These results, namely the combination of statistically different and non-statistically different changes to node distribution, suggest a situation more complex than the coherent increasing mean rainfall projection for the East African short rains under a climate change scenario that has emerged in the scientific literature thus far (e.g. Solomon et al, eds, 2007, Shongwe et al, 2011, Cook and Vizzy, 2013).

Interestingly, of two of the models which emerged as having some of the most realistic annual rainfall cycles during the twentieth century MRI shows no statistically significant change in the distribution of types of daily rainfall whereas CSIRO does see a statistically significant change in the distribution of daily rainfall states under the chi-squared test.

The twenty nodes of the five by four SOM array to which the data is matched represent a wide continuum of different spatial patterns of daily rainfall over the geographically vast East African region; with nodes representing daily rainfall centres over Tanzania, Kenya, Uganda, Ethiopia and the Congo of varying strengths, the Indian Ocean coastline and a number of dry states where suppressed convection dominates. A significant change to any one, or more, of these nodes could impact greatly on the mean rainfall profile for an individual country or region within East Africa even while the overall mean trend for rainfall is for a short rains increase.

Figures 7.3 (a) to (e) show the percentage change in occurrence of each daily rainfall state within the 5 by 4 ERA-40 SOM array for the ensemble rainfall series and the four models which show a statistically significant change in distribution of states between the climatology period and the end of the twenty-first century. It is not necessary to examine again at this point the distribution of those AOGCMs (GFDL, MIUB and MRI) for which there is no significant change to the distribution of daily rainfall types; these follow the same distribution of daily states as outlined in Chapter 6, Section 6.2.2.

-2.03	-0.33	0.32	-1.81	-1.34
-0.26	-0.91	1.30	-0.80	1.52
1.19	-0.18	0.90	1.12	1.30
0.87	-1.74	-0.98	1.12	0.72

Figure 7.3 (a) Percentage change in node occurrence from 1971-2000 to the 2090s - ensemble

0.83	-1.52	0.62	1.41	-1.16
0.83	1.05	-0.36	0.22	-1.01
0.69	-1.63	0.83	-0.43	-0.80
-4.09	2.72	0.58	0.00	1.23

Figure 7.3 (b) Percentage change in node occurrence from 1971-2000 to the 2090s - CCCMA

0.14	-1.16	-0.07	0.33	0.36
-0.83	0.94	-0.33	0.07	-0.72
1.52	2.21	0.51	-0.98	0.07
-2.90	0.91	-0.07	-0.04	0.04

Figure 7.3 (c) Percentage change in node occurrence from 1971-2000 to the 2090s - CNRM

4.06	1.12	-0.80	-0.29	1.12
0.18	0.76	-0.80	0.36	-1.81
-1.20	0.80	-0.54	0.94	-0.11
-0.65	-2.75	-0.69	1.20	-0.91

Figure 7.3 (d) Percentage change in node occurrence from 1971-2000 to the 2090s - CSIRO

-1.34	-0.94	-0.43	-0.14	0.22
0.14	-0.40	0.22	-0.33	0.25
2.75	-1.30	0.62	0.72	1.27
-1.01	-0.07	1.52	-0.81	-0.94

Figure 7.3 (e) Percentage change in node occurrence from 1971-2000 to the 2090s - MPI

The ensemble series sees a decrease in nodes located in the top left hand corner of the array (0 0, 1 0 and 0 1) representing a rainfall centre over the Congo (10mm/day mean) with dry conditions elsewhere. There is an increase of 2.06% in the occurrence of nodes 0 2 and 0 3, states depicting strong suppressed convection over the East African land region, though conversely a decrease in other states representing relatively dry conditions over the whole of the domain (0 1, 1 2 and 1 3, -2.18%). This suggests projections in the 2090s of a greater occurrence of very dry days within the short rains and fewer days with the encroachment of light rainfall originating from a rainfall centre on the fringes of East Africa. In terms of changes to states representing rainfall regions centred over the East

African domain there is a decrease in the amount of days when rainfall is centred over the Indian Ocean of -0.98% (node 2 3), the Kenyan-Ethiopian border of -1.81% (3 0) and when more generally widespread rainfall over the northern half of the domain of -2.14% (3 1 and 4 0).

However, the biggest increase in the occurrence of any state is for node 4 1 which sees a 1.52% increased occurrence in the 2090s relative to the 1971-2000 period. It is also for the node with the strongest rainfall over the domain; the bi-centred pattern over the Uganda-southern Ethiopian border, associated with mean rainfall of 12mm/day in the ensemble. This equally reflects the spatial pattern of the mean projected increase for the short rains for the ensemble with this node being co-located over the region where the largest mean daily increases are projected over the short rains as a whole of 1.2 to 1.8mm/day. Some of the next largest increases in frequency of occurrence of days within the short rains occur in the nodes located in the bottom right hand corner of the array (4 2, 4 3 and 3 3), totalling a 3.14% increase for these nodes combined. These nodes depict rainfall centres over Tanzania where the smallest projected increases in mean rainfall occur in the ensemble, this means that a more complex situation than a straightforward increase in rainfall is present, which will be investigated further in Section 7.3 when the changes to the amounts of rainfall associated with different daily modes of rainfall events over East Africa are investigated.

Aside from the ensemble, four of the seven individual models (CCCMA, CNRM, CSIRO and MPI) also show a statistically significant change in the distribution of daily rainfall modes over East Africa during the short rains under the high emissions scenario by the end of the twenty-first century. There are both similarities and differences in the patterns of change of occurrence of different nodes that these AOGCMs present. A similarity between all models is the redistribution of days assigned to nodes representing suppressed

convection over East Africa. All of the AOGCMs with statistically significant changes to node distribution except the CSIRO show an increase in the proportion of days within the short rains allocated to node 0 2. This node is the driest of the nodes representing suppressed convection over the domain, with dry conditions present over the four focus countries of Tanzania, Kenya, Uganda and Ethiopia and no light rainfall falling in the Indian Ocean affecting coastal regions of East Africa. The increase in the occurrence of this node (of 0.69%, 1.52% and 2.75% for CCCMA, CNRM and MPI respectively) is combined with decreases in the occurrence of other nodes representing suppressed convection during the short rains of up to 4.09%. Overall for three of the models, the ensemble, CCCMA and CSIRO, there is a decrease in the percentage of days allocated to nodes with dry conditions over the majority of the East African land region. This suggests overall the dominant projection for dry days over East Africa is for there to be fewer dry days during the short rains by the end of the twenty-first century, but that for those days which are dry, the dry conditions will be more widespread over a larger part of the East African region with less chance of fringe regions near the borders of Kenya, Tanzania and Uganda getting residual rainfall originating either in the Congo or Indian Ocean.

From closer examination of the individual models for which a change in distribution of daily short rains rainfall events over East Africa is present, whilst there are projected changes in the individual types of daily rainfall events which match neatly to the overall short rains projections for the region, there are also discrepancies which contradict the plots shown in Figure 7.1. To highlight some of the projected changes which map logically on to the projections of Figure 7.1, the MPI model sees a change in distribution which has the key features of an increase in nodes characterising daily rainfall events in the northern half of the domain, for example nodes 4 0 and 4 1, and a decrease in nodes representing daily rainfall events to the south of the East African region, particularly evident in nodes 3

3 and 4 3 with a combined projected decrease of 1.75%. This ties in with the mean projected increase in rainfall to be greatest north of 2°S, reaching a maximum of 2.1mm/day over Lake Victoria with lower increases of between 0.0 and 0.3mm/day relative to the climatology present over Tanzania by the 2090s under a high emissions scenario. Furthermore, this 0.3mm/day increase in short rains rainfall over Tanzania is very much a reasonable projection even in conjunction with the projection of a decrease in rainy days in this region in MPI due to the widely hypothesised likelihood of high atmospheric emissions and subsequently heating leading to a more active hydrological cycle and more heavy rainfall events, for which quantification is occurring (Polson et al, 2013). For example, there is robust evidence of an intensification of the global hydrological cycle to the magnitude of 8% ($\pm 5\%$) per degree of warming over the second half of the twentieth century, obtained through combining observed ocean salinity changes with global climate model projections. It is expected that this trend will continue through the twenty-first century (Durack et al, 2012). In contrast, the CNRM AOGCM projections are for decreases of between 0.11 and 0.91% in the occurrence of many nodes with rainfall centres over the East African land domain. These include days with widespread rainfall over the north of the domain (4 0), widespread Tanzanian rainfall (4 2 and 4 3), Indian Ocean rainfall centres encroaching on coastal regions of East Africa (2 2) and light rainfall over eastern Kenya (3 0).

There are also considerable discrepancies between the mean projection plots presented in Section 7.1.2 and the changes in daily distribution of rainfall events for the same time period and SRES emissions scenario. For example, the CCCMA and CNRM models both see a decrease (1.01% and 0.72% respectively) in the occurrence of node 4 1, representing the bi-centred pattern of heavy rainfall across the Ugandan-Ethiopian border. This is despite being the location for which the greatest increase in short rains rainfall is projected

for these two models. Evidently the projections are of greater complexity than it would initially appear. There could be plausible reasons for such disparity; for example if rainfall in this region in the future is projected to become more widespread, this could lead to these days mapping more effectively on to a different node even if the maximum increase in rainfall is seen over the same region over the short rains as a whole. Alternatively, rainfall intensity could be projected to increase, leading to more rainfall falling in this region over fewer days. The CCCMA model also projects an increase in the number of rainfall days where the main rainfall centre is located over Tanzania (0.43% over nodes 4 2, 4 3 and 3 3) in contrast to the overall projection for this area of the domain which in CCCMA is for a projected decrease in rainfall of up to 0.3mm/day by the end of the twenty-first century over the coastal regions of Tanzania. The source of these disparities is investigated in Section 7.3.

In addition, there are three AOGCMs that do not show a significant change in the projected future distribution of daily rainfall events over East Africa; GFDL, MIUB and MRI. Despite this, all three of these models project short rains rainfall increases over the majority of the East African domain under the high SRES A2 scenario by the 2090s period. It could therefore be said that for these models, the key to the increased rainfall projections is likely to arise within the intensity of rainfall projected within the same spatial distribution of rainfall over the domain as is present in the 1971-2000 period. These changes in intensity however, will not be consistent over the domain given that there is variation in the amplitude in projected rainfall change across the domain, with the greatest projected increases found over Tanzania for MIUB (of 1.8mm/day) but outside Tanzania for the other two models (GFDL and MRI). These features add to the intricacy of the emerging picture of climate change projections of climate change over East Africa.

To summarise, for the model rainfall series which project a statistically significant change in the distribution of the spatial pattern and amplitude of daily rainfall occurrences over the East African region a complex picture arises as to the nature of projected climate change across the region, despite the coherent model agreement on a general increase in short rains rainfall in East Africa. In addition, the absence of a significant change in node distribution in the remaining three of the selected AOGCMs (GFDL, MIUB and MRI) contributes to this complexity. Projected changes to the intensity of daily rainfall events are examined in Section 7.3, allowing a fuller picture to be built of likely climate change scenarios in East Africa.

7.3 Quantification of projected daily rainfall changes between 1971-2000 and 2090-2099

In conjunction with understanding the distribution of daily types of rainfall over the East African domain (Section 7.2), a key part of the analysis of projected future rainfall is to quantify how the projected mean increases in rainfall seen in each AOGCM are distributed across the domain through the clustered types of daily rainfall. Table 7.4 shows the projected mean annual increase in rainfall for the ensemble and each individual AOGCM during the October to December season.

Table 7.4 Mean annual short rains increase (mm) between 1971-2000 and 2090-2099 under the SRES A2 scenario

AOGCM	Mean Short Rains Total (mm) 1971- 2000	Mean Short Rains Total (mm) 2090- 2099	Mean annual rainfall change (mm)
Ensemble	356	431	75
CCCMA	355	430	75
CNRM	444	553	109
CSIRO	266	305	39
GFDL	334	432	98
MIUB	372	435	63
MPI	272	340	68
MRI	452	519	67

To reiterate, each individual model shows a positive rainfall anomaly between the end of the twentieth century and the end of the twenty-first century during the short rains, one of the most coherent model rainfall projections for any region or season globally (Solomon et al, eds, 2007). However, to fully understand this projection in more detail and to add useful information to the current climate literature as is the aim of this thesis, it is vital to be able to quantify the projected change in terms of daily rainfall events over different parts of the domain. For example, what proportion of the projected rainfall increase is associated with which type of daily rainfall events and does this vary between the models? This addresses the fourth research question of this thesis to establish which types of daily rainfall are responsible for the change in future modelled rainfall. Whilst the earlier work of this chapter revealed a complex picture in the spatial changes to rainfall over East Africa through the distribution of different rainfall types, it is imperative to be able to allocate the overall projected rainfall increase to particular types of daily rainfall over East Africa. This question is addressed through quantifying the change in contribution of each node type to

the total short rains rainfall between the two time periods. First, the mean annual short rains rainfall (in mm) associated with each node from 1971-2000 under the 20C3M model run is calculated. Second, the mean annual short rains rainfall associated with each node from 2090-2099 under the SRES A2 scenario is calculated. Following this, the mean annual difference (mm) in the mean rainfall contribution of each node to the overall short rains mean total is calculated through finding the difference between the mean contribution of the control and 2090s time periods. From this, it is possible to attribute the projected change in rainfall to particular types of rainfall, quantifying the percentage of projected change occurring due to different types of daily rainfall events over the domain.

Figure 7.4 shows the mean annual short rains rainfall in mm associated with each node for the control period (top line in each node) and the 2090s (middle line in each node) when the ensemble data is applied to the normalised ERA-40 rainfall SOM. The figures in parentheses depict the anomaly in mm of the annual mean contribution to the total short rains rainfall of each node. Plots for individual models are not shown here. These figures illustrate that there are substantial differences between the contribution of different spatial rainfall patterns to the overall projected short rains rainfall increase over the course of the twenty-first century.

34.99	10.25	10.31	20.50	29.53
34.49	10.61	13.66	16.99	28.97
(-0.5)	(+0.36)	(+3.35)	(-3.51)	(-0.56)
12.25	9.67	8.48	15.60	16.36
13.27	8.32	16.17	14.17	27.18
(+1.02)	(-1.35)	(+7.69)	(-1.43)	(+10.82)
37.27	5.92	8.38	10.64	16.82
47.83	6.61	13.66	19.07	26.23
(+10.56)	(+0.69)	(+5.28)	(+8.43)	(+9.41)
21.76	24.60	13.30	16.80	32.80
29.33	22.73	12.81	25.80	42.64
(+7.57)	(-1.87)	(-0.49)	(+9.0)	(+9.84)

Figure 7.4 Mean rainfall associated with each node for the ensemble when applied to the ERA-40 normalised rainfall SOM. Top row of each node is mm contribution of that node to OND for 1971-2000, middle row is mm contribution of that node for the 2090s and the bottom row is the anomaly for that node contribution in mm for the 2090s.

In order to fully understand these variations, changes to different types of daily events across the ensemble and the seven models are evaluated. For ease of analysis, the twenty rainfall states comprehensively covering the observed daily rainfall patterns of the twentieth century were divided in to sub-groups, clustered to represent the six main patterns that emerged under the climatology conditions of the control period, as seen in earlier chapters. These groups can be seen in Figure 7.5 and were allocated as follows:

Group A – These nodes cover states depicting rainfall down the western part of the domain, originating in the Congo and covering the western parts of Ethiopia, Uganda and Tanzania along with South Sudan, Rwanda and Burundi.

Group B – This group contains dry nodes, that is, those nodes where suppressed convection covers the whole of the East African land region within the domain. Node 13 has been allocated to this grouping as opposed to Group D even though it has a light Indian Ocean rainfall centre. This is because in this node conditions over land are very dry (less than 2mm/day mean rainfall in both ERA-40 and TRMM series), following the same spatial patterns of suppressed convection as other nodes in group B.

Group C – Nodes clustered in group C depict a rainfall event centred over the Kenyan-Ethiopian border, with rainfall focused along the east side of these two countries. Sometimes, light rainfall also occurs over South Sudan.

Group D – This group contains Indian Ocean rainfall events. The rainfall centre is located off the East African coast, adjacent with Tanzania. Some rainfall encroaches on coastal East Africa mainly through Tanzania but also through coastal Kenya.

Group E – Group E contains rainfall events characterised by the bi-centred rainfall pattern centred over northern Uganda and south-western Ethiopia with widespread rainfall over much of the north of the domain and north-west Tanzania.

Group F – These nodes are characterised by widespread rainfall over Tanzania, with the rainfall centred over this country. There is some encroachment of rainfall northwards through the domain over Kenya and Uganda.

0 0	1 0	2 0	3 0	4 0
0 1	1 1	2 1	3 1	4 1
0 2	1 2	2 2	3 2	4 2
0 3	1 3	2 3	3 3	4 3

Figure 7.5 Sub groups of daily rainfall types for the quantification of projected rainfall projections at the end of the twenty-first century under the SRES A2 scenario

Table 7.5 shows the percentage of the projected mean short rains increase for the ensemble and each individual AOGCM associated with each node grouping. The percentage increase associated with each node is calculated by dividing the mean daily rainfall change per node in the 2090s by the mean short rains total change in the 2090s and then multiplying this figure by 100. The percentage increases for each grouping are then calculated by summing the percentage increase in the rainfall allocated to each node in that group. As is evident from this table, the projected increase in short rains rainfall over the twenty-first century under a high emission scenario is not distributed equally over the different types of daily rainfall present over the domain for either the ensemble or any of the individual AOGCMs.

Taking the ensemble, the most notable feature is the contribution of just two rainfall state groupings of more than 62% of the overall projected short rains rainfall increase. These two groups are group B, representing suppressed convection over the domain, and group F, characterised by daily rainfall events centred over Tanzania and encroaching northwards through the domain. Group A, clustering events of light rainfall down the western side of

the domain accounts for 10% of days during the 2090s but does not contribute at all to the projected positive short rains rainfall anomaly relative to the control, suggesting that in the ensemble these events see either little change or days with heavier rainfall are countered by days with lighter rainfall than seen during the end of the twentieth century. The remaining three groups contribute approximately 40% of the projected short rains increase between them. These groups include the daily events over the Kenya-Ethiopia border, the bi-centred Ugandan-southern Ethiopian pattern and Indian Ocean rainfall centres. To conclude, the quantified daily increases in rainfall accounting for the mean positive change in the East African short rains rainfall total at the end of the twenty-first century originate from a variety of locations over the domain, although this is skewed to events located over the eastern part of the domain and the increase is very much dominated by wetter ‘dry’ days (24.2%) and increased Tanzanian based rainfall (38.0%) in the ensemble.

Table 7.5 Percentage of projected short rains East African rainfall increase accounted for by each grouping for the ensemble and individual AOGCMs.

Group	Ensemble	CCCMA	CNRM	CSIRO	GFDL	MIUB	MPI	MRI
A	-0.2	9.6	6.4	48.7	14.8	16.3	5.4	0.2
B	24.2	31.0	24.6	6.3	23.8	31.2	20.5	34.5
C	8.3	27.9	21.3	7.7	15.6	22.2	11.3	22.3
D	17.8	9.4	2.9	6.0	19.4	10.2	22.9	25.7
E	11.9	6.4	24.4	9.3	21.2	-5.9	30.4	9.4
F	38.0	15.7	20.4	22.0	5.2	26.0	9.6	8.0

There is significant variation in the contributions of different node types to the overall projected rainfall increases over the seven individual models and the ensemble.

Nonetheless, some primary similarities do emerge. The primary feature which emerges is the high proportion of rainfall increase associated with group B; days with suppressed

convection over the land portion of the East African domain. East Africa has a relatively low annual rainfall total for its tropical latitude, as demonstrated through the global GPCP rainfall climatology shown in Chapter 1, Figure 1.2. These dry days are associated with at least 20% of the projected short rains increase by the 2090s for the ensemble and six of the seven individual AOGCMs and over 30% in three of the models; CCCMA, MIUB and MRI. The only exception to this is CSIRO, an AOGCM which additionally has the smallest projected short rains increase of all the examined models, measuring only +38.95mm relative to the period 1971-2000 by the 2090s, approximately 25mm less than MIUB, the model with the next smallest projected rainfall increase during the October to December season.

The contributions of days which are characterised by suppressed convection to the projected short rains increase over East Africa presents an interesting result for further investigation. The climate change literature contains some contrasting results as to the possibilities of changes to dry days under increasingly warm conditions over East Africa. Lott et al (2013) state that dry days, manifested as the failure of the rainy seasons, are likely to become drier in the future whereas in contrast Shongwe et al (2011) suggest it is likely that dry days over East Africa will become less dry by the end of the twenty-first century, with drought periods decreasing in severity. Therefore, further investigation to the characteristics of these dry days is necessary to fully address research questions three and four of this thesis.

In terms of the other groupings of different daily rainfall states over East Africa during the short rains, there is considerable variation in how the projected increases are accounted for between different models. CSIRO is the only model which does not see at least one fifth of the projected short rains increase allocated to dry days. Instead, the smaller short rains rainfall increase seen in this model is dominated through group A, which accounted for

48.7% of the projected increase under the A2 scenario. Group A is a cluster of events characterised by light rainfall down the western side of the domain, largely outside of the focus countries of this study. CSIRO is the only model for which a large proportion of the rainfall increase is attributed to this type of rainfall event. Of the remaining projected increases, the second most dominant grouping for the CSIRO model is group F; the bi-centred rainfall pattern centred on Uganda and southern Ethiopia. This grouping represents 22% of the projected short rains increase.

With regards to groups C to F, the clustering of daily rainfall events centred over the four focus countries and the Indian Ocean, they all feature in various combinations over the other six individual models to contribute 20% or more to the projected short rains increases by the end of the twenty-first century, alongside the large contributions by the dry states of group B. For CCCMA, the daily rainfall events over the Kenyan-Ethiopian border of group C contribute significantly to the overall rainfall increase (27.9%). Group C also contributes greatly to the short rains increase for CNRM, MIUB and MRI. A notable feature of four of the six models with a bias towards dry days contributing significantly to projected short rains increases is the distribution of remaining increase in rainfall over groupings covering events in both the north and south of the East African domain.

CNRM has between a 20 and 25% contribution to the projected 108.87mm short rains increase for events located over the Kenyan-Ethiopian border, Ugandan-south Ethiopian border and also in the south of the domain centred over Tanzania. The MIUB model also sees large contributions (22.2% and 26.0% respectively) for group C, Kenyan-Ethiopian border events, and group F, Tanzanian centred events. Two models also have significant contributions in the south of the domain from the Indian Ocean centred events of group D. These two models are MPI (22.9%) and MRI (25.7%). The rainfall centre over the north of the domain that contributes significantly to the rainfall increase alongside group D for MPI

is group E, the bi-centred rainfall pattern mapped on to Uganda and southern Ethiopia, and contributing 30.4% of the projected short rains increase. For MRI the second group is group C, representing rainfall events on the Kenyan-Ethiopian border and contributing 22.3% of the 66.71mm projected increase.

The quantification of the projected increase in rainfall over East Africa under a high emissions scenario by the end of the twenty-first century produces some interesting similarities and differences across the ensemble and seven individual AOGCMs. One model stood out as an exception for two main reasons; CSIRO. First, it is the only model to not see at least 20% of the contribution to the projected short rains rainfall increase come from days of suppressed convection across the domain. Second, it has by far the smallest projected mean annual short rains increase of the models examined here of just 38.95mm by the 2090s under a high emissions scenario. Aside from this model, the agreement across the ensemble and six remaining models for at least 20% of the projected short rains increase to originate from dry days where suppressed convection dominates the domain is remarkable. Additionally, many of the models also see important contributions to the projected increases from rainfall events located in both the north and south of the domain. The quantification of the projected rainfall increases over East Africa reveals contrasting results to those of Section 6.2. Whilst an examination of the change in distribution of different daily events over East Africa points to an increase in occurrence of the driest days over the region, the accompanying changes to rainfall indicated that it is these days which are responsible for a large part of the projected short rains wetting over East Africa under a climate change scenario. Consequently, the key question that arises from these results is why do dry days show the accompanying increase in overall rainfall associated with these states? Understanding this is imperative to evaluating how realistic the increasing short rains rainfall projections are for this region and this is addressed in

Chapter 8, forming a key contribution to the overall thesis aim of addressing how sensible climate model projections during the short rains for East Africa are.

7.4 Chapter summary

The key conclusion that has emerged from this chapter is related to the complexity of the rainfall projections over East Africa that emerged from examination of the daily rainfall field in a subset of CMIP3 AOGCMs, contributing to the third and fourth research questions of this thesis. In relation to the third research question, the mean rainfall statistics for the East African short rains showed a consistent, and statistically significant, increasing projection under rising greenhouse gas emissions. The majority of models and the ensemble showed the greatest projected short rains rainfall increases spatially located over Uganda and southern Ethiopia, peaking at 3mm/day in GFDL. Only one model, MIUB, had the region of greatest projected increases over southern Tanzania, measuring 1.8m/day.

Attention was then turned to the fourth and final individual research question, to identify which daily rainfall patterns are responsible for the overall change in model rainfall.

Notably, all of the models except CSIRO attributed a large proportion (20-35%) of the projected rainfall increase to days associated with a state of suppressed convection within the SOM node. Given the weaknesses that emerged in Chapter 6 relating to the simulation of the daily rainfall field in AOGCMs, namely the inability of the models to remain dry on days of suppressed convection, and the unexpected result of such a large proportion of the projection being attributed to days characterised by the suppression of rainfall, rather than widespread rainfall further investigation into this projection is necessitated in order to achieve the overall aim of this thesis. This is obtained in Chapter 8.

Chapter 8: Changes to daily rainfall regimes during the short rains in a CMIP3 AOGCM subset by the 2090s under the SRES A2 scenario - Part II – Spatial changes and moisture flux projections

Chapter 8 aims to build on the quantification of the projected rainfall increase, established in Chapter 7. Chapter 7 has shown that for six of the seven models and the ensemble mean, between 20 and 35% of the projected rainfall increase for the 2090s results from an increase in rainfall generated from days characterised by widespread dry conditions in the contemporary climate classification.

In order to explain the regimes responsible for the remaining 65-80% of the projected rainfall increase, the focus in this chapter is initially directed to how the projected mean short rains increase of the 2090s relative to the period 1971-2000 is distributed across the domain in terms of spatial anomalies in daily rainfall means associated with each node of the SOM array. This analysis reveals the nodes responsible for rainfall increases.

Following this, moisture flux divergence patterns, the most consistent circulation indicators of rainfall events and dry conditions over East Africa during the short rains in the contemporary climate are examined in order to establish whether the rainfall changes occur for physically discernible reasons. The AOGCMs showed a consistent weakness in their ability to reproduce contemporary events of suppressed convection in the rainfall field under contemporary conditions. Given the importance of days of suppressed convection to the explanation of why East Africa becomes wetter in the models in the twenty-first century, it is important to investigate the moisture flux field associated with

the rainfall types in the models. A coherent physical association with the moisture flux field will therefore aid interpretation of the model wetting response.

8.1 Ensemble rainfall – 2090 to 2099 under the SRES A2 high emissions scenario

In Chapter 7, it emerged that the only common mechanism by which the models become wetter in the 2090s is through the 20-35% proportional increase in rainfall on days associated with suppressed convection. This result is evident in six of the seven models and the ensemble mean under the SRES A2 scenario. The remainder of the increase (65-80% dependent on the model) was divided between different types of daily rainfall event over East Africa, with no agreement amongst the models as to which types of events were responsible for the projected increase. To fully understand these results it is necessary to explore the origins of the increase in different types of daily rainfall events further.

The daily ensemble rainfall series is used as a starting point to investigate further the daily rainfall projections, under a climate change scenario, that emerged in Chapter 7. The reasons for using the ensemble as a starting point are twofold. First the ensemble, as compiled of the seven individual AOGCMs, provides a ‘best estimate’ of the mean change projected across the subset of models examined here. Second, the ensemble is one of the data series (alongside CCCMA, CNRM, CSIRO and MPI) that shows a statistically significant change to the proportional occurrence of different daily rainfall types during the short rains under a high emissions climate change scenario. Additionally, the ensemble is characteristic of the dominant pattern across the AOGCMs, namely for the states of suppressed convection, which contribute significantly to the overall projected short rains wetting for the region, comprising 24.2% of the ensemble increase.

The same groupings as for Section 7.3 for the clustering of daily rainfall types are continued here. These are repeated briefly below for clarity and shown in Figure 8.1 (a). Figure 8.1 (b) shows the ERA-40 normalised rainfall SOM (based on the 1971-2000 data series) forming the basis of these groupings.

Group A – Rainfall along the western part of the domain, originating in the Congo and covering the western parts of Ethiopia, Uganda and Tanzania along with South Sudan, Rwanda and Burundi.

Group B – Dry nodes, where suppressed convection covers the whole of the East African land region within the domain.

Group C – Rainfall centred over the Kenyan-Ethiopian border.

Group D – Indian Ocean rainfall events. Some rainfall encroaches on coastal Kenya and Tanzania.

Group E – Bi-centred rainfall over northern Uganda and south-western Ethiopia with widespread rainfall over the north of the domain and north-west Tanzania.

Group F – Widespread rainfall over Tanzania, with the rainfall centred over this country. There is some encroachment of rainfall northwards through the domain over Kenya and Uganda.

0 0	1 0	2 0	3 0	4 0
0 1	1 1	2 1	3 1	4 1
0 2	1 2	2 2	3 2	4 2
0 3	1 3	2 3	3 3	4 3

Figure 8.1 (a) Sub groups of daily rainfall types for the quantification of projected rainfall projections at the end of the twenty-first century under the SRES A2 scenario

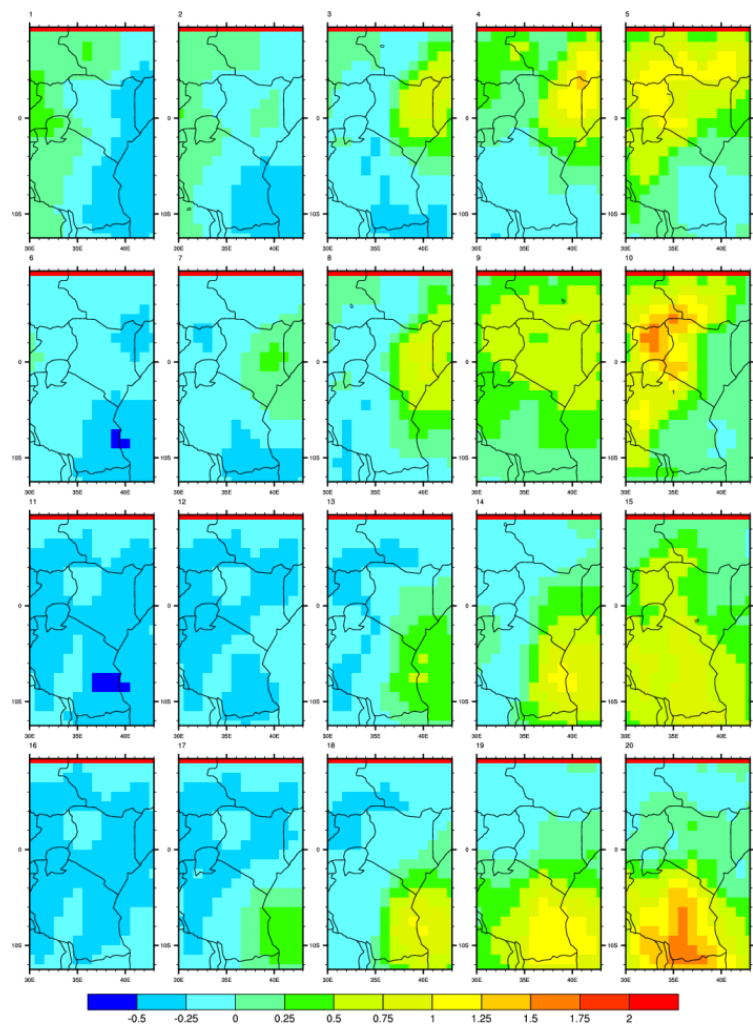


Figure 8.1 (b) 5 by 4 ERA-40 normalised rainfall SOM used as a basis for evaluation

Figure 8.2 shows the rainfall anomaly in mm per day for each node of the SOM array for the 2090s period when the ensemble data is applied to the ERA-40 climatology SOM. This plot scale is constrained to a projected change range of -2.0 to +3.0 mm/day to reflect the smaller range of projected changes present in the ensemble mean in comparison to individual models and for clarity of spatial patterns in the ensemble mean. Table 8.1 (originally Table 7.5, repeated here as a recap) shows the percentage increases of each grouping of nodes responsible for the overall projected short rains increases for the ensemble mean and each model.

Table 8.1 Percentage of projected short rains East African rainfall increase accounted for by each grouping for the ensemble and individual AOGCMs.

Group	Ensemble	CCCMA	CNRM	CSIRO	GFDL	MIUB	MPI	MRI
A	-0.2	9.6	6.4	48.7	14.8	16.3	5.4	0.2
B	24.2	31.0	24.6	6.3	23.8	31.2	20.5	34.5
C	8.3	27.9	21.3	7.7	15.6	22.2	11.3	22.3
D	17.8	9.4	2.9	6.0	19.4	10.2	22.9	25.7
E	11.9	6.4	24.4	9.3	21.2	-5.9	30.4	9.4
F	38.0	15.7	20.4	22.0	5.2	26.0	9.6	8.0

There is not a uniform increase in rainfall projected across the domain associated with each mode of daily rainfall seen during the short rains as visualised in the SOM (Figure 8.2). For some types of daily rainfall patterns, parts of East Africa are associated with projected rainfall decreases, often over Tanzania, in particular where a large increase in rainfall intensity has been projected elsewhere in the region. For example, nine of the twenty nodes within the array have projections for rainfall decreases of between 0.5 and 2mm/day over Tanzania relative to the twentieth century but these decreases are accompanied by projections for increased rainfall elsewhere in the domain.

The greatest positive anomalies in the ensemble series are associated with rainfall centres over Kenya, Uganda and southern Ethiopia, in particular nodes 4 0, 3 1 and 4 1 comprising group E; days assigned to these nodes have a positive rainfall anomaly of 3mm/day. Indeed, none of the states in the SOM array see a projection of a negative rainfall anomaly over this region (Uganda through to southern Ethiopia) regardless of the strength or exact location of the daily rainfall centre, which makes an increase in daily short rains rainfall over these countries a particularly robust projection in the seven model ensemble. Nonetheless, despite the strong positive anomalies in the rainfall centres for Group E, negative rainfall anomalies of 1mm/day over Tanzania for days associated with these nodes means that group E is responsible for just 11.9% of the projected mean short rains increase by the 2090s when the complete domain is considered.

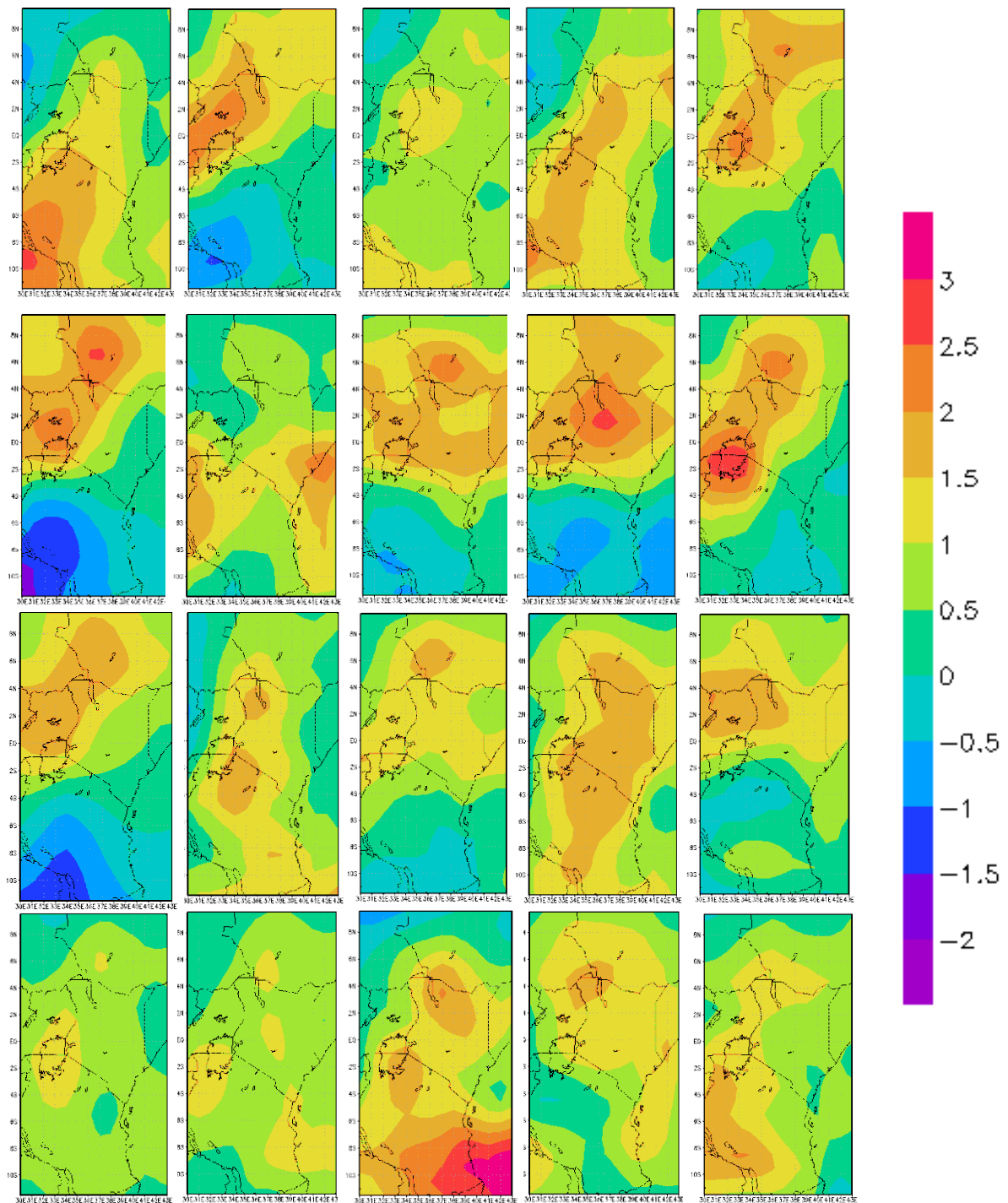


Figure 8.2 Rainfall anomaly (mm/day) for the ensemble 2090s relative to 1971-2000 for each node of the SOM array during the short rains

Projections for both daily rainfall increases and decreases are seen over Tanzania throughout the SOM array. Of nodes with rainfall events located over Tanzania (groups D, Indian Ocean centre with encroachment and F, rainfall centred on Tanzania) some have

projections for drying and others for wetting. Specifically, nodes 2 3 and 4 3 have an ensemble projection for increased rain over Tanzania on those days peaking at 3mm/day and 2mm/day respectively. Conversely, where the dominant rainfall centre is located north of the equator, including in nodes 2 1 and 3 1, there is a mean negative rainfall anomaly projected over Tanzania of 0.5-1mm/day every time this rainfall pattern occurred during the short rains. These contrasting projections in Tanzanian rainfall, dependent on rainfall centres elsewhere in East Africa, create a mixed picture. Overall for the ensemble, group F with rainfall centred over Tanzania contributes 38.0% of the rainfall increase; negative anomalies over Tanzania during days where rainfall was centred elsewhere reduces the impact of such days on the projection for East Africa as a whole.

Whilst there is a decrease in occurrence in the number of days with widespread rain over the north of the East African domain, these days are projected to have increased intensity of rainfall, accounting for the increased mean short rains rainfall increase projected for the north of the region as demonstrated in Chapter 7, Section 7.2. In the ensemble, over the Indian Ocean there is a decrease in occurrence of days associated with heavy positive rainfall anomalies (node 2 3, -0.98% occurrence) and an increase in the number of days projected to have a negative rainfall anomaly by the end of the twenty-first century (node 2 2, +0.9% occurrence). The large projected increase in rainfall in this region for some days (for example, node 2 3, rainfall anomalies of +3mm/day) means that these states have a net positive contribution to the rainfall increase even though many days in this region are projected to be drier. The nodes within the SOM array representing the suppression of convection over much of the East African land region (Group B) see a slight projected increase in rainfall of approximately 1mm/day but also there is drying over Tanzania in the nodes 0 1 and 0 2 which see a net increase in occurrence. The rainfall increases associated with these days seem muted in comparison to some days. Nonetheless, the high occurrence

of these days, with group B accounting for 34.9% of all short rains days during the 2090s, means that the impact of this increase is disproportionate in comparison to other groupings and has a large influence on the overall short rains projected increase. More generally, the projected anomalies for the ensemble for each daily rainfall state in the 2090s under a high emissions scenario are much more muted than in the individual models, which see projected mean daily increase of more than 10mm, as would be expected when a variety of different projections are combined, although this did not impact on the overall projected short rains wetting.

The mean node rainfall anomalies for the ensemble reinforce the impact that a frequently occurring node with just a small positive rainfall anomaly (group B, 0.5-2mm/day) can have on the overall projection. Individual models are now examined to investigate these findings further.

8.2 Individual AOGCM rainfall – 2090 to 2099 under the SRES A2 high emissions scenario

The combination of projected changes to daily rainfall states in the ensemble in terms of mean daily rainfall totals associated with each node and the quantified contribution they made to the overall projected rainfall increase highlights considerable variability in projections to different types of rainfall events over East Africa. The individual model projections of mean rainfall changes associated with each node in the SOM array are now examined here, both for models which show a statistically significant projected change in distribution of daily rainfall states in East Africa under a climate change scenario and those with no statistically significant projected change in the distribution of daily rainfall types during the short rains. Again, this aims to reconcile the quantified projected increases of daily rainfall to different event types to the spatial changes projected for these days. This

will increase understanding of projected changes to each model outside of the coherent projection across the models of a projected increase allocated to days of suppressed convection. First, models with statistically significant changes to the distribution of daily rainfall types are examined followed by those with no significant change in distribution.

Figure 8.3 (a) to (g) show the mean projected rainfall change for days assigned to each node during the October to December 2090-2099 period for each of the seven selected AOGCMs. The scale for these plots has been set for a greater range of projected change than for the ensemble to allow for illustration of the full range of projections found within the individual models, but is consistent across all the individual AOGCMs, ranging from -10mm/day to 10mm/day for the 2090s relative to 1971-2000. Table 8.1 shows the proportional increase of each rainfall grouping within the SOM array to the overall projected short rains increase for each model.

8.2.1 CCCMA

CCCMA is one of the models with a statistically significant change in the distribution of the daily rainfall types over East Africa under the high emissions SRES A2 scenario between 1971-2000 and the 2090s. In terms of the projected increase this model is placed in the middle of all the models with a projected short rains increase of 75.09mm by the 2090s relative to 1971-2000. The key projected change evident in this plot (Figure 8.3 (a)) is that whilst the majority of the nodes show a projected rainfall anomaly of between -2mm/day and 4mm/day, there are some nodes associated with large changes of greater than 10mm/day. The nodes with the particularly high projections of climate change are all related to rainfall centres over the north of the domain, specifically for group E nodes of 4 0, 3 1 and 4 1, located in the top right corner of the SOM array. Node 3 1 has a projected increased occurrence under increasing emissions and nodes 4 0 and 4 1 have decreased

occurrence by 1.16 and 1.01% respectively. Interestingly, despite the very large positive rainfall anomalies associated with this group ($>10\text{mm/day}$), these nodes contribute only 6.4% of the projected short rains increase over the domain as a whole. This is due to a combination of a low frequency of occurrence of these nodes alongside negative rainfall anomalies over Tanzania for these nodes of up to 6mm/day . However, on a more localised spatial scale, the balance is still for a large positive rainfall anomaly ($4\text{--}10\text{mm/day}$) in the 2090s for rainy days over Uganda and southern Ethiopia, in alignment with the mean short rains projection for this model.

There is an increase in the occurrence of nodes where the daily state is for an Indian Ocean rainfall centre (group D), but all of these nodes are additionally associated with a projected decrease in rainfall over the Indian Ocean of between 2 and 4mm/day . More widely, in the CCCMA AOGCM the Indian Ocean is projected to see a decrease in rainfall across fifteen of the twenty nodes in this SOM array of 2mm/day . In contrast, over the land region of this East African domain there is always a positive rainfall anomaly projected for each daily state of at least 2mm/day . This indicates a potential shift in convection away from the Indian Ocean to the land domain of East Africa. The increased rainfall projection over Tanzania seen in the daily rainfall states seems to arise as a result of positive rainfall anomalies originating over the Congo but exerting a wider spatial influence; the positive rainfall anomalies for the twenty-first century predominantly spread eastwards from the western half of the domain. In particular this pattern dominates the anomaly seen in western Tanzania whereas during the control period there is a greater amount of rainfall over eastern Tanzania, controlled by Indian Ocean rainfall. This suggests that in this model, Tanzania will see potentially very different changes to the mean rainfall under a high emissions scenario with each half of the country being controlled by influences adjacent to it.

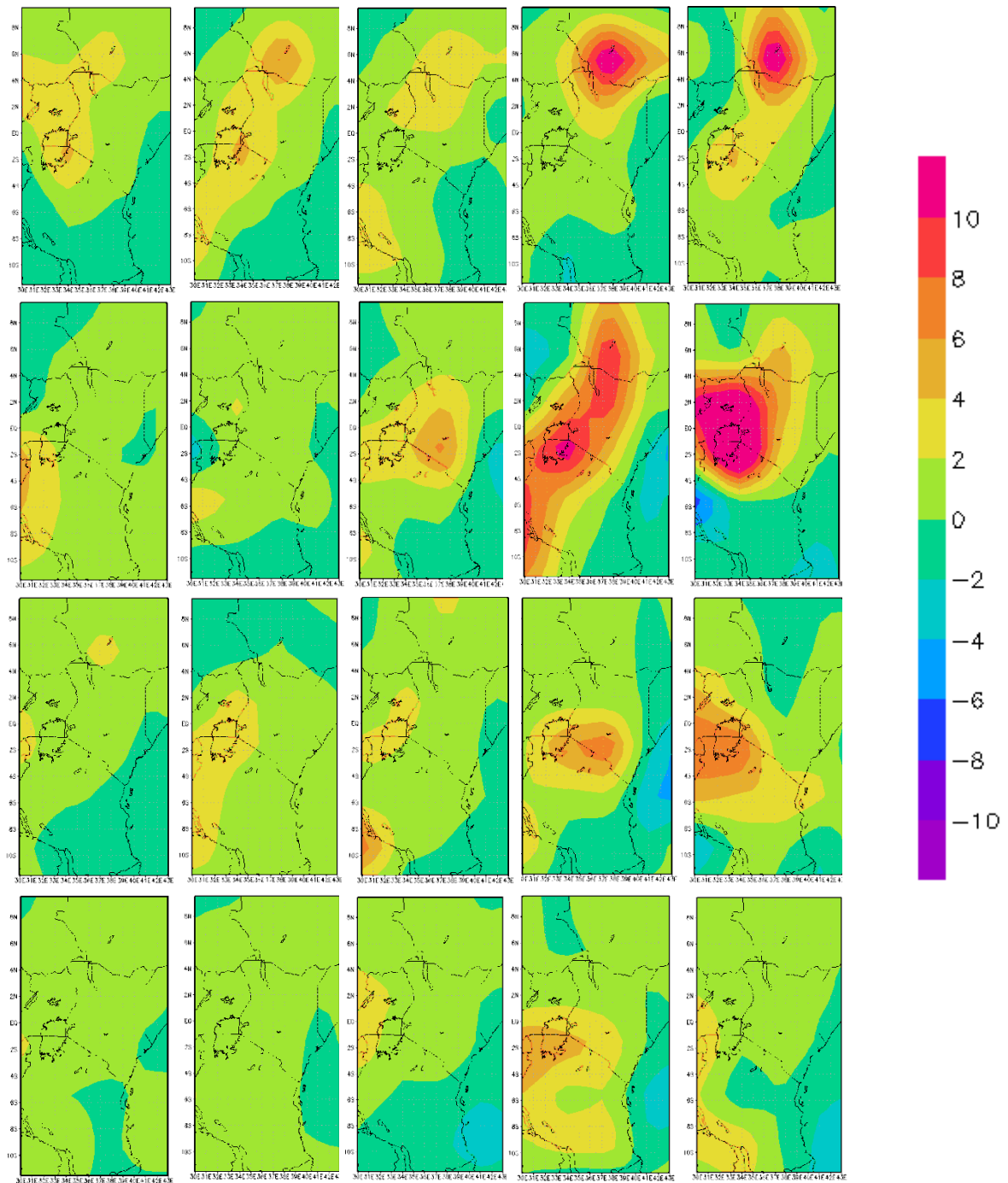


Figure 8.3 (a) Rainfall anomaly (mm/day) for CCCMA 2090s relative to 1971-2000 for each node of the SOM array during the short rains

8.2.2 CNRM

The CNRM model also has a projection for a statistically significantly different distribution of nodes under the SRES A2 scenario and is the wettest model of those examined with a projected positive short rains anomaly of 108.87mm by the 2090s. As with CCCMA, the key feature of this model is that although the majority of nodes are projected to see small to moderate rainfall increases of up to 4mm/day a few see much larger projected positive rainfall anomalies of over 10mm/day. Three such nodes are 4 0, 3 1 and 4 1 of group E, the bi-centred rainfall patterns focused over Uganda and southern Ethiopia. Overall these three nodes are projected to remain approximately the same in terms of frequency of occurrence, so the net effect is for large positive rainfall anomalies for rainy days over the north of the domain including over the Ugandan-southern Ethiopian border. Combined with no regions of drying for these nodes, this results in these nodes having a significant contribution to the overall rainfall increase during the short rains of 24.4%.

In addition to this strong similarity with the CCCMA projection, there is also a notable difference which relates to the Indian Ocean response to increased emissions. In the CNRM model no nodes see a projected decrease in Indian Ocean rainfall and indeed several, including 3 3, 3 2, 2 1 and 2 2, see projections for significant positive rainfall anomalies of up to 10mm/day. Although some of these states are projected to occur slightly less, with the total difference over all states approximately -1.5%, the increase in associated rainfall is large still. A further difference can be seen in Congo rainfall, group A; only 6.4% of the projected rainfall increase is attributed to rainfall in this region. It appears that influence from this region on the focus region of East Africa is decreasing which has led to slight negative rainfall anomalies being projected over western Tanzania in many spatial configurations of daily short rains rainfall patterns over the region,

totalling twelve of the twenty states and measuring a decrease in rainfall of 2mm/day. This is confirmed by the mean climate projection for this model.

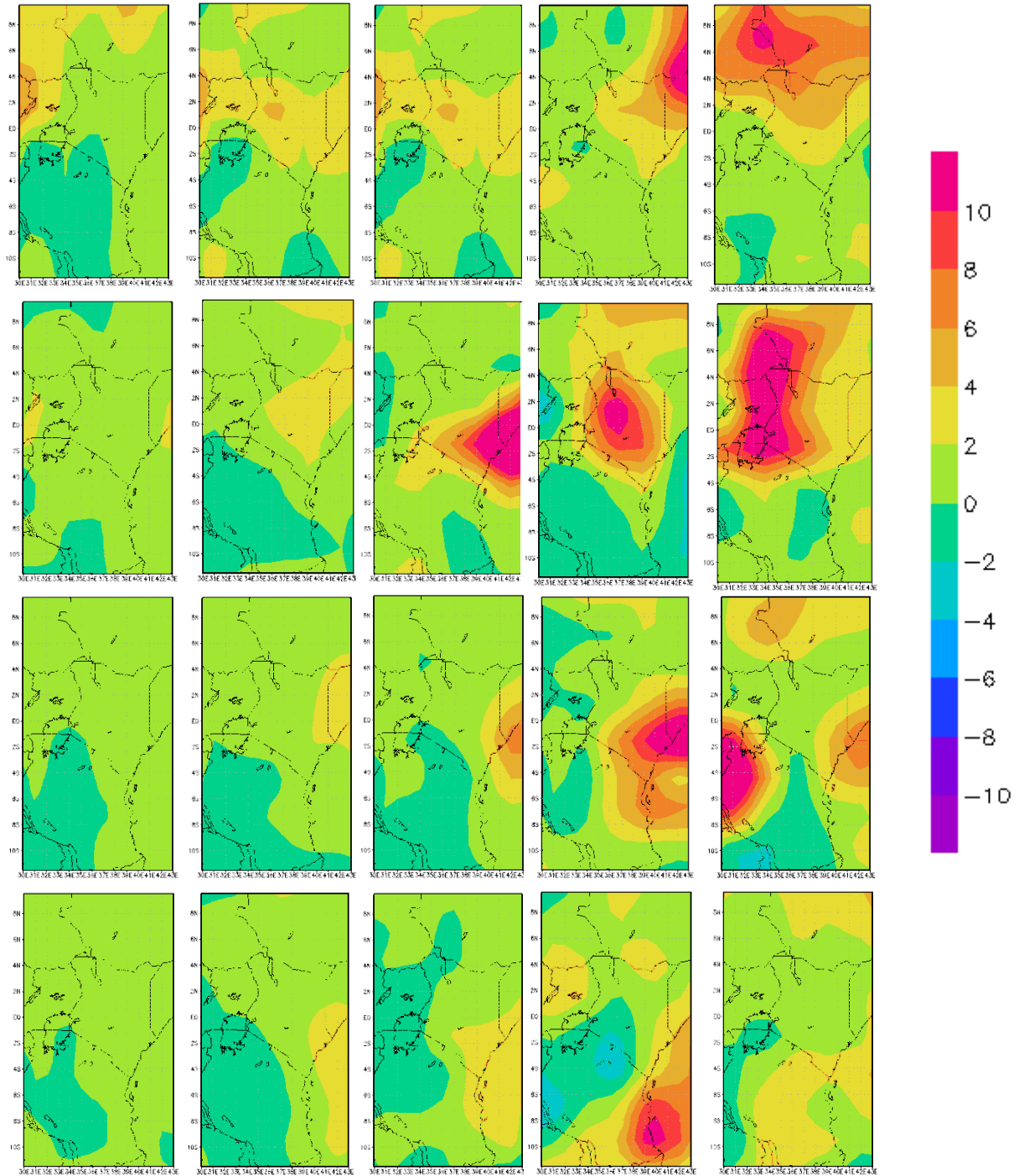


Figure 8.3 (b) Rainfall anomaly (mm/day) for CNRM 2090s relative to 1971-2000 for each node of the SOM array during the short rains

8.2.3 CSIRO

As with the two previously examined individual AOGCMs, CSIRO projects a statistically significant change in node distribution over East Africa by the 2090s. Overall this model projects the most moderate change to daily rainfall patterns in the region; the mean projected increase during the short rains is 38.95mm (next lowest projection is 63.38mm). This is also evident in the projected changes to rainfall associated with individual nodes with all of the twenty nodes within the SOM array seeing a projected rainfall anomaly of -2mm/day to 4mm/day with two nodes seeing a change of up to 6mm/day, both over Tanzania. Notably, the dry states of group B see very neutral projected rainfall changes, with regions of anomalies either side of 0. This explains why in CSIRO just 6.4% of the projected rainfall increase originates in this region even though these states account for 36.7% of all days in the short rains in the 2090s.

The two nodes with the greatest positive anomalies over East Africa are 2 3 and 4 3 where there is a positive anomaly of 6mm/day over a Tanzanian rainfall centre. These nodes are both projected to decrease in occurrence, explaining the smaller increase projected over Tanzania during the short rains as a whole in comparison to the northern half of the domain. It is projected that there will be fewer wet days over Tanzania but that rainfall events occurring will be of greater intensity, with the states of group F contributing 22.0% to the overall CSIRO increase in rainfall. Outside of this region CSIRO does not project any notable changes to daily rainfall states.

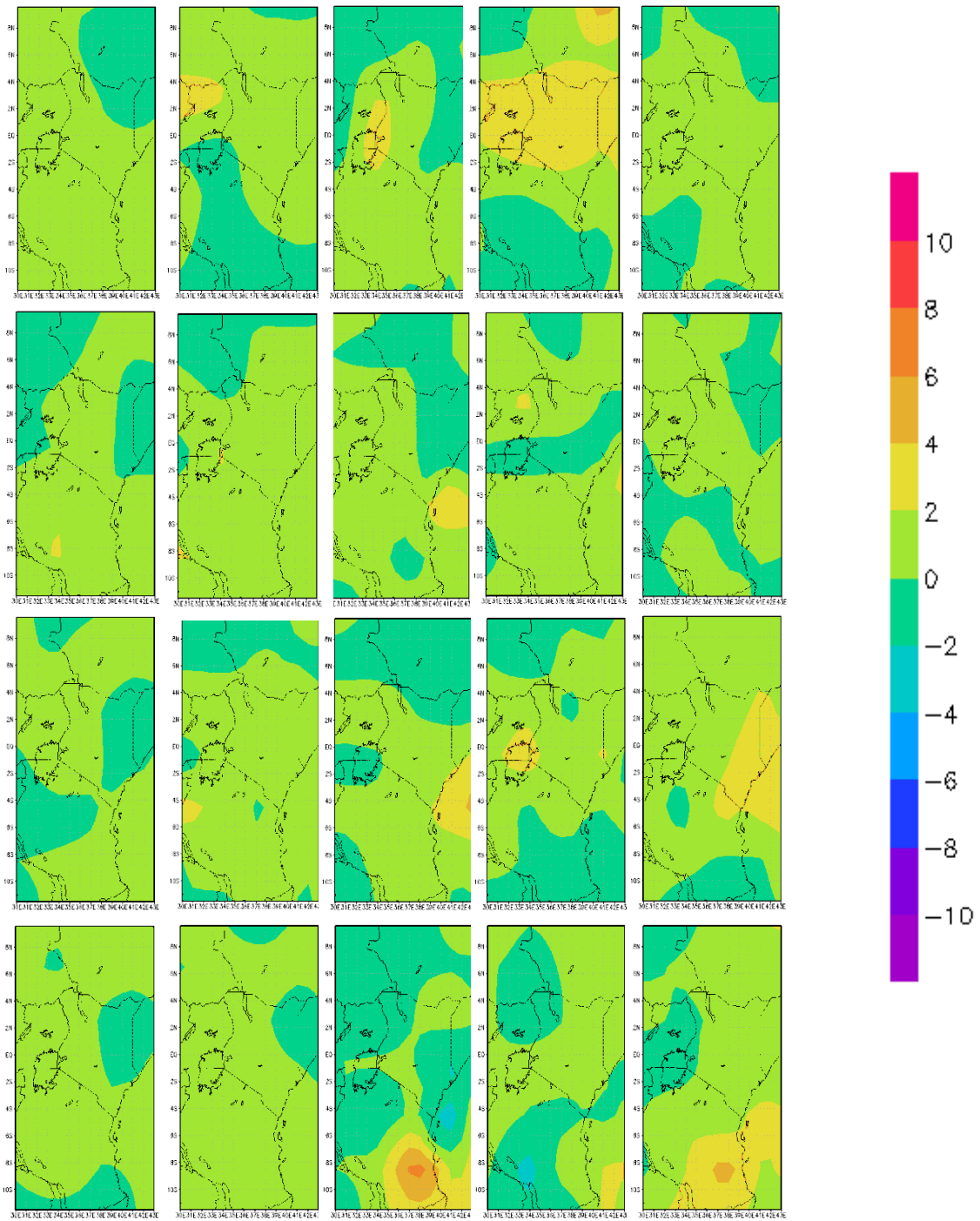


Figure 8.3 (c) Rainfall anomaly (mm/day) for CSIRO 2090s relative to 1971-2000 for each node of the SOM array during the short rains

8.2.4 MPI

MPI is the final model of the seven examined in more detail with a statistically significant change in node distribution over East Africa by the 2090s and has a short rains anomaly of 68.32mm projected in the 2090s. Over the short rains the greatest projections for increased rainfall are seen outside of Tanzania and indeed nodes with the rainfall centre over Tanzania (Group F) contribute only 9.6% of the total projected short rains increase. This model does not have any states with very large (10mm+/day) projected rainfall increases as the other models do; the greatest mean daily increases are of 8mm/day for node 4 1 with no projected increases of greater than 6mm/day for other nodes. Nonetheless, some interesting points emerge; states representing suppressed convection over much of the domain (group B) see no large projected increases, with the increase in occurrence of the driest of these states suggesting a tendency to a larger number of dry days in the future under a climate change scenario. On the other hand, the quantification indicates that still these states contribute 20.5% of the 68.32mm mean short rains anomaly, a significant amount. The greatest projected rainfall increases are found in nodes 4 0, 3 1 and 41 (group E), with node 4 1 in particular having projected rainfall increases of up to 8mm/day over the bi-centred Ugandan-southern Ethiopian rainfall node. Combined with the positive anomalies elsewhere over the domain for these states this group contributes more than 30% to the overall short rains wetting under the A2 scenario. Total occurrence of these few nodes is approximately the same in the 2090s projections, highlighting why this region is co-located with the area of maximum rainfall in the mean short rains plot for MPI due to the increased rainfall amplitude in this region relative to the rest of East Africa. Also reflected in the mean rainfall plot is the above average projected rainfall increase for Kenyan rainfall centres, although these states contribute less to the overall domain projected increase due to negative rainfall anomalies in some parts of the domain. The

picture that emerges over the Indian Ocean in this model is of greater contrast. Node 2 3 shows a projected increase in rainfall (4mm/day) associated with an Indian Ocean rainfall centre but conversely node 3 3 shows a projected decrease to a similarly placed rainfall centre (-4mm/day). Node 2 3 is projected to increase in frequency, meaning the net projection for this region is still for a rainfall increase but, as with previous models, this disparity in projections when some other regions in this relatively small domain show significant agreement raises interesting questions on potential controls on Indian Ocean rainfall off the East African coast by the end of the twenty-first century and how this impacts on East African land based rainfall.

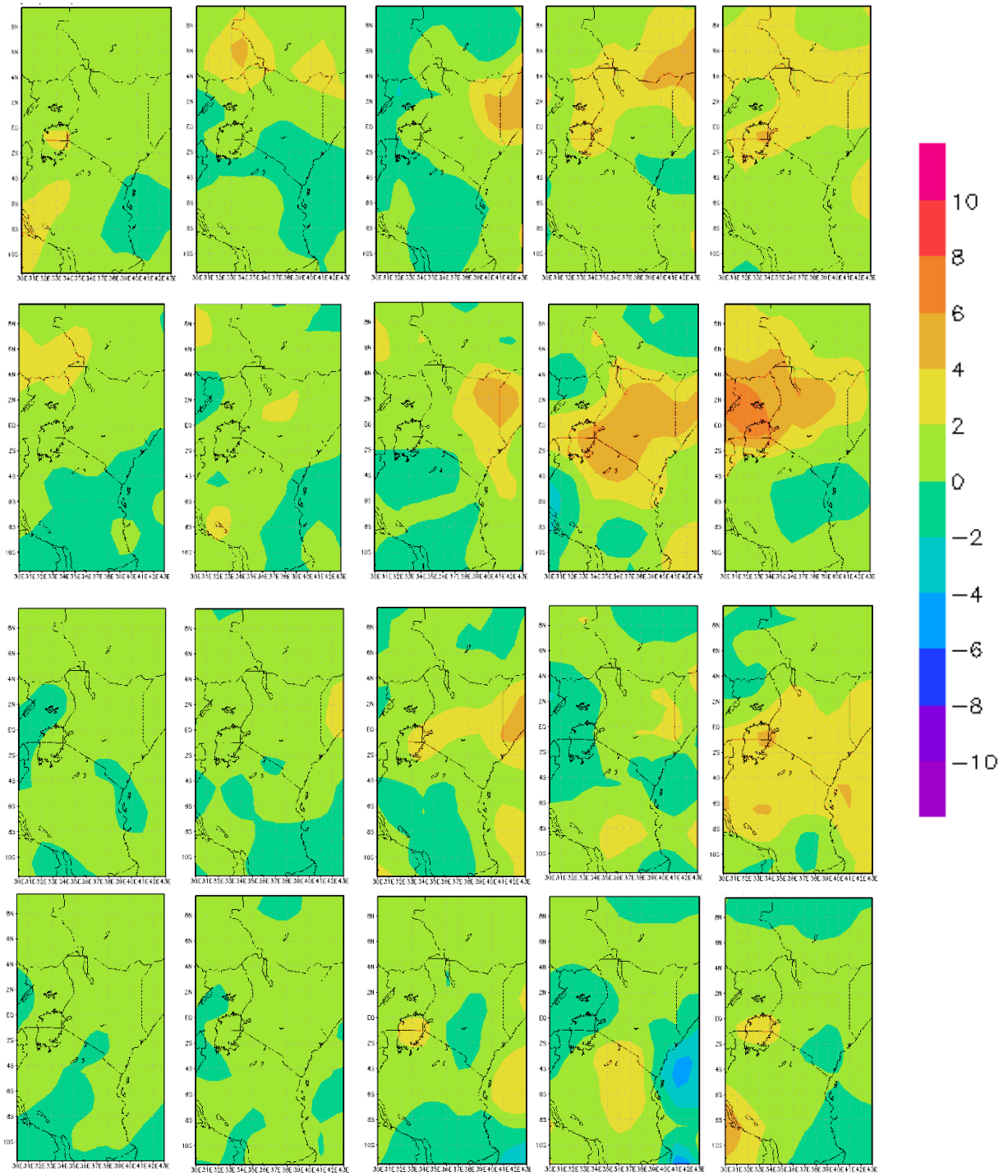


Figure 8.3 (d) Rainfall anomaly (mm/day) for MPI 2090s relative to 1971-2000 for each node of the SOM array during the short rains

8.2.5 GFDL, MIUB and MRI

GFDL, MIUB and MRI are AOGCMS with no projected significant change in the occurrence of different types of daily rainfall over East Africa under a high emissions climate change scenario. They have projected short rains anomalies of 98.16mm, 63.38mm and 66.71mm respectively. For GFDL, all the nodes are generally associated with a projected increase in rainfall which is greatest over Uganda and through to Ethiopia, aligning neatly with the spatial location of the biggest mean change of the short rains season during the 2090s relative to the 1971-2000 climatology and reflected in the 21.2% contribution of the states of group E to the overall projected rainfall increase in this model.

MIUB is slightly unusual in terms of its mean projection over East Africa; it is the only one of the selected AOGCMs to show the greatest projected increase in rainfall over Tanzania rather than further north. This can be explained through the large increase in projected rainfall (of up to 10mm/day) located over the Tanzanian rainfall centre in node 3 3, with large projections of positive Tanzanian rainfall anomalies of up to 8mm/day also found in nodes 2 3 and 1 3. Indeed, this is reinforced further through the second largest contribution to the overall mean increase coming from these states (group F) and comprising 26.0% of the short rains increase over the domain as a whole. In contrast to CCCMA where rainfall is projected to decrease over the Indian Ocean on many days, it appears that the projected intensity increase in the Tanzanian rainfall centres arises from a connected increase in the intensity of Indian Ocean rainfall centres of up to 8mm/day (group D) which causes rainfall to encroach significantly over Tanzania, especially the coastal regions. MIUB is the model with greatest projected increases in the magnitude of Tanzanian rainfall centre, raising interesting questions as to why this region is so different in this model to some of the others. Outside of Tanzania and the adjacent Indian Ocean, projected rainfall increases are relatively even over much of the domain, largely measuring

2-4mm/day and nodes with these centres contribute less to the overall rainfall increase than other patterns.

The overriding pattern for MRI is that there is very little difference in the anomalies in terms of spatial variance of projected change patterns between the individual nodes. All projected changes for individual nodes are very moderate, falling in the range of -2mm/day to 2mm/day for each node with some small regions projecting changes of 4mm/day for some rainfall patterns over East Africa. Notably, some nodes see a small projected decrease (2mm/day) in Tanzanian rainfall, specifically nodes 0 1, 1 1, 0 3 and 1 3 associated with suppressed convection. The largest projected changes (4mm/day) are found over western Kenya to southern Ethiopia, with nodes with this centre (group C) also having a significant contribution to the overall increase in the short rains rainfall in this AOGCM, recorded at 22.3%.

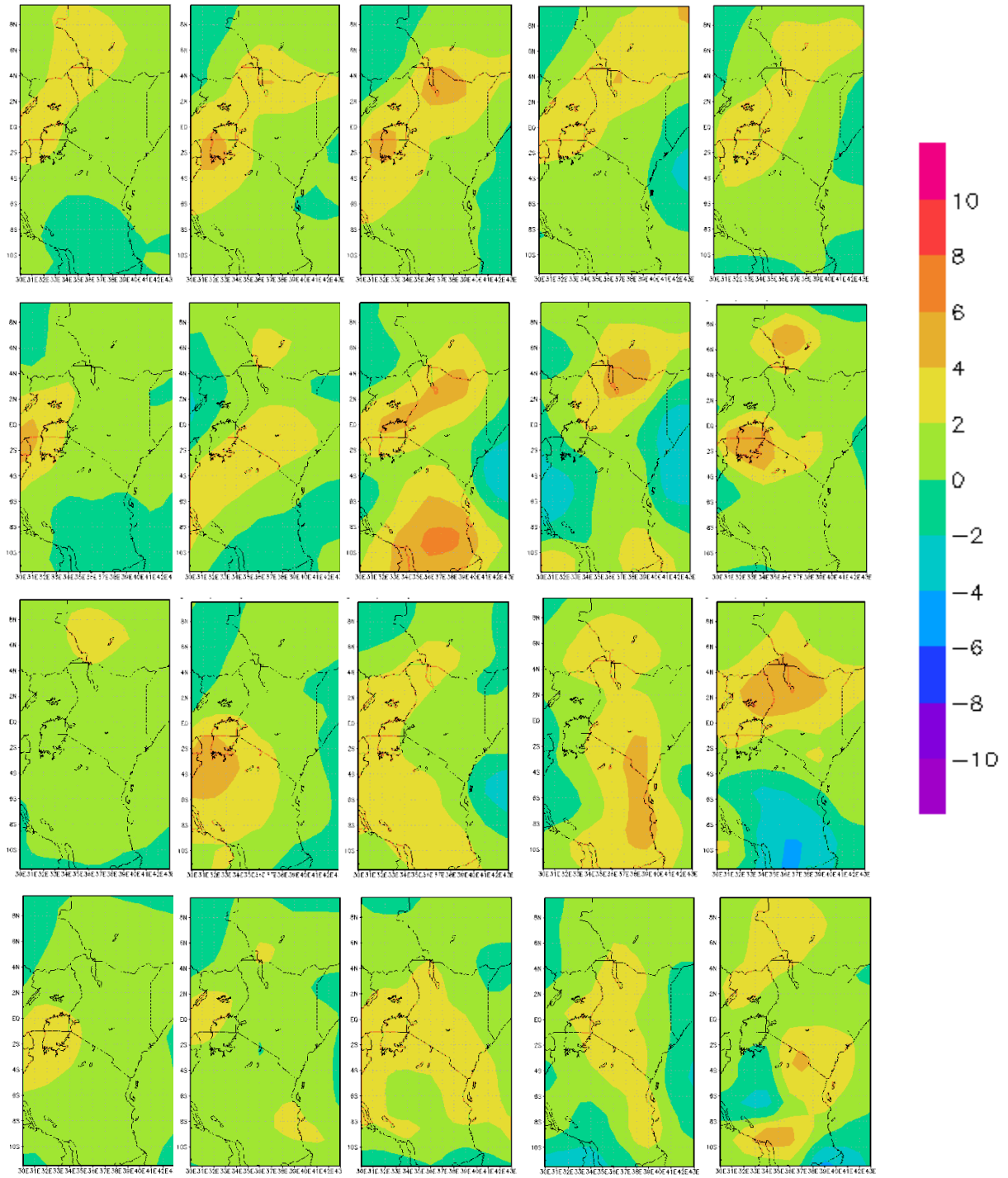


Figure 8.3 (e) Rainfall anomaly (mm/day) for GFDL 2090s relative to 1971-2000 for each node of the SOM array during the short rains

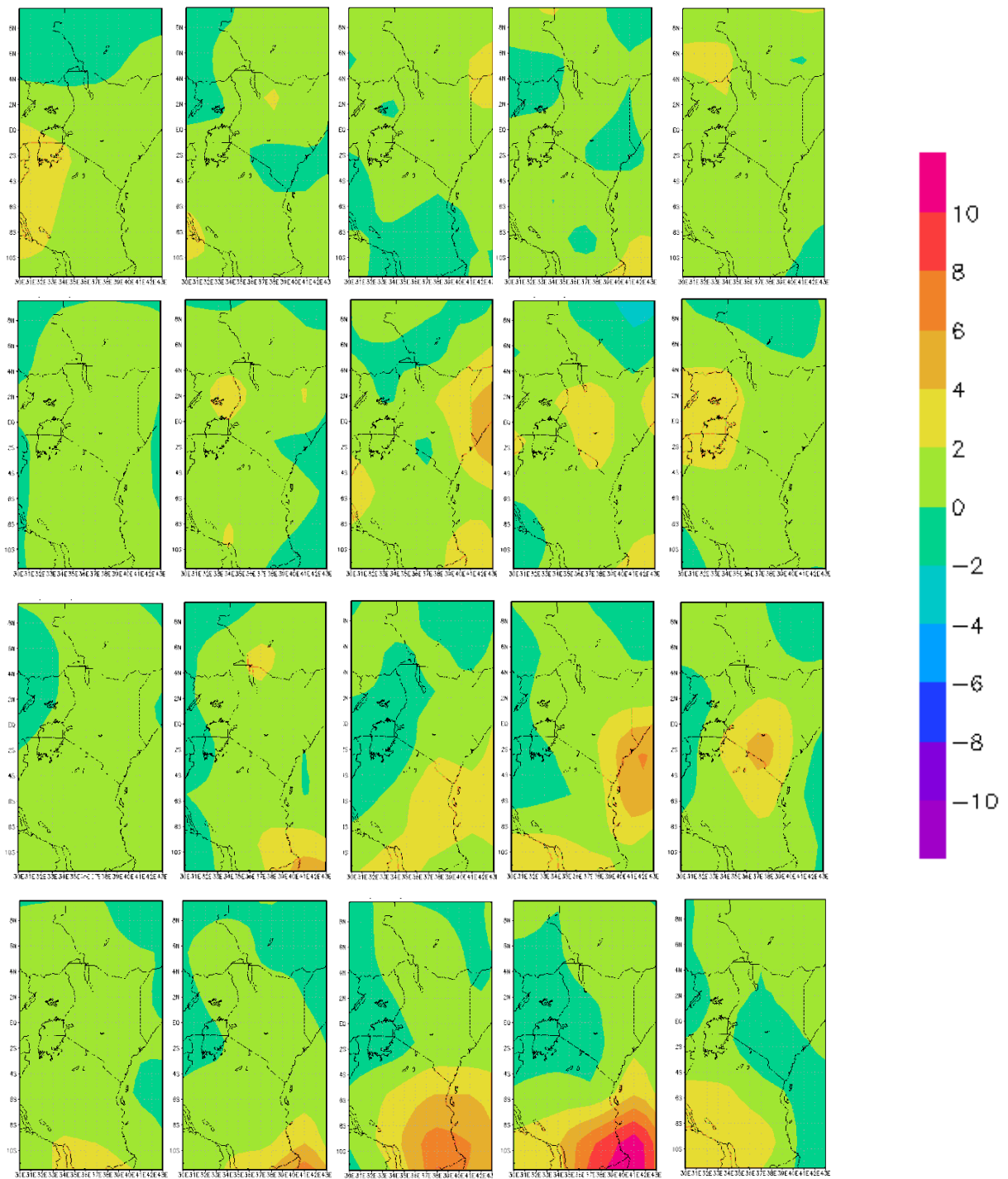


Figure 8.3 (f) Rainfall anomaly (mm/day) for MIUB 2090s relative to 1971-2000 for each node of the SOM array during the short rains

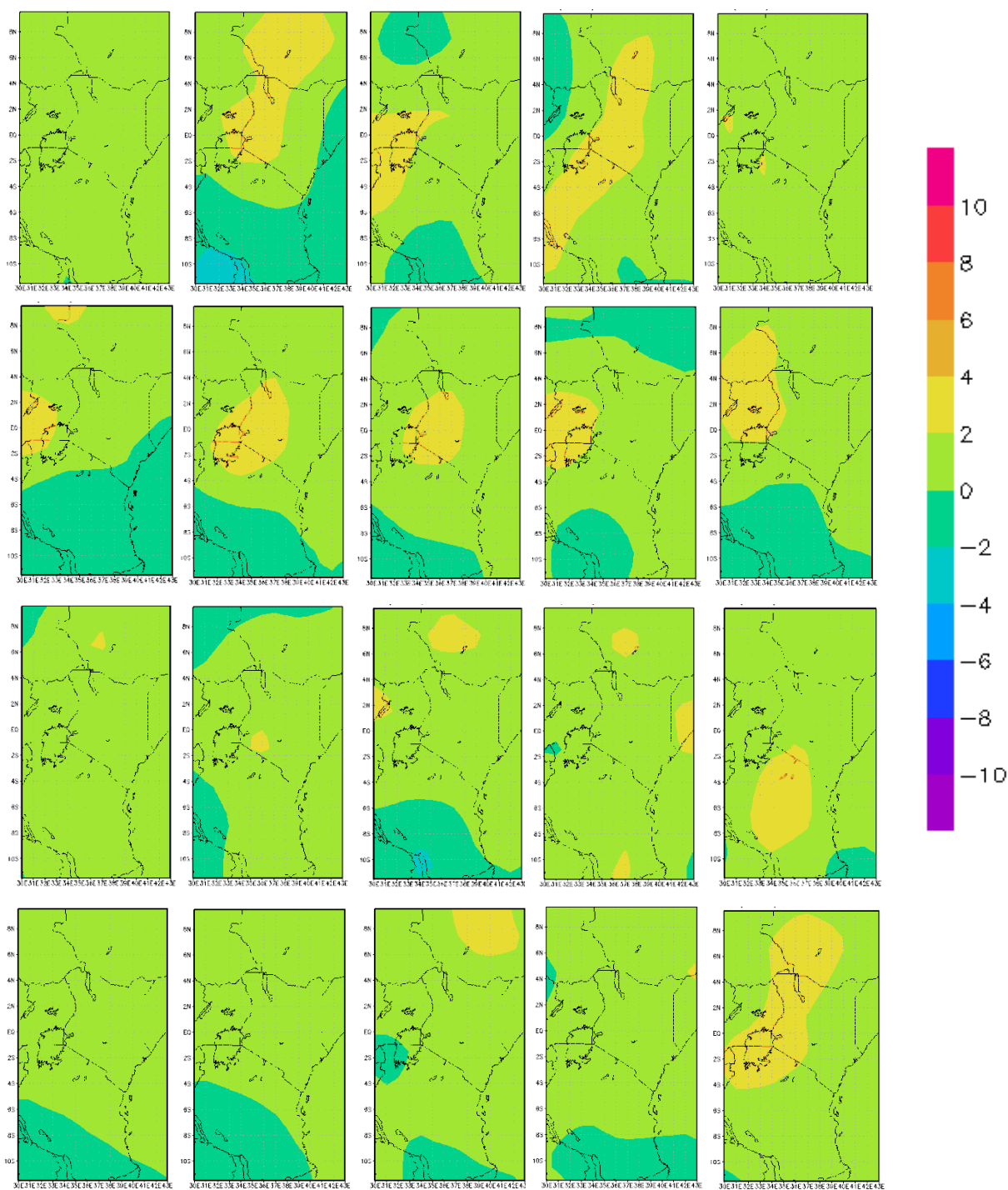


Figure 8.3 (g) Rainfall anomaly (mm/day) for MRI 2090s relative to 1971-2000 for each node of the SOM array during the short rains

To summarise, it is clear that the uniform mean model projection for a wetter East Africa is more complex than it first appeared when potential changes to daily rainfall states across the regions during the short rains are considered. The key result across the models and the ensemble mean (with the exception of CSIRO) is the proportion of the rainfall increase linked to days characterised by suppressed convection in the SOM classification. However, the remainder of the projected increase derives from rainfall patterns which are particular to each of the models that have been investigated. Most agreement is nevertheless seen in the north of the domain, in particular over Uganda and southern Ethiopia where the greatest increases are projected in daily rainfall centres in this region, peaking at over 10mm/day in CCCMA and CNRM and up to 8mm in GFDL and MPI. In contrast, centres over Tanzania and the Indian Ocean show less agreement, and it is in this region that more analysis is required through evaluating projected changes to daily patterns of moisture flux over East Africa under an increasing emissions scenario by the end of the twenty-first century. Additionally, focus needs to be given to the nodes of group B, suppressed convection across the domain, given the errors in the model rainfall field for these days.

8.3 Moisture Flux Changes

Between 20 and 35% of the projected rainfall increase for East Africa in the 2090s results from days classified as suppressed convection and identified as dry days in the contemporary analysis of the rainfall types. Comparison between the coupled model and the reanalysis and satellite rainfall suppressed convection mode for the years 1971 to 2000 (see Chapter 6, Section 6.4) showed that the coupled models produced more rainfall than the reanalysis and satellite datasets for this mode. That the coupled models then go on to produce a 20-35% increase in rain for this mode which is not well simulated in the contemporary climate, is a cause for concern. This result is therefore investigated further

by means of inspection of the water vapour flux associated with the suppressed convection mode. The moisture flux field in the models was in agreement with the suppressed convection classified for these days in the ERA-40 SOM (see Chapters 4 and 6), with anomalous moisture flux divergence across the focus countries of East Africa for each model and the ensemble mean.

This section examines the moisture flux divergence field at the 850hPa level in the 2090s. Coherence between rainfall projections and moisture flux fields for days associated with group B would increase confidence in this projection, with the converse also holding true. It is important to establish if the rainfall field and the moisture flux divergence field show coherence in their projections, both for the rainfall increase assigned to days of suppressed convection and for the different rainfall events responsible for the remainder of the increase.

The larger East African domain, measuring 20°E to 60°E, 15°S to 20°N, is used to examine the moisture flux dynamics outside of the immediate region of interest (Tanzania, Kenya, Uganda and Ethiopia) and to identify any strong controls on the focus countries.

All models except CSIRO produce more rain in the 2090s from the suppressed convection type. The formation of rainfall centres over Tanzania and the Indian Ocean, a region which emerged as showing the most disparity of projections of daily rainfall patterns under a high emissions scenario by the 2090s, will be evaluated through the daily moisture flux field.

The ensemble as the mean of the seven individual AOGCMs is examined first, followed by each individual model in turn.

The moisture flux divergence anomaly values are plotted to a set scale for direct comparison of the two time periods. Figures 8.4 (a) and (b) show the mean moisture flux divergence anomaly for the ensemble rainfall series for days assigned to each node within

the 5 by 4 SOM array for the control period and 2090s respectively. The days allocated for the calculation of the mean anomaly for each node are determined through matching each of the two rainfall daily series, the 20C3M 1971-2000 and SRES A2 2090-2099, to the ERA-40 climatology SOM used as the basis for establishing projected changes to East African rainfall under a climate change scenario. The moisture flux anomalies are then calculated from the mean moisture flux values for each of the respective series (that is the mean moisture flux values for the time period as a whole are subtracted from the moisture flux values associated with each node in the SOM array).

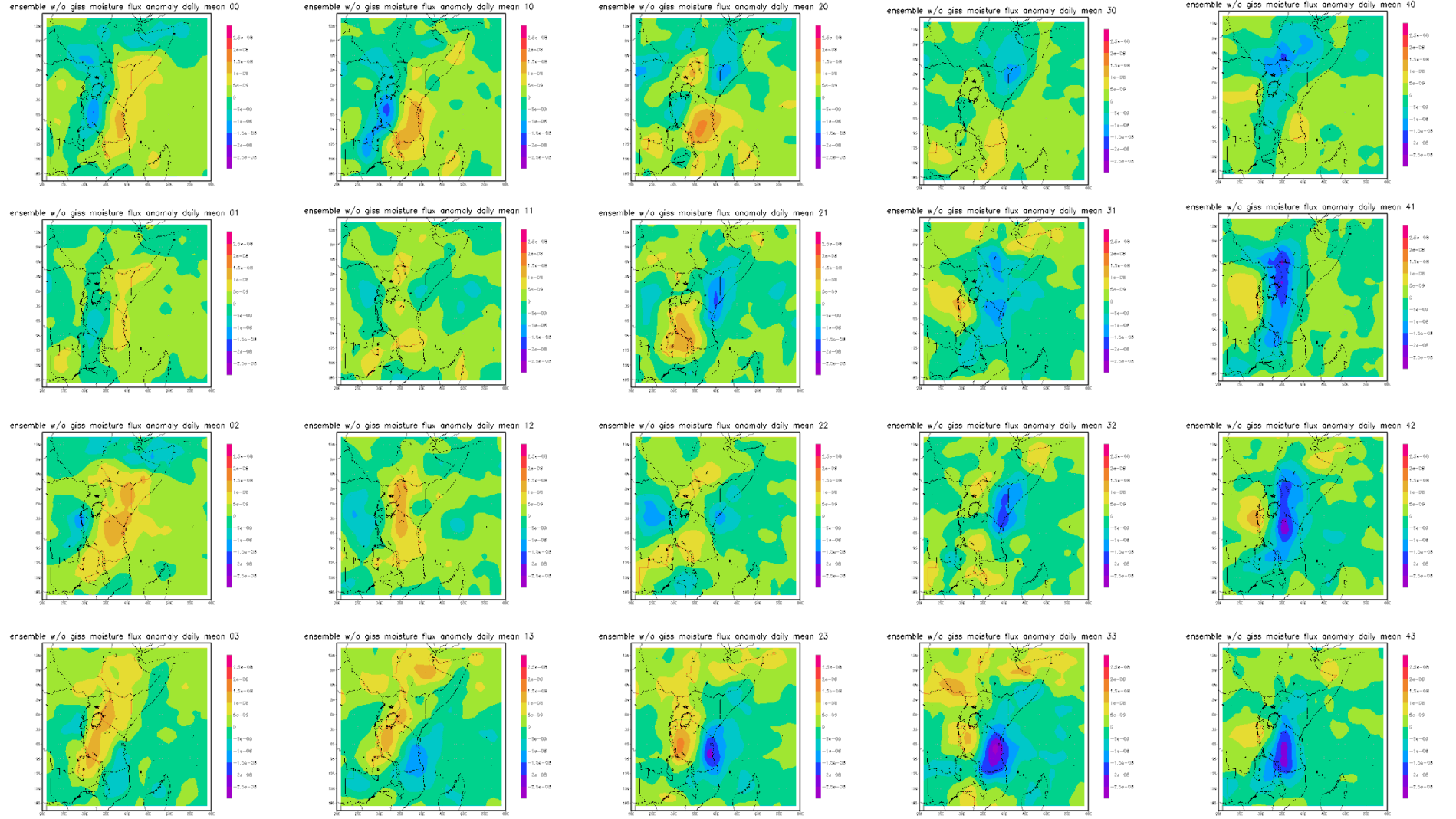


Figure 8.4 (a) 1971-2000 moisture flux divergence anomaly associated with each node of the ERA-40 normalised SOM array, ensemble series.

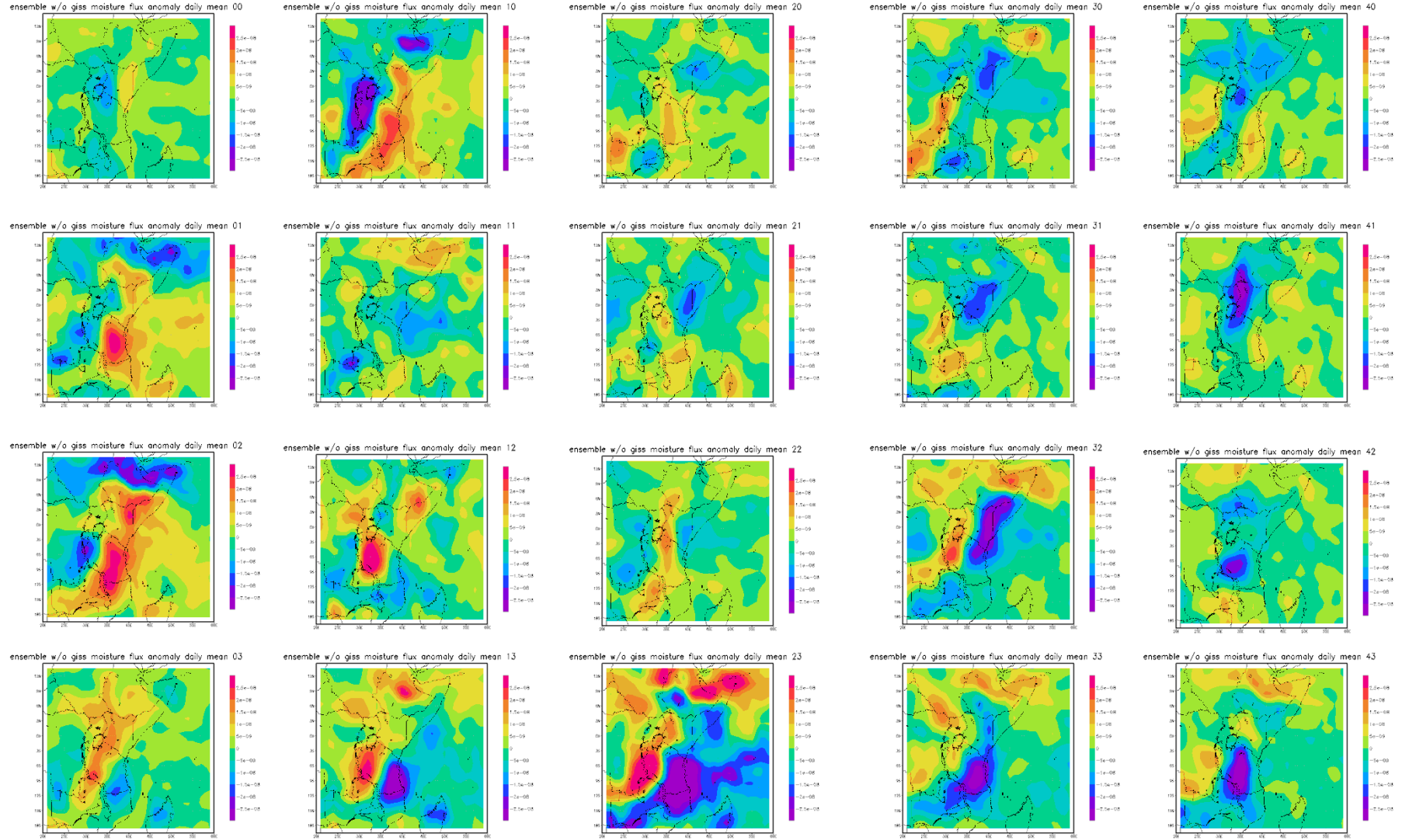
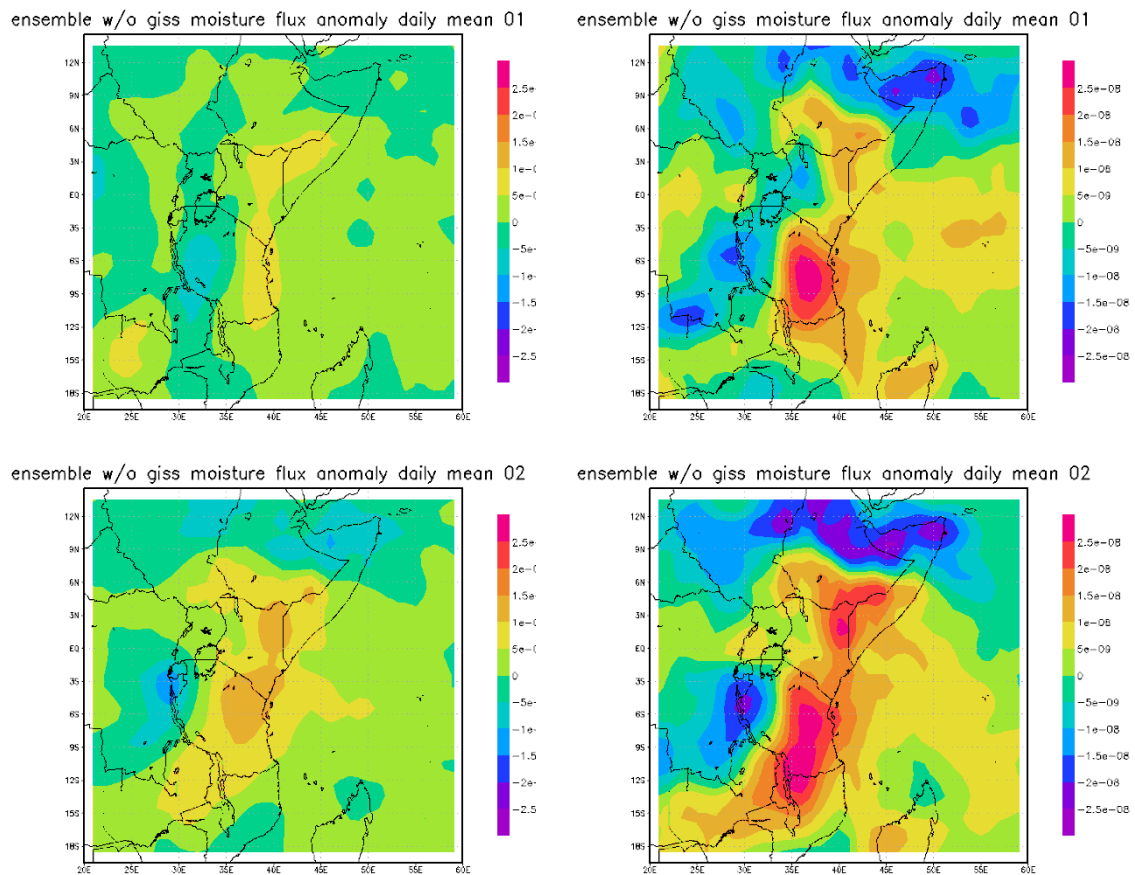


Figure 8.4 (b) 2090-2099 moisture flux divergence anomaly associated with each node of the ERA-40 normalised SOM array, ensemble series.

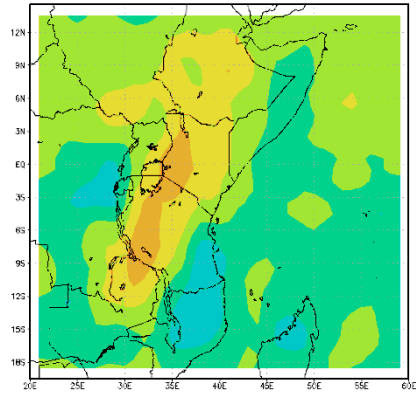
8.3.1 Group B – events of suppressed convection over the East African domain

Group B, comprising the nodes 0 1, 0 2, 0 3, 1 2 and 1 3, was a cluster that contained nodes representing suppressed convection over the East African land domain, in particular over the four focus countries of Ethiopia, Uganda, Kenya and Tanzania. The presence of suppressed convection over the land domain would not normally be associated with large amounts of rainfall. However, these nodes are responsible for between 20% and 35% of the projected short rains increase by the 2090s under the high emissions A2 scenario for the ensemble and all models except one, CSIRO.

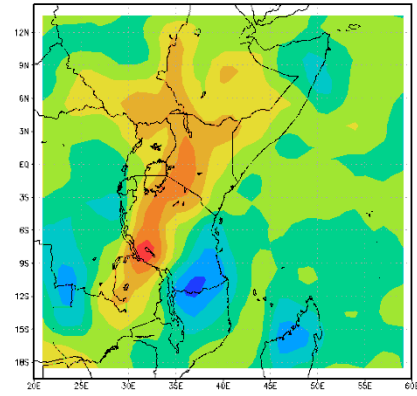
To illustrate the results in further detail, Figure 8.5 (a) and (b) show the mean moisture flux anomalies for the 1971-2000 and 2090s for group B nodes, in a larger size than full array allowed, for the ensemble and CSIRO.



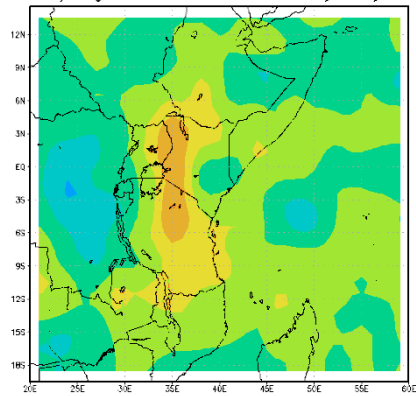
ensemble w/o giss moisture flux anomaly daily mean 03



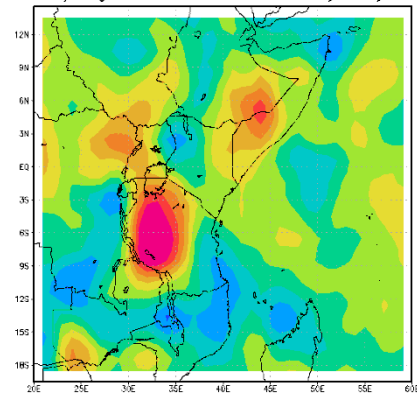
ensemble w/o giss moisture flux anomaly daily mean 03



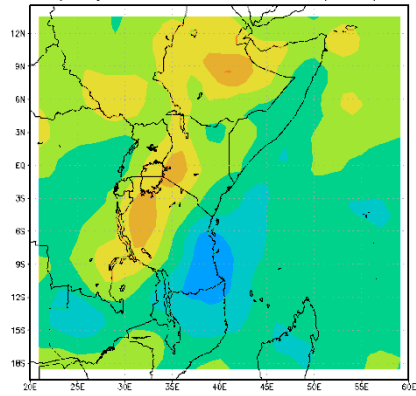
ensemble w/o giss moisture flux anomaly daily mean 12



ensemble w/o giss moisture flux anomaly daily mean 12



ensemble w/o giss moisture flux anomaly daily mean 13



ensemble w/o giss moisture flux anomaly daily mean 13

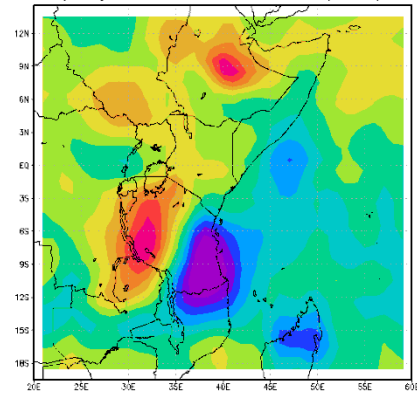
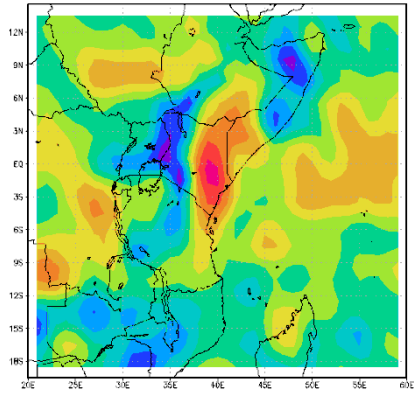
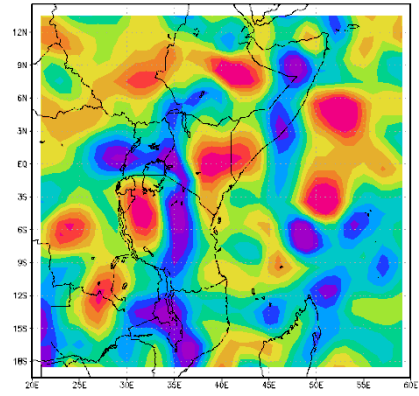


Figure 8.5 (a) Mean moisture flux divergence anomaly associated with nodes of Group B, left column 1971-2000, right column 2090-2099, ensemble series to ERA-40 normalised rainfall SOM, set scale

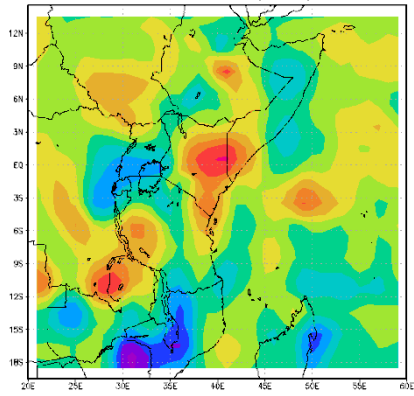
csiro moisture flux anomaly daily mean 01



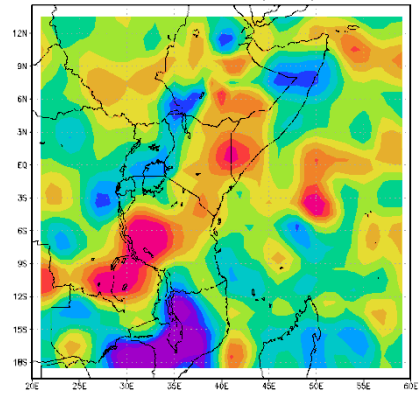
csiro moisture flux anomaly daily mean 01



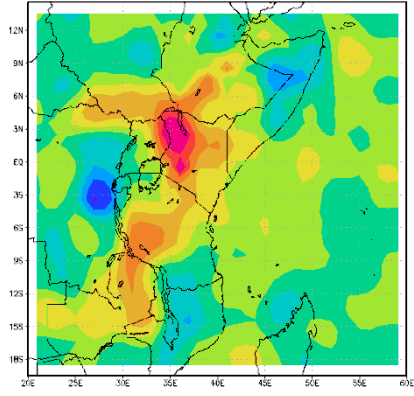
csiro moisture flux anomaly daily mean 02



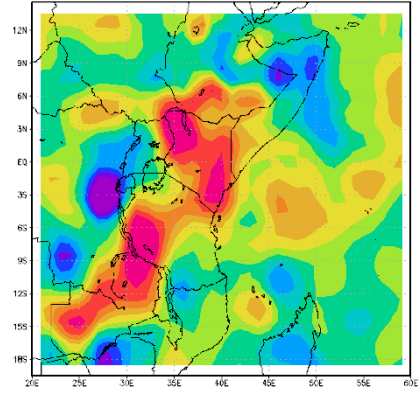
csiro moisture flux anomaly daily mean 02



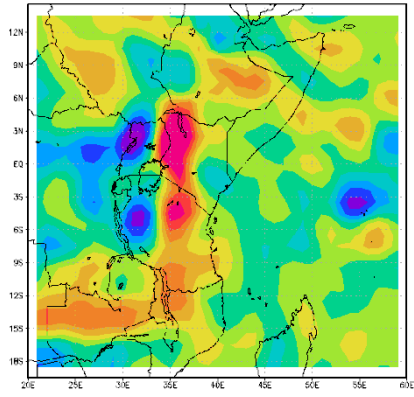
csiro moisture flux anomaly daily mean 03



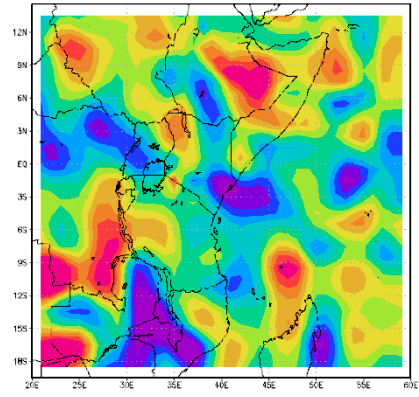
csiro moisture flux anomaly daily mean 03



csiro moisture flux anomaly daily mean 12



csiro moisture flux anomaly daily mean 12



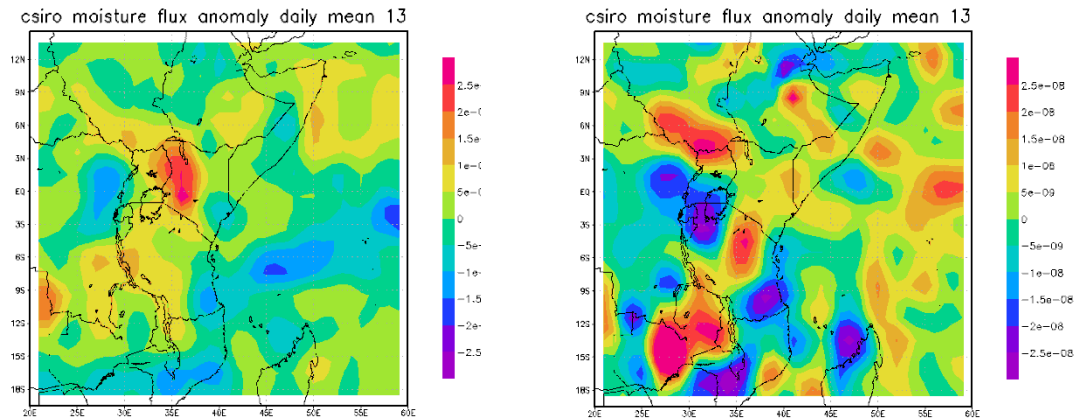


Figure 8.5 (b) Mean moisture flux divergence anomaly associated with nodes of Group B, left column 1971-2000, right column 2090-2099, CSIRO series to ERA-40 normalised rainfall SOM, set scale

Taking the ensemble first, for group B there are strong divergent anomalies in the moisture flux field, over the focus countries in particular, by the 2090s. This would not normally be indicative of the presence of large amounts of rainfall, even with the small regions of anomalous convergence that are present in some of these nodes. Crucially, the regions of anomalous divergence are even stronger in the 2090s than they are for the same nodes at the end of the twentieth century, raising questions as to where the projected increase in short rains rainfall is originating for days assigned to these nodes of suppressed convection. The only AOGCM that does not have at least 20% of the projected increase attributed to group B was CSIRO, where days assigned to group B are responsible for just 6.3% of the projected short rains increase over the twenty-first century. In this model, group B nodes have strong moisture flux divergence anomalies over much of the land domain, interspersed with some small pockets of convergence. For this model, the dry nodes present no evidence for moisture flux anomalies that would be demonstrative of a large increase in rainfall which tallies with the small amount of the overall projected

rainfall increase for the short rains for CSIRO, in contrast to the conflicting results that have emerged from the ensemble.

The ensemble and CSIRO show two conflicting pictures as to the projected changes to moisture flux divergence during the short rains over East Africa, with moisture flux in CSIRO reinforcing the rainfall projection for the group B nodes and the ensemble showing contrasting results for these nodes between the moisture flux divergence and the rainfall projections for the same group B nodes of suppressed convection. The remainder of the six individual AOGCMs (not shown) all show a significant proportion of the short rains rainfall attributed to the suppressed convection in the nodes of group B. CCCMA projects strong moisture flux divergence over the focus countries in the 2090s, with divergent anomalies at least as strong as at the end of the twentieth century. In this model, however, there are regions of anomalous convergence in group B nodes in the 2090s sometimes present over the Indian Ocean and Congo regions. The Indian Ocean does not show a projected rainfall increase but convergence over the Congo could hold some responsibility for the projected rainfall increase for this model. In CNRM, there is a large amount of moisture flux divergence over the focus countries in the 2090s, with large regions of strong divergence presence across the domain. The only exception to this pattern in group B is the node 0 1 which has some convergence presence through central Tanzania.

GFDL shows much stronger moisture flux divergence anomalies for group B in the 2090s under the A2 scenario than during the climatology period. There are only a few small areas of convergence and these anomalies are equivalent to those present during the climatology period, making it very unlikely that these regions would have been responsible for the large rainfall increase associated with group B. In MIUB, the left most column in particular of the group B nodes, has stronger moisture flux divergence and less regions of convergence in the 2090s period than during the control period. The only exception to this

pattern is node 1 2 which has several of regions of convergence. This node is responsible for only 1.46% of the projected short rains increase, with the majority of the increase associated with group B coming from other nodes. In general, the MPI model has similar moisture flux convergence and divergence patterns in both time periods, with stronger anomalies for both at the end of the twenty-first century. There is a possibility that stronger regions of convergence could have led to intense convection in this model, contributing to an overall increase in the short rains projection, even though divergence is the dominant pattern over the focus countries. The final model, MRI, has very similar patterns of moisture flux divergence anomalies for group B nodes. The strength of the regions of divergence is equivalent, or increased, for each node in the group in the 2090s A2 projections, with moisture flux convergence projected to decrease in strength for these nodes. One exception is node 0 1 where there is strong convergence in the domain over the Congo and the north of the East Africa region. However, this node is responsible for only 2.6% of the projected short rains increase in this model, with other group B nodes connected with a greater percentage of the short rains increase even where moisture flux divergence is prevalent over the domain.

Overall, there is very little evidence in the 850hPa moisture flux field to support the increased rainfall projection in the nodes of group B in the ensemble and six of the seven individual AOGCMs. This raises the question as to how much confidence can be placed in the rainfall projections of the AOGCMs under a high emissions scenario by the end of the twenty-first century. In particular, the moisture flux field is a strong and reliable indicator of the presence of model rainfall during the control period and the deviation from this relationship for dry days by the 2090s introduces considerable uncertainty to the rainfall projections as a whole. Section 8.3.2 examines the consistency of the relationship between moisture flux and rainfall projections for other areas of the SOM array.

8.3.2 Projected moisture flux divergence changes to other nodes within the SOM array

Section 8.3.1 found that evidence presented in the daily moisture flux divergence field for each individual AOGCM and the ensemble did not support the increased rainfall projection for nodes associated with group B, characterised largely by divergent moisture flux across the domain. It is important to contextualise this result through investigation of the moisture flux fields for other nodes within the SOM array; is it just the dry nodes of group B for which the rainfall and moisture flux projections are even more divergent than the current climate, or does this apply for rainfall centres present elsewhere in the SOM array over East Africa? This is done by inspection of whether the projections in moisture flux divergence for the other rainfall groupings are consistent with the model projections for these nodes. Particular attention is given to moisture flux divergence projections associated with the rainfall centres of the nodes of groups D and F, as this is where there are the greatest disparities in daily rainfall projections across the subset of models. The ensemble and each AOGCM are considered in turn.

8.3.2.1 Ensemble

The mean moisture flux divergence state for the ensemble shows very similar spatial patterns over East Africa for the short rains for both the control period of 1971-2000 and the 2090s under a high emissions scenario as shown in Figure 8.6. This pattern is for moisture flux convergence over the western side of the focus countries and divergence over coastal East Africa, with the longitude of 35°E marking the dividing point, with the 2090s showing increased strength and spatial extent of the regions of both convergence and divergence.

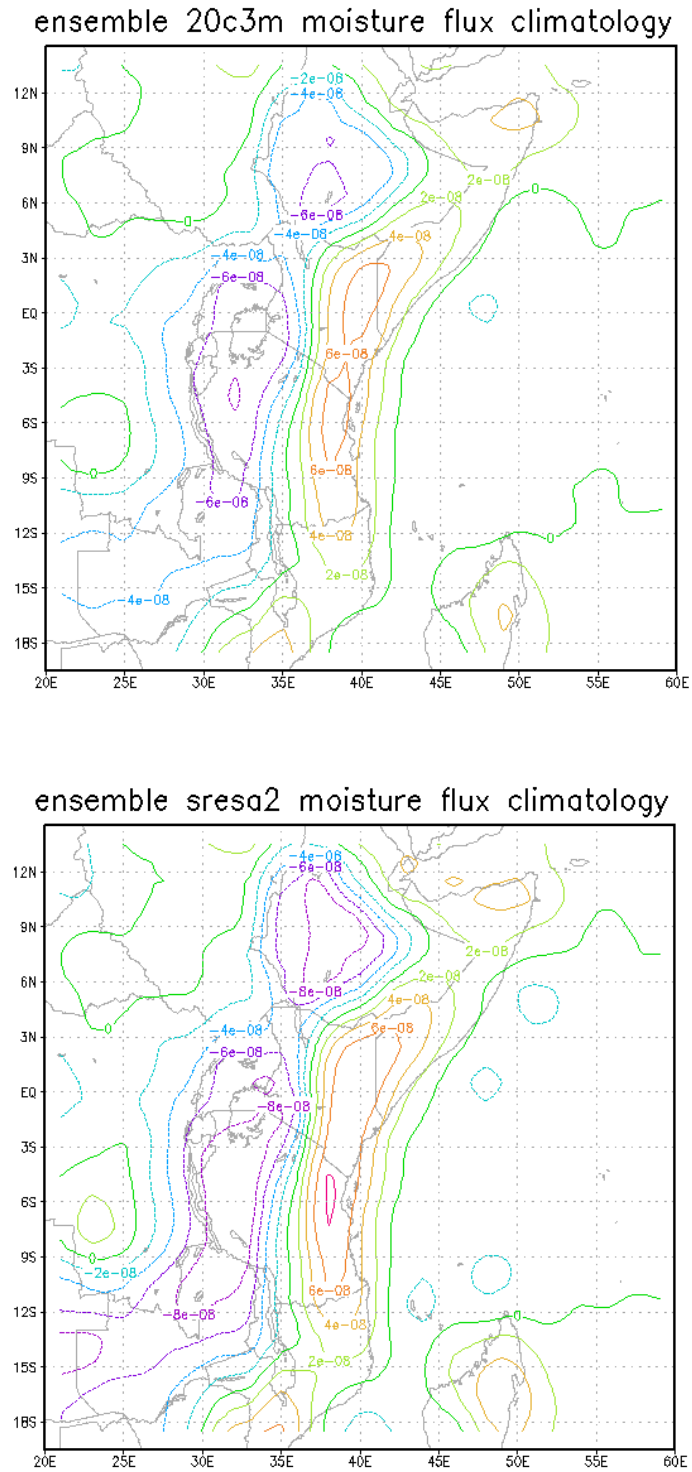


Figure 8.6 Moisture flux divergence climatology for the periods 1971-2000 (top) and 2090s (bottom) for the short rains season (OND)

Between the control and the 2090s, there are both similarities and differences in the mean anomaly moisture flux divergence patterns for individual nodes within the daily rainfall SOM array. Nodes with rainfall centres over the Kenyan-Ethiopian border are very similar in terms of moisture flux divergence anomalies for both the control and 2090s time slices. Specifically, this includes nodes 1 1 and 2 0 of group C, representing a rainfall state over coastal northern Kenya and southern Ethiopia. This is the expected result with this grouping responsible for only 8.3% of the projected increase in rainfall under increased emissions during the twenty-first century. However, other nodes show considerable differences in projected moisture flux anomalies for the 2090s in comparison to the climatology. Some rainfall centres see a projected increase in moisture flux convergence with node 1 3 seeing strong convergence over an Indian Ocean rainfall centre in opposition with strong divergence of moisture seen over the Congo. This pattern is replicated in node 2 3. Along with nodes 2 2 and 3 2 the Indian Ocean rainfall centres of group D make up 17.8% of the projected rainfall increase. The greatest proportion of rainfall increase in the ensemble is associated with Tanzanian rainfall centres (group F), accounting for 38.0%. In terms of the accompanying moisture flux divergence patterns, they show a slight increase in the strength of the convergent moisture flux anomalies centred over Tanzania in addition to an expanding of the region of convergence. This is especially evident in nodes 3 3 and 4 3.

To summarise, in the ensemble series, the greatest differences in mean moisture flux anomalies over each node from the control period to the 2090s are seen in the Indian Ocean rainfall centres, with a notable strengthening present in the Tanzanian rainfall centres of group F also present. This significant variation in the strength of moisture flux anomalies over the domain associated with rainfall anomalies under a climate change scenario follows the patterns that are expected for the daily rainfall projections for these

rainfall centres that have emerged. Therefore, for the ensemble the outlier in the otherwise good agreement in the rainfall and moisture flux fields is for nodes of suppressed convection (group B). This result will be examined now in the individual AOGCMs.

8.3.2.2 CCCMA

The mean moisture flux divergence plots for the CCCMA (not shown) show very similar spatial patterns to the ensemble, again with larger regions of both convergence and divergence in the 2090s in comparison to the control period, with convergence the mean state over the west of the focus countries during the short rains and divergence to the east along coastal East Africa. 35°E marks the division between convergence and divergence.

In terms of projected changes to the mean moisture flux divergence anomaly conditions associated with each daily rainfall pattern under a high emissions scenario, CCCMA projects stronger moisture flux divergence anomalies over the Indian Ocean, which are not confined to the nodes of rainfall centres over the Indian Ocean located in group D. These can be seen, for example, in nodes 1 3, 1 2 and 1 1 and are also present in nodes located in the left most column of the SOM array, representing suppressed convection over the land region of the domain (all group B). In these nodes, moisture flux anomalies over the Congo are projected to be of much the same magnitude in the 2090s as during the control period at the end of the twentieth century. However, moisture flux divergence anomalies over the Indian Ocean are much more active by the end of the twenty-first century, suggesting a projection for increased moisture flux activity here and an increasing influence of Indian Ocean dynamics over the East African region in comparison to dynamics in the Congo under a climate change scenario. These regions of convergence have contributed to the high rainfall projections coming from these nodes in CCCMA (31.0%). However, given that the largest projected rainfall increase does not originate over

the Indian Ocean but over the land region of East Africa and that the land region shows strong divergent anomalies for these nodes it is more likely that the rainfall field is a poor representation of states of suppressed convection.

Nodes 2 0 and 2 1 (group C), representing a rainfall centre over Kenya and Ethiopia sees a significant expansion of the region of strong convergence co-located with the rainfall centre although the spatial positioning remains the same. This is a pattern repeated in all the nodes within the SOM array representing rainfall centres over the four focus countries, that is, the moisture flux indicates an increase in strength of daily rainfall events over East Africa, which is reinforced through the daily rainfall field for this model. The strength of moisture flux anomalies for both regions of convergence and divergence are projected to increase by the 2090s relative to the control period, with regions of convergence expanding where rainfall anomalies are projected to be positive.

In contrast, the emerging picture over the Indian Ocean is more complex. Some Indian Ocean rainfall centres in this model are projected to see less change in the 2090s including nodes 2 3 and 2 2 of group D. The emerging moisture flux pattern is for increased moisture flux activity in this region in comparison to the climatology but the pattern is not for universal increased convergence but for far more mixed anomalies. This is representative of the projected proportion of annual short rains rainfall increase allocated to these nodes which stood at just 9.4%.

8.3.2.3 CNRM

The mean moisture flux plots for this AOGCM follows the same patterns as for the ensemble and CCCMA for both time periods and the spatial location of divergence and convergence over the short rains as a whole.

Nodes 0 0 and 1 0 (group A), which show rainfall to the western edges of the focus countries originating in the Congo see a strong increase in the strength of the convergence in this region accompanied by a mixture of anomalous divergence and convergence over the focus countries and Indian Ocean where previously the situation was very neutral, although this group does not contribute greatly to the projected rainfall increase. This suggests an increase in moisture flux dynamics in the region by the end of the twenty-first century, something that is reinforced through the projected changes to moisture flux patterns in other nodes.

For example, the Kenyan-Ethiopian rainfall centres (group C) see spatially more widespread anomalous convergence in the 2090s time period, which explains the fifth of the projected rainfall increase assigned to this group. Equally, centres of rainfall over the Indian Ocean are accompanied by stronger moisture flux convergence anomalies. These nodes also have more pronounced divergence elsewhere in the domain and equally in the Indian Ocean remote from the rainfall centre when this is the case, as seen in node 2 3 of group D. It may have been this divergence elsewhere that prevents this group being responsible for any more than 2.9% of the annual short rains projected increase. For nodes 4 0, 3 1 and 4 1 of group E, which show daily rainfall events over the northern part of the domain and the greatest projected rainfall increases over land in this model, the convergence in this region becomes more widespread with strong divergence elsewhere in the domain.

Again, aside from the nodes of suppressed convection examined in section 8.2, the moisture flux anomaly patterns and daily rainfall projections show a coherent picture as to the origins of the projected short rains rainfall increase.

8.3.2.4 CSIRO

The mean moisture flux patterns during the short rains are very similar to the other plots already described, with one exception being the region of divergence along coastal East Africa which is less pronounced in this AOGCM.

Projected changes to the moisture flux divergence anomalies associated with this model show a variety of patterns depending on the node. Some rainfall states have very similar moisture flux anomalies for both the control period and the 2090s. On the other hand, other nodes show projections for considerable changes to moisture flux divergence anomalies including for increased moisture flux anomalies over the Indian Ocean where convective activity in the control period is limited. The Kenyan-Ethiopian rainfall centres seen in nodes 2 0 and 2 1 have very similar moisture flux anomaly patterns during the end of both the twentieth and twenty-first centuries over the land regions of the domain accompanied by much larger Indian Ocean anomalies, a pattern which is visible in the majority of models examined thus far. Group C nodes are responsible for only a small proportion of the rainfall increase in CSIRO (7.7%). Increased Indian Ocean moisture flux activity is also projected for nodes 2 2 and 2 3, where the main daily rainfall centre is located over the Indian Ocean adjacent to Tanzania. The projection here is not just for increased convergence co-located with the rainfall centre although some strong convergence is projected, but instead there are many small regions of divergence as well, which explains the small contribution (6.0%) of group D nodes to the rainfall increase.

The land based rainfall states seen in the two right most columns of the domain see very little change to moisture flux anomalies in the projections for the 2090s under the SRES A2 scenario over the land regions aside from a small increase in convergence over the location of the rainfall. This reinforces earlier indications that the key influence of this

region which emerged in many of the AOGCMs examined so far under a high emissions projection is that of the Indian Ocean as opposed to the Congo which has also exerted considerable influence over East Africa during the control period.

The largest share of a single group of nodes of the rainfall increase (48.7%) for CSIRO comes from node 0 0 and 1 0 (group A), which represent a pattern of rainfall running the length of the western part of the domain. The anomaly patterns of moisture flux convergence are very similar for both time periods, though the convergence in the 2090s is slightly stronger and more widespread which may account for the rainfall projection assigned to these days.

8.3.2.5 GFDL

As with the models analysed above, the same mean moisture flux patterns are present, with the moisture flux in the 2090s being stronger than during the control. However, in terms of moisture flux anomaly patterns of the individual nodes, there are some differences which emerge in comparison to the AOGCMs examined so far. Most notably, there is not the same increase in anomalous moisture flux convergence and divergence over the Indian Ocean that is seen in the other models. In fact, the majority of nodes in the GFDL array see no projected change to the spatial patterns of the moisture flux anomalies with the only change being increased strength of convergence and divergence associated with some nodes.

There is just a single node that is an exception to this; node 2 0 representing a Kenyan-Ethiopian rainfall centre. In the 2090s this region is projected to see anomalous moisture flux divergence over the previous rainfall centre. However, in conjunction with this the second node which depicts this pattern (2 1) continues to show anomalous convergence in the region of the rainfall centre and therefore the change to node 2 0 may be related to a

redistribution of rainfall days between these two nodes; indeed group C contributes approximately 15% of the projected short rains increase so this rainfall pattern is generally expected to become stronger in future.

Therefore, GFDL shows the same patterns as previous models, which allows the conclusion that it is becoming more apparent that it is in the simulation of dry days over East Africa during the short rains that caution has to be taken in considering the reliability of projections. The remaining three models are now evaluated for the same patterns before a final conclusion is made as to the reliability of rainfall projections over East Africa during the short rains.

8.3.2.6 MIUB

The mean moisture flux plots for this model show much weaker regions of convergence and divergence than the previous models and the ensemble, with the patterns also unchanged from the control period of 1971-200 to the 2090s under the high emissions scenario.

Node 1 1, with a light rainfall centre over Kenya sees convergence in this region projected to become more focused in the 2090s. This is accompanied by the emergence of divergence to the east in the Indian Ocean and an increase in strength of the convergence seen over the Congo. Group C nodes, of which node 1 1 was a part, are projected to account for 22.2% of the short rains increase for MIUB, reinforced by the increased convergence over the rainfall centre seen in this node. The majority of the other nodes representing rainfall centres over the land region of the small East African domain also see the accompanying regions of moisture flux convergence projected to both expand and increase in strength by the end of the twenty-first century under a climate change scenario;

this is in agreement with the projections for moisture flux anomalies co-located with rainfall centres in many of the AOGCMS for this region.

On the other hand, MIUB differs from the many of the previously examined models and the ensemble and shows agreement with GFDL when it comes to projected changes to moisture flux divergence anomalies over the Indian Ocean across all nodes of the SOM array, not just those associated with Indian Ocean rainfall centres in group D. Whereas the predominant emerging pattern has been for increasing moisture flux divergence anomalies over the Indian Ocean, MIUB instead sees little projected change to the Indian Ocean anomalies. Instead, this AOGCM sees the strength of moisture flux anomalies over the Congo projected to increase consistently across the nodes of the SOM array. This is the first time this projection has emerged in the subset of models examined here, and provides an alternative scenario for the region which could have greatest dynamical influence on East African rainfall by the end of the twenty-first century. In terms of the Indian Ocean rainfall centres of Group D, the region of anomalous convergence expands and strengthens in the 2090s in comparison to the control period and is indicative of the increased rainfall associated with these nodes of 10.2% of the short rains projected increase.

8.3.2.7 MPI

The mean moisture flux plots for the short rains for MPI are very similar in both spatial patterns and the magnitude of divergence for both the control and 2090s. The pattern is comparable to that seen in the ensemble with divergence down the East African coast and convergence to the west of 35°E.

In common with the majority of models, there is a significant projected increase in the variety of moisture flux patterns seen over the ocean whilst spatial patterns and magnitude of anomalies over the land remain much the same. This is especially evident in the nodes

of group E, 4 0, 4 1 and 3 1. These three nodes account for 30.4% of the projected short rains increase and this rainfall projection is accompanied by strong and widespread anomalous moisture flux convergence in the 2090s in comparison to the anomalies of the control period. The projection is for the dominance of anomalous moisture flux convergence over the whole of the region of the focus countries, often extending in to the Indian Ocean.

8.3.2.8 MRI

The final of the seven selected AOGCMS also sees very similar spatial patterns and magnitude of moisture flux divergence for both time periods. Therefore, all of the models and the ensemble have been seen to produce very similar patterns of moisture flux over the East African domain, which increases confidence in the quality of this variable in the AOGCMs over this region.

In a similar manner to GFDL, and to some extent MIUB, the MRI AOGCM sees moisture flux anomalies over the Indian Ocean remain neutral during the 2090s, although a dominance of the moisture flux over the Congo does not emerge in opposition as it did in the MIUB. MRI rainfall projections see more than 80% of the short rains increase accounted for by three clustering of daily types, states of suppressed convection, Indian Ocean rainfall centres and rainfall events centred over the Kenyan-Ethiopian coastal border. The lack of any notable rainfall increase associated with rainfall centre over Tanzania and Uganda-Ethiopia is reinforced through the projections to moisture flux associated with days assigned to these nodes. For these rainfall centres over land, located in the right hand side of the array, there is not any drastic change to the moisture flux patterns projected although there is expansion and strengthening of the moisture flux convergence anomaly co-located with the rainfall centre, consistent with the small

projected rainfall increases associated with these days. In contrast mean moisture flux patterns associated to days assigned to groups C and D, responsible for 48% of the rainfall increase, show a marked increase in the strength and extent of anomalous moisture flux convergence, which is consistent with strong rainfall events and a significant contribution to the overall short rains increase.

8.4 Chapter summary

This chapter aimed to build on the results of Chapter 7 in order to answer the overall question of this thesis as to the reliability of model projections over East Africa during the short rains under a high emissions scenario.

As emerged in Chapter 7, between 20 and 35% of the rainfall increase in all models (except CSIRO) and the ensemble was quantified as originating from days assigned to nodes of suppressed convection. Moisture flux divergence was found to be a key indicator of the presence of rainfall when investigated in the ERA-40 and ERA-Interim datasets in conjunction with the ERA-40 and TRMM rainfall series respectively. The supply of water vapour, through the moisture flux variable, is a vital stage in the creation of rainfall (Washington et al, 2013). Moisture flux is therefore a suitable variable for assessing confidence in the rainfall projections under a climate change scenario in this chapter. In particular, the moisture flux divergence field indicated in Chapter 6 that it is days of suppressed convection that the models consistently struggled to replicate in comparison to the reanalysis data. This made the coherence in the projected increase for these nodes striking, inviting further investigation.

In contrast to the rainfall increases associated with the nodes of group B under a high emissions scenario, the moisture flux divergence field showed contrasting results. For all models and the ensemble, the moisture flux field showed increasing divergence over East

Africa for days associated with the group B nodes of suppressed convection, in direct opposition to the rainfall fields which show an increase. This unusual result was accompanied by projections to the moisture flux field associated with other nodes that were in line with the presence of the increasing rainfall over these nodes. Indeed, there was plentiful evidence that moisture flux convergence anomalies in the 2090s increased in conjunction with the node types responsible for large parts of the projected short rains rainfall increase for each model and the ensemble, that is, that the nodes with the largest increases in projected rainfall concurrently showed the greatest strengthening of moisture flux convergence relative to other nodes of the SOM array.

Given that the East African projected rainfall increase is cited as a coherent model response (Solomon et al, eds, 2007) and given that at least part of the rainfall increase arises from a mode which the models do not simulate well in comparison with reanalysis data, the strength of evidence for the wetting response in the case of East Africa requires much further analysis, including a closer analysis of the way convection is simulated in the region and the way in which the models are able to simulate more rainfall despite the increasing moisture flux divergence across the region.

The IPCC AR5 reinforces the results of the IPCC AR4 in stating that it is '*likely*' that the short rains will see increased rainfall by the end of the twenty-first century (Christensen, J. H. and Kanikichana, K. K. eds, 2013). These results can be further seen in work by Otieno and Anyah (2013). Their work on the CMIP5 models show that projected increases over East Africa during the short rains are much greater over Kenya and Uganda than over the southern GHA region of Tanzania (Otieno and Anyah, 2013). This reinforces the spatial patterns of rainfall found in this work in the CMIP3 model subset. Further agreement for the increasing projection over East Africa is found in the work of Cook and Vizy (2013) who present results in an NOAA/NCAR regional climate model showing a projected

increase to the short rains due to a lengthening of the short rains season. Neither of these approaches examined the patterns within the daily rainfall model field. Whilst results when examining the daily field do show a projected short rains increase in the CMIP3 model subset employed here, the only common allocation of the increase across the models in terms of daily rainfall patterns is to days of suppressed convection, which were the most poorly simulated days in the models. The combination of the variety of daily rainfall types responsible for the remainder of the projected increase and the disparity between moisture flux convergence and rainfall projections for days of suppressed convection suggests that further research is necessary to fully understand the coherence of the increasing rainfall projection over East Africa during the short rains.

This chapter has addressed the overall aim of the thesis through the examination of the moisture flux field in relation to the rainfall field at the daily timescale. The key result is that rainfall projections during the short rains seem to be overstated, as characterised through the 20-35% of the projected increase assigned to daily rainfall regimes for which the moisture flux field does not align with rainfall values, a weakness that occurred consistently through the control period and the 2090s.

The results of this thesis will be synthesised in Chapter 9; conclusions and directions for further research.

Chapter 9: Conclusions and directions for future research

Chapter 9 will consider the results and findings of this thesis. The justification and aims for this study on daily rainfall regimes within the October to December East African short rains will be restated. A summary of results from earlier chapters will be synthesised to address the conclusions relating to both the individual research questions and the overall hypothesis of this thesis. Finally, recommendations for further directions of research will be suggested.

9.1 Justification for study and thesis aims

Study of anthropogenic climate change, of which the establishment of reliable climate projections is a major component, signals a research area of vital importance for coming decades. This is particularly the case for Africa, where the effects of rainfall variability and change are expected to be keenly felt (Zhao et al, 2005). Specifically, East Africa, the focus of this study, has been identified as a climate ‘hotspot’ that is especially vulnerable to climate change (Thornton et al, 2009). Climate change scenarios are typically developed from coupled climate models (atmosphere-ocean global climate models, AOGCMs). The all-important variable of rainfall is parameterised in climate models and this makes the realistic simulation of the field difficult. Additionally, there is no simple, widely accepted, method of how to evaluate this field and weight model projections accordingly (Giannini et al, 2008). Moreover, Rowell (2012) has suggested that model uncertainty increases over time throughout the twenty-first century.

Clearly there is a need for focused additional work on model rainfall and how rainfall changes in models forced with increasing greenhouse gases. Rowell (2012) suggests that this could lead to targeted model improvement, especially in the tropics of which East Africa is a part. The short rains of East Africa show remarkable coherence in projected mean changes, when coherence is defined by the sign of change projected across the vast majority of models being the same. For East Africa this projection is for increased rainfall during the short rains at the end of the twenty-first century (Solomon et al, eds, 2007).

The overall aim of this thesis was to evaluate the fidelity of model rainfall projections for the East African short rains from the perspective of changes in daily rainfall regimes. Is the increasing projection spread evenly over individual rainfall types or do a small number of systems dominate the projected changes? Do the climate models achieve the coherence in rainfall increase through an adjustment to the same individual daily rainfall regimes or do they achieve that increase by changes in daily rainfall regimes which are different from one model to the next? What do these results mean for confidence in climate change projections?

To address these questions, the strategy chosen was to establish whether AOGCMs can accurately simulate contemporary rainfall over East Africa during the short rains; which was approached here in an original manner in comparison to previous research through the focus on the daily temporal scale. The majority of studies over East Africa that address model evaluation do so either through the mean monthly rainfall field or the simulation of known teleconnections to East Africa in AOGCMS (for example Shongwe et al, 2011, Anyah and Semazzi, 2007, Camberlin et al, 2004, Black et al, 2003, Mutai and Ward, 2000).

In this thesis, work was undertaken on the daily rainfall field during the short rains. The monthly rainfall field is an accumulation of daily rainfall and therefore the monthly totals are dependent on the fidelity of the systems that deliver rainfall on daily timescales. However, research on the structure of daily rainfall regimes in AOGCMs during this season has been neglected, compromising evaluation of model rainfall over East Africa. Work on daily timescales in Africa as a whole has so far also been generally limited, with studies focused on southern Africa (Rocha et al, 2008 and Williams et al, 2010), the Sahel (Sylla et al, 2010) and West Africa (Sylla et al, 2012).). However, it has been suggested that in Africa the diversity of topography and land-sea contrasts necessitates model evaluation of the fine temporal scales of rainfall due to the impact of daily rainfall statistics on natural systems and human impacts (Sylla et al, 2013). The setting is complex and deserves a closer examination than can be achieved by the analysis of monthly totals alone. In this work, evaluation has focused on whether the daily rainfall fields in AOGCMs are a feasible representation of observed conditions in order to evaluate the overall question of the reliability of model rainfall projections over East Africa during the short rains. This has been addressed through four distinct research questions.

1. What is the contemporary structure of daily rainfall regimes over East Africa during the October to December short rains season?
 - Is the self-organising map an appropriate tool for this investigation?
2. To what extent can coupled climate models replicate the contemporary structure of the East African daily rainfall regimes of the short rains season?
 - To what extent is the mean rainfall climatology replicated in AOGCMs?
3. How do the East African short rains change under the high emissions SRES A2 scenario at the end of the twenty-first century in coupled climate models?

4. What patterns of daily rainfall are responsible for the change in future modelled rainfall?

9.2 Conclusions for each research question

9.2.1 Question 1 and methodology sub question

1. What is the contemporary structure of daily rainfall regimes over East Africa during the October to December short rains season?
 - Is the self-organising map an appropriate tool for this investigation?

In order to address the contemporary structure of daily East African rainfall regimes during the short rains it was decided to first establish the mean climatology over the region through the use of gridded reanalysis, station and statistical datasets introduced in Chapter 2, Section 2.1 to 2.6. The main aim of this work was to contextualise the overall climate envelope from which the daily rainfall events of the short rains would be evaluated in greater detail. This work was undertaken during Chapter 3; a preliminary investigation of the background climate of East Africa during a control period of 1971-2000 focusing on the annual climatology of the region and encompassing the climate means, 95th percentile extreme rainfall and temperature, interannual variability and trends over the second half of the twenty-first century.

Hence, it was possible to address the first research question, through the examination of the daily short rains fields in two different datasets representing conditions over the domain; the ERA-40 reanalysis dataset and TRMM satellite derived rainfall fields. Whilst station data for Tanzania, Kenya, Uganda and Ethiopia was obtained with the intention of using it to characterise daily rainfall regimes in the short rains, analysis undertaken in

Chapters 2 and 3 identified that the spatial and temporal coverage was inadequate to interpolate to a gridded field. In particular, data coverage was poorest for the years after 1973, corresponding to the 1971-2000 control period for this study. A peak coverage of approximately 300 stations recording between 1960 and 1970 dropped to less than 200 in the mid 1970s with a continual decline to less than 50 recording stations by the final years of the twentieth century, inadequate coverage for this geographically diverse domain.

Chapter 4 addressed the first research question in the ERA-40 dataset, with Chapter 5 employing the TRMM datasets for further analysis. Throughout Chapters 4 and 5 a detailed picture of the characteristics of the short rains of East Africa in the daily field was obtained that could be used to evaluate the coupled models when investigating the later research questions pertaining to the simulation of daily short rains characteristics in AOGCMs and the reliability of rainfall projections originating from these models.

To effectively evaluate daily rainfall events, a method which could classify these daily regimes objectively and in a reproducible way was required. This posed the sub question of this aim; is the self-organising map an appropriate tool for classification of typical daily rainfall patterns? Most previous applications of the SOM were confined to mid-latitude regions where the synoptic scale dominates (see Chapter 2, Section 2.9). With the control of convection over tropical rainfall events it was necessary to establish whether the SOM could classify daily rainfall events in East Africa in an objective and reproducible way.

There were three stages undertaken to address the sub aim of the first research question. First, the self-organising maps method was tested for its ability to cluster atmospheric variables into distinct states from which individual dates in a second, independent, dataset could be matched. A 7 by 5 zonal winds SOM was trained over the East Africa domain on the ERA-40 dataset and a list of dates matched to each node obtained. For each node, six

dates of occurrence were selected at random and the zonal wind field in the independent NCEP reanalysis dataset was plotted. A good qualitative similarity between the NCEP individual plots and the ERA-40 node to which they were matched suggested the method could cluster daily states over East Africa in a reproducible way, allowing further analysis of rainfall SOMs to begin.

The second stage was to focus on the rainfall field of interest in this study. Three realisations of the SOM were evaluated; an anomaly series, a normalised series and a normalised series with a constrained domain. The tendency without the constrained domain was for large scale rainfall events over the Congo and Indian Ocean to dominate and for detailed rainfall events over the four focus countries to not feature. This was rectified by constraining the domain to cover only the land region of Tanzania, Kenya, Uganda and Ethiopia, excluding the large scale features and leading to a SOM array containing distinct rainfall events over these four focus counties, coastal Indian Ocean rainfall events and events of suppressed convection.

The final test of using the SOM to analyse daily rainfall events was to establish if the clustering of daily rainfall events meant anything physically; was there a reason for the rainfall being clustered spatially that could be identified in an independent variable? A suite of variables was examined for reasons of thoroughness, comprising moisture flux, moisture flux divergence, zonal winds, geopotential height and specific humidity fields.

Moisture flux divergence at the 850hPa level was confirmed as the most reliable indicator of the presence of rainfall events or dry conditions over an East African domain; the other circulation indicators were also useful but not to the same extent as the effective combination of specific humidity and zonal and meridional wind fields. For each node within the SOM array trained on the ERA-40 normalised rainfall series the accompanying

mean moisture flux divergence plot showed a spatially co-located (with the rainfall centre) convergent moisture flux anomaly. Equally nodes showing suppressed convection are supported by a physically consistent anomalous moisture flux divergence centre over the focus countries. The provision of water vapour, through moisture flux, is a key stage in the generation of rainfall. The quantity of this provision is controlled by the wind fields, which are better represented in climate models than the parameterised rainfall field (Washington et al, 2013). This accounted for the close relationship between the moisture flux and rainfall fields observed in the reanalysis datasets in this chapter and justified the utilisation in later chapters of the moisture flux and moisture flux divergence fields to evaluate model rainfall.

As a result of these three stages, it was concluded that the self-organising map could achieve this goal, with effective clustering results pertinently characterising distinct daily rainfall events within the short rains once the domain was constrained to exclude the considerable influence and domination of rainfall events originating over either the Congo or the Indian Ocean and that this was supported by physically meaningful states in the moisture flux divergence field for days assigned to each rainfall node.

Focus was then directed towards the research question; defining the contemporary structure of the daily rainfall field during the short rains from the period 1971-2000. This research question was explored further, building upon the initial analysis of Chapter 4, in Chapter 5, an examination of the characteristics of the daily rainfall field over the East African domain during the short rains of October to December in a second dataset. The absence of station data with sufficient spatial and temporal coverage that allowed a meaningful interpolation to a gridded field over the region necessitated the use of two different datasets in order to evaluate daily rainfall characteristics given the known weaknesses of the available data products (see Chapter 2, Sections 2.1 and 2.5). Chapter 5

focused on the TRMM satellite derived rainfall product in comparison with the ERA-40 reanalysis dataset. The overall aim of this work was to see if the two datasets produced the same structure of daily rainfall regimes, in particular in the clustering of spatial rainfall patterns and the proportional occurrence of these patterns. These were the two primary markers from which AOGCM daily rainfall regimes would be evaluated; the twentieth century runs of the models do not replicate the calendar dates, so similarities between TRMM and ERA-40 in calendar representations of SOM nodes were less important for this work.

The ERA-40 and TRMM rainfall series showed a high degree of qualitative similarity in the spatial patterns and amplitude of rainfall in the characteristic nodes within the SOM array clustered independently from the two data daily series, supported by the high statistical correlations between independently trained nodes in the two datasets. Each node in the independently trained TRMM normalised rainfall SOM was statistically significantly correlated at the 99% level with at least one node in the independently trained ERA-40 normalised rainfall SOM, with r values of 0.709 to 0.889 for the key rainfall events characterised during the short rains. This analysis supported conclusions that the short rains of East Africa were comprised of six main types of daily regimes, namely rainfall events centred over Uganda-southern Ethiopia, the eastern border of Kenya-Ethiopia, Tanzania, the Indian Ocean adjacent to the Tanzanian coast, the western length of the domain originating in the Congo and dry events of suppressed convection across the four focus countries.

Small variations in nodes depicting similar rainfall events led to the decision to reduce the number of states characterising daily rainfall events during the short rains from thirty five to twenty. The smaller number of nodes facilitated more focused analysis whilst still having an array size (five by four) that adequately covered the continuum of daily rainfall

events during the short rains. In order to assess the proportional occurrence of the different types of daily rainfall regimes identified within the SOM array the TRMM data series was applied to the ERA-40 normalised rainfall SOM. The proportion of dates assigned to each node was calculated for both the TRMM and ERA-40 series and compared through six groupings of qualitatively similar nodes, representing the main patterns that occurred during the short rains.

Three of the groupings, representing light rainfall down the western side of the domain, Uganda-southern Ethiopian bi-centred rainfall and Tanzanian rainfall centres showed almost identical occurrences in both datasets, with differences of less than 1% in occurrence over the two different control periods. Rainfall events over eastern Kenya and Uganda also showed very similar occurrences with only a 2.55% difference in allocations to these nodes. The main origins of difference were in the occurrence of dry conditions over the land domain and rainfall centres in the Indian Ocean, adjacent to Tanzania. Dry conditions occurred 8.32% more frequently in TRMM than ERA-40, with Indian Ocean rainfall events occurring 7.12% more frequently in ERA-40. Given the differences in time period of the two datasets (1971-2000 for ERA-40 and 1998-2010 for TRMM) and in the different methodologies of calculating daily rainfall values over East Africa, the similarities in the occurrence of each type of rainfall event is good.

The qualitative similarities of daily clustered rainfall regimes, supported by the high statistical correlation of nodes between the two datasets, confirmed the use of the ERA-40 SOM as an approximation of observed rainfall over East Africa to use as a basis for evaluating the daily rainfall in the AOGCM subset. Equally, when the key rainfall patterns were grouped together both rainfall datasets showed good similarities in the occurrence of the different events within the short rains as a whole, in particular for the four distinct land based rainfall events characterised in the SOM array. When the two series were matched to

the same SOM, there were differences in the days matched to each node which remained even when criteria for matching days were expanded. However, in Chapter 6, Section 6.4, qualitative examination of the two datasets reproduced the same amplitude and spatial pattern of daily mean rainfall centres associated with each node of the ERA-40 normalised SOM array, minimising these differences. Moreover, the 20C3M model runs do not correspond with calendar dates so the reproduction of rainfall types and occurrences of each was more important to this investigation than TRMM and ERA-40 matching each calendar date to the same node. The length and date availability of the TRMM data made it an unsuitable choice for comparison with the 20C3M model run, but the mean rainfall associated with nodes of the TRMM rainfall series were considered as part of the Chapter 6 AOGCM evaluation of the model rainfall field.

To summarise the characteristics of daily rainfall regimes within the short rains, six main regimes were clustered within the self-organising map, four rainfall centres over the focus countries, one in the western Indian Ocean adjacent to Tanzania and finally the condition of suppressed convection over the land region. Both TRMM and ERA-40 datasets clustered similar spatial patterns, reinforced by statistically significant correlations between the two arrays at the 99% level. The most frequently occurring states, accounting for 37.42% of days in the ERA-40 series were those associated with suppressed convection. States representing rainfall along the western side of the array accounted for approximately 15.5% of days in both series, with Indian Ocean rainfall centres having a similar frequency of occurrence in the ERA-40 series. Rainfall centres over Tanzania, Uganda-southern Ethiopia and eastern Kenya-Ethiopia accounted for a lesser proportion of days, measuring 8.99%, 8.41% and 12.71% of occurrences in ERA-40. All states had a physically consistent co-located moisture flux anomaly (divergent or convergent,

dependent on dry or wet conditions respectively) in both datasets, providing an atmospheric reason for the clustering of these six key types of events to have occurred.

9.2.2 Question 2 and sub question

2. To what extent can coupled climate models replicate the contemporary structure of the East African daily rainfall regimes of the short rains season?
 - To what extent is the mean rainfall climatology replicated in AOGCMs?

The sub question of the second research question was addressed in Chapter 3, Section 3.5. The ability of a subset of CMIP3 AOGCMs to replicate the annual climatology of East Africa was examined, comprising simulation of climate means, 95th percentile extremes, interannual variability and trends over the second half of the twenty-first century.

For the key variable of interest for this study, the rainfall field, the AOGCMs replicated the ITCZ related movement of the rainfall band of the annual cycle throughout the year, as demonstrated through comparison of mean daily rainfall at each latitude throughout the year. However, in five of the eight AOGCMs the rainfall band moved too far north over East Africa during the boreal summer months of June to August, as seen in CCCMA, CNRM and MIUB to the greatest extent, with ITCZ related rainfall moving to a peak latitude of 16°N in comparison to the peak of 13°N seen in the observational datasets. The rainfall band was too latitudinally widespread (specifically in CCCMA and CNRM) and simulated peak rainfall which was muted in intensity in comparison to observed datasets but contributed to an overestimated annual rainfall total. MPI and MRI were found to have the most realistic annual cycles according to monthly model rainfall data in terms of temporally and latitudinally correct mean daily rainfall amounts. GISS had daily mean rainfall of more than 8mm throughout the year at all latitudes from 20°S to 15°N, with no

spatial variation or annual cycle visible, making it a poor simulation of the annual cycle. This result was confirmed through analysis of the daily regimes of the short rains period, leading to the conclusion that under control conditions GISS had a highly unrealistic rainfall field under all markers of this variable.

Chapter 6 addressed the replication of the contemporary daily structure of East African short rains in coupled climate models during the control period of 1971-2000. Daily rainfall characteristics in the AOGCMs were established through appraisal of the following features; quantisation error when model data were applied to an observed SOM, the distribution of daily rainfall types within the models in comparison to the observed datasets and circulation indicators in the moisture flux field of the presence of rainfall.

There were several key results to emerge from this chapter. The quantisation error is a numerical measure of the ‘closeness of fit’ of the daily data when each date is matched to the trained SOM, calculated through the mean distance of each data point of the input series to the equivalent data point in the node to which the data are most closely fitted. It was found that the daily data from the AOGCMs could be applied to the SOM trained on the ERA-40 data set with a comparable quantisation error over the control period of 1971-2000, ranging from 11.603 to 13.991 in comparison to the error of the fit of the ERA-40 data on to the SOM it trained of 13.239. This relationship indicated a comparable level of variability in the daily rainfall regimes in terms of how well they could be associated with the trained ERA-40 SOM, but did not imply that the AOGCMs could replicate other aspects of the observed daily rainfall characteristics over East Africa. This was exemplified by the result that none of the models or the ensemble simulated the observed distribution of the different types of daily rainfall events within the months of October to December. Every model and the ensemble had a statistically different distribution of nodes types during the short rains to the ERA-40 data at the 99% level.

The models showed a disparity between the rainfall and moisture flux fields in their depiction of days assigned to the states of suppressed convection over the East African land domain. The moisture flux field associated with days characterised by suppressed convection in observed datasets reinforced the dry conditions seen in the rainfall field, with a mean state of widespread anomalous divergence for days assigned to these nodes. Conversely, the rainfall field in AOGCMS did not replicate these dry conditions with both the ensemble and individual AOGCMS characterised by widespread rainfall across the domain, far in excess of the 0-2mm daily mean seen in both the TRMM and ERA-40 datasets for these days. For days assigned to these nodes the ensemble had a daily rainfall mean of 2-8mm over East Africa when GISS is excluded. The mean moisture flux results for these days was for divergence over East Africa, as was expected for dry days but in direct conflict with the mean rainfall field in models for these days, as evidenced in Chapter 6, Section 6.5.

In contrast, for all nodes within the ERA-40 SOM array characterised by a rainfall event, when model data was applied, the rainfall and moisture flux fields agreed on a rainfall event co-located with that in the assigned SOM node. Model rainfall centres were accompanied by a centre of co-located anomalous moisture flux convergence for all of the models and the ensemble. Given the influence of the simulated wind field on the moisture flux field, this agreement between moisture flux and rainfall for rainfall events increased confidence in the ability of AOGCMS to reproduce daily rainfall events of the East African short rains during the control period. This confidence was tempered by the issues arising in the simulation of days which were expected to be dry over East Africa; the CMIP3 model subset and the ensemble could not accurately simulate dry days across the domain during the short rains. This issue was considered in Chapter 8.

9.2.3 Question 3

3. How do the East African short rains change under the high emissions SRES A2 scenario at the end of the twenty-first century in coupled climate models?

Chapter 7 addressed the third research question of this thesis, to examine daily rainfall events under a high greenhouse gas emissions scenario at the end of the twenty-first century. Rainfall projections in the CMIP3 model subset and the ensemble were examined. Each of the seven models and the ensemble mean examined here had a statistically significant increase in short rains rainfall between 1971-2000 and 2090-2099, confirming the IPCC conclusions on coherence for the East African short rains projected increase under a climate change scenario (Solomon et al, eds, 2007).

The projected increase was not uniform across the domain. The ensemble and all individual models except MIUB showed the greatest projected rainfall increase in the north of the East African domain, centred either over Uganda or Lake Victoria and southern Ethiopia. The ensemble had a peak projected increase of 1.8mm/day during the short rains centred over Lake Victoria followed by increases of 1.5mm/day over western Kenya and eastern Uganda. The individual AOGCM with the greatest projected increase of 3mm/day was GFDL, centred in Uganda to the west of Lake Victoria and in south-west Ethiopia. The model with the lowest projected increase was CSIRO; the greatest daily projection of 1.2mm was centred over northern Uganda.

To conclude, the projected increase in rainfall during the short rains was a statistically robust result in the CMIP3 model subset. The remainder of this work, under the fourth research question, examined how these increases manifested in the daily rainfall regimes of the short rains.

9.2.4 Question 4

4. How do projected changes to the East African short rains occur in coupled climate models in the daily rainfall field?

The fourth research question of this thesis dealt with the projected rainfall increase during the short rains over East Africa, specifically how it was manifested in the daily field in the AOGCMs. In Chapter 7, Section 7.3, the percentage of the projected rainfall increase by the 2090s in comparison to 1971-2000 associated with different spatial locations of rainfall was quantified. Most striking was the coherence across the ensemble and six of the seven individual models (excluding CSIRO) as to the proportion (between 20% and 35%) of the rainfall increase associated with days that were characterised in the control period by the suppression of convection across East Africa. This was important given the weakness of the model simulation of dry events during the control period (identified in Chapter 6) in the rainfall field in comparison to rainfall events across the domain.

The quantification of between 20 and 35% of the projected rainfall increase associated with days characterised by dry conditions was developed in Chapter 8, having been identified as significant in assessing confidence in the increasing rainfall projection across East Africa during the short rains.

Chapter 8 focused on establishing the origins of the projected rainfall increase assigned to days characterised by dry conditions over East Africa due to the suppression of convection across Tanzania, Kenya, Uganda and Ethiopia. This work, under the fourth research aim, contributed to the overall aim of evaluating the fidelity of East African model projections. Differences in rainfall projections that are not reinforced in the moisture flux field could

indicate that the rainfall is responding to differences in convection schemes (Washington et al, 2013).

Under twentieth century control conditions there was qualitative agreement between the moisture flux divergence field and rainfall centres for all nodes of the SOM array except those representing suppressed convection where the rainfall field was unable to reproduce the expected dryness associated with the anomalous divergence across the land regions of the domain.

Given the proportion of the projected rainfall increase assigned to these nodes, both the rainfall projections and moisture flux projections associated with these nodes of suppressed convection were examined. The 2090s projections differed between moisture flux and rainfall for days associated with nodes of suppressed convection. Without exception, the models showing more than 20% of the rainfall projection increase being assigned to nodes of suppressed convection did not have a moisture flux projection that supported this increase. Specifically, all of the models and the ensemble showed a strengthening of moisture flux divergence over continental East Africa under a high emissions scenario for dry nodes; suggesting that days assigned to these nodes would be drier under a climate change scenario. This was determined through mean moisture flux divergence anomaly plots for the 1971-2000 and 2090-2099 periods, set to the same scale for direct comparison.

Following the same methodology, nodes with rainfall centres over the domain had increasing rainfall projections that were coherent with the moisture flux projections, with the strongest increases in anomalous moisture flux convergence co-located with the rainfall centres of those nodes which project the greatest proportion of the rainfall increases for each model. The proportion of the increase assigned to days shown to be

poorly simulated in the models suggested that weaknesses in rainfall parameterisation schemes could be responsible for at least one fifth of twenty-first century rainfall increases under the A2 scenario.

9.3 Final Conclusions

The overall aim of this work was to assess the fidelity of rainfall projections for the East African short rains under a high emissions scenario. A new approach to model evaluation was undertaken, focused on redressing previous neglect of the structure of daily rainfall regimes over East Africa and how this structure was depicted in climate models. The premise of this work was that weaknesses in model representation of daily rainfall events during the short rains had the potential to undermine the coherent increasing rainfall projection for this region. If models were unable to realistically simulate daily rainfall regimes, in amplitude and proportional occurrence of events, during the twentieth century it would introduce uncertainty in their future projections. Equally, if future changes to model rainfall arose from states which were atypical, including daily events which were not found to exist in the observed datasets, this would erode confidence in the coherence of the wetting projection.

This work uncovered complexity in the daily rainfall projections for East Africa under a high emissions climate change scenario that was not immediately apparent through a focus on monthly mean rainfall. This was established through the use of the self-organising map, an automated clustering tool not liable to overtraining, which facilitated investigation of multiple large and heterogeneous reanalysis and model datasets over Tanzania, Kenya, Uganda and southern Ethiopia. The ability to analyse in detail the daily rainfall events during the East African short rains provided insight into both the composition of the short rains from individual daily events within this three month season and also the previously

unstated complexity within the increasing rainfall projection for the region under a climate change scenario by the end of the twenty-first century.

Most notable, was the large proportion of the projected short rains increased (20-35%) assigned to days of suppressed convection, for which the models could not simulate the rainfall field well. Rainfall was overestimated on such days in the coupled models despite the moisture flux field being characterised by anomalous moisture flux divergence, indicative of dry conditions. The one model which showed less than 20% of rainfall increase associated with days of suppressed convection, CSIRO, was also that with the most conservative rainfall projection of all the models under the SRES A2 scenario. This measured 38.95mm/year for the short rains season by the end of the twenty-first century in comparison with an ensemble increase across all seven individual CMIP3 models of 75.09mm/year and the next most conservative model, MRI, projecting a 66.71mm/year increase for the same period for the October-December season. The CSIRO model only allocated 6.3% of the projected increase to these days and has the lowest short rains anomaly by the 2090s in comparison to the control period. This model is less likely to have an overstated rainfall projection than those with larger amount allocated to days where the simulation of the rainfall field was poor.

This result was reinforced through the disparity in the moisture flux divergence and rainfall fields at the daily level for days characterised by the suppression of convection over East Africa. It was only for these dry days, with a mean of less than 2mm of rainfall a day in both the ERA-40 and TRMM datasets, that this result existed and was consistently found to be present. For all other nodes and models, the moisture flux divergence field and the rainfall characteristics showed coherent results; the rainfall centre was qualitatively co-located with moisture flux convergence.

Given that the East African projected rainfall increase is cited as a coherent model response (Solomon et al, eds, 2007) and given that at least part of the rainfall increase arises from a mode which the models do not simulate well in comparison with reanalysis data, the strength of evidence for the wetting response in the case of East Africa requires much further analysis. It could be that simulation of rainfall on days which should be dry arose in the way the AOGCMs create a threshold for the production of rainfall, a parameterised variable, and that the rainfall projection for the 2090s for East Africa was overstated. This result has implications for the region; the number of reported hydrometeorological disasters reported in East Africa has increased significantly since the 1980s with almost ten events a year reported in the 2000s in comparison to only three a year in the 1980s. These events affect more than two million people a year, impacting on economic and poverty issues (Shongwe et al, 2011). Therefore, projected changes under a climate change scenario, to the rainfall field in particular, over the twenty-first century need careful attention for adaptation and mitigation policy considerations in East Africa, including a closer analysis of the way convection is simulated in the region and the way in which the models are able to simulate more rainfall despite the increasing moisture flux divergence across the region.

9.4 Directions for future research

There are directions in which further research into East Africa daily rainfall events would be of benefit to the climate research community. Three suggestions as to areas of research that would directly complement the work undertaken in this thesis are outlined below.

9.4.1. A thorough examination of the convective scheme in AOGCMs

The primary unresolved question of this research related to the creation of daily rainfall values within the AOGCMs. Without exception, the models struggled to replicate small amounts of rainfall present on days of suppressed convection. Rainfall amounts for these days diverged from the associated moisture flux field, which was found to be the best indicator of rainfall for observed datasets on all days and the models on all days with a distinct rainfall centre located either over the land region or Indian Ocean.

Rainfall is a parameterised variable within AOGCMs, meaning that it is not explicitly resolved within the model; typically, such parameterisations originate through a simplification of physical models of the unresolved process. Parameterisations are often ‘tuned’ in order to maximise the possibility of a realistic simulation of a particular variable (Randall and Wood et al, 2007). Over East Africa, convection dominates and it is the convection schemes in the AOGCMs that are responsible for the simulation of rainfall. An examination of these schemes and the thresholds for the creation of rainfall would increase understanding of the reliability of the rainfall projections under a climate change scenario, in particular for days of suppressed convection which are poorly simulated in the climate models.

9.4.2 Investigation of daily rainfall regimes during the East African short rains in regional climate models

The work in this thesis has been undertaken on global climate models, focused on an East African domain. Regional climate models could provide further insight. Regional climate models (RCMs) are dynamical numerical prediction models run on a localised scale, covering the domain of interest, not a global domain. Use of a RCM is undertaken through ‘nesting’ it inside a global climate model, utilising a technique previously used in weather

forecasting then becoming widespread in examining climate change projections after it was demonstrated they could be successfully employed in this manner (Deque et al, 2005). RCMs use boundary conditions as simulated by an AOGCM to force a high resolution simulation within a limited domain, therefore the skill level of an RCM relies on the ability of the AOGCM to simulate appropriate boundary conditions in order to obtain a skilful high resolution RCM projection (Herrera et al, 2006).

Results from this thesis could be used to select AOGCMs to utilise in forcing boundary conditions for an RCM over the East African domain to obtain further insight into characteristics of the daily fields at a higher resolution. It has been suggested that due to the complex topography in East Africa and regionality in rainfall seasons a high resolution gridded dataset is needed to capture the rainfall characteristics of this region (Cook and Vizzy, 2013). As outlined in Chapter 3, Sections 3.5 there have been a limited number of studies undertaken over an East African domain using RCMs. These studies involved either the customisation of a RCM to accurately reproduce rainfall climatologies or the evaluation of a model over the region. Examples include the evaluation of different parameterisation schemes in the RCM CLM (Kaspar and Cubasch, 2008) and the examination of the ability of MM5 to simulate the rainy season onset (Riddle and Cook, 2008).

Alternatively, there is justification for using RegCM3 as a starting point for the in-depth evaluation of daily rainfall events over East Africa, using the SOM as a clustering tool. First, Sylla et al (2013) suggest that RegCM3 has daily rainfall statistics over East Africa within the range of values calculated in four observational datasets. Additionally, this model has been customised for use over the Horn of Africa by the ICTP (International Centre for Theoretical Physics), with a convective scheme selected for its performance during both a very dry and very wet year (1984 and 1986 respectively), allowing the model

to be employed to investigate climate dynamics in this domain with confidence in the ability of the RCM to reproduce mean and interannual rainfall realistically (Segele et al, 2009).

A specific direction would be to employ the SOM methodology on daily data series for rainfall and accompanying variables, namely moisture flux divergence, in a regional climate model. It is suggested that this could be done using the RegCM3 regional climate model accompanied by boundary conditions provided from one of the AOGCMs used in this study. For example, a comparison could be made between results under the SRES A2 scenario when the same RCM was forced by two different AOGCMs with contrasting results. The CSIRO model, with a conservative rainfall projection for the short rains and a small proportion aligned with days of suppressed convection could be compared to MRI, characterised by a realistic mean annual cycle of rainfall and a high proportion of the projected 66.71mm short rains rainfall increase associated with ‘dry’ days. Such a model setup would provide further insight to the daily rainfall regimes of the short rains over East Africa, both under control conditions and a climate change scenario.

9.4.3 Extending investigation of daily rainfall regimes to other seasons over East Africa

Taking other rainfall seasons over East Africa, there have already been conflicting projections and theories for the mechanisms behind potential changes, notably during the long rains of March to May.

The amplitude of the East African long rains have been in decline over Kenya and Ethiopia for the past thirty years (Williams and Funk, 2010). It is hypothesised that this is due to the interrelations in atmospheric dynamics between the Indian Ocean and East Africa. With increasing sea surface temperatures in the Indian Ocean as a result of the warming climate,

the concurrent release of latent heat induces the westward movement of dry air towards continental East Africa, decreasing convection, and therefore rainfall, in this region. The authors suggest that this is the most likely response of the long rains under climate change conditions; that the drying seen over the last decades of the twentieth century will continue through the twenty-first century (Williams and Funk, 2010). This result is reinforced through further confirmation of the decline in the East African long rains in recent decades, associated with changes to sea surface temperatures globally, but particularly in the tropical Pacific Ocean (Lyon and Dewitt, 2012). In contrast, it has been suggested that a higher coefficient of variation in AOGCM rainfall at the end of the twenty-first century for the Horn of Africa points towards a high likelihood of increased interannual rainfall variability, manifested through both dry and wet periods relative to the rainfall mean (Anyah and Qiu, 2012). A third perspective on potential changes to the long rains is offered by Shongwe et al (2011) who project increases to this rainfall period in the majority of East Africa, although they counter this projection with a lower certainty due to known model deficiencies in simulating control conditions for the long rains. Nonetheless, AOGCM rainfall projections for East Africa are of greater magnitude than the zonal mean rainfall increase for the tropics. This suggests that this region has additional circulation dynamics impacting on rainfall over and above the well understood increase in atmospheric water vapour in conjunction to a warming climate, even though for the long rains external controls are not well understood or documented (Shongwe et al, 2011).

These conflicting rainfall projections for the long rains season over East Africa present a dynamic research area. An extension of the examination of daily rainfall events in observed and model datasets over this domain to the long rains of March to May could yield results that could contribute to the clarification of rainfall projections over the region. The self-organising map has proven to be a useful and insightful tool enabling the

evaluation of daily rainfall events through classification of individual rainfall events in a reproducible manner. Prioritisation of the second rainy season of March to May would concentrate the study of the complex dynamics on daily timescales affecting the long rains over East Africa, which could lead to a convergent rainfall projection for the region.

9.5 Final Words

This thesis has researched daily characteristics of rainfall events within the short rains of October to December over an East African domain. It became apparent that differences were present between different datasets representing observed conditions over the domain and also in how AOGCMs were able to replicate these conditions on the daily timescale. There was agreement between the models, and the seven model ensemble, in the weakness of the AOGCMs in their rainfall response to days characterised by suppressed convection during the short rains over East Africa. This weakness was present in both the control period of 1971-2000 and at the end of the high emissions SRES A2 run in the 2090s. Days of suppressed convection were consistently associated with projected increased model rainfall despite the projected moisture flux divergence field indicating these days would become drier in future. Therefore, the conclusion of this work was that whilst AOGCMs over East Africa did simulate increased rainfall under a climate change scenario, it was possible that this was overstated.

It is hoped that this work has demonstrated the importance of considering the daily field in the examination of the rainfall field, and rainfall projections, over East Africa. Equally, it is intended that the work presented here could provide the basis for further examination of daily events over East Africa in alternative seasons and regional climate models, but in particular a thorough examination of the threshold for convection and rainfall in AOGCMs over East Africa.

References

- Abe, T. et al (2006) Self-Organizing Map (SOM) unveils and visualizes hidden sequence characteristics of a wide range of eukaryote genomes, *Gene*, 365 (1-2 SPEC. ISS.), pp27-43.
- Adeyewa, Z. D. and Nakamura, K. (2003) Validation of TRMM radar rainfall data over major climatic regions in Africa, *Journal of Applied Meteorology*, 42 (2), pp331-347.
- Ambroise, C. et al (2000) Hierarchical clustering of self-organizing maps for cloud classification, *Neurocomputing*, 30 (1-4), pp47-52.
- Anyah, R. O. and Semazzi, F. H. M. (2006) Climate variability over the Greater Horn of Africa based on NCAR AGCM ensemble, *Theoretical and Applied Climatology*, 86 (1-4), pp39-62.
- Anyah, R. O. and Semazzi, F. H. M. (2007) Variability of East African rainfall based on multiyear RegCM3 simulations, *International Journal of Climatology*, 27 (3), pp357-371.
- Anyah, R. O. and Qiu, W. (2012) Characteristic 20th and 21st century precipitation and temperature patterns and changes over the Greater Horn of Africa, *International Journal of Climatology*, 32 (3), pp347-363.
- Basalirwa, C. P. K. et al (1999) The climatological regions of Tanzania based on the rainfall characteristics, *International Journal of Climatology*, 19 (1), pp69-80.
- Behera, S. K. et al (2005) Paramount impact of the Indian Ocean dipole on the East African short rains: A CGCM study, *Journal of Climate*, 18 (21), pp4514-4530.
- Betts, A. K. and Beljaars, A. C. M. (2003) ECMWF ISLSCP-II near-surface dataset from ERA-40, *ECMWF ERA-40 Project Report Series, No. 8*, European Centre for Medium-Range Weather Forecasts, Shinfield, Reading, UK, available from www.ecmwf.int/publications, [accessed August 2012].
- Björnsson, H. and Venegas, S. A. (1997) A manual for EOF and SVD analyses of climate data, *Report No 97-1*, Department of Atmospheric and Oceanic Science and Centre for Climate and Global Change Research, McGill University, 52.

- Black, E. et al (2003) An observational study of the relationship between excessively strong short rains in in coastal East Africa and Indian Ocean SST, *Monthly Weather Review*, 113 (1), pp74-94.
- Blackwell, P. (2010) East Africa's Pastoralist Emergency: is climate change the straw that breaks the camel's back? *Third World Quarterly*, 31 (8), pp1321-1338.
- Bowden, J. H. and Semazzi, F. H. M (2007) Empirical analysis of intraseasonal climate variability over the Greater Horn of Africa, *Journal of Climate*, 20 (23), pp5715-5731.
- Bunzel, F. and Schmidt, H. (2013) The Brewer-Dobson circulation in a changing climate: impact of the model configuration, *Journal of the Atmospheric Sciences*, 70 (5), pp1437-1455.
- Camberlin, P. and Planchon, O. (1997) Coastal precipitation regimes in Kenya, *Geografiska Annaler, Series A: Physical Geography*, 79 (1-2), pp109-119.
- Camberlin, P. et al (2004) Simulated ENSO-tropical rainfall teleconnections in present-day and under enhanced greenhouse gases conditions, *Climate Dynamics*, 23 (6), pp641-657.
- Camberlin, P. et al (2009) Components of rainy seasons' variability in Equatorial East Africa: onset, cessation, rainfall frequency and intensity, *Theoretical and Applied Climatology*, 98 (3-4), pp237-249.
- Camberlin, P. et al (2010) Climate adjustments over Africa accompanying the Indian monsoon onset, *Journal of Climate*, 23 (8), pp2047-2064.
- Caminade, C. and Terray, L. (2006) Influence of increased greenhouse gases and sulphate aerosols concentration on diurnal temperature range over Africa at the end of the 20th century, *Geophysical Research Letters*, 33 (15), art no L15703.
- Caminade, C. and Terray, L. (2010) Twentieth century Sahel rainfall variability as simulated by the ARPEGE AGCM, and future changes, *Climate Dynamics*, 35 (1), pp75-94.
- Cavazos, T. (1999) Large scale circulation anomalies conducive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southwestern Texas, *Journal of Climate*, 12 (5 II), pp1506-1523.

- Cavazos, T. (2000) Using self-organizing maps to investigate extreme climate events: An application to wintertime precipitation in the Balkans, *Journal of Climate*, 13 (10), pp1718-1732.
- Christensen, J. H. and Kanikichana, K. K. (eds, 2013) Chapter 14: Climate phenomena and their relevance for future regional climate, in *Working Group I Contribution to the IPCC Fifth Assessment Report (AR5), Climate Change 2013: The Physical Science Basis*, Qin, D. and Stocker, T. et al (eds, 2013).
- Christy, J. R. et al (2009) Surface temperature variations in East Africa and possible causes, *Journal of Climate*, 22 (12), pp3342-3356.
- Chu, J. et al (2010) Spatial and temporal variability of daily precipitation in Haihe River basin. 1958-2007, *Journal of Geographical Sciences*, 20 (2), pp248-260.
- Clark, C. O. et al (2003) Interdecadal variability of the relationship between the Indian Ocean zonal mode and East African coastal rainfall anomalies, *Journal of Climate*, 16 (3), pp548-554.
- CMIP5 Coupled Model Intercomparison Project, *online information available from* <http://cmip-pcmdi.llnl.gov/cmip5/> [accessed September 2013].
- Conway, D. et al (2007) GCM simulations of the Indian Ocean dipole influence on East African rainfall: Present and future, *Geophysical Research Letters*, 34 (3), art no. L03705.
- Cook, K. H. and Vizy, E. K. (2013) Projected changes in East African rainy seasons, *Journal of Climate*, 26 (16), pp5931-5948.
- Copsey, D. et al (2006) Recent trends in sea level pressure in the Indian Ocean region, *Geophysical Research Letters*, 33 (19), art no. L19712.
- Crane, R. G. and Hewitson, R. G. (2003) Clustering and upscaling of station precipitation records to regional patterns using self-organizing maps (SOMs), *Climate Research*, 25 (2), pp95-107.
- Dee, D. P. et al (2011) The ERA-Interim reanalysis: configuration and performance of the data assimilation system, *Quarterly Journal of the Royal Meteorological Society*, 137 (656), pp553-597.

- Deque, M. et al (2005) Global high resolution versus Limited Area Model climate change projections over Europe: Quantifying confidence level from PRUDENCE results, *Climate Dynamics*, 25 (6), pp653-670.
- Diro, D. T. et al (2009) Evaluation of reanalysis rainfall estimates over Ethiopia, *International Journal of Climatology*, 29 (1), pp67-78.
- Doherty, R. M. et al (2010) Implications of future climate and atmospheric CO₂ content for regional biogeochemistry, biogeography and ecosystem services across East Africa, *Global Change Biology*, 16 (2), pp617-640.
- Durack, P. et al (2012) Ocean salinities reveal strong global water cycle intensification during 1950-2000, *Science*, 336 (6080), pp455-458.
- Dytham, C. (2003) Choosing and Using Statistics: A Biologist's Guide, *Blackwell Publishing*.
- ECMWF Data Server, information at <http://www.ecmwf.int/products/data/archive/>, [Accessed multiple times, 2008-2013].
- Faisal, T. et al (2010) Reexamination of risk criteria in dengue patients using the self-organizing map, *Medical and Biological Engineering and Computing*, 48 (3), pp293-301.
- Funk, C. et al (2008) Warming of the Indian Ocean threatens eastern and southern African food security but could be mitigated by agricultural development, *Proceedings of the National Academy of Sciences in the United States of America*, 105 (32), pp11081-11086.
- Gebril, M. et al (2011) Detecting environmental change using Self-Organising Map techniques applied to the ERA-40 database, *Data Science Journal*, 10, pp1-12.
- Gerdes, R. and Koberle, C (2007) Comparison of Arctic sea ice thickness variability in IPCC climate of the 20th century experiments and in ocean-sea ice hindcasts, *Journal of Geophysical Research C: Oceans*, 112 (4), art no C04S13.
- Giannini, A. et al (2008) A global perspective on African climate, *Climate Change*, 90 (4), pp359-383.

- Gleckler, P. J. et al (2008) Performance metrics for climate models, *Journal of Geophysical Research*, 113 (6), art no. DO6104.
- Griebeler, E. M. and Seitz, A. (2006) The use of Markovian metapopulation models: Reducing the dimensionality of transition matrices by self-organizing Kohonen networks, *Ecological Modelling*, 192 (1-2), pp271-285.
- Hagemann, S., Arpe, K. and Bengtsson, L. (2005) Validation of the hydrological cycle of ERA-40, *ECMWF ERA-40 Project Report Series No. 24*, European Centre for Medium-Range Weather Forecasts, Shinfield, Reading, UK, available from www.ecmwf.int/publications [accessed August 2012].
- Hannachi, A. et al (2007) Empirical orthogonal functions and related techniques in atmospheric sciences: A Review, *International Journal of Climatology*, 27 (9), pp1119-1152.
- Henson, S. A. et al (2010) Detection of anthropogenic climate change in satellite records of ocean chlorophyll and productivity, *Biogeosciences*, 7 (2), pp621-640.
- Herrera, E. et al (2006) Downscaling methods applied to atmosphere-ocean general circulation models (AOGCM). [Méthodes de désagrégation Appliquées aux Modèles du Climat Global Atmosphère-Océan (MCGAO)], *Revue des Sciences De L'Eau*, 19(4), pp297-312
- Herrmann, S. M. and Mohr, K. I (2011) A continental-scale classification of rainfall seasonality regimes in Africa based on gridded precipitation and land surface temperature products, *Journal of Applied Meteorology and Climatology*, 50 (12), pp2504-2513.
- Hession, S. L and Moore, N. (2011) A spatial regression analysis of the influence of topography on monthly rainfall in East Africa, *International Journal of Climatology*, 31 (10), pp1440-1456.
- Hewitson, B. C. and Crane, R. G (1994, eds) *Neural Nets: Applications in Geography*, Kluwer Academic Publishers.
- Hewitson, B. C. and Crane, R. G. (2002) Self-organising maps: applications to synoptic climatology, *Climate Research*, 22 (1), pp13-26.

- Hewitson, B. C. and Crane, R. G. (2006) Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa, *International Journal of Climatology*, 26 (10), pp1315-1337.
- Hirabayashi Y. et al. (2008) A 59-year (1948-2006) global near-surface meteorological data set for land surface models. Part I: Development of daily forcing and assessment of precipitation intensity, *Hydrological Research Letters*, 2, pp36-40.
- Hirabayashi, Y. et al (2005) A 100-year (1901-2000) global retrospective estimation of terrestrial water cycle, *Journal of Geophysical Research D: Atmospheres*, 110 (19), art no. D19101, pp1-23.
- Hsu, Y. -T. et al (2010) Forecast of financial time series based on grey self-organizing maps, *International Journal of Innovative Computing, Information and Control*, 6 (2), pp 475-486.
- Hulme, M. et al (2001). African climate change: 1900-2100, *Climate Research*, 17 (2), pp145-168.
- Kalnay, E. et al (1996) The NCEP/NCAR 40-Year Reanalysis Project, *Bulletin of the American Meteorological Society*, 77 (3), pp437-471.
- Kaspar, F. and Cubasch, U. (2008) Simulation of East African precipitation patterns with the regional climate model CLM, *Meteorologische Zeitschrift*, 17 (4), pp511-517.
- Kenya's Climate Change Action Plan – National Adaptation Plan (2013), *Republic of Kenya Ministry of Environment and Mineral Resources*, available at <http://www.kccap.info/>. [Accessed 15/09/2013].
- Kharin, V. V. et al (2007) Changes in temperature and precipitation extremes in the IPCC ensemble of global coupled model simulations, *Journal of Climate*, 20 (8), pp1419-1444.
- Kiktev, D. et al (2003) Comparison of modelled and observed trends in indices of daily climate extremes, *Journal of Climate*, 16 (22), pp3560-3571.
- Kijazi, A. L. and Reason, C. J. C. (2005) Relationships between intraseasonal rainfall variability of coastal Tanzania and ENSO, *Theoretical and Applied Climatology*, 82 (3-4), pp153-176.

- Kizza, M. et al (2009) Temporal rainfall variability in the Lake Victoria Basin in East Africa during the twentieth century, *Theoretical and Applied Climatology*, 98 (1-2), pp119-135.
- Kniveton, D. R. et al (2009) Trends in the start of the wet season over Africa, *Journal of Climatology*, 29 (9), pp1216-1225.
- Kohonen T. (1990) The self-organizing map, *Proceedings of the IEEE*, 78, pp1464–1480.
- Legutke, S. and Voss, R. (1999) The Hamburg Atmosphere-Ocean Couple Circulation Model echo-G, Technical Report 18, *Hamburg: Deutsches Klimarechenzentrum*.
- Lessman, B. et al (2007) Multiscale analysis of MR-mammography data, *Zeitschrift fur Medizinische Physik*, 17 (3), pp166-171.
- Liskha-Hackzell, J. J. and Martin, D. P. (2005) Analysis of nighttime activity and daytime pain in patients with chronic back pain using a self-organizing map neural network, *Journal of Clinical Monitoring and Computing*, 19 (6), pp411-414.
- Liu, Y. et al (2006) Performance evaluation of the self-organizing map for feature extraction, *Journal of Geophysical Research C: Oceans*, 111 (5), art no C05018.
- Love, T.B et al (2004) A 20-year daily Africa precipitation climatology using satellite and gauge data, preprints, 14th Conference on Applied Climatology, Seattle, WA, American Meteorological Society, P5.4.
- Lott, F. et al (2013) Can the 2011 East African drought be attributed to human-induced climate change? *Geophysical Research Letters*, 40 (6), pp1177-1181.
- Lyon, B. and Dewitt, D. G. (2012) A recent and abrupt decline in the East African long rains, *Geophysical Research Letters*, 39 (2), art no. L02702.
- Maidment, R. I. et al (2012) Evaluation of satellite-based and model re-analysis rainfall estimates for Uganda, *Meteorological Applications*, in press, accessed online at <http://www.met.reading.ac.uk/~sgs02rpa/PAPERS/Maidment12MA.pdf> [accessed August 2012].

- Malmgren, B. A. and Winter, A. (1999) Climate zonation in Puerto Rico based on principal components analysis and an artificial neural network, *Journal of Climate*, 12 (4), pp977-985.
- Manatsa, D. et al (2012) Shifts in IOD and their impacts on association with East African rainfall, *Theoretical and Applied Climatology*, 110 (1-2), pp115-128.
- Mapande, A. T. and Reason, C. J. C. (2005) Links between rainfall variability on intraseasonal and interannual scales over western Tanzania and regional circulation and SST patterns, *Meteorology and Atmospheric Physics*, 89 (1-4), pp215-234.
- McHugh, M. J. (2006) Impact of South Pacific circulation variability on East African rainfall, *International Journal of Climatology*, 26 (4), pp505-521.
- McSweeney, C. F. (2009) Series of 353 stations with daily precipitation data located over Ethiopia, Kenya, Tanzania and Uganda, *personal correspondence passed on with permission from D. Conway*.
- Meehl, G. et al (2007) The WCRP CMIP3 Multimodel Dataset: A New Era In Climate Change Research, *Bulletin of the American Meteorological Society*, 88 (9), pp1383-1394.
- Mitchell, T. D. et al (2004) A comprehensive set of high-resolution grids of monthly climate for Europe and the globe: the observed record (1901-2000) and 16 scenarios (2001-2100), *Tyndall Working Paper 55*, Tyndall Centre, UEA, Norwich, UK, www.tyndall.ac.uk.
- Mitchell, T. D. and Jones, P. D. (2005) An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *International Journal of Climatology*, 25 (6), pp693-712.
- Mölg, T. et al (2009) Quantifying climate change in the tropical midtroposphere over East Africa from glacier shrinkage on Kilimanjaro, *Journal of Climate*, 22 (15), pp4162-4181.
- Mpeta, E. J. and Jury, M. R. (2001) Intra-seasonal convective structure and evolution over tropical East Africa, *Climate Research*, 17 (1), pp 83-92.

- Mutai, C. C. and Ward, M. N. (2000) East African rainfall and the tropical circulation/convection on intraseasonal to interannual timescales, *Journal of Climate*, 13 (22), pp3915-3939.
- Mworia, J. K and Kinyamario, J, I (2008) Traditional strategies used by pastoralists to cope with La Nina induced drought in Kajiado, Kenya, *African Journal of Environmental Science and Technology*, 2 (1), pp10-14.
- Nakicenovic, N. and Swart, R. (eds, 2000) IPCC Special Report on Emissions Scenarios, IPCC, Cambridge University Press, UK, pp570.
- National Center for Atmospheric Research Staff (Eds) The Climate Data Guide: TRMM: Tropical Rainfall Measuring Mission, *retrieved from* <https://climatedataguide.ucar.edu/climate-data/trmm-tropical-rainfall-measuring-mission>, [Accessed multiple times].
- NCEP online composite plotting tool, <http://www.esrl.noaa.gov/psd/data/composites/day/index.html>, [Accessed multiple times].
- New, M. et al (1999) Representing Twentieth-Century Space-Time Climate Variability. Part I: Development of a 1961-1990 Mean Monthly Terrestrial Climatology, *Journal of Climate*, 12 (3), pp829-856.
- New, M et al (2000) Representing Twentieth-Century Space-Time Climate Variability. Part II: Development of 1901-1996 Monthly Grids of Terrestrial Surface Climate, *Journal of Climate*, 13(13), pp2217-2238.
- Nicholson, S. E. (1996) A review of climate dynamics and climate variability in eastern Africa, in *The limnology, climatology and paleoclimatology of the East African Lakes*, Johnson, T. C. and Odada, E. eds, Gordon and Breach, Amsterdam, pp25-56.
- Nishiyama, K. et al (2007) Identification of typical synoptic patterns causing heavy rainfall in the rainy season in Japan by a self-organising map, *Atmospheric Research*, 83 (2-4 special issue), pp185-200.
- Oettli, P. and Camberlin, P. (2005) Influence of topography on monthly rainfall distribution over East Africa, *Climate Research*, 28 (3), pp199-212.

- Omondi, P. et al (2012) Decadal rainfall variability modes in observed rainfall records over East Africa and their relations to historical sea surface temperature changes, *Journal of Hydrology*, 464-465, pp140-156.
- Osman, Y. Z. and Shamseldin, A. (2002) Qualitative rainfall prediction models for central and southern Sudan using El Nino-Southern Oscillation and Indian Ocean sea surface temperature indices, *International Journal of Climatology*, 22 (15), pp1861-1878.
- Otieno, V. O. and Anyah, R. O. (2013) CMIP5 simulated climate conditions of the Greater Horn of Africa (GHA). Part II: Projected climate, *Climate Dynamics*, 41 (7-8), pp2099-2113.
- Philippon, N. et al (2002) Empirical predictability study of October-December East African rainfall, *Quarterly Journal of the Royal Meteorological Society*, 128 (585 Part A), pp2239-2256.
- Polson, D. et al (2013) Have greenhouse gases intensified the contrast between wet and dry regions, *Geophysical Research Letters*, 40, art no 50923.
- Randall, D. A. and Wood, R. A et al (2007) Climate Models and Their Evaluation, 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. et al, (eds)]. Cambridge University Press. Cambridge, United Kingdom and New York, NY, USA.
- Raupach, M. R. et al (2007) Global and regional drivers of accelerating CO₂ emissions, *Proceedings of the National Academy of Sciences of the United States of America*, 104 (24), pp10288-10293.
- Reljin, I. S. et al (2002) Clustering of climate data in Yugoslavia by using the SOM neural network, *Neural Network Applications in Electrical Engineering*, NEUREL '02 6th Seminar, pp202-206.
- Riddle, E. E. and Cook, K. H. (2008) Abrupt rainfall transitions over the Greater Horn of Africa: Observations and regional model simulations, *Journal of Geophysical Research-Atmospheres*, 113 (D15), D15109.

- Rocha, A. et al (2008) High-frequency precipitation changes in southeastern Africa due to anthropogenic forcing, *International Journal of Climatology*, 28 (9), pp1239-1253.
- Rowell, D. P. (2012) Sources of uncertainty in future changes in local precipitation, *Climate Dynamics*, 39 (7-8), pp1929-1950.
- Schreck III, C. J. and Semazzi, F. H. M. (2004) Variability of the recent climate of eastern Africa, *International Journal of Climatology*, 24 (6), pp681-701.
- Segele, Z. T. et al (2009) Evaluation and adaptation of a regional climate model for the Horn of Africa: Rainfall climatology and interannual variability, *International Journal of Climatology*, 29 (1), pp47-65.
- Severin, E. (2010) Self organizing maps in corporate finance: Quantitative and qualitative analysis of debt and leasing, *Neurocomputing*, 73 (10-12), pp2061-2067.
- Shongwe, M. E. et al (2006) Performance of recalibration systems for GCM forecasts for southern Africa, *International Journal of Climatology*, 26 (12), pp1567-1585.
- Shongwe, M. E. et al (2011) Projected changes in mean and extreme precipitation in Africa under Global Warming. Part II: East Africa, *Journal of Climate*, 24 (14), pp3718-3733.
- Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor, M., and Miller, H. L. (eds, 2007) Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, *Cambridge University Press*, Cambridge, United Kingdom and New York, NY, USA, 996pp.
- SOM Programming Team (1995) SOM_PAK, The Self-Organizing Map Program Package, *Helsinki University of Technology Laboratory of Computer and Information Science*, Rakentajanaukio 2 C, SF-02150 Espoo, Finland.
- Song, Y. et al (2004) A coupled regional climate model for the Lake Victoria basin of East Africa, *International Journal of Climatology*, 24 (1), pp57-75.
- Sun, L. et al (1999) Application of the NCAR regional climate model to eastern Africa-2. Simulation of interannual variability of short rains, *Journal of Geophysical Research-Atmospheres*, 104 (D6), pp6459-6562.

- Sylla, M. B. et al (2010) Consistency of projected drought over the Sahel with changes in the monsoon circulation and extremes in a regional climate model projections, *Journal of Geophysical Research Atmospheres*, 115, art no D16108.
- Sylla, M. B. et al (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*, 2012, art no 981649.
- Sylla, M. B. et al (2013) Uncertainties in daily rainfall over Africa: Assessment of gridded observation products and evaluation of a regional climate model simulation, *International Journal of Climatology*, 33 (7), pp1805-1817.
- Taylor, S. et al (2004) An evaluation of two spatial interpolation techniques in global sea-surface temperature reconstructions: Last Glacial Maximum and Pliocene case studies, *Quaternary Science Reviews*, 23 (9-10), pp1041-1051.
- Thornton, P. K. et al (2009) Spatial variation to crop yield response to climate change in East Africa, *Global Environmental Change- Human and Policy Dimensions*, 19 (1), pp54-65.
- Tozuka, T. et al (2008) Tropical Indian Ocean variability revealed by self-organizing maps, *Climate Dynamics*, 31 (2-3), pp333-343.
- Trenberth, K. E and Guillemot, C. J. (1998) Evaluation of the atmospheric moisture and hydrological cycle in the NCEP/NCAR reanalyses, *Climate Dynamics*, 14 (3), pp213-231.
- Trenberth, K. E. et al (2001) Quality of Reanalyses in the Tropics, *Journal of Climate*, 14 (7), pp1499-1510.
- Trenberth, K. E. et al (2003) The changing character of precipitation, *Bulletin of the American Meteorological Society*, 84 (9), pp1205-1217.
- TRMM info taken from <http://trmm.gsfc.nasa.gov> and <http://disc.sci.gsfc.nasa.gov>. Accessed on [12/01/2012].
- Uppala, S. M. et al (2005) The ERA-40 re-analysis, *Quarterly Journal of the Royal Meteorological Society*, 131 (612), pp2961-3012.

- Verdin, J. et al (2005) Climate science and famine early warning, *Philosophical Transactions of the Royal Society B- Biological Sciences*, 360 (1463), pp2155-2168.
- Waldstreicher, J. S. (1989) A guide to utilizing moisture flux convergence as a predictor of convection, *National Weather Association Training Papers*, 14 (4), pp20-35.
- Wang, J. et al (2002) Clustering of the SOM easily reveals distinct gene expression patterns: results of a reanalysis of lymphoma study, *BMC bioinformatics [electronic resource]*, 3 (1), pp36.
- Washington, R. and Pearce, H. (2012) Climate Change in East African Agriculture: Trends, current projections, crop-climate suitability, and prospects for improved climate model information. *CGIAR Research Program on Climate Change, Agriculture and Food Security (CCAFS)*. Copenhagen, Denmark. Available online at www.ccafs.cgiar.org
- Washington, R. et al (2013) Congo Basin rainfall climatology: Can we believe the climate models? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368 (1625), art no 20120296.
- Webster, P. J. (1983) Large-scale structure of the tropical atmosphere, in *Large-Scale Dynamical Processes in the Atmosphere*, Hoskins, B. and Pearce, R. (1983, eds), pp235-275, New York, NY: Academic Press.
- Wilby, R. L. et al (2002) SDSM – A decision support tool for the assessment of regional climate change impacts, *Environmental Modelling and Software*, 17 (2), pp147-159.
- Williams, A. P. and Funk, C. (2010) A westward extension of the Tropical Pacific warm pool leads to March through June drying in Kenya and Ethiopia, *U.S Geological Survey Open File Report*, 2010-1199.
- Williams, A. P. et al (2012) Recent summer precipitation trends in the Greater Horn of Africa and the emerging role of Indian Ocean sea surface temperature, *Climate Dynamics*, 39 (9-10), pp2307-2328.
- Williams, C. J. R. et al (2010) Assessment of a climate model to reproduce rainfall variability and extremes over southern Africa, *Theoretical and Applied Climatology*, 99 (1-2), pp9-27.

- Xiao, Y. and Lei, L. (2008) Research on comparison of credit risk evaluation models based on SOM and LVQ neural networks, *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*, art no. 4593270, pp2230-2235.
- Yang, X. and Hodler, T. (2000) Visual and statistical comparisons of surface modelling techniques for point-based environmental data, *Cartography and Geographical Information Science*, 27, pp165-175.
- Yarnal, B. (1993) Synoptic climatology in environmental analysis: a primer, Wiley, ISBN 978-0-471-94796-7.
- Yarnal, B. et al (2001) Developments and prospects in synoptic climatology, *International Journal of Climatology*, 21 (15), pp1923-1950.
- Yin, L. et al (2013) How well can CMIP5 simulate precipitation and its controlling processes over tropical South America? *Climate Dynamics*, 41 (11-12), pp3127-3143.
- Zhao, Y. et al (2005) Validation of a coupled GCM and projection of summer rainfall change over South Africa, using a statistical downscaling method, *Climate Research*, 28 (2), pp109-122.