IMPROVED CONVECTIVE PARAMETRIZATION FOR PRECIPITATION MODELING OVER THE HORN OF AFRICA

By

Temesgen Gebremariam

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE IN PHYSICS
AT
ADDIS ABABA UNIVERSITY
ADDIS ABABA, ETHIOPIA
JUNE 2011

ADDIS ABABA UNIVERSITY

DEPARTMENT OF PHYSICS

IMPROVED CONVECTIVE PARAMETRIZATION FOR PRECIPITATION MODELING OVER THE HORN OF AFRICA

By Temesgen Gebremariam

Approved by the Examination Committee

Supervisor: _____

Dr. Gizaw Mengistu

Examiners: _____

Dr. Elias Lewi

_____

Dr. Lemi Demeyu

ADDIS ABABA UNIVERSITY

Date: **June 2011**

Author: **Temesgen Gebremariam**

Title: **IMPROVED CONVECTIVE PARAMETRIZATION FOR PRECIPITATION MODELING OVER THE HORN OF AFRICA**

Department: **Physics**

Degree: **M.Sc.** Convocation: **June** Year: **2011**

Permission is herewith granted to Addis Ababa University to circulate and to have copied for non-commercial purposes, at its discretion, the above title upon the request of individuals or institutions.

Signature of Author

Dedicated to My Father (deceased), who is behind every success of mine.

Table of Contents

List of Figures

List of Tables

Acronyms

ABE Available buoyant energy

AIRS Alliance Icing Research Study

AVHRR Advanced Very High Resolution Radiometer

BATS Biosphere-Atmosphere Transfer Scheme

BATS1e Biosphere-Atmosphere Transfer Scheme version 1e

CCM Community Climate Model

CCM3 Community Climate Model version 3

CCSM Community Climate System Model

CLARA Clustering Large Applications

CLM Community Land Model

CLM3.5 Community Land Model version 3.5

CMAP CPC Merged Analysis of Precipitation

CPC Climate Prediction Center

ECMWF European Centre for Medium-Range Weather Forecasts

ENSO El Niño-Southern Oscillation

ERA40 ECMWF 40-year Reanalysis

ERA-Intrim ECMWF Reanalysis

GCM General Circulation Model

GISST Global Sea Surface Temprature

GLCC Global Land Cover Characterization

GOES Geostationary Satellite

GPCP Global Precipitation Climatology Project

GPI GOES Precipitation Index

Grell-AS Grell scheme with Arkawa-Schubert closure

Grell-FC Grell scheme Fritsch and Chappell closure

GTOPO30 Global 30 arc second elevation dataset

ICTP The Abdus Salam International Centre for Theoretical Physics

ICTZ Inter Tropical Convergence Zone

LCB Level of Cloud Base

LNB Level of Neutral Buoyancy

MIT Massachusetts Institute of Technology

MM4 Mesoscale Model version 4

MSU Microwave Sounding Unit

NCAR National Center for Atmospheric Research

NCEP National Centers for Environmental Prediction

NNRP1 NCEP/NCAR Reanalysis Product version 1

NNRP2 NCEP/NCAR Reanalysis Product version 2

NOAA National Oceanic and atmospheric Administration

OISST Optimum Interpolation Sea Surface Temprature

OLR Outgoing Long wave Radiation

OPI Outgoing Longwave Radiation Precipitation Index

PAM Partitioning Around Medoids

PSU Pennsylvania State University

PSUNCAR PSU National Center for Atmospheric Research

RCM Regional Climate Model

RegCM REGional Climate Model

RegCM3 REGional Climate Model version 3

RegCM4 REGional Climate Model version 4

RMSE Root Mean Square Error

SSM/I Special Sensor Microwave/Imager

SST Sea Surface Temperature

SUBEX the SUB-grid EXplicit moisture scheme

USGS United States Geological Survey

TIROS Television Infrared Observation Satellite Program

TOVS TIROS Operational Vertical Sounder

Abstract

The study compared the performance of four different cumulus parametrization schemes in the ICTP's Regional Climate Model RegCM4 for the period 1989-2008 over the Horn of Africa. The Horn of Africa has been delineated in to 12 homogeneous region based on clustering method that minimizes a sum of dissimilarities. Grell-FC scheme best captured the mean climatology. Within the 12 clustered regions positive/negative precipitation biases are obtained in Emanuel/Grell-AS and Kuo convective schemes while intermediate values closer to zero are obtained in Grell-FC. Grell-FC scheme is then selected for customization over the Horn of Africa. To improve the simulation of rainfall using the selected Grell-FC scheme, several sensitivity experiments are performed by adjusting the minimum and maximum precipitation efficiency, the minimum and maximum shear effect on precipitation efficiency, the minimum and maximum convective heating and the available buoyant energy removal time scale. Setting the minimum and maximum shear effect on precipitation efficiency to 0.10 and 0.30 respectively improve the correlation, which implies a better simulation of precipitation. Setting the minimum and maximum convective heating to -150 and 400 respectively minimize the positive bias over most of the regions. 17 years long period run revealed that using the new values of shear effect on precipitation efficiency improve the correlation and representation of the inter-annual variability.

Acknowledgments

I would like to thank Dr. Gizaw Mengistu, my adviser, for his consistent support and friendly approach during this research. He showed me different ways to approach a research problem and the need to be persistent to accomplish any goal.

I sincerely appreciate Dr. Gulilat Tefera for his support in suggesting solution during RegCM runs and for his quick response to my e-mails. I would also like to thank Addisu Gezahegne for his tireless help to make the AddisHPC work properly.

I am also thankful to Tamrat Bekele and Endris Tilahun for their patient support in collecting and sending me my salary on time. I also thank all of my friends for being with me when I need their help.

Finally, I wish to thank my family especially my mother W/ro Kahssa Gebremecheal for their regular support throughout the course of my study.

June, 2011
Addis Ababa, Ethiopia

Chapter 1

Introduction

Rainfall is one of the least well-simulated parameters in numerical models. This is, in part, due to the fact that the spatial scale of the physical mechanisms which generate rainfall is much finer than can be represented by the models (Diro *et al.*, 2008).

The adequate representation of convective processes is particularly important in RCMs, but there is no universally accepted framework for representing convection in numerical simulation models operating with grid scales that prohibit fully explicit representation (Gochis *et al.*, 2001).

Segele *et al.* (2008) assessed the ability of the four convective schemes, using RegCM3, to capture extreme interannual rainfall variations over the Horn of Africa by comparing the July-August simulated rainfall for widespread very dry 1984 and locally very wet 1996 seasons and suggested that Emanuel Scheme captured the 1984-1996 interannual variability over most of the Horn of Africa better than the other schemes. Davis *et al.* (2009) performed customization of RegCM3 for East Africa and Tropical Indian Ocean domain by focuses on the short rains season [October, November, December (OND)] using 3 year (2000, 2001, 2002) climatology and suggested that Emanuel scheme to be the best scheme, based on the criteria examined in their

study; It provided a more realistic partitioning of convective and stratiform rainfall, and it has consistency, produced large convective rainfall rates over the Horn of Africa (Davis *et al.*, 2009).

The Horn of Africa exhibits high interannual and seasonal rainfall variability ranging from unimodal, bimodal and trimodal rainfall distributions (Indeje *et al.*, 2000). In this study, we have used 20 years long period simulation and the annual climatology instead of particular rain seasons. We also have performed cluster analysis based on observational precipitation dataset, using CLARA to identify homogeneous regions of climate variability.

The general objective of this study is to evaluate the four alternative convective precipitation schemes on the performance of regional climate modeling (RegCM4), in particular their role on improving the accuracy of the simulated precipitation field over the Horn of Africa.

The specific objectives of the study are:

- Asses the RegCM4 performance in representing the climatology, annual cycle and interannual variability of precipitation over the Horn of Africa.
- Diagnose the spatial, seasonal and interannual variability of the Horn of Africa precipitation based on 20 years RegCM4 simulations.
- Identify the best convective precipitation scheme used to simulate precipitation over the Horn of Africa.
- Improve the selected convective scheme for better representation of precipitation over the Horn of Africa.

In view of the above stated objectives, this study discuses comparison of the precipitation distribution from RegCM4 simulations using Emanuel, Grell-AS, Grell-FC, and Kuo convective precipitation schemes with respect to CMAP and GPCP datasets over the Horn of Africa for the period of 1989-2008. Spatial, seasonal and inter-annual variability of rainfall will be evaluated for the Horn of Africa at a spatial scale that reflects the local variability of the rainfall climate.

After selecting one of the convective schemes, several sensitivity experiments are performed to RegCM4 by adjusting parameters which were believed to improve precipitation simulations over the Horn of Africa. Quantitative evaluation of the precipitations obtained using the different parameters is then performed and further analysis of the selected set of parameters will be discussed based on 1989-2005 simulation.

This thesis is organized as follows. Following the introductory chapter, Chapter 2 discuses the basic atmospheric dynamics equations and their implementation on RegCM4. Chapter 3 provides information on basic aspects of regional climate modeling, including brief descriptions of the model physics and different convective precipitation schemes. Chapter 4 describes the validation Datasets and methods of study. In Chapter 5, the results of the multiyear RegCM4 simulations of rainfall over the Horn of Africa using the four alternative convective schemes are discussed. Chapter 6 presents the results of sensitivity experiments involving the adjustment of parameters in the Grell-FC scheme. Finally, in Chapter 7, the overall conclusions of the thesis are provided.

Chapter 2

Dynamics of the Atmosphere

This chapter describes the basic conservation laws of atmospheric dynamics, the vertical and horizontal coordinate systems, the map projection and map scale factors used in the model and the basic model equations.

2.1 Conservation Laws of Atmospheric Motion

Atmospheric motions are governed by the fundamental physical laws of conservation of momentum, conservation of mass and conservation of energy which can be used in driving the momentum equation, the continuity equation and the thermodynamic energy equation (Holton, 2004; Saha, 2008).

2.1.1 Conservation of Momentum

One of the fundamental principels governing atmospheric motion is the conservation of momentum (Newton's second law of motion) which can be expressed mathematically by the momentum equation.

Newtons second law of motion states that the rate of change of momentum of an object, as measured relative to inertial reference frame, equals the sum of all forces

acting on it. i.e.,

$$\frac{Da\vec{U}_a}{Dt} = \sum \vec{F} \tag{2.1.1}$$

The right hand side of Eq. (2.1.1) is the sum of all forces acting on the air parcel per unit mass. While, the left hand side of the equation is the rate of change of the absolute velocity following the motion.

For most applications in meteorology it is desirable to refer the motion to a reference frame rotating with the earth. So, we need to transform the momentum equation to a rotating coordinate system. The total derivative of any arbitrary vector ($\vec{A}$) in an inertial reference frame is related to the corresponding total derivative in a rotating system as

$$\frac{D_a\vec{A}}{Dt} = \frac{D\vec{A}}{Dt} + \vec{\Omega} X \vec{A}, \tag{2.1.2}$$

where $\vec{\Omega}$ is the velocity of the rotating system with respect to the fixed system.

If we apply Eq. (2.1.2) to the position vector $\vec{\Omega}$ for an air parcel on rotating Earth with constant angular velocity, we obtain

$$\frac{D_a\vec{R}}{Dt} = \frac{D\vec{R}}{Dt} + \vec{\Omega} X \vec{R} \tag{2.1.3}$$

But $D_a\vec{R}/Dt \equiv \vec{U}_a$ and $D\vec{R}/Dt \equiv \vec{U}$; therefore Eq. (2.1.3) may be written as

$$\vec{U}_a = \vec{U} + \vec{\Omega} X \vec{R} \tag{2.1.4}$$

Now we apply Eq. (2.1.2) to the velocity vector $\vec{U_a}$ and obtain

$$\frac{D_a\vec{U}_a}{Dt} = \frac{D\vec{U}_a}{Dt} + \vec{\Omega} X \vec{U}_a \tag{2.1.5}$$

Substituting from Eq. (2.1.4) into the right-hand side of Eq. (2.1.5) gives

$$\frac{D_a\vec{U}_a}{Dt} = \frac{D\left(\vec{U} + \vec{\Omega} X \vec{R}\right)}{Dt} + \vec{\Omega} X \left(\vec{U} + \vec{\Omega} X \vec{R}\right) = \frac{D\vec{U}}{Dt} + \vec{\Omega} X \frac{D\vec{R}}{Dt} + \vec{\Omega} X \vec{U} + \vec{\Omega} X \left(\vec{\Omega} X \vec{R}\right) \tag{2.1.6}$$

With the aid of a vector identity, $\vec{\Omega}X\left(\vec{\Omega}X\vec{R}\right)$ = -$\vec{\Omega}^2\vec{R}$, Eq. (2.1.6) can be written as

$$\frac{D_a\vec{U}_a}{Dt} = \frac{D\vec{U}}{Dt} + 2\vec{\Omega}X\vec{U} - \vec{\Omega}^2\vec{R}. \tag{2.1.7}$$

Here $\vec{R}$ is a vector perpendicular to the axis of rotation, with magnitude equal to the distance to the axis of rotation.

Eq. (2.1.7) states that the acceleration following the motion in an inertial system equals the rate of change of relative velocity following the relative motion in the rotating frame plus the Coriolis acceleration due to relative motion in the rotating frame plus the centripetal acceleration caused by the rotation of the coordinates.

For large scale atmospheric motion, the only real forces that are of primary concern are the pressure gradient force, the gravitational force, and friction. So we can rewrite Eq. (2.1.1) with the aid of Eq. (2.1.7) as

$$\frac{D\vec{U}}{Dt} = -2\vec{\Omega}X\vec{U} - \frac{1}{\rho}\nabla P + \vec{g} + \vec{F}_r \tag{2.1.8}$$

where $\frac{1}{\rho}\nabla P$ is the pressure gradient force acting on the surface which arises as a result of pressure difference, $\vec{g}$ is the effective gravity that includes the gravitational force and centrifugal force, and $\vec{F}_r$ is the frictional force because of the normal and tangential stress on the surface.

Eq. (2.1.8) is the statement of Newtons second law for motion relative to a rotating coordinate frame which is known as momentum equation. It states that the acceleration following the relative motion in the rotating frame equals the sum of the Coriolis force, the pressure gradient force, effective gravity, and friction.

2.1.2 Conservation of Mass

Atmospheric motion is also governed by the conservation of mass, which is expressed mathematically by the continuity equation.

Let us consider a volume element $\delta x\ \delta y\ \delta z$ that is fixed in a Cartesian coordinate frame as shown in Fig 2.1.

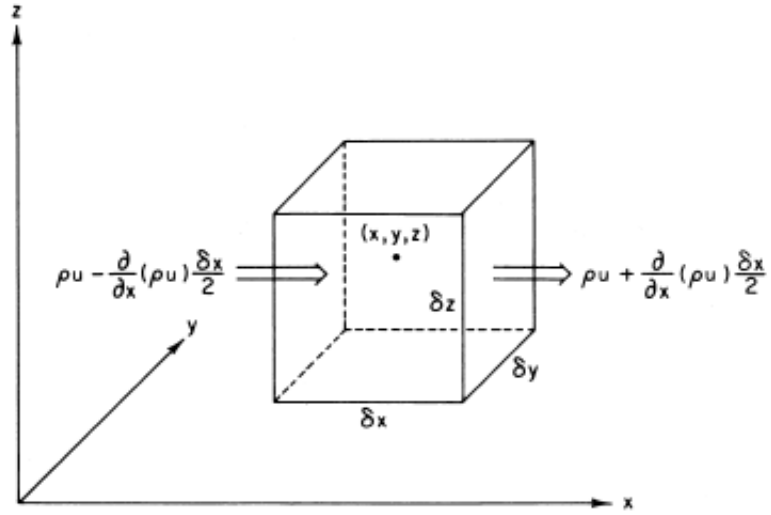


Figure 2.1: Mass inflow and mass out flow of a fixed volume element due to motion parallel to the x axis.

According to the law of conservation of mass, the net rate of mass inflow through the sides must equal the rate of accumulation of mass within the volume.

The rate of mass inflow per unit area through the left side is [ρu - $\frac{\partial(\rho u)}{\partial x}\ \frac{\delta x}{2}$], where u is velocity in the x-direction and ρ is density of air mass. And the rate of mass outflow per unit area through the right side is [ρu + $\frac{\partial(\rho u)}{\partial x}\ \frac{\delta x}{2}$].

Since the area of each of these faces is $\delta y\ \delta z$, the net rate of flow into the volume due to the x velocity component is

$$\left\{ \left[\rho u - \frac{\partial (\rho u)}{\partial x} \frac{\delta x}{2} \right] - \left[\rho u + \frac{\partial (\rho u)}{\partial x} \frac{\delta x}{2} \right] \right\} \delta y \delta z = - \frac{\partial (\rho u)}{\partial x} \delta x \delta y \delta z \tag{2.1.9}$$

Including y and z components and dividing by the volume $[\delta x \delta y \delta z]$ gives the rate of change of density

$$\frac{\partial \rho}{\partial t} = - \left[\frac{\partial (\rho u)}{\partial x} + \frac{\partial (\rho v)}{\partial y} + \frac{\partial (\rho w)}{\partial z} \right] \tag{2.1.10}$$

Therefore,

$$\frac{\partial \rho}{\partial t} = - \nabla \cdot (\rho \vec{U}). \tag{2.1.11}$$

This equation is the mass derivative form of the continuity equation, and states that the local rate of change of density is equal to minus the mass divergence.

2.1.3 Conservation of Energy

The third fundamental principle of atmospheric dynamics is the conservation of energy which can be described in terms of the first law of thermodynamics. The first law of thermodynamics states that "the change in internal energy of the system is equal to the difference between the heat added to the system and the work done by the system". i.e.,

$$dU = dQ - dW \tag{2.1.12}$$

The above equation relates the change in temperature of a parcel of air to energy transfer between the parcel of air and the environment and work done by or on a parcel.

If energy is added to the system, then some of it is used to change the internal energy (and temperature) and the rest is used by the parcel of air to do work. Energy is released to the air during condensation of water vapor, deposition of water vapor, freezing of liquid water. Energy is removed from the air during melting of ice, sublimation of ice and evaporation of liquid water (Jackobson, 2005).

The thermodynamic state of the atmosphere at any point is determined by the values of pressure, temperature, and density (or specific volume) at that point. These field variables are related to each other by the equation of state for an ideal gas.

A real gas is ideal only to the extent that intermolecular forces are small, which occurs when pressures are low enough or temperatures are high enough for the gas to be sufficiently dilute. Under typical atmospheric temperature and pressure conditions, the ideal gas law gives an error of less than 0.2% for dry air and water vapor in comparison with an expanded equation of state. Thus, the ideal gas law can reasonably approximate the equation of state.

The equation of state for dry air is expressed as

$$P\alpha = RT \quad or \quad P = \rho RT, \tag{2.1.13}$$

where R is the gas constant for dry air ($R = 287 Jkg^{-1}K^{-1}$) and $\alpha \equiv \rho^{-1}$ is specific volume.

The energy exchange equations in the atmosphere can be derived by combining the first law of thermodynamics with equation of state for an air parcel.

The first law of thermodynamics given in Eq. (2.1.12) can be rewritten as

$$dQ = C_v dT + Pd\alpha \tag{2.1.14}$$

Since d(Pα) = Pdα + αdP, we can rewrite the above equation as

$$dQ = C_v dT + d\left(P\alpha\right) - \alpha dP \tag{2.1.15}$$

From equation of state, d$(P\alpha) = RdT$ and using $C_p = C_v + R$, we obtain

$$dQ = \left(C_v + R\right) dT - \alpha dP = C_p dT - \alpha dP \tag{2.1.16}$$

By substituting $\alpha = \frac{1}{\rho}$, and rearranging gives the thermodynamic energy equation

$$dT = \frac{1}{C_p} dQ + \frac{1}{\rho C_p} dP \tag{2.1.17}$$

2.2 Conservation Laws as Implemented in RegCM4

The dynamical core of RegCM4 is based on the hydrostatic version of the fifth-generation Pennsylvania State University National Center for Atmospheric Research (PSUNCAR) Mesoscale Model (Grell *et al.*, 1994).

2.2.1 Horizontal and Vertical Grid

The modeling system usually gets and analyzes its data on pressure surfaces, but these have to be interpolated to the models vertical coordinate before input to the model. The vertical coordinate is terrain-following meaning that the lower grid levels follow the terrain while the upper surface is flatter (tend to approximate isobaric surfaces). Intermediate levels progressively flatten as the pressure decreases toward the top of the model.

A dimensionless σ coordinate is used to define the model levels

$$\sigma = \frac{(p - p_t)}{(p_s - p_t)}, \tag{2.2.1}$$

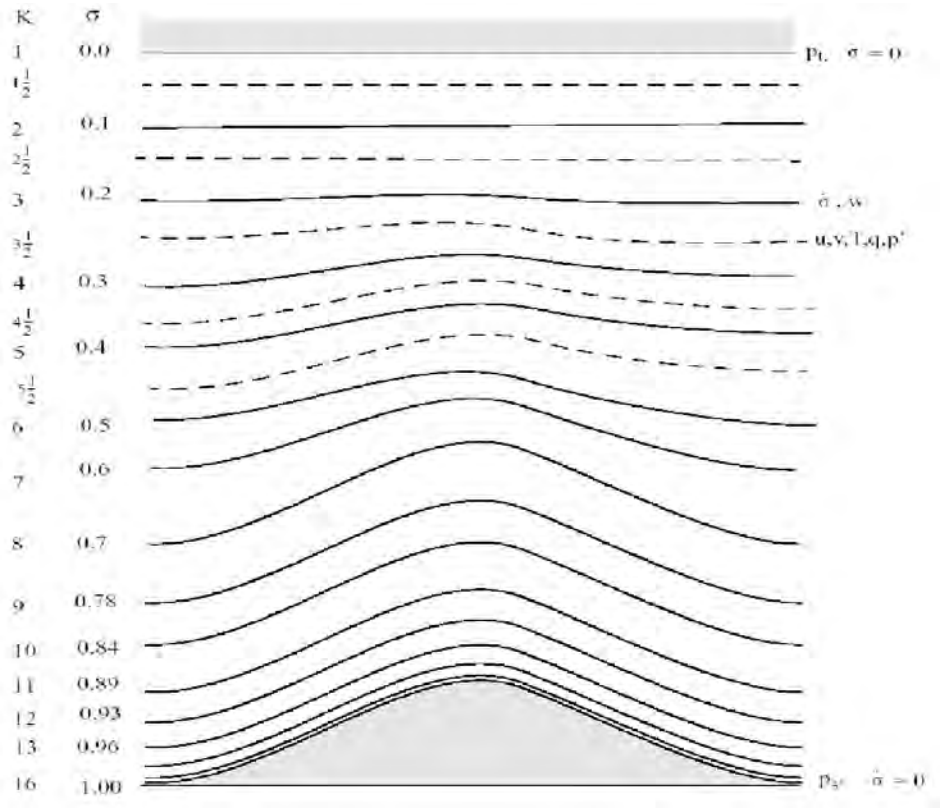


Figure 2.2: Schematic representation of the vertical structure of the model. This example is for 16 vertical layers. Dashed lines denote half-sigma levels, solid lines denote full-sigma levels.

where p is the pressure, p_t is a specified constant top pressure, p_s is the surface pressure.

It can be seen from Eq. (2.2.1) that σ is zero at the top and one at the surface, and each model level is defined by a list of σ values between zero and one that do not necessarily have to be evenly spaced.

The horizontal grid has an Arakawa-Lamb B-staggering of the velocity variables with respect to the scalar variables. The scalars (T, q, p, etc) are defined at the center of the grid box, while the eastward (u) and northward (v) velocity components are

collocated at the corners (Elguindi *et al.*, 2010).

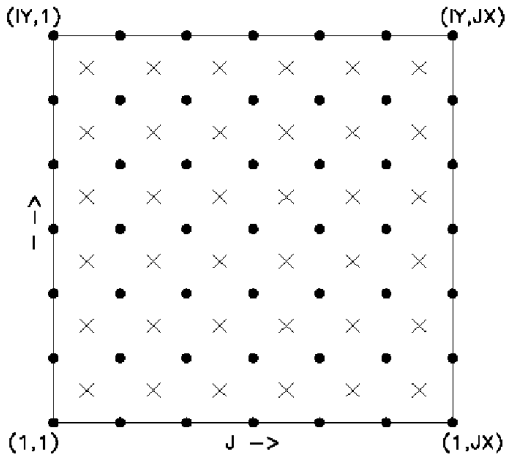


Figure 2.3: Schematic representation showing the horizontal Arakawa B-grid staggering of the dot and cross grid points. The center points of grid squares are cross points, and the corner points are dot points.

2.2.2 Map Projections and Map-scale Factors

The modeling system has a choice of four map projections. Lambert Conformal is suitable for mid-latitudes, Polar Stereographic for high latitudes, Normal Mercator for low latitudes, and Rotated Mercator for extra choice. The x and y directions in the model do not correspond to west-east and north-south except for the Normal Mercator projection, and therefore the observed wind generally has to be rotated to the model grid, and the model u and v components need to be rotated before comparison with observations.

The map scale factor is defined by

$$m = \frac{distance\ on\ grid}{actual\ distance\ on\ earth} \tag{2.2.2}$$

and its value is usually close to one, varying with latitude.

The projections in the model preserve the shape of small areas, so that dx=dy everywhere, but the grid length varies across the domain to allow a representation of a spherical surface on a plane surface (Elguindi *et al.*, 2010).

2.2.3 Horizontal Momentum Equations

The horizontal momentum equations in spherical coordinate system is given by

$$\frac{\partial p^* u}{\partial t} = -m^2 \left(\frac{\partial p^* uu/m}{\partial x} + \frac{\partial p^* vu/m}{\partial y} \right) - \frac{\partial p^* u\dot{\sigma}}{\partial \sigma} - mp^* \left[\frac{RT_v}{p^* + p_t/\sigma} \frac{\partial p^*}{\partial x} + \frac{\partial \phi}{\partial x} \right] + fp^* v + F_H u + F_V u \quad (2.2.3)$$

$$\frac{\partial p^* v}{\partial t} = -m^2 \left(\frac{\partial p^* uv/m}{\partial x} + \frac{\partial p^* vv/m}{\partial y} \right) - \frac{\partial p^* v\dot{\sigma}}{\partial \sigma} - mp^* \left[\frac{RT_v}{p^* + p_t/\sigma} \frac{\partial p^*}{\partial y} + \frac{\partial \phi}{\partial y} \right] + fp^* u + F_H v + F_V v \quad (2.2.4)$$

where u and v are the eastward and northward components of velocity, T_v is virtual temperature, ϕ is geopotential height, f is the Coriolis parameter, R is the gas constant for dry air, m is the map scale factor for either the Polar Stereographic, Lambert Conformal, or Mercator map projections, $\dot{\sigma} = \frac{d\sigma}{dt}$, and F_H and F_V represent the effects of horizontal and vertical diffusion, and $p^* = p_s - p_t$.

2.2.4 Continuity and Sigma dot ($\dot{\sigma}$) Equations

The continuity equation in spherical coordinate system is given by

$$\frac{\partial p^*}{\partial t} = -m^2 \left(\frac{\partial p^* u/m}{\partial x} + \frac{\partial p^* v/m}{\partial y} \right) - \frac{\partial p\dot{\sigma}}{\partial \sigma} \quad (2.2.5)$$

The vertical integral of Eq. (2.2.5) is used to compute the temporal variation of the surface pressure in the model

$$\frac{\partial p^*}{\partial t} = -m^2 \int_0^1 \left(\frac{\partial p^* u/m}{\partial x} + \frac{\partial p^* v/m}{\partial y} \right) d\sigma \tag{2.2.6}$$

After calculation of the surface-pressure tendency $\frac{\partial p^*}{\partial t}$, the vertical velocity in sigma coordinates ($\dot{\sigma}$) is computed at each level in the model from the vertical integral of Eq. (2.2.5)

$$\dot{\sigma} = -\frac{1}{p^*} \int_0^\sigma \left[\frac{\partial p^*}{\partial t} + m^2 \left(\frac{\partial p^* u/m}{\partial x} + \frac{\partial p^* v/m}{\partial y} \right) \right] d\sigma' \tag{2.2.7}$$

where σ' is a dummy variable of integration and $\dot{\sigma}(\sigma = 0) = 0$.

2.2.5 Thermodynamic Equation and Equation for Omega (ω)

The Thermodynamic equation in spherical coordinate system is given by

$$\frac{\partial p^* T}{\partial t} = -m^2 \left(\frac{\partial p^* uT/m}{\partial x} + \frac{\partial p^* vT/m}{\partial y} \right) - \frac{\partial p^* T\dot{\sigma}}{\partial \sigma} + \frac{RT_V \omega}{c_{pm}(\sigma + p_t/p^*} + \frac{p^* Q}{c_{pm}} + F_H T + F_V T \tag{2.2.8}$$

where c_{pm} is the specific heat capacity for moist air at constant pressure, Q is the diabatic heating, $F_H T$ represents the effect of horizontal diffusion, $F_V T$ represents the effect of vertical mixing and dry convective adjustment, and ω is

$$\omega = p^* \dot{\sigma} + \sigma \frac{dp^*}{dt} \tag{2.2.9}$$

where,

$$\frac{dp^*}{dt} = \frac{dp^*}{dt} + m \left(u \frac{\partial p^*}{\partial x} + v \frac{\partial p^*}{\partial y} \right) \tag{2.2.10}$$

The expression for $c_{pm} = c_p(1 + 0.8q_v)$, where c_p is the specific heat at constant pressure for dry air and q_v is the mixing ratio of water vapor.

Chapter 3

Model Description and Configuration

This chapter describes briefly the features of Regional Climate Model version 4 (RegCM4) and the model configuration of our simulations. In addition to a detailed description of RegCM4's schemes, overview of the input datasets and the study area are provided.

3.1 Model Description

The Regional Climate Model used in this thesis is the ICTP's Regional Climate Model version 4 (updated in June 2010). The model is a development from its previous versions with improvements in the representation of precipitation physics, surface physics, atmospheric chemistry and aerosols, model input fields and user interface (Pal *et al.*, 2007; Elguindi *et al.*, 2010).

As improvements of the previous version (RegCM3), RegCM4 is coupled with CLM3.5 to represent land surface, a diurnal ocean SST scheme is included, sea ice is introduced in the SST boundary and an improved simple aerosol scheme including dust, SO_4, black carbon and organic carbon is implemented.

3.1.1 Model Physics

Radiation Scheme

RegCM4 uses the radiation scheme of the National Center for Atmospheric Research (NCAR) Community Climate Model version 3 (CCM3) radiative transfer package (Kiehl *et al*, 1996).

This scheme breaks the solar spectrum into 18 spectral intervals from 0.2 to 5 μm, which allows for gaseous absorption by O_3, O_2, H_2O, and CO_2.

Consistent with the vertical structure of RegCM4 (Fig 2.2), each layer of the atmosphere is modeled as a homogeneous blend of all radiatively significant elements. For each layer, reflectivity and transmissivity are calculated, and subsequently combined together explicitly allowing scattering between layers.

Land Surface Model

BATS

Biosphere-Atmosphere Transfer Scheme (BATS) version 1e (Dickinson *et al*, 1993) is a surface package designed to describe the role of vegetation and interactive soil moisture in modifying the surface-atmosphere exchanges of momentum, energy, and water vapor. The model has a vegetation layer, a snow layer, a surface soil layer 10 cm thick, root zone layer 1-2 m thick, and a third deep soil layer 3 m thick.

The soil hydrology calculations include predictive equations for the water content of the soil layers. These equations account for precipitation, snow melt, canopy foliage drip, evapotranspiration, surface runoff, infiltration below the root zone, and diffusive exchange of water between soil layers. BATS has 20 vegetation types; soil textures

ranging from coarse (sand) to intermediate (loam), to fine (clay); and different soil colors (light to dark) for the soil albedo calculations.

BATS account for the sub grid variability of topography and land cover using a mosaic-type approach (Giorgi *et al.*, 2003). This approach adopts a regular fine-scale surface sub grid for each coarse model grid cell. Meteorological variables are disaggregated from the coarse grid to the fine grid based on the elevation differences. The BATS calculations are then performed separately for each sub grid cell, and surface fluxes are reaggregated on to the coarse grid cell for input to the atmospheric model.

CLM

The Community Land Model (Oleson *et al.*, 2008) is the land surface model developed by the National Center of Atmospheric Research (NCAR) as part of the Community Climate System Model (CCSM). CLM contains five possible snow layers with an additional representation of trace snow and ten unevenly spaced soil layers with explicit solutions of temperature, liquid water and ice water in each layer.

To account for land surface complexity within a climate model grid cell, CLM uses a mosaic approach to capture surface heterogeneity. Each CLM grid cell contains up to four different land cover types (glacier, wetland, lake, and vegetated), where the vegetated fraction can be further divided into 17 different plant functional types. Hydrological and energy balance equations are solved for each land cover type and aggregated back to the grid cell level.

Ocean Flux Parametrization

The air-sea energy fluxes from open bodies (e.g Ocean) are computed from prescribed SSTs with no two-way interaction. That is, the Ocean affects the Atmosphere, but

the Atmosphere does not affect the Ocean (Pal *et al.*, 2007).

In RegCM4, there are two parametrization options for computing fluxes from Ocean: the BATS formulation and the newly implemented Zeng scheme (Zeng *et al.*, 1998). BATS computes the fluxes with no special treatment of convective and very stable conditions. In addition, the roughness length is set to a constant; that is, it is not a function of wind and stability. The Zeng scheme describes all stability conditions and includes a gustiness velocity to account for the additional flux induced by boundary layer scale variability (Pal *et al.*, 2007).

Prognostic Sea Surface Skin Temperature Scheme

RegCM4 contains an option which allows for the computation of sea surface skin temperature as a prognostic variable. This allows for a realistic representation of the diurnal sea surface skin temperature, leading to improvements in the surface fluxes thus air-sea interactions. The scheme is based on a two-layer model which includes warm layer/cool skin effects as described by Fairall et al., 1996. Temperatures in the two layers are calculated using a one-dimensional heat transfer equation and boundary conditions determined by surface to atmosphere fluxes (latent, sensible and radiative) and a 3 m depth sea surface temperature taken from the prescribed SSTs.

Aerosols and Dust Chemistry Model

The representation of dust emission processes is a key element in a dust model and depends on the wind conditions, the soil characteristics and the particle size (Elguindi *et al.*, 2010).

The Chemistry Model accounts for affection by atmospheric winds, diffusion by turbulence, vertical transport by deep cumulus convection, dry and wet removal processes,

and gas and aqueous phase chemical conversion mechanisms. Wet removal both by resolvable scale and sub grid scale precipitation is parametrized as a function of the rate of conversion from cloud-water to rainwater and the hydrophilicity of the different aerosols. Surface dry deposition is calculated according to prescribed deposition velocities over land and water surfaces (Pal *et al.*, 2007).

Large-scale Precipitation Scheme

The formation of precipitation in RegCM4 is represented in two forms: resolvable (or large scale) and convective (or sub grid). The resolvable precipitation is generally associated with large-scale systems that move relatively slowly in the vertical direction, and is most common in the winter hemisphere.

In RegCM4, the resolvable-scale precipitation is represented using the sub grid explicit moisture (SUBEX) scheme (Pal *et al.*, 2000). SUBEX accounts for the sub grid-scale variability of clouds and includes formulations for the auto conversion of cloud water into rain water, the accretion of cloud droplets by falling raindrops, and the evaporation of falling raindrops.

The fraction of the grid cell covered by clouds, FC, at a given model level varies based on the average grid cell relative humidity RH and it is determined by

$$FC = \sqrt{\frac{RH - RH_{min}}{RH_{max} - RH_{min}}}, \tag{3.1.1}$$

where RH_{min} is the relative humidity threshold at which clouds begin to form, and RH_{max} is the relative humidity where FC reaches unity. FC is assumed to be zero when RH is less than RH_{min} and unity when RH is greater than RH_{max}.

Precipitation ,P, in a given model level is formed when the cloud water content exceeds the auto conversion threshold Q_c^{th} according to the relation

$$P = C_{ppt}\left(Q_c/FC - Q_c^{th}\right) FC, \tag{3.1.2}$$

where $1/C_{ppt}$ can be considered the characteristic time for which cloud droplets are converted to raindrops.

When precipitation is formed, it is assumed to fall instantaneously. The threshold is obtained by scaling the median cloud liquid water content equation according to:

$$Q_c^{th} = C_{acs} 10^{-0.49+0.013T}, \tag{3.1.3}$$

where T is temperature in degrees Celsius, and C_{acs} is the auto conversion scale factor (the amount of cloud water converted to rain) .

Convective Precipitation Schemes

Convective precipitation typically occurs in the summer hemisphere and tropics at scales finer than 1 km. Because of its fine scale, convective precipitation must be parametrized in most climate models and its parametrization still remains one of the most important sources of errors (Pal *et al.*, 2007).

RegCM4 has four options to compute convective precipitation. These options are the Modified-Kuo scheme (Anthes, 1977); Grell scheme (Grell, 1993), which can use either the Arakawa-Schubert (Grell-AS; (Arakawa and Schubert, 1974) or the Fritsch-Chappell (Grell-FC; (Fritsch and Chappell, 1980) closure assumptions; and MIT-Emanuel scheme (Emanuel, 1991); (Emanuel and Zivkovic-Rothman, 1999).

1. Kuo Scheme

In the Modified AnthesKuo scheme (Anthes, 1977), precipitation is initiated when the rate of moisture convergence in a column exceeds a given threshold and the column is convectively unstable. A fraction of the total moisture convergence precipitates, depending on the mean columnar relative humidity, while the remaining fraction is redistributed throughout the column in proportion to the dryness of the column. A schematic diagram of the moisture cycle in a column which contains convection is shown in Fig 3.1.

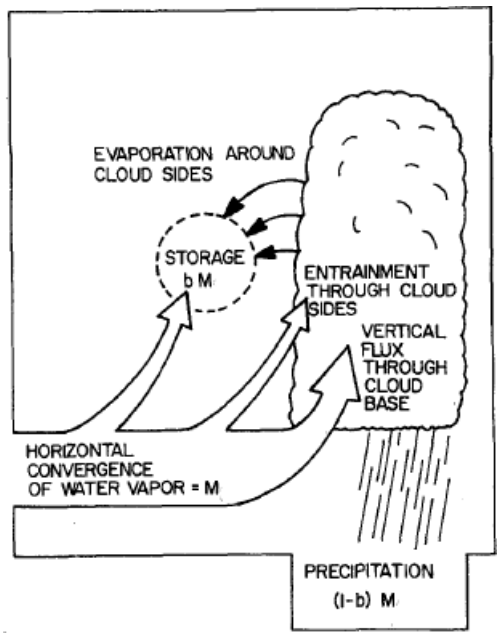


Figure 3.1: Schematic diagram of the Kuo scheme (Anthes, 1977).

A fraction of the moisture convergence converted into rainfall, P^{CU}, is given by

$$P^{CU} = M(1-\beta), \tag{3.1.4}$$

where M is moisture convergence and β is the fraction of moisture convergence which moistens the column.

β is a function of the average relative humidity $\overline{RH}$ of the column as

$$\beta = \begin{cases} 2(1-\overline{RH}) & \overline{RH} \geq 0.5 \\ 1.0 & \text{otherwise} \end{cases} \tag{3.1.5}$$

The moisture convergence term includes only the advective tendencies for water vapor. However, evapotranspiration from the previous time step is indirectly included in M since it tends to moisten the lower atmosphere. Hence, as the evapotranspiration increases, more and more of it is converted into rainfall assuming the column is unstable.

2. Grell Scheme

In the Grell scheme (Grell, 1993), convection is represented by an updraft and downdraft pair in steady state circulations with no direct mixing between the environment and convective clouds except at the top and bottom of the circulations (Fig 3.2). The mass flux in the updraft and downdraft is assumed constant with height and no entrainment or detrainment occurs along the cloud edges. The originating levels of the updraft and downdraft are given by the levels of maximum and minimum moist static energy, respectively. The scheme is activated when a lifted parcel attains moist convection.

The downdraft mass flux (m_0) depends on the updraft mass flux (m_b) according to

$$m_0 = \frac{\beta I_1}{I_2} m_b, \tag{3.1.6}$$

where I_1 is the normalized updraft condensation, I_2 is the normalized downdraft evaporation, and β is the fraction of updraft condensation that re-evaporates in the

downdraft. (1-β) is the precipitation efficiency. β depends on the wind shear and typically varies between 0.3 and 0.5.

Rainfall is given by

$$P^{CU} = I_1 m_b (1 - \beta). \tag{3.1.7}$$

Heating and moistening in the Grell scheme are determined both by the mass fluxes and the detrainment at the cloud top and bottom. In addition, the cooling effect of moist downdrafts is included.

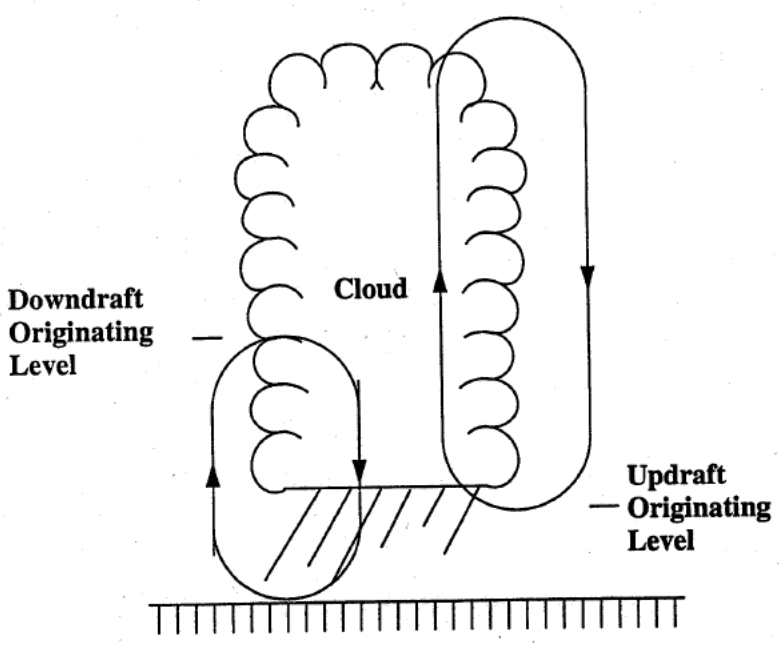


Figure 3.2: Conceptual picture of the Grell scheme (Grell *et al.*, 1994).

Due to the simplistic nature of the Grell scheme, several closure assumptions can be used to relate the mass flux at the bottom of the updraft to the large-scale forcing. RegCM4 implements the environmental stability closure assumptions of Arakawa and Schubert (1974) and Fritsch and Chappell (1980).

(i) Arakawa and Schubert Closure Assumption

It assumes that convective clouds stabilize the environment as fast as non-convective processes destabilize it as

$$m_b = \frac{ABE'' - ABE}{NA\Delta t}, \tag{3.1.8}$$

where ABE is the buoyant energy available for convection, ABE'' is the amount of buoyant energy available for convection in addition to the buoyant energy generated by some of the non-convective processes during the time interval Δt, and NA is the rate of change of ABE per unit m_b. The difference $ABE'' - ABE$ can be thought of as the rate of destabilization over time Δt. ABE'' is computed from the current fields plus the future tendencies resulting from the affection of heat and moisture and the dry adiabatic adjustment.

(ii) Fritsch and Chappell Closure Assumption

It assumes that convection removes the ABE over a given time scale as follows:

$$m_b = \frac{ABE}{NA\tau}, \tag{3.1.9}$$

where τ is the ABE removal time scale.

3. MIT-Emanuel Scheme

The Emanuel scheme (Emanuel, 1991; Emanuel and Zivkovic-Rothman, 1999) assumes that the mixing in clouds is highly episodic and inhomogeneous and attempts to reproduce the inhomogeneity of convective clouds by considering convective fluxes on the basis of an idealized model of subcloud-scale entities (updrafts and downdrafts), mixing, and buoyancy sorting.

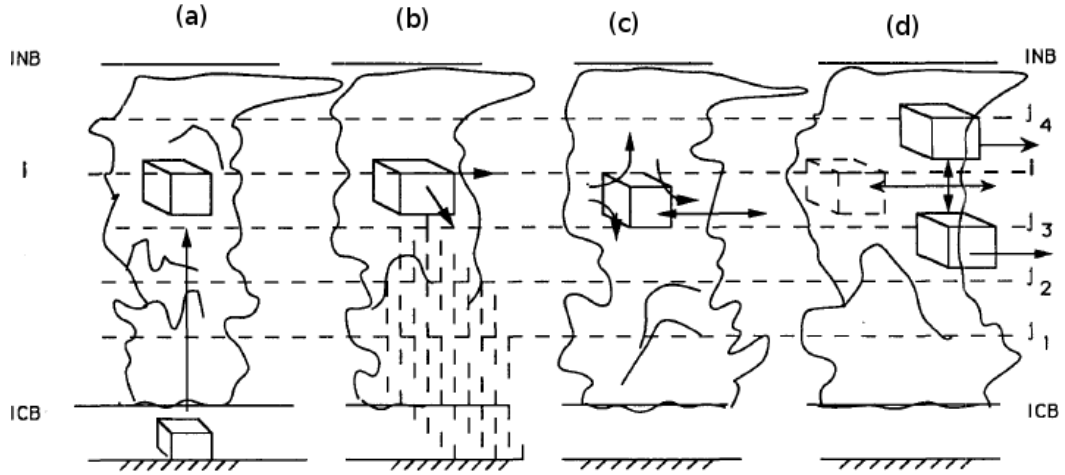


Figure 3.3: Idealized model of the convection of subcloud-scale parcles in the Emanuel scheme. (a) Reversible ascent from subcloud layer to arbitrary level i between cloud base (LCB) and level of neutral buoyancy (LNB). (b) A fraction of condenced water is converted to precipitation, which is added to a single unsaturated downdraft. (c) Remaining cloudy air is mixed according to an equal probability distribution with the environment at level i. (d) Mixtures then ascend or descend to levels at which their liquid water potential temprature is equal to that of the environment (Emanuel, 1991).

Convection is mainly driven by buoyancy and is triggered when the level of neutral buoyancy is greater than the cloud base level. Between these two levels, air is lifted and fraction of the condensed moisture forms precipitation, while the remaining fraction moistens the environment and forms cloud. The cloud is assumed to mix with the air from the environment according to a uniform spectrum of mixtures that ascend or descend to their respective levels of neutral buoyancy.

The fraction of the total cloud base mass flux that mixes with its environment at each level is proportional to the undiluted buoyancy rate of change with altitude. The cloud base upward mass flux is relaxed towards the sub-cloud layer quasi equilibrium. Idealization of convective process is illustrated in Fig 3.3.

3.1.2 Initial and Boundary Conditions

RegCM4 requires initial conditions and time-dependent lateral boundary conditions for the wind components, temperature, surface pressure, and water vapor. In addition, SSTs must be specified over oceans.

Several global reanalysis products and GCMs have provided boundary conditions to RegCM. Currently a total of four datasets are available for use as boundary conditions: The European Center for Medium-Range Weather Forecasts (ECMWF) reanalysis (ERA-Interim) dataset, the ECMWF 40-year reanalysis (ERA40) dataset, the National Centers for Environmental Prediction (NCEP) reanalysis 1 dataset (NNRP1), and the NCEP reanalysis 2 dataset (NNRP1).

Two land surface datasets are used to initialize RegCM: land cover and Elevation. The Global Land Cover Characterization (GLCC) dataset for the vegetation/landuse data is derived from 1 km Advanced Very High Resolution Radiometer (AVHRR) data, and is based on the vegetation/land cover types defined by BATS (Biosphere Atmosphere Transfer Scheme). The elevation data used is Global 30 arc second elevation dataset (GTOPO30). Both the landuse and elevation data are provided by the U.S. Geological Survey (USGS) Earth Resources Observation System Data Center and can be used at 60, 30, 10, 5, 3, and 2 minute resolutions.

Since RegCM lacks an interactive ocean model, SSTs are prescribed using the National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) dataset or the Hadley Center Meteorological Office Global Sea Surface Temperature (GISST) dataset.

3.2 Model Configuration

The selection of the model domain, its size and the resolution are important issue for regional climate modeling (Salzmann, 2006; Anyah, 2005). The model domain should be sufficiently small that the synoptic circulation does not depart from that of the driving GCM. On the other hand, the model domain should be large enough to avoid an increased role of the lateral boundary effect, which prevent the development of small-scale details of the local climate (Krichak, 2008). This means that the optimum domain size should be one where large-scale circulation in RCM is constrained to follow the driving GCM fields, but the finer mesoscale systems also have enough space to fully develop. The domain should also includes all areas where forcing and processes are dominant elements for the climate of the specific region.

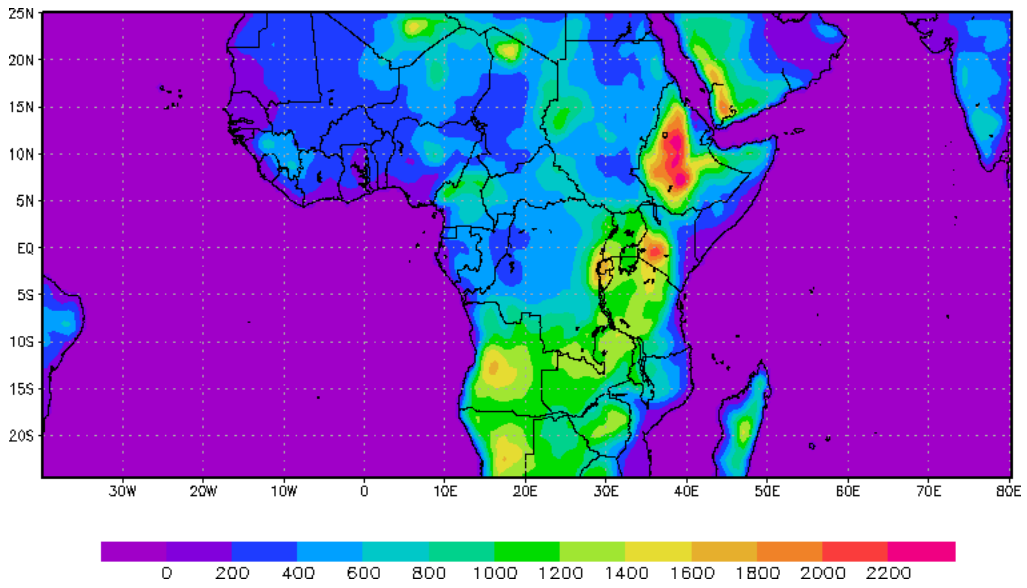


Figure 3.4: Model domain with surface elevation (m, scale at the bottom)

In this thesis, the model is configured with horizontal resolution of 60 km with its standard configuration of 18 vertical σ-levels and model top at 50 hpa. The domain

size and location is selected so that the main large scale forcing like ICTZ, the nearby subtropical high pressure systems and the main monsoon systems are included. It is centered at 0^0N, 20^0E and covers 96 grid points in the y-direction and 224 grid points in the x-direction, giving a total area of 5760 $km \times 13400$ km (Fig 3.4).

Observed sea surface temperature (SST) derived from the Optimum Interpolation sea surface temperature (OISST) one-degree gridded data from National ocean and Atmosphere Administration (NOAA) at weakly scale is used in the Simulation. Constant (in time) surface parameters (topography, land use, vegetation, soil type etc) are determined from a 10-min archive. Meteorological initial and time-evolving boundary conditions (for the wind components, temperature, surface pressure, and water vapor) are taken from global reanalysis (ERA-Interim) data developed by European Center for Medium range weather Forecast (ECMWF) and the lateral boundary conditions are relaxed exponentially.

Four model runs were performed, one for each convective precipitation Scheme, and the results were compared. 20 year runs was performed for each convective precipitation Scheme from 00:00 UTC 1 January 1989 to 31 December 2008 and calculations were performed with a 60 km space increment using Normal Mercator map projection and a time step of 150 seconds for Atmospheric model and 600 seconds for land Surface model.

3.3 Area of Study

The area of the study is the Horn of Africa. The Greater Horn of Africa (Fig 3.5) comprises 10 countries: Burundi, Djibouti, Eritrea, Ethiopia, Kenya, Rwanda, Somalia, Sudan, Uganda, and Tanzania (Bowden and Semazzi, 2007).

Rainfall is by far the most important climatic element over the Horn of Africa region since the local economies are largely sustained by rain-fed agriculture and agro-based manufacturing (Anyah and Semazzi, 2007).

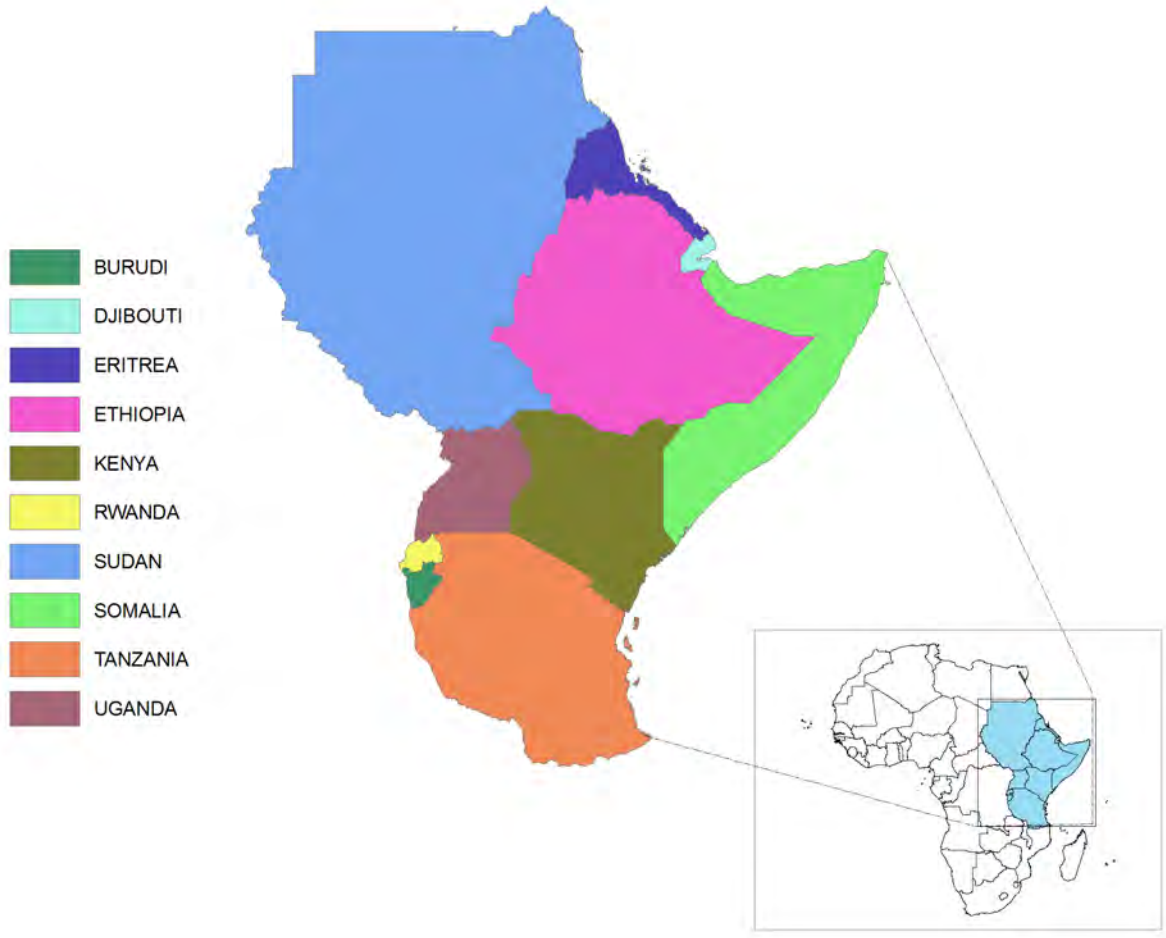


Figure 3.5: Study area (the Horn of Africa) showing geographical location and national boundaries.

The Horn of Africa has a distinct climate regime compared to the rest of the continent. The region is characterized by complex terrain and highly heterogeneous land surface. The Rift Valley system, with a chain of high mountains lined in north-south direction on both sides also hosts a family of large fresh water lakes that significantly modulate

regional climate variability. The inter-tropical convergence zone (ITCZ) also sweeps across the region twice a year and determines the seasonal cycle and climatological rainfall pattern (Anyah, 2005).

The seasonal cycle of the Horn of Africa rainfall is mainly controlled by the north-south migration of ITCZ across the region, which itself follows the interhemispheric migration of the overhead sun. The majority of the GHA region experiences a bi-modal rainfall pattern associated with the progression of the inter-tropical convergence zone (ITCZ) across the region (Anyah and Semazzi, 2007; Indeje *et al.*, 2000; Bowden and Semazzi, 2007).

The temporal pattern of rainfall over the Horn of Africa displays a strong inter-annual variability, punctuated with extreme events. The inter-annual variability is mainly associated with perturbations in the global SSTs, especially over the equatorial Pacific and India Ocean basins. SST perturbations over the tropical Atlantic also have important influence on the regional climate. In particular, El Niño-Southern Oscillation (ENSO) anomaly patterns are closely linked to the anomalous dry and wet seasons (Anyah, 2005, and references therein).

The complex topographical patterns, the existence of large lakes, variations in vegetation type and land-ocean contraction also give rise to high spatial and temporal variation in precipitation over the region. Other factors known to influence precipitation over the Horn of Africa include tropical storms, easterly waves, jet streams, the continental low-level trough, and extra-tropical weather systems (Indeje *et al.*, 2000, and reference therein) .

Chapter 4

Data and Methodology

This chapter gives overview of data sets used for validation and data analysis methods. After describing bias, RMSE and correlation coefficient, Clustering technique used in this study is briefly discussed.

4.1 Validation Data

In most cases, gridded data sets of observations are used for the validation of the present- day statistical mean climatology of a RCM (Salzmann, 2006). In this study, we used CMAP and GPCP data sets after regrided them, using bilinear interpolation, to the spatial resolution of our RegCM simulation.

4.1.1 CPC Merged Analysis of Precipitation (CMAP)

The Climate Prediction Center Merged Analysis of Precipitation (CMAP) global monthly precipitation have been constructed by merging several kinds of information sources including gauge observations, estimates inferred from 5 different satellite observations (GOES Precipitation Index (GPI), outgoing long wave radiation (OLR) based precipitation index (OPI), Microwave Sounding Unit (MSU), and microwave

(MW) scattering and emission from the Special Sensor Microwave Imager(SSM/I)), and the NCEP-NCAR reanalysis precipitation values (Xie and Arkin, 1998).

There are two merged precipitation data sets; The standard file that includes rain gauge and satellite estimates, and the enhanced file that includes blended NCEP/NCAR Reanalysis Precipitation values in addition to rain gauge and satellite estimates. In this study the version that combines only gauge and satellite estimates is used to avoid introducing any model bias associated with the data set while comparing it with our simulations.

CMAP data set has temporal coverage of monthly/pentad averaged precipitation rate values from 1979/01 through 2009/09 (to near present) and spatial coverage of 2.5 degree latitude $\times$ 2.5 degree longitude global grid (144$\times$72) in latitude and longitude range of $88.75^0N - 88.75^0S$, $1.25^0E - 358.75^0E$.

4.1.2 Global Precipitation Climatology Project Precipitation (GPCP)

The Global Precipitation Climatology Project(GPCP) version 2.1 data combines precipitation estimates from Special Sensor Microwave Imager (SSM/I) emission and scattering algorithms, Geostationary Satellite Precipitation Index (GPI), Outgoing Longwave precipitation Index (OPI), rain gauges, and TOVS sounders on NOAA polar orbiting satellites. The TOVS estimate inputs end in April 2005, and are replaced beginning in May 2005 with estimates from the Alliance Icing Research Study (AIRS) (Adler *et al*, 2003).

GPCP data set has temporal coverage of monthly values from January 1979 through the present (with some delay) and spatial coverage of 2.5 degree latitude $\times$ 2.5 degree

longitude global grid (144×72) in latitude and longitude range of $88.75^0N - 88.75^0S$, $1.25^0E - 358.75^0E$.

CMAP and GPCP Precipitation data sets are provided by the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA, from their Web site at http://www.esrl.noaa.gov/psd.

4.2 Evaluation Methodology

4.2.1 Analysis of Bias, RMSE and Correlation

The performance of the four options of convective schemes in RegCM4 is evaluated quantitatively by analyzing the bias, root mean square error and correlation of the simulations relative to the observation datasets using the methods found in (Wang *et al*, 2003; Diro *et al.*, 2008; Segele *et al.*, 2008).

The primary measure of simulation skill is the model bias (a time average of the error), defined as

$$bias = \frac{1}{N}\left[\sum_i \sum_j \left(a_{i,j}^M - a_{i,j}^O\right)\right], \tag{4.2.1}$$

where N is the total number of grid points within a given region; subscripts i, j are the horizontal grid point indices in the zonal and meridional directions, respectively; a can be any meteorological parameters either daily mean or monthly mean; superscripts O and M refer to the observed and model simulated quantities, respectively.

The root mean square error (RMSE) is defined as

$$RMSE = \sqrt{\frac{1}{N}\left[\sum_i \sum_j \left(a_{i,j}^M - a_{i,j}^O\right)^2\right]} \tag{4.2.2}$$

A correlation coefficient measures the strength and direction of a linear association between two variables. The correlation coefficient between simulated and observed quantity a is defined as

$$corrcoef = \frac{\sum_i \sum_j \left(a_{i,j}^M - \overline{a_{i,j}^M}\right)\left(a_{i,j}^O - \overline{a_{i,j}^O}\right)}{\sqrt{\sum_i \sum_j \left(a_{i,j}^M - \overline{a_{i,j}^M}\right)^2 \sum_i \sum_j \left(a_{i,j}^O - \overline{a_{i,j}^O}\right)^2}}, \tag{4.2.3}$$

where the over bar denotes spatial mean.

4.2.2 Cluster Analysis

Clustering is the process of grouping the data into classes or clusters so that objects within a cluster have high similarity in comparison to another, but are very dissimilar to objects in other clusters (Zhijie Xu *et al*, 2005).

k-medoid is a classical partitioning technique of clustering that clusters the data set of n objects into k clusters. A medoid is an object that is selected from the data set representing a cluster. It can also be defined as object of a cluster, whose average dissimilarity to all the objects in the cluster is minimal i.e. it is a most centrally located object within the cluster.

k-medoid method selects k medoids to represent the k clusters. Clusters are then created by assigning each of the remaining objects to the nearest medoid. The most common k-medoid algorithm is the Partitioning Around Medoids (PAM) algorithm of Kaufman and Rousseeuw (1990).

Partitioning Around Medoids (PAM)

To find k clusters, PAM's approach is to determine a medoid. Once the medoids have been selected, each non-selected object is grouped with the medoid to which it

is the most similar. More precisely, if O_j is a non-selected object, and O_i is a selected medoid, we say that O_j belongs to the cluster represented by O_i, if $d(O_j, O_i) = min_{Oe} d(O_j, O_e)$, where the notation min_{Oe}, denotes the minimum over all medoids O_e, and the notation $d(O_a, O_b)$ denotes the dissimilarity or distance between objects O_a, and O_b. All the dissimilarity values are given as inputs to PAM. Finally, the quality of a clustering (i.e. the combined quality of the chosen medoids) is measured by the average dissimilarity between an object and the medoid of its cluster.

To find the k medoids, PAM begins with an arbitrary selection of k objects. Then in each step, a swap between a selected object O_i and a non-selected object O_h (which is diffrent from O_j) is made, as long as such a swap would result in an improvement of the quality of the clustering. To calculate the effect of such a swap between O_i and O_h, PAM computes the contribution to swap between O_i and O_h for all non-selected objects O_j.

The total swap contribution to swap between O_i and O_h is given by

$$T_{ih} = \sum_j C_{jih}. \tag{4.2.4}$$

where C_{jih} is the contribution to swap between O_i and O_h from all non-selected objectes O_j.

There are four possibilities to consider when calculating C_{jih}.

1. If O_j currently belongs to the cluster represented by O_i. Furthermore, if O_j be more similar to O_l than O_h, i.e. $d(O_j, O_h) > d(O_j, O_l)$, where O_l is the second most similar medoid to O_j. Then the contribution from O_j to the swap is

$$C_{jih} = d\left(O_j, O_l\right) - d\left(O_j, O_i\right). \tag{4.2.5}$$

Thus, if O_i is replaced by O_h as a medoid, O_j would belongs to the cluster represented by O_l.

2. If O_j currently belongs to the cluster represented by O_i. But this time, O_j is less similar to O_l than O_h i.e. $d(O_j, O_h) < d(O_j, O_l)$. Then, the contribution from O_j to the swap is

$$C_{jih} = d\left(O_j, O_h\right) - d\left(O_j, O_i\right). \tag{4.2.6}$$

Thus, if O_i is replaced by O_h, O_j would belong to the cluster represented by O_h.

3. If O_j currently belongs to a cluster other than the one represented by O_i. Let O_l be the representative object of that cluster. Furthermore, if O_j be more similar to O_l than O_h i.e. $d(O_j, O_h) > d(O_j, O_l)$. Then, the contribution from O_j to the swap is

$$C_{jih} = d\left(O_j, O_h\right) - d\left(O_j, O_l\right) = 0. \tag{4.2.7}$$

Thus, even if O_i is replaced by O_h, O_j would stay in the cluster represented by O_l.

4. If O_j currently belongs to the cluster represented by O_l. But this time O_j is less similar to O_l than O_h i.e. $d(O_j, O_h) < d(O_j, O_l)$. Then, the contribution from O_j to the swap is

$$C_{jih} = d\left(O_j, O_h\right) - d\left(O_j, O_l\right). \tag{4.2.8}$$

Thus, if O_i is replaced by O_h, O_j would jump to the cluster of O_h from that of O_l.

Algorithm PAM The PAM algorithm can be described as follows (Ng and Han, 1994; Kaufman and Rousseeuw, 1990):

1. Select k representative objects arbitrarily.
2. Compute T_{ih}, for all pairs of objects O_i, O_h where O_i is currently selected, and O_h is not.
3. Select the pair O_i, O_h which corresponds to $min_{oi,oh}$, T_{ih}. If the minimum T_{ih} is negative, replace O_i with O_h, and go back to Step (2).
4. Otherwise, for each non-selected object, find the most similar representative object.

Experimental results reprorted in (Kaufman and Rousseeuw, 1990) show that PAM works satisfactorily for small data sets (e.g. 100 objects in 5 clusters). But it is not efficient in dealing with medium and large datasets. In Steps (2) and (3), there are altogether k(n - k) pairs of O_i, O_h. For each pair, computing T_{ih} requires the examination of (n - k) non-selected objects. Thus, Steps (2) and (3) combined is of $O(k(n-k)^2)$, where k is the number of clusters, and n is the total number of objects. And this is the complexity of only one iteration. Thus, it is obvious that PAM becomes too costly for large values of n and k. This analysis motivates the development of CLARA.

Clustering Large Applications (CLARA)

Kaufman and Rousseeuw (1990) suggested the CLARA algorithm for tackling large applications. CLARA is a combination of a sampling procedure and the classical PAM algorithm (Ng and Han, 1994; Chih-Ping *et al*, 2000).

Instead of finding representative objects for the entire dataset, CLARA draws a sample of the data set, applies PAM on the sample, and finds the medoids of the sample. The point is that if the sample is drawn in a sufficiently random way, the medoids of the sample would approximate the medoids of the entire data set. To come up with better approximations, CLARA draws multiple samples and gives the best clustering as the output. Here, for accuracy, the quality of a clustering is measured based on the average dissimilarity of all objects in the entire dataset, and not only of those objects in the samples.

Algorithm CLARA The CLARA algorithm, briefly, is as follows (Ng and Han, 1994; Kaufman and Rousseeuw, 1990):

1. Draw a sample of 40 + 2k objects randomly from the entire dataset, and call Algorithm PAM to find k medoids (k number of clusters) of the sample.

2. For each object O_j in the entire dataset, determine which of the k medoids is the most similar to O_j.

3. Calculate the average dissimilarity of the clustering obtained in the previous step. If this value is less than the current minimum, use this value as the current minimum, and retain the k medoids found in Step (2) as the best set of medoids obtained so far.

4. Return to Step (1) five times to start the next iteration.

Complementary to PAM, CLARA performs satisfactorily for large datasets (e.g. 1000 objects in 10 clusters). Experiments reported in (Kaufman and Rousseeuw, 1990) indicate that 5 samples of size $40 + 2k$ give satisfactory results. The complexity of each iteration now becomes $O(k(40 + k)^2 + k(n - k))$.

We have used CLARA analysis and found 12 homogeneous regions based on 1989-2008 rainfall climatology from the CMAP dataset.

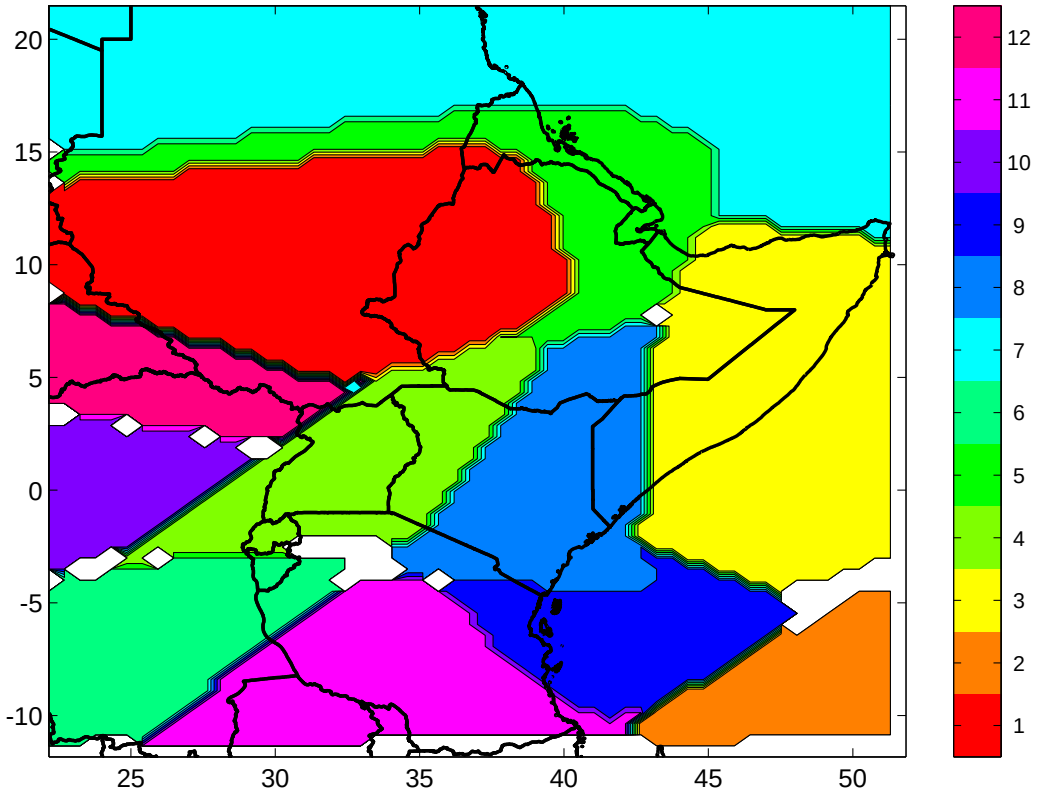


Figure 4.1: The 12 homogeneous rainfall regions over the Horn of Africa obtained from CLARA cluster analysis.

Chapter 5

Comparison of the Alternative Convective Schemes

This chapter presents the results of simulations of RegCM4 with the four different convective precipitation schemes. Quantitative evaluation of the precipitations obtained using the different convective schemes were carried out by comparison of bias, root mean square error(RMSE) and correlation coefficient with respect to both CMAP and GPCP data sets for individual grid squares.

5.1 Mean Climatology

Fig 5.1 shows long-term mean annual precipitation over the Horn of Africa averaged over a 20-year (1989-2008) period for the observed (CMAP and GPCP) and simulated using each of the convective schemes. The spatial distribution of the CMAP and GPCP rainfall climatology are in agreement over the Horn of Africa except at central Ethiopia, some parts of Kenya and Somalia, where GPCP shows relatively higher rainfall amount; and Tanzania (near Kenya boarder), where CMAP shows relatively higher rainfall amount. However, the difference is not significant.

Comparison of the simulated and observed 20-year mean precipitation climatology reveals that simulation using Grell-FC scheme exhibit better agreement with observation. However, it overestimate the amount of rainfall over some parts of Kenya and Ethiopia high lands and underestimate over Kenya and Somalia boarder. Over the remaining parts of the Horn of Africa, the simulated (using Grell-FC) precipitation amount and spatial distribution are fairly consistent with observation.

Emanuel scheme highly overestimates precipitation over central Ethiopia, Western Kenya and Uganda. It also overestimates over the coast of Somalia, Kenya and Tanzania. Where as, Gell-AS and Kuo schemes underestimate over most parts of Kenya, Somalia and Tanzania. In particular, Kuo scheme underestimates the mean rainfall over Uganda and Southern part of Sudan. Gell-AS and Kuo schemes are better in capturing the rainfall patterns over central Ethiopia.

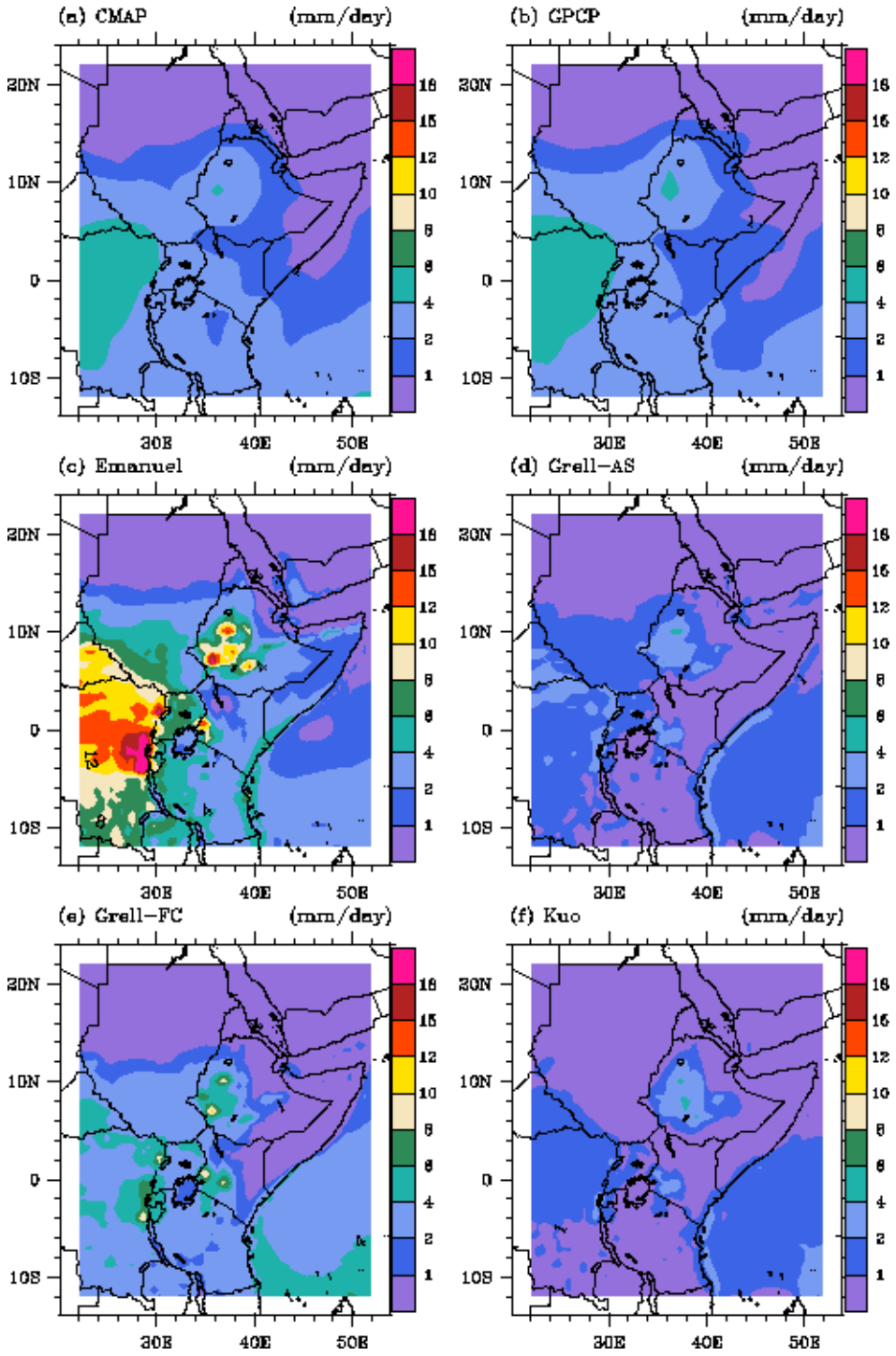


Figure 5.1: Long-term mean annual climatological Precipitation over the Horn of Africa (1989-2008).

5.2 Spatial Variability

The performance of each convective scheme over different homogeneous regions (Fig 4.1) is examined quantitatively by analyzing the bias, RMSE and correlation of the simulated rainfall relative to the observations (CMAP and GPCP).

Fig 5.2 shows the correlation coefficient between the simulated and observed precipitation (top panel) and the correlation significance level (bottom panel) of the 20 years temporal averages of the simulations with respect to CMAP (left panel) and GPCP (right panel).

There is high correlation between the precipitation by Emanuel scheme and observation for all of the clusters. There is also high correlation between the precipitation by Grell-FC scheme and observation for most of the clusters except clusters 3 and 7. (The correlation coefficient of Grell-AS with respect to CMAP and the correlation coefficients of Kuo and Grell-AS schemes with respect to GPCP for cluster 2 is excluded since the significance level of the correlation is bellow 50%). Emanuel scheme has the highest correlation coefficient for almost all of the clusters. Grell-FC and Emanuel schemes have the highest correlation coefficient for clusters 1, 6, 11 and 12 with respect to both CMAP and GPCP. Over Indian Ocean (cluster 2) and the coastal region of the Indian Ocean (clusters 3, 8 and 9) all schemes except Emanuel scheme shown relatively lower correlation which is in agreement with the suggestion by Davis *et al.* (2009).

Compared to Grell-AS and Kuo schemes, Grell-FC has the highest correlation coefficient except for clusters 5, 7 and 9. Grell-FC has also the same amount of correlation coefficient with Emanuel scheme for clusters 1, 2, 6 and 11.

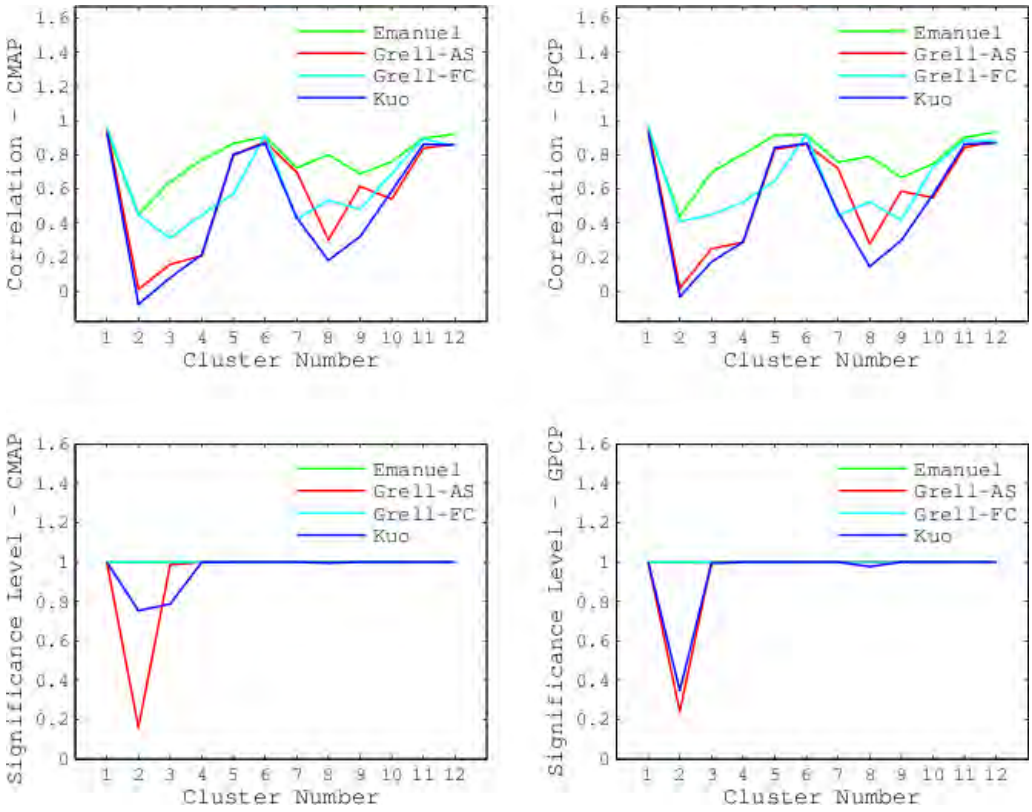


Figure 5.2: Correlation between Observation and RegCM4 simulated precipitation with different convective precipitation schemes over 12 homogeneous regions. Top-left: Correlation with CMAP; Top-right: correlation with GPCP; Bottom-left: significance level with CMAP; Bottom-right: significance level with GPCP.

Fig 5.3 shows the RMSE (top panel) and bias (bottom panel) of the 20 years Temporal averages of the simulations with respect to CMAP (left panel) and GPCP (right panel).

The RMSE is $0.1-0.9$ mm/day for all of the 12 clusters. Generally, the RMSE and bias of all convective schemes are large for cluster 12 even if they are larger for Emanuel convective scheme. With respect to both CMAP and GPCP, Grell-FC exhibits the lowest RMSE for clusters 6, 11 and 12 and Grell-AS has the lowest RMSE for clusters 4 and 9. Where as Grell-AS and Kuo have the lowest RMSE for cluster

8. All convective schemes perform the same for clusters 1, 3, 5 and 7. Generally, Grell-FC performs well for almost all of the clusters including for clusters in which other schemes have the lowest RMSE.

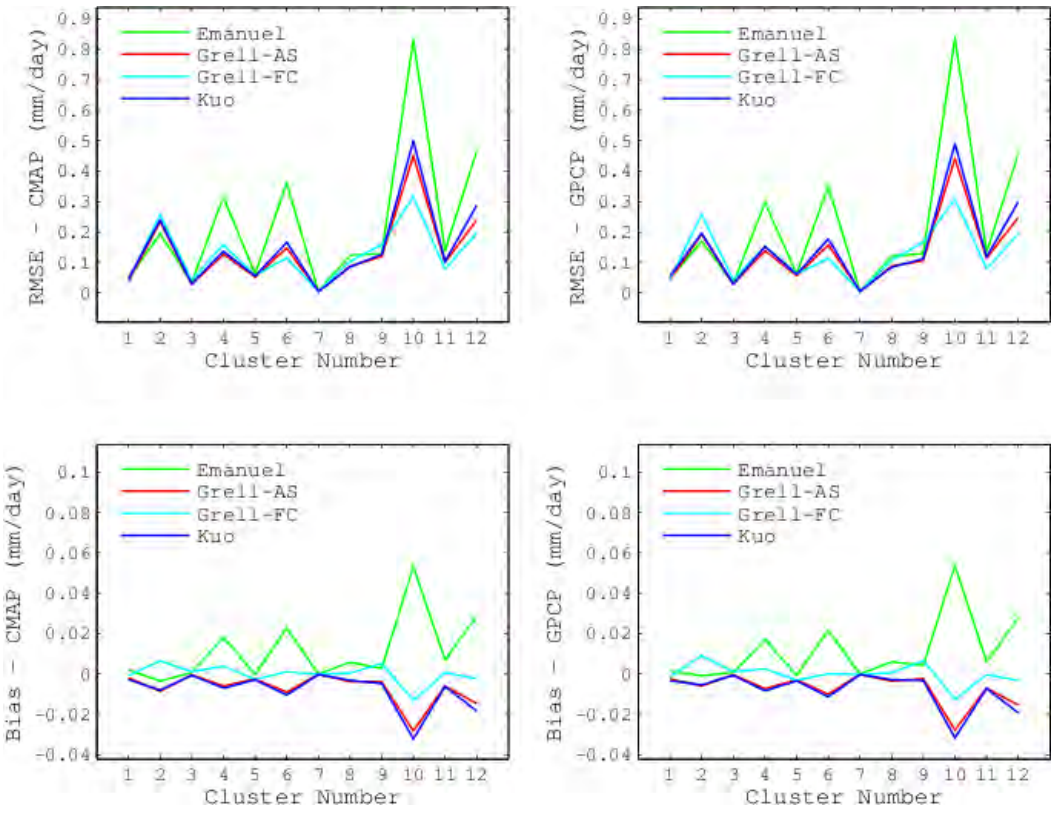


Figure 5.3: The error estimation of RegCM4 Simulated precipitation with different convective precipitation schemes in the 12 clustered regions. Top-left: RMSE with respect to CMAP; Top-right: RMSE with respect to GPCP; Bottom-left: bias with respect to CMAP; Bottom-right: bias with respect to GPCP.

The bias is mostly positive for Emanuel and Grell-FC convective schemes and mostly negative for Grell-AS and Kuo convective schemes except for clusters 3 and 7 where the bias of all schemes is nearly zero (see Fig 5.3 bottom panel). In terms of magnitude, the bias by Grell-FC is minimum for all of the cluster regions except cluster 2. For clusters 4, 6 and 12, we can observe strong overestimation by Emanuel scheme

and strong underestimation by Grell-AS and Kuo schemes. Grell-FC slightly overestimates for clusters 2, 4, 9 and 11. Grell-FC performs very well compared to the other schemes, having the lowest bias values for all clusters except cluster 2.

5.3 Annual Cycle

To show the performance of the four different convective precipitation schemes to capture annual cycle of East Africa Precipitation, monthly mean values were calculated for the period of 1989 to 2008 and the correlation and the error estimations are presented on Fig 5.4 and Fig 5.5.

Fig 5.4 presents the correlation coefficient between the simulated and observed precipitation (top panel) and the correlation significance level (bottom panel) of the 20 years temporal averages of the simulations with respect to CMAP (left panel) and GPCP (right panel).

The correlation coefficient shows high correlation between the simulated and observed precipitation for all of the months except December by Grell-AS convective scheme and from June to September by Kuo convective scheme. The highest correlation coefficient with respect to CMAP is achieved by Emanuel scheme except for months through January to March. Emanuel scheme also achieves the highest correlation coefficient with respect to GPCP expect for months January and February. In terms of correlation coefficient in general, Emanuel scheme performs well almost through out the entire year.

When compared to the Grell-AS and Kuo convective schemes, Grell-FC has the highest correlation coefficient except for months of February, April, June and September with respect to CMAP and GPCP. Even for those months, the correlation coefficient

is slightly less than that of the Grell-AS and Kuo convective schemes. The significance level is 100% (see Fig 5.4, bottom panel) through the entire year and for all of the convective schemes.

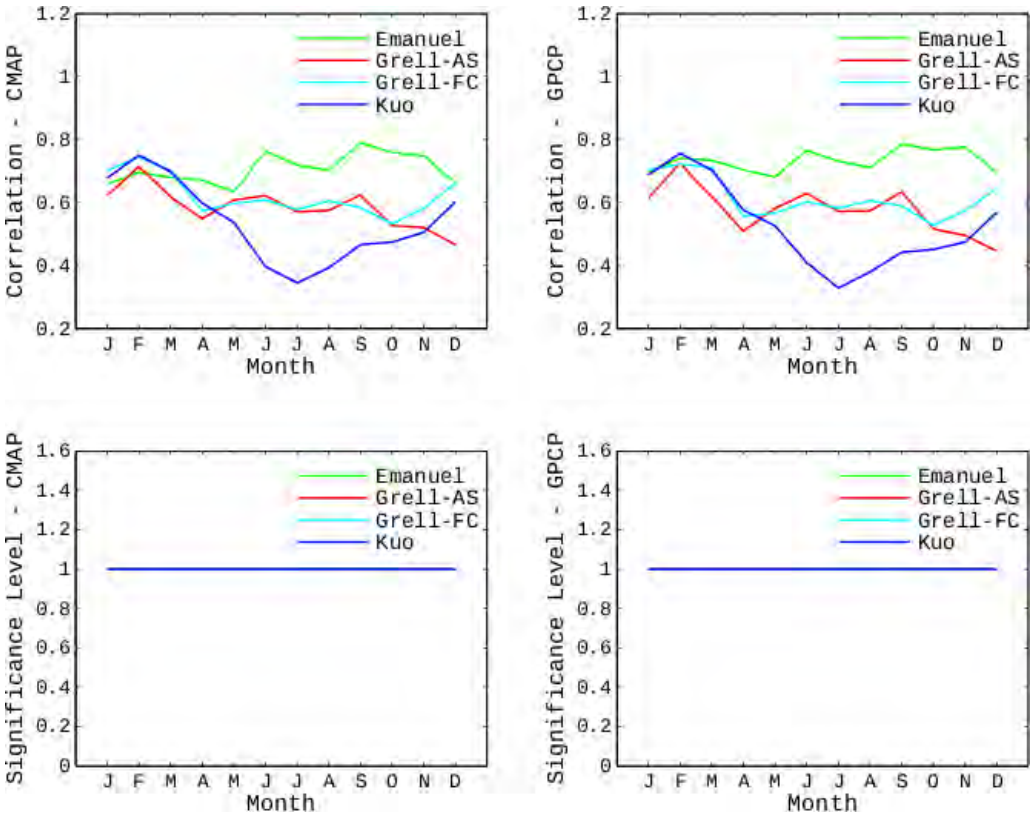


Figure 5.4: Monthly correlation between Observation and RegCM4 Simulated precipitation with different convective precipitation schemes. Top-left: Correlation with CMAP; Top-right: correlation with GPCP; Bottom-left: significance level with CMAP; Bottom-right: significance level with GPCP.

Fig 5.5 shows the RMSE (top panel) and bias (bottom panel) of the 20 years monthly averages of the simulations with respect to CMAP (left panel) and GPCP (right panel).

From the analysis of RMSE and bias we can tell that Emanuel scheme overestimates for all months. The overestimation is even worse for months through February to May

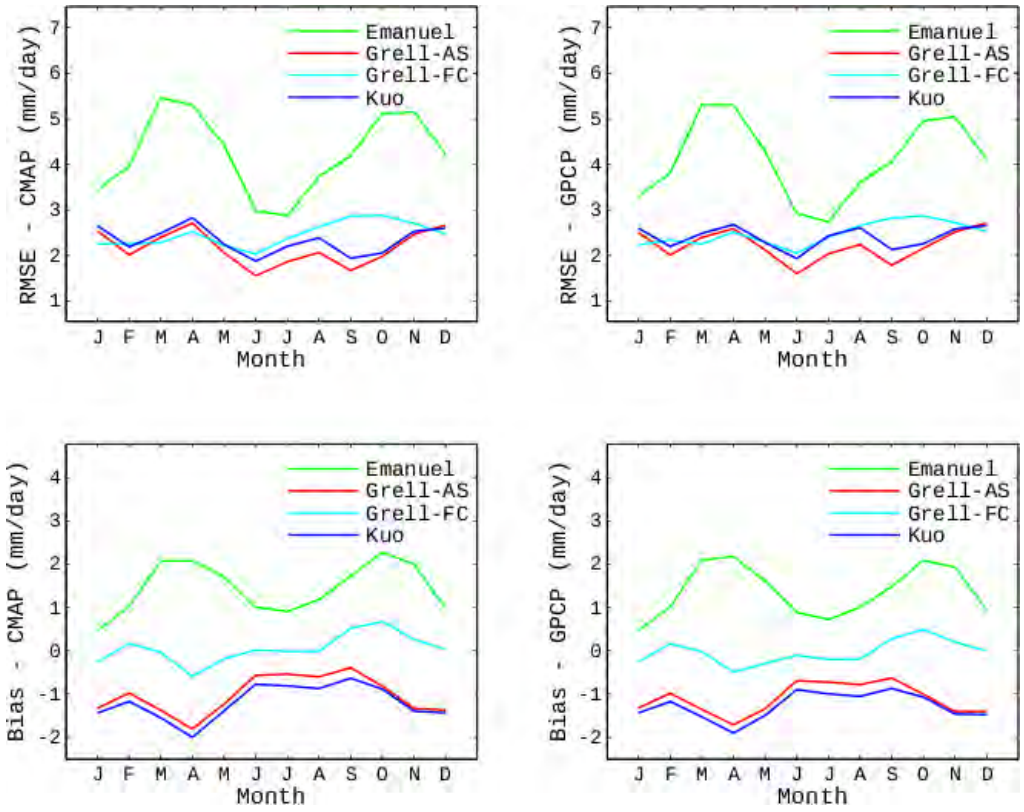


Figure 5.5: Monthly error estimation of RegCM4 Simulated precipitation with different convective precipitation schemes. Top-left: RMSE with respect to CMAP; Top-right: RMSE with respect to GPCP; Bottom-left: bias with respect to CMAP; Bottom-right: bias with respect to GPCP.

and through September to November. Even if Grell-AS has the lowest RMSE except for months January, March, April and December, it has negative bias (-1 in average) throughout the year. Grell-FC performs well as the bias is minimum in magnitude when compared to the other schemes throughout the entire year. Where as Grell-AS and Emanuel schemes give the largest negative and positive bias respectively.

5.4 Inter Annual Variability

The performance of each convective scheme to capture the inter annual variation of precipitation over different homogeneous regions (Fig 4.1) is examined quantitatively by comparing the time series of areal mean monthly precipitations obtained from the simulations using the convective schemes and observation data sets for the period of 1989 - 2008.

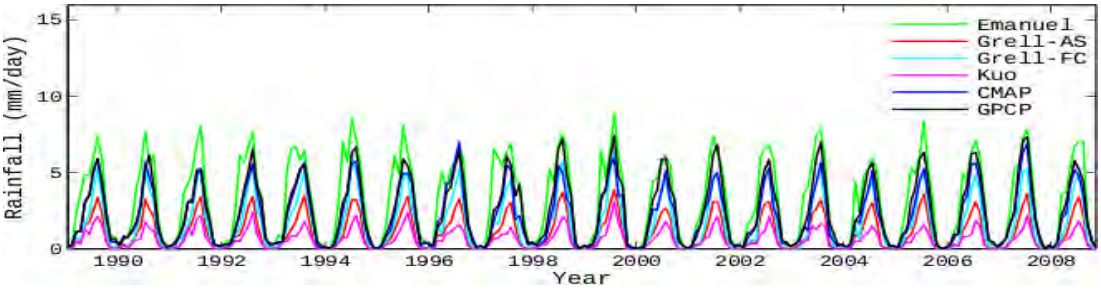
Figs 5.6 - 5.8 show the time series of monthly areal mean precipitation for both the simulations and the observation data sets for the 12 clustered regions.

From the plots in Figs 5.6 - 5.8, we can observe in general that, the magnitude of the areal mean precipitation by CMAP is larger than that of GPCP for clusters 1, 2, 3, 5, 7 and 9; and smaller than that of GPCP for clusters 4 and 6. They provide almost equal amount of precipitation for clusters 8, 11 and 12. There is better agreement in capturing the precipitation pattern for clusters 1, 4, 5, 6 and 8. The worst agreement is observed for cluster 7. We can also observe particular years in which the two data sets differ extremely both in magnitude and in pattern of precipitation.

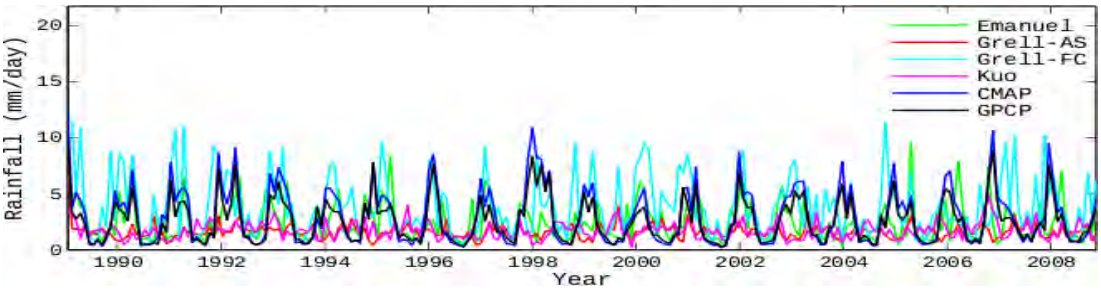
Emanuel convective scheme highly overestimates for all of the clusters except clusters 2 and 7. The overestimation is greater than 100% for cluster 4, 6 and 12. Grell-AS and Kuo convective schemes underestimate the areal mean precipitation for all clusters, even if they perform better for cluster 3 and 5. The underestimation is greater than 100% for clusters 6-12. Grell-FC convective scheme overestimates for clusters 3, 4 and 9; underestimates for cluster 7; and simulates well for the rest of the clusters. All of the convective schemes underestimate for cluster 7.

For clusters 1 and 8, Grell-AS and Grell-FC captured well the pattern of observed precipitation. Emanuel scheme and Kuo scheme in contrast have deficiencies in representing the observations. Grell-AS and Grell-FC also represent the precipitation on cluster 11 and cluster 12 respectively. Cluster 3 is better represented by Kuo convective scheme. Neither of the schemes capture the patterns of the observed precipitation for clusters 2, 4, 5 and 7. Where as, all of the schemes are good in representing the precipitation pattern for cluster 6.

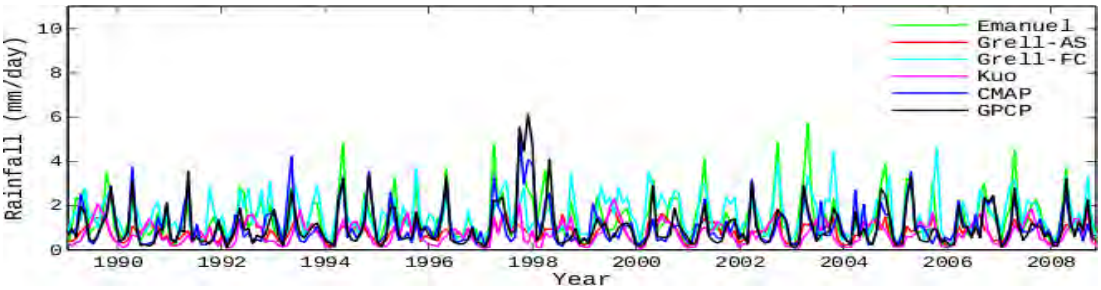
Generally, Kuo scheme is poor in representing the magnitude and pattern of inter annual variation of the observed precipitation. The other schemes, in contrast, have better performance in representing the observed inter annual variations (even though Grell-AS and Emanuel schemes are weak in representing the magnitude of the precipitation for most of the clusters). Grell-FC has the better performance with respect to both magnitude and pattern of the observed precipitations when compared to the other convective schemes. High inconsistency between CMAP and GPCP is observed on cluster 7 (see Fig 5.7(c)). All of the convective schemes also underestimate for this cluster.



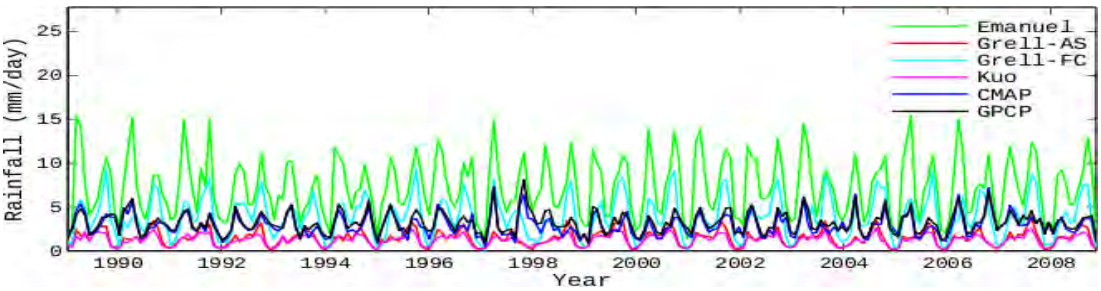
(a) Cluster 1



(b) Cluster 2

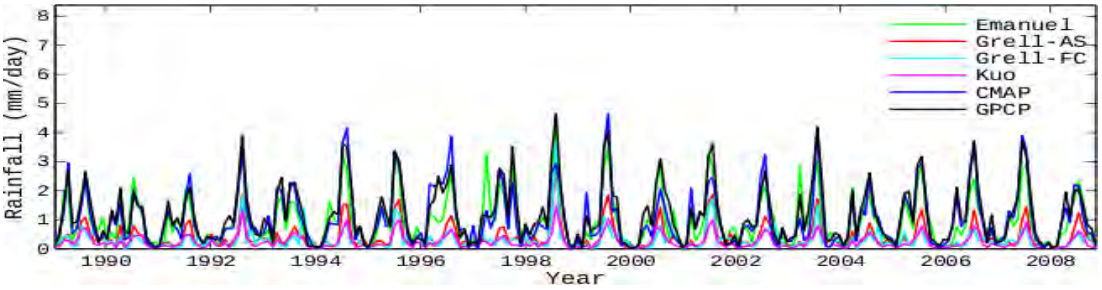


(c) Cluster 3

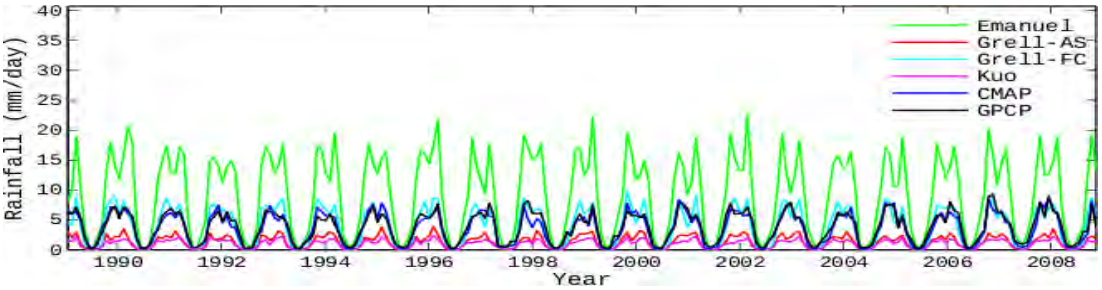


(d) Cluster 4

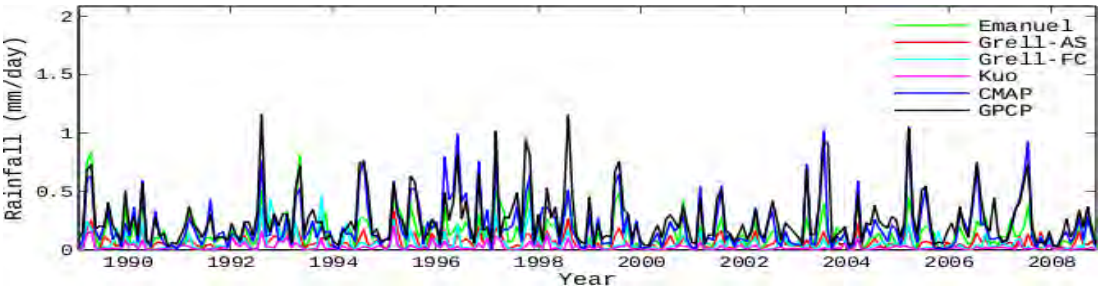
Figure 5.6: The time series of areal mean of simulated and observed monthly precipitations for clusters 1-4.



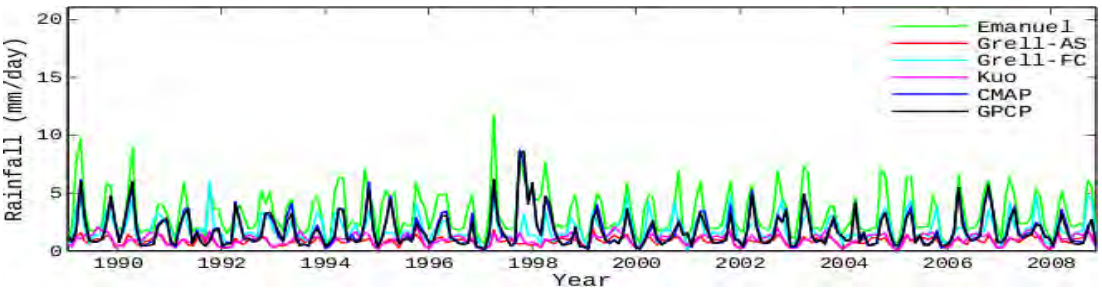
(a) Cluster 5



(b) Cluster 6

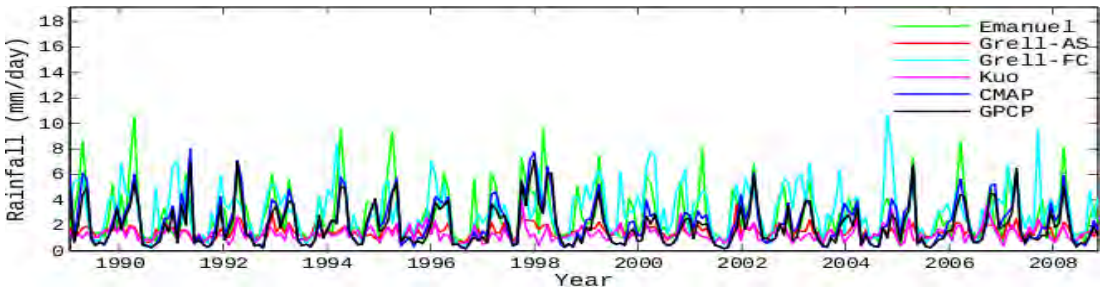


(c) Cluster 7

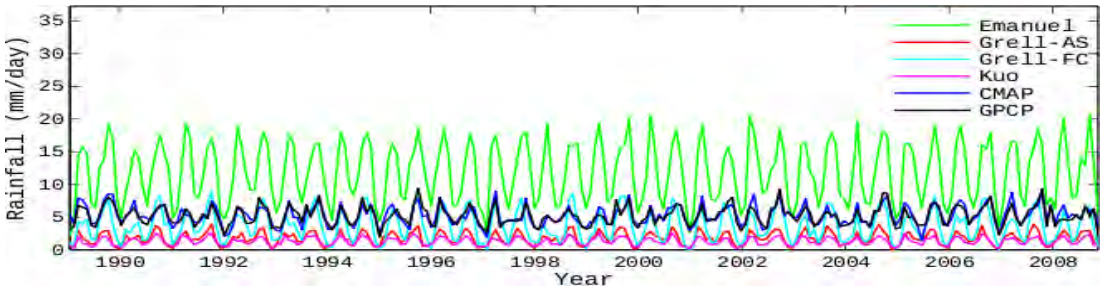


(d) Cluster 8

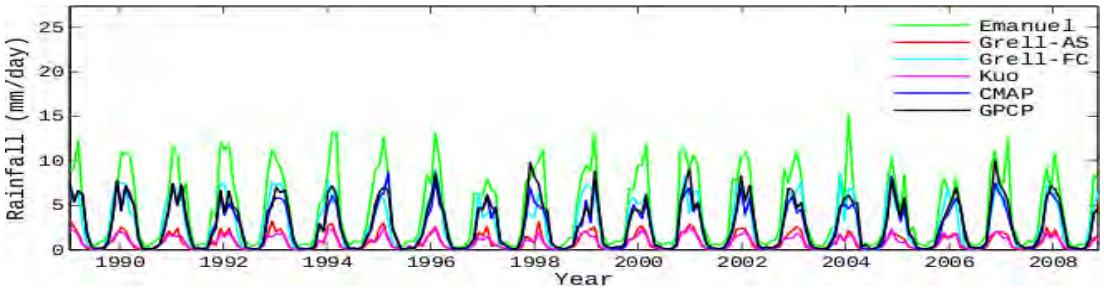
Figure 5.7: The time series of areal mean of simulated and observed monthly precipitations for clusters 5-8.



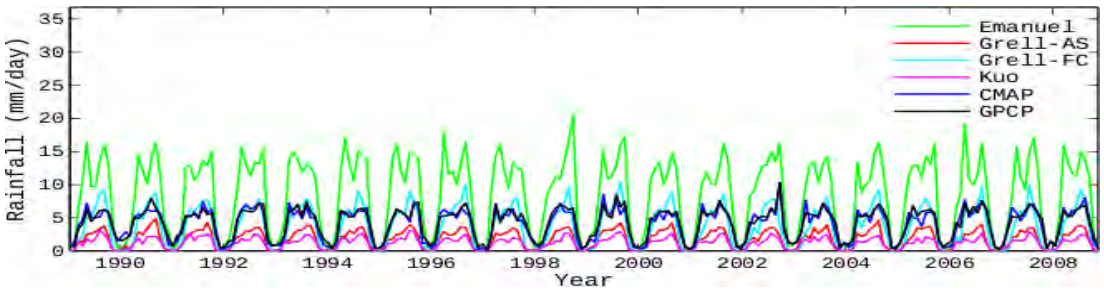
(a) Cluster 9



(b) Cluster 10



(c) Cluster 11



(d) Cluster 12

Figure 5.8: The time series of areal mean of simulated and observed monthly precipitations for clusters 9-12.

Chapter 6

Customization of RegCM4 to the Horn of Africa

This chapter presents the results of sensitivity experiments involving the adjustment of key parameters in Grell-FC scheme. After quantitative evaluation of the precipitations obtained using the different parameters, further analysis of the selected set of parameters will be discussed based on 1989-2005 simulation.

6.1 Sensitivity Experiments in Grell-FC Scheme

Based on the results of the previous chapter, Grell-FC convective precipitation scheme is chosen for further customization over the Horn of Africa.

6.1.1 New Set of Parameters in Grell-FC Scheme

In Grell-FC Scheme, the amount of simulated precipitation is affected by several, adjustable (by the user), parameters including:

1. Minimum and maximum Precipitation efficiency;

2. Minimum and maximum shear effect on precipitation efficiency;

3. Minimum and maximum convective heating; and

4. Available buoyant energy removal time scale.

The above key parameters were varied around the default values in four sets of experiments to improve the simulated precipitation over the Horn of Africa (Table 6.1). In Experiment I (E I), the minimum and maximum Precipitation efficiency (edtmin and edtmax) are varied separately and in combination; In Experiment II (E II), the minimum and maximum shear effect on Precipitation efficiency (shrmin and shrmax) are varied separately and in combination; for Experiment III (E III), the minimum and maximum convective heating (htmin and htmax) are varied separately and in combination; and in Experiment IV (E IV), available buoyant energy removal time scale (dtauc) is varied. In each experiment, all other parameters in the Grell-FC scheme are held to their default values.

Table 6.1: Values of key Grell-FC Scheme parameters used in sensitivity experiments. Default values for each parameter are included for comparison.

	Experiment I		Experiment II		Experiment III		Experiment IV
	edtmin	edtmax	shrmin	shrmax	htmin	htmax	dtauc
Default run	0.25	1.00	0.25	0.50	-250	500	30
Run 1	0.00	1.00	0.10	0.50	-150	500	35
Run 2	0.20	1.00	0.20	0.50	-200	500	40
Run 3	0.30	1.00	0.30	0.50	-300	500	45
Run 4	0.25	0.80	0.25	0.30	-250	400	-
Run 5	0.25	0.90	0.25	0.35	-250	450	-
Run 6	0.20	0.80	0.25	0.40	-250	550	-
Run 7	-	-	0.25	0.45	-150	400	-
Run 8	-	-	0.25	0.60	-	-	-
Run 9	-	-	0.10	0.30	-	-	-

After Analyzing the individual effect of each parameters in Grell-FC Scheme, further experiments have been performed to evaluate the combined effect of the parameters.

Table 6.2: Values of key Grell-FC Scheme parameters used in further sensitivity experiments.

	edtmin	edtmax	shrmin	shrmax	htmin	htmax
Default run	0.25	1.00	0.25	0.50	-250	500
Experiment V	0.20	0.80	0.25	0.30	-250	500
Experiment VI	0.20	0.80	0.10	0.30	-250	500
Experiment VII	0.25	1.00	0.25	0.30	-150	400
Experiment VIII	0.20	0.80	0.25	0.50	-150	400
Experiment IX	0.25	1.00	0.10	0.30	-150	400

6.1.2 Evaluation of the Sensitivity Experiments

The correlation coefficient, the RMSE and the bias of the simulated precipitation with respect to CMAP and GPCP are evaluated for the 12 homogeneous regions in Fig 4.1. In this section, the results of each experiment with respect to CMAP are presented. Similar results are found with respect to GPCP and the results with respect to both CMAP and GPCP are included in the Appendix for comparison.

Precipitation Efficiency

The Precipitation efficiency is defined as the ratio of rainout to water vapor inflow. We have already discussed in Section 3.1.1 that Rainfall in Grell scheme is given by

$P^{CU} = I_1 m_b(1-\beta)$, where (1-$\beta$) is the precipitation efficiency [Eq. (3.1.7)].

From Eq. (3.1.7) we can see that decreasing/increasing the precipitation efficiency will decrease/increase the rainfall directly.

Fig 6.1 shows the correlation coefficient, RMSE and bias of simulated precipitation, using different values of minimum and maximum precipitation efficiency, with respect to CMAP over the 12 homogeneous regions .

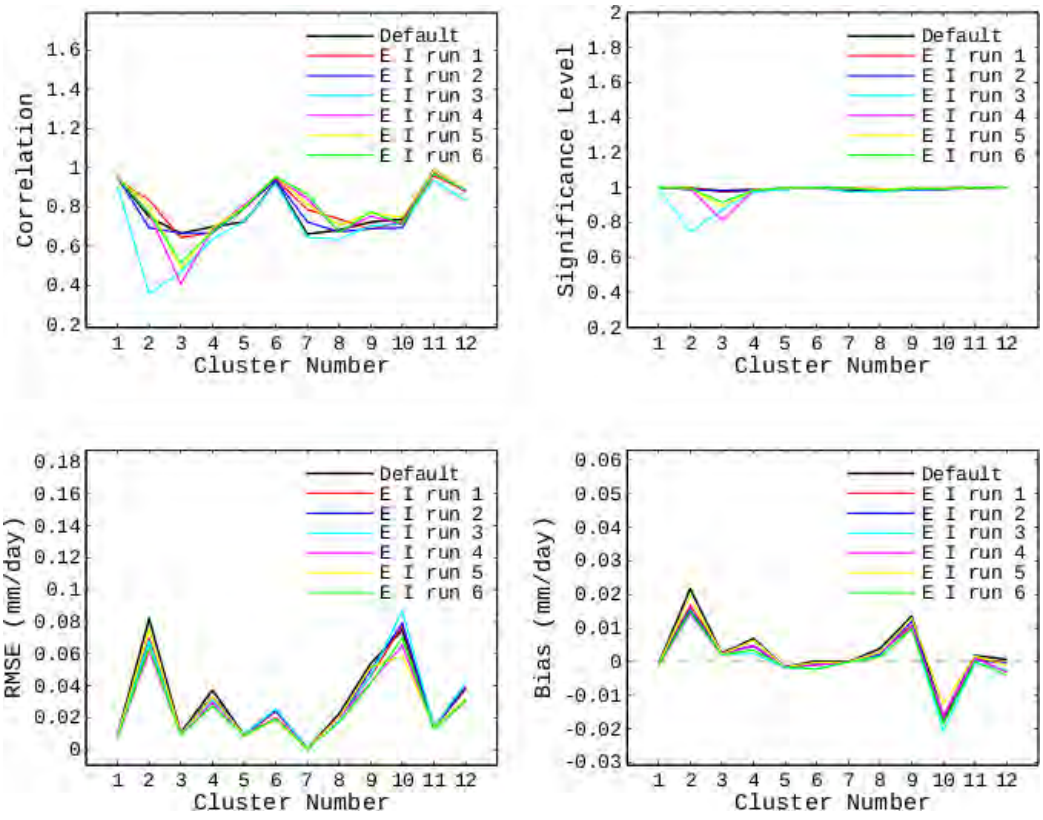


Figure 6.1: Correlation, RMSE and bias of simulated precipitation with different values of minimum and maximum precipitation efficiency over 12 homogeneous regions. Top-left: Correlation ; Top-right: significance level; Bottom-left: RMSE; Bottom-right: bias.

Changing the values of minimum and maximum precipitation efficiency alters the the rainfall amount and intrn the correlation coefficient, RMSE and bias over the 12 homogeneous regions.

Increasing the minimum precipitation efficiency to 0.30 (run 3) reduces the correlation coefficient over most of the clusters especially over cluster 2 (Indian Ocean). Whereas decreasing it to 0.00 (run 1) improve the correlation over clusters 2, 6, 7 and 8. Reducing the maximum precipitation to 0.80 (run 4) also increase the correlation over clusters 7 and 9 but decrease the correlation over cluster 3.

The RMSE also reduced over all of the clusters by decreasing the minimum and maximum precipitation efficiency to 0.20 and 0.80 respectively (run 6). The bias is not changed sufficiently in this experiment even though there is a little improvement by run 3 in reducing the positive bias over clusters 2 and 4.

Shear Effect on Precipitation Efficiency

At the ground, wind speeds are zero. In the surface layer, which is a 50-300m thick region of the atmosphere adjacent to the surface, wind speeds increase logarithmically with increasing height, creating wind shear. The change of wind speed with height (i.e., $\frac{\partial u}{\partial z}$) due to interactions among air molecules sliding over each other is wind shear (Jackobson, 2005).

Precipitation efficiency of the convective clouds is related to the vertical wind shear across the cloud depth and approximated by

$$E = 1.591 - 0.639\left(\frac{\Delta v}{\Delta z}\right) + 0.0953\left(\frac{\Delta v}{\Delta z}\right)^2 - 0.00496\left(\frac{\Delta v}{\Delta z}\right)^3,$$
$$and \tag{6.1.1}$$
$$E = 0.9 \qquad for \;\; \frac{\Delta v}{\Delta z} \leq 1.53.$$

where $\frac{\Delta v}{\Delta z}$ is the vertical shear of the horizontal wind evaluated over the cloud depth and has the units $10^{-3}s^{-1}$ (Fritsch and Chappell, 1980).

Fig 6.2 shows the correlation coefficient, RMSE and bias of RegCM4 simulated precipitation, using different values of minimum and maximum shear effect on precipitation efficiency, with respect to CMAP over the 12 homogeneous regions .

The correlation coefficient also altered by varying the minimum and maximum shear effect on precipitation efficiency.

Increasing the minimum shear effect to 0.30 (run 3) and the maximum shear effect to 0.60 (run 7) reduce the correlation almost over the whole region. On the other hand, decreasing the minimum shear effect to 0.10 (run 1) improve the correlation over clusters 2, 4 and 8; but reduce the correlation over cluster 9. Decreasing the maximum shear effect to 0.30 (run 4) also improve the correlation over clusters 5, 7, 8 and 9.

The combination of changing the minimum and maximum shear effect to 0.10 and 0.30 (run 9) also leads to improvement in correlation over most of the clusters except cluster 3 where the correlation is reduced a little bit.

There is small improvement over clusters 4 and 6 in reducing the positive bias by run 7 . The RMSE, in general, is not improved significantly. Note that cluster 10 is out side the Horn of Africa.

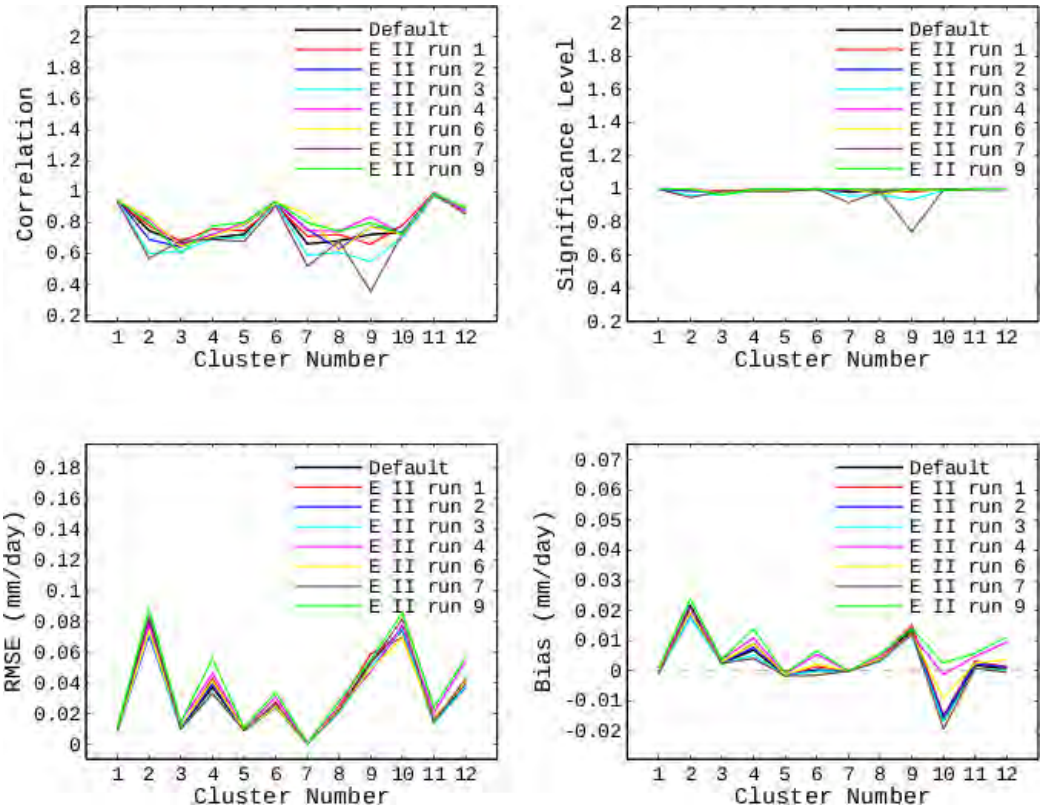


Figure 6.2: Correlation, RMSE and bias of simulated precipitation with different values of minimum and maximum shear effect on Precipitation efficiency over 12 homogeneous regions. Top-left: correlation ; Top-right: significance level; Bottom-left: RMSE; Bottom-right: bias.

Convective Heating

convection affects (modifies) the model-scale environment in two ways. First, when saturated air containing liquid water detrains (escapes) from cloud tops and evaporates, it cools the model-scale environment. Evaporation increases the water-vapor content in the environment. Second, cumulus convection, which occurs when clouds grow vertically, induces subsidence between clouds. During subsidence, the modelscale temperature increases and relative humidity decreases (Jackobson, 2005). In Grell scheme the convective cloud have a purely steady-state character. Consequently,

convection influences the environment through subsidence and detrainment at the top of the updraft and the downdraft (Grell, 1993).

In RegCM4 the convective scheme is called before the SUBEX scheme. This implies that some SUBEX precipitation could be generated by moisture moved aloft by the convection that moves moisture aloft but that the convective scheme itself does not rain out (Davis *et al.*, 2009). Grell (1993) also showed that for simulation without downdraft the intense heating and drying by this mass-flux-type scheme in the low level (through subsidence) hindered the explicit moisture scheme from becoming active.

Fig 6.3 shows the correlation coefficient, RMSE and bias of RegCM4 simulated precipitation, using different values of minimum and maximum convective heating, with respect to CMAP over the 12 homogeneous regions .

Varying the minimum and maximum convective heating affect satisfactorily the correlation coefficient, RMSE and bias. Increasing the minimum convective heating to -200 (run 2) reduces the correlation over most of the regions but further increasing it to -150 (run 1) shows improvement over clusters 2, 3 and 8 and poor correlation relative to the default over clusters 4, 11 and 12. Decreasing the minimum convective heating to -300 (run 3) also shows poor correlation relative to the default over most of the clusters but we can also observe improvement in correlation over clusters 5 and 7.

Decreasing the maximum convective heating to 400 (run 4) improves the correlation over clusters 1-7 and increasing it to 550 (run 6) improve the correlation over clusters 3, 4 and 7 and shows poor correlation over cluster 9. The combined effect of increasing the minimum convective heating to -150 and decreasing the maximum convective

heating to 400 (run 7) yields strong improvement in correlation over clusters 3 and 8; and poor correlation relative to the default over clusters 4, 9 and 12.

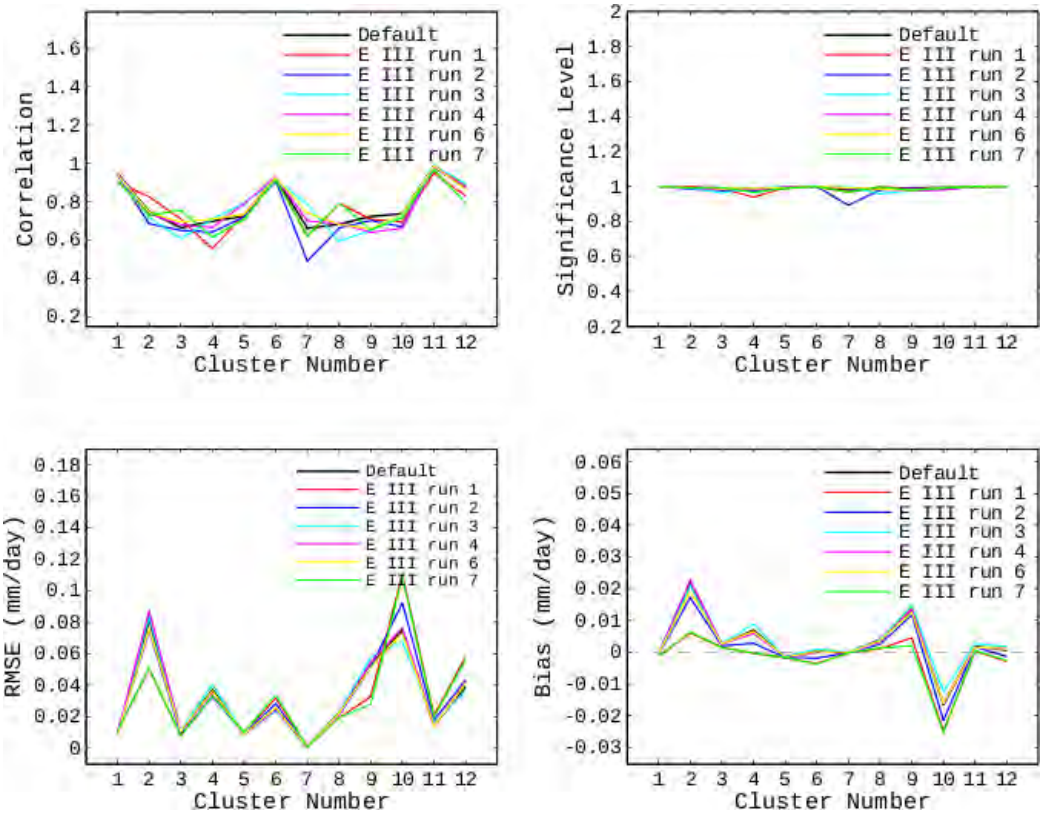


Figure 6.3: Correlation, RMSE and bias of simulated precipitation with different values of minimum and maximum convective heating over 12 homogeneous regions. Top-left: Correlation ; Top-right: significance level; Bottom-left: RMSE; Bottom-right: bias.

Setting the minimum and maximum convective heating to -150 and 400 respectively (run 7) reduces precipitation production. This leads to strong improvement by reducing the overestimation (positive bias) over all of the clusters except clusters 6 and 12, where small amount of negative bias by run 7 is observed as shown in Fig 6.3. The RMSE is also reduced by run 7 over all of the clusters except clusters 6 and 12. Run 1 also reduce the RMSE over cluster 9.

Available Buoyant Energy Removal Time Scale

In FC closure assumption the updrift (m_b) is given by

$m_b = \frac{ABE}{NA\tau}$, where τ is the ABE removal time scale [Eq. (3.1.9)].

The grid-scale stabilization rate (ABE removal) by moist convective processes is given by the ratio between ABE and the time for convective cells to move across the grid element. This time period τ is obtained by dividing the grid length by the mean environmental wind speed over the cloud depth; it has a lower limit defined by the average lifetime of individual cells (30 minute), and an upper limit (60 minute) to allow large-scale changes to alter the characteristics of the convective clouds (Fritsch and Chappell, 1980).

Fig 6.4 shows the correlation coefficient, RMSE and bias of RegCM4 simulated precipitation, using different values of available buoyant energy removal time scale, with respect to CMAP over the 12 homogeneous regions .

Increasing the available buoyant energy removal time scale to 35 (run1) improves the correlation over clusters 5, 7 and 9. Increasing it to 40 (run 2) and 45 (run 3) also improve the correlation over clusters 2, 5 and 7. Nevertheless, the correlation becomes worse over clusters 3 and 8.

The RMSE is reduced by run 2 and run 3 over the entire region. Run 2 and run 3 also reduced the positive bias over clusters 2, 9 and 12.

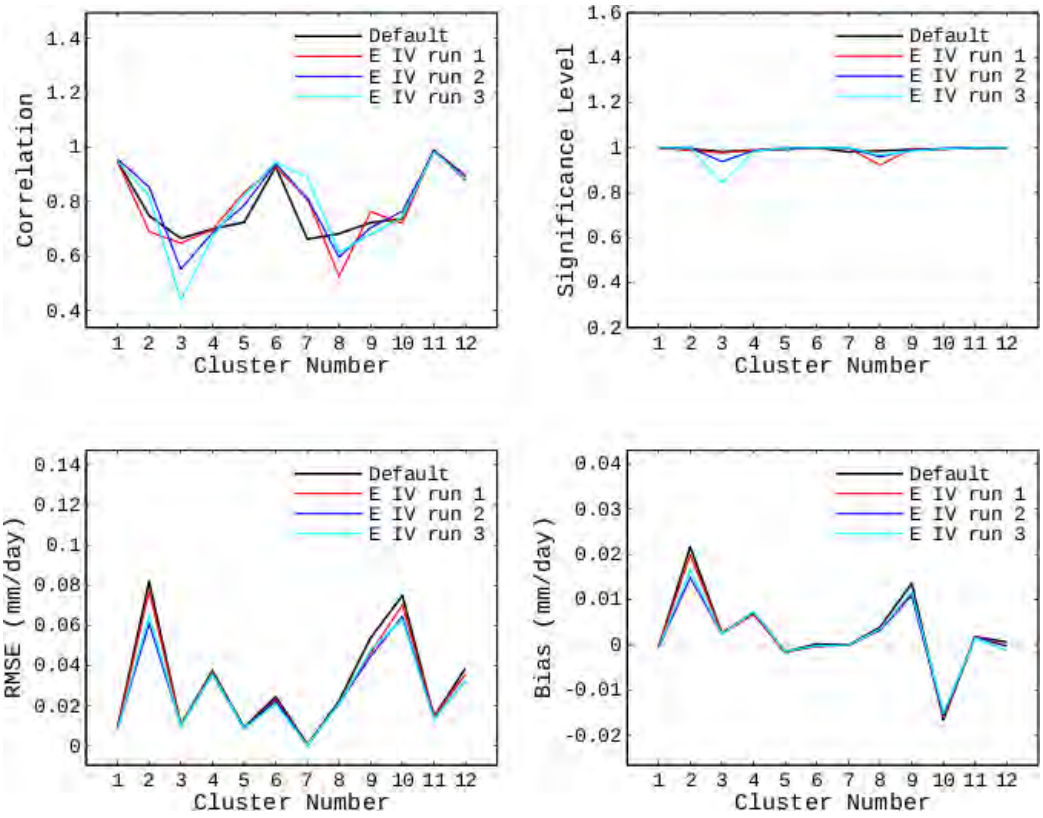


Figure 6.4: Correlation, RMSE and bias of simulated precipitation with different values of available buoyant energy removal time scale over 12 homogeneous regions. Top-left: Correlation ; Top-right: significance level; Bottom-left: RMSE; Bottom-right: bias.

Combination of the Adjusted Parameters

Fig 6.5 shows the correlation coefficient, RMSE and bias of RegCM4 simulated precipitation, using different values of the above parameters, with respect to CMAP over the 12 homogeneous regions .

The correlation is improved in Experiment VI and the RMSE and bias are improved in Experiment VII and VIII. Nevertheless, the improvement is not significant compared to the modification by the individual parameters.

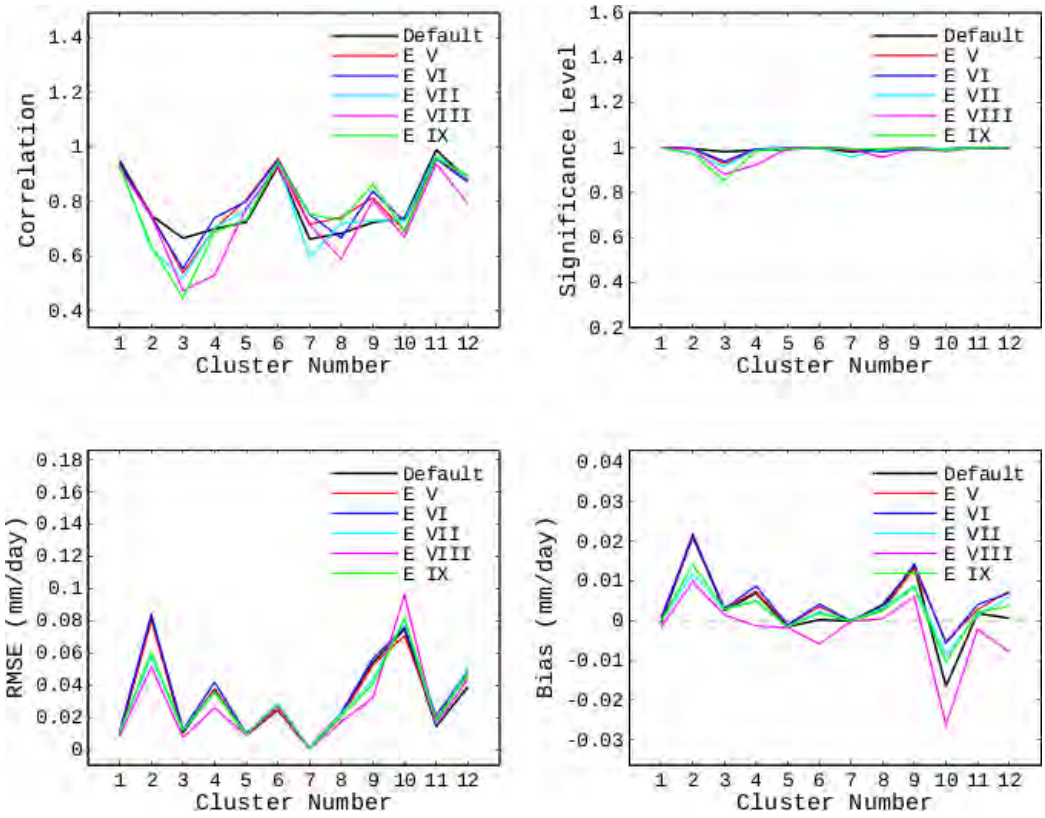


Figure 6.5: Correlation, RMSE and bias of simulated precipitation with different values of the above parameters over 12 homogeneous regions. Top-left: Correlation ; Top-right: significance level; Bottom-left: RMSE; Bottom-right: bias.

6.2 Long Period Run Using the New Set of Parameters in Grell-FC Scheme

To see the improvement of setting minimum and maximum shear effect on precipitation efficiency to 0.10 and 0.30 respectively, 17 years run is performed and the correlation, RMSE and bias are analyzed.

Fig 6.6 shows the correlation coefficient between the simulated and observed precipitation (top panel) and the correlation significance level (bottom panel) of the 17 years temporal averages of the simulations using the new set of parameters with respect to

CMAP (left panel) and GPCP (right panel).

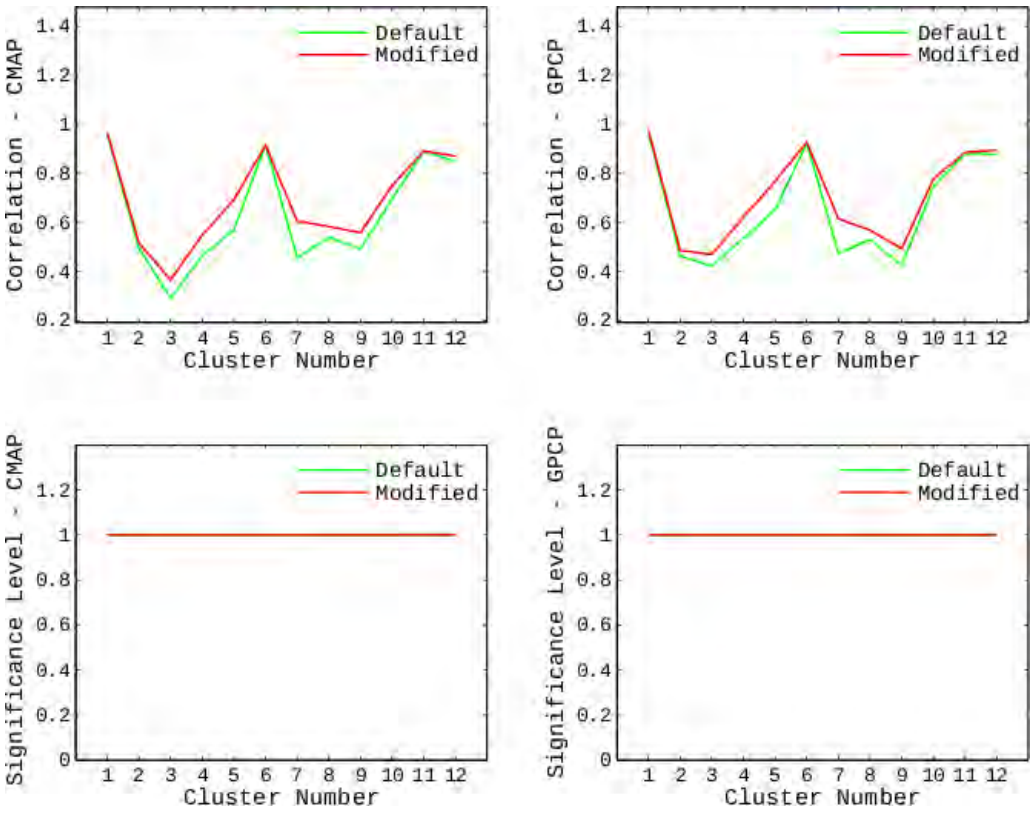


Figure 6.6: Correlation between Observation and simulated precipitation using the default and modified Grell-FC schemes over 12 homogeneous regions. Top-left: Correlation with CMAP; Top-right: correlation with GPCP; Bottom-left: significance level with CMAP; Bottom-right: significance level with GPCP.

Grell-FC scheme with the selected new set of parameters, minimum and maximum shear effect on precipitation efficiency of 0.10 and 0.30 respectively, have better correlation when compared to the default Grell-FC scheme. The improvement is significant for clusters 3, 4, 5, 7 and 9.

Fig 6.7 shows the RMSE (top panel) and bias (bottom panel) of the 17 years Temporal averages of the simulations using the default and modified Grell-FC schemes with respect to CMAP (left panel) and GPCP (right panel).

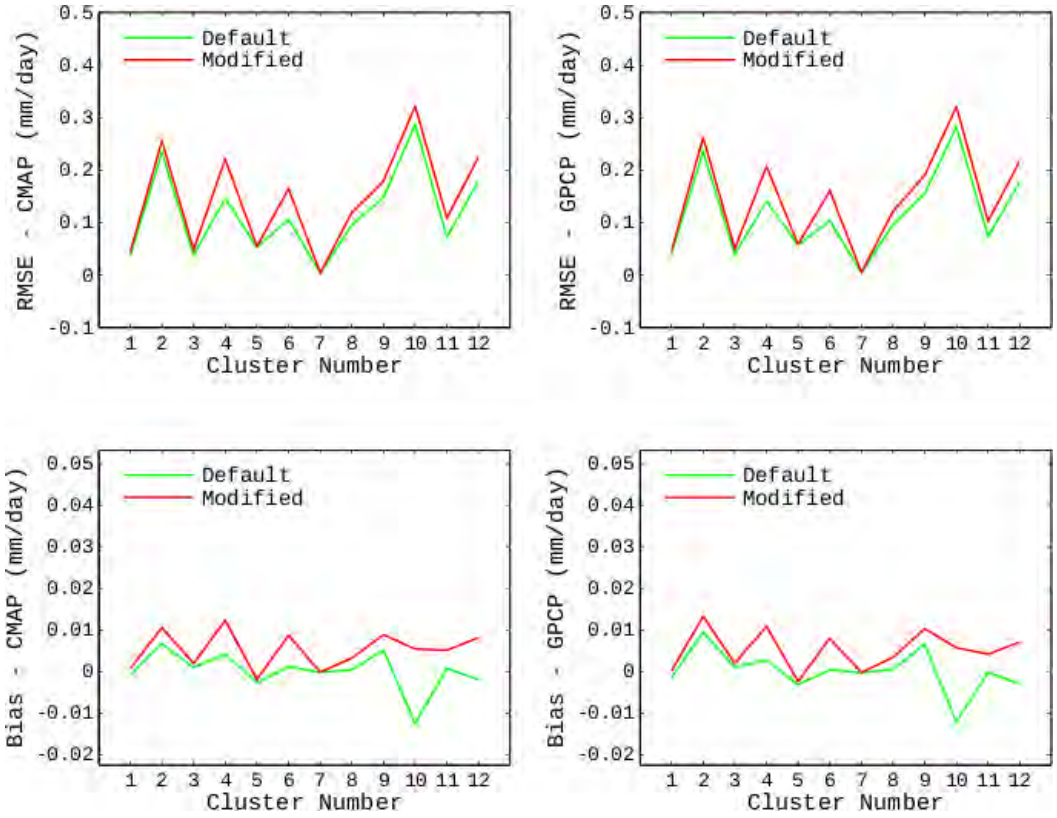


Figure 6.7: The error estimation of RegCM4 Simulated precipitation with the default and modified Grell-FC schemes in the 12 clustered regions. Top-left: RMSE with respect to CMAP; Top-right: RMSE with respect to GPCP; Bottom-left: bias with respect to CMAP; Bottom-right: bias with respect to GPCP.

The modified Grell-FC scheme has positive bias which is larger than (in the order of 0.01 mm/day) that of the default Grell-FC scheme over the clusters except clusters 3, 5 and 7. It also has larger RMSE over those clusters.

Fig 6.8 presents the correlation coefficient between the simulated and observed precipitation (top panel) and the correlation significance level (bottom panel) of the 17 years temporal averages of the simulations using the default and modified Grell-FC schemes with respect to CMAP (left panel) and GPCP (right panel).

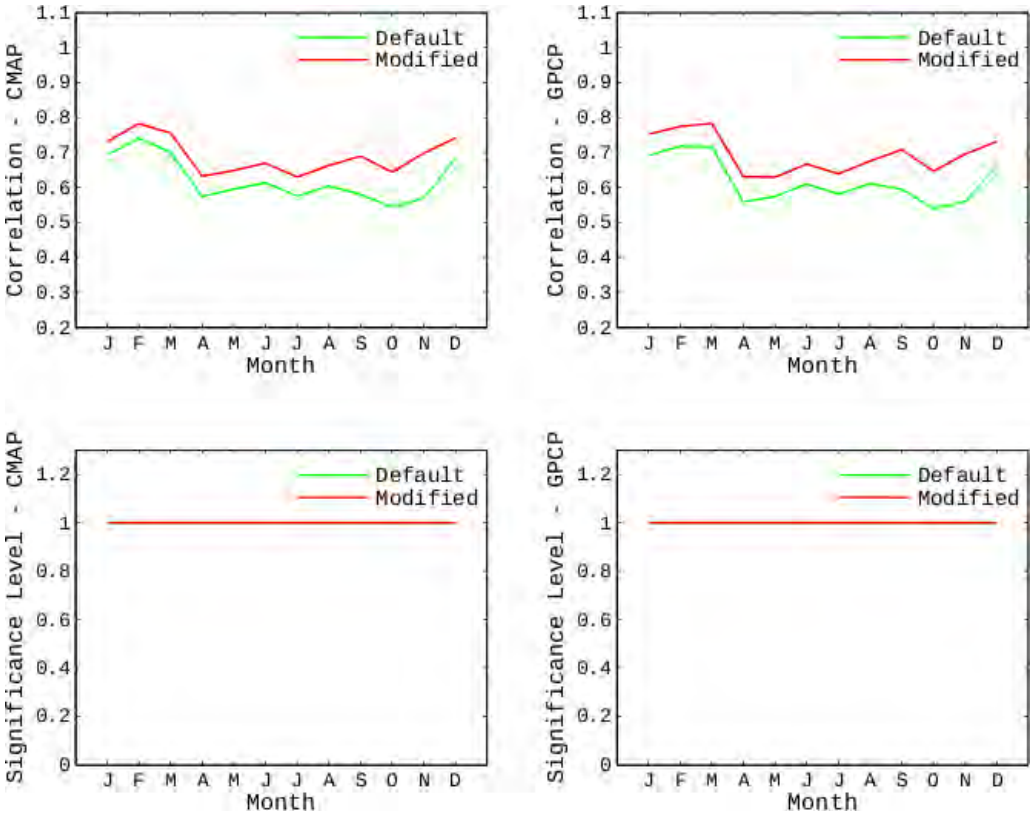


Figure 6.8: Monthly Correlation between Observation and RegCM4 Simulated precipitation with the default and modified Grell-FC schemes. Top-left: Correlation with CMAP; Top-right: correlation with GPCP; Bottom-left: significance level with CMAP; Bottom-right: significance level with GPCP.

The modified Grell-FC scheme exhibits high correlation throughout the entire months. Particularly, it outperform the default Grell-FC scheme for months September, October, and November.

Fig 6.9 shows the RMSE (top panel) and bias (bottom panel) of the 17 years monthly averages of the simulations using the default and modified Grell-FC schemes with respect to CMAP (left panel) and GPCP (right panel).

The RMSE of the modified and default Grell-FC schemes are similar with respect to CMAP and the modified Grell-Fc scheme has lower RMSE for August, September

and October with respect to GPCP. The modified Grell-FC scheme also reduced the positive bias of September, October and November. It has higher negative bias for May, June, July and August.

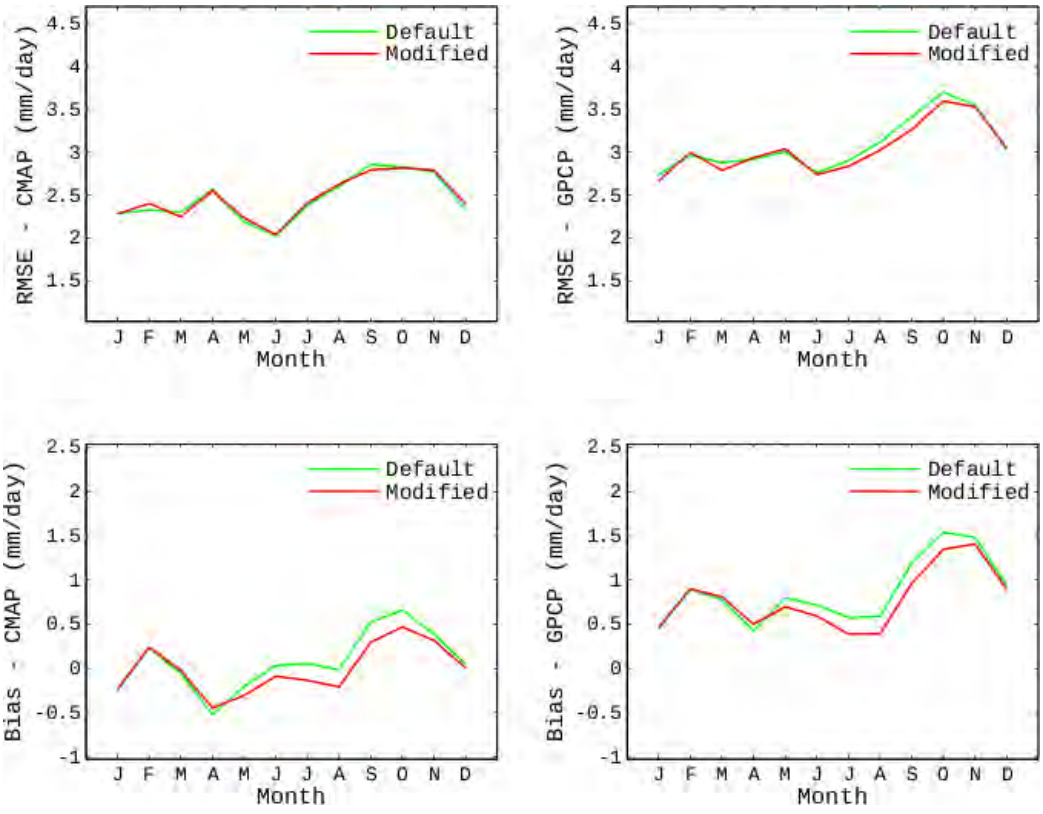
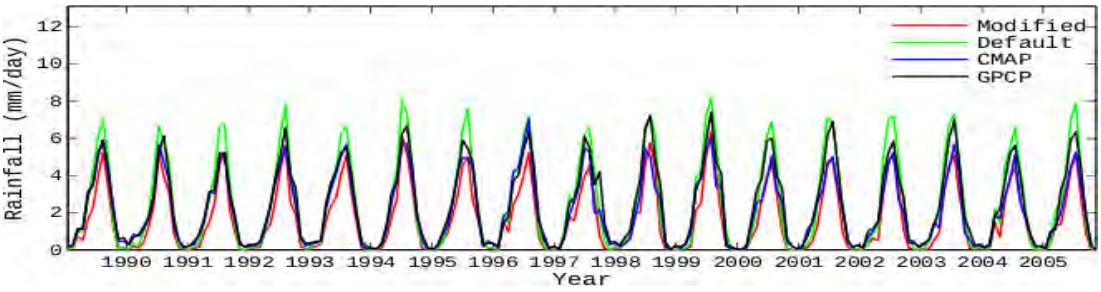


Figure 6.9: Monthly error estimation of RegCM4 Simulated precipitation with the default and modified Grell-FC schemes. Top-left: RMSE with respect to CMAP; Top-right: RMSE with respect to GPCP; Bottom-left: bias with respect to CMAP; Bottom-right: bias with respect to GPCP.

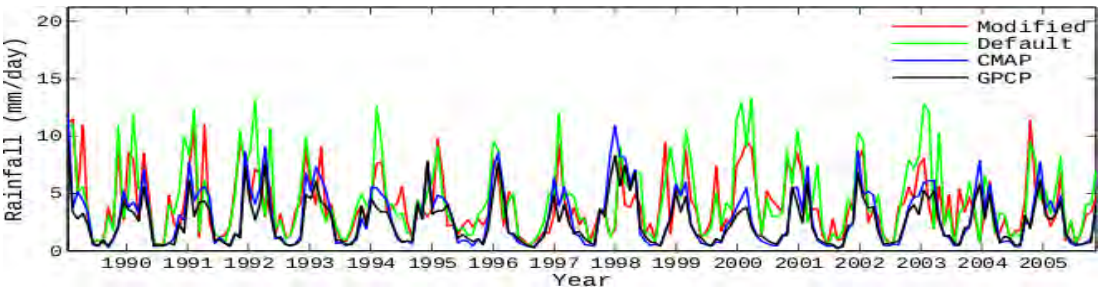
Figs 6.10 - 6.12 show the time series of monthly areal mean precipitation for both the simulations with the default and modified Grell-FC schemes and the observation data sets for the 12 clustered regions.

Setting the minimum and maximum shear effect on precipitation efficiency to 0.10 and 0.30 respectively is effective in reducing the overestimation by Grell-FC scheme

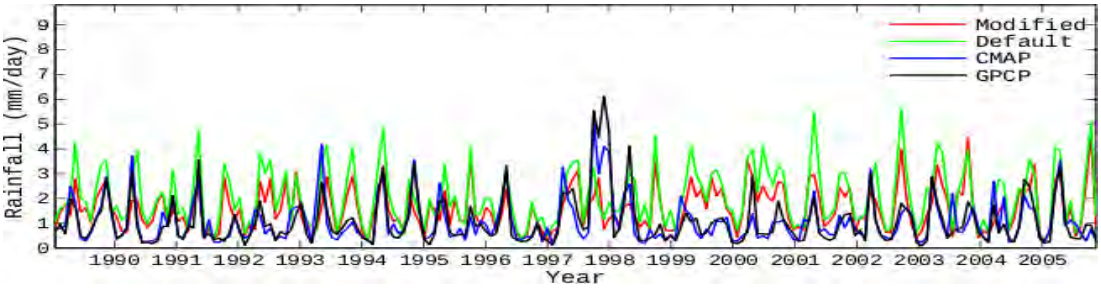
over clusters 2, 3, 4, 6, 9, 11 and 12. Where as it makes worse the underestimation over clusters 5, and 7. Particularly, it represent very well the time series and the amount of precipitation over clusters 6, 11 and 12 throughout the 17 years period.



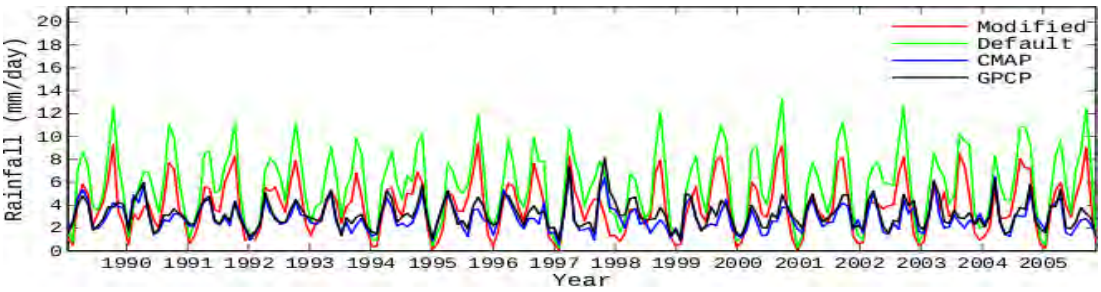
(a) Cluster 1



(b) Cluster 2

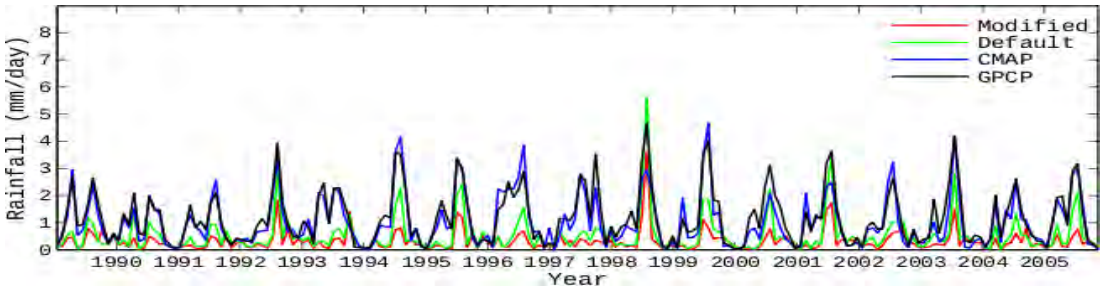


(c) Cluster 3

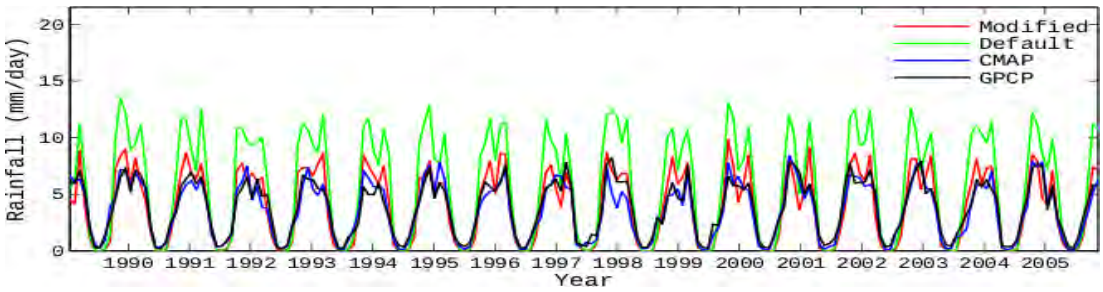


(d) Cluster 4

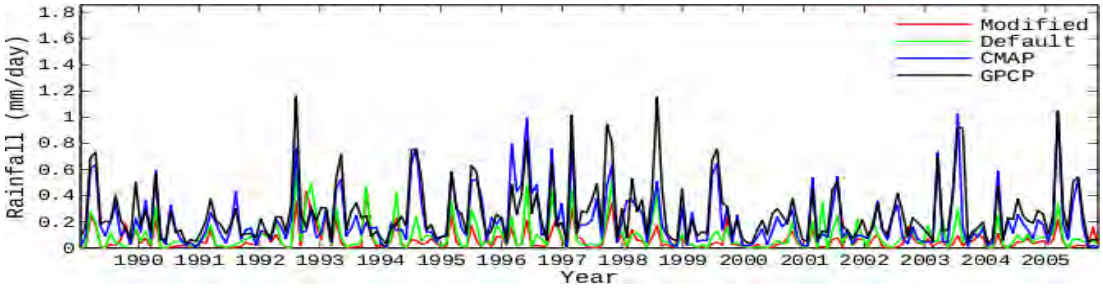
Figure 6.10: The time series of areal mean of simulated and observed monthly precipitations for clusters 1-4.



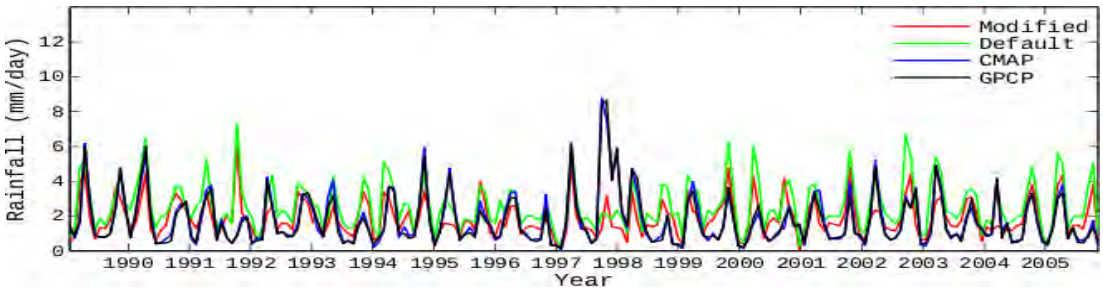
(a) Cluster 5



(b) Cluster 6

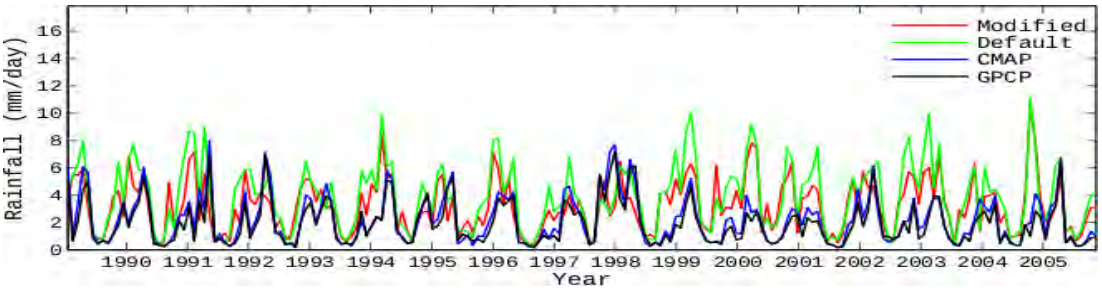


(c) Cluster 7

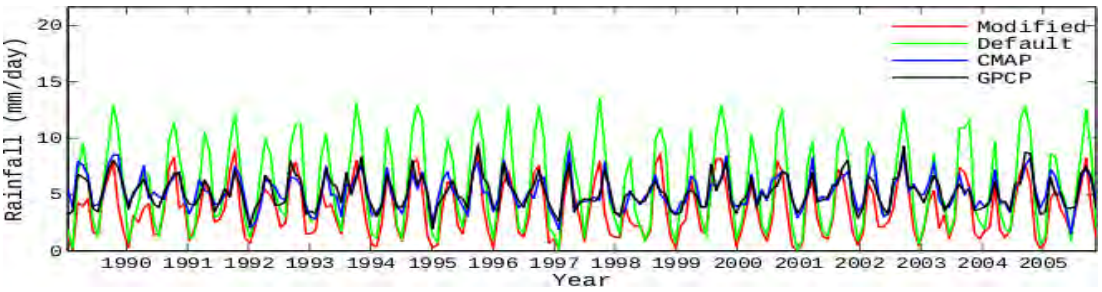


(d) Cluster 8

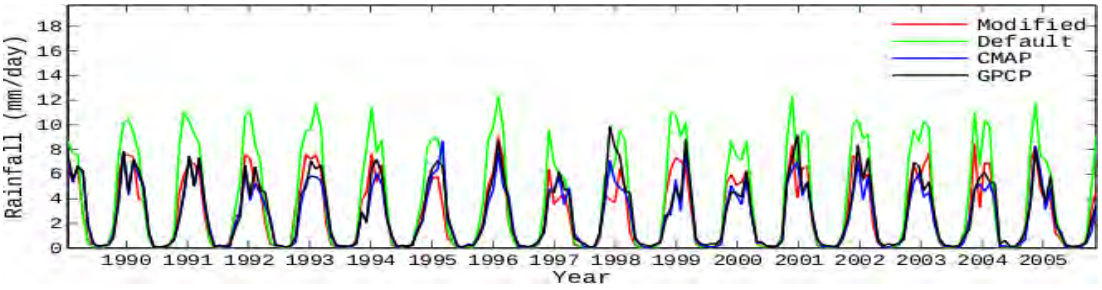
Figure 6.11: The time series of areal mean of simulated and observed monthly precipitations for clusters 5-8.



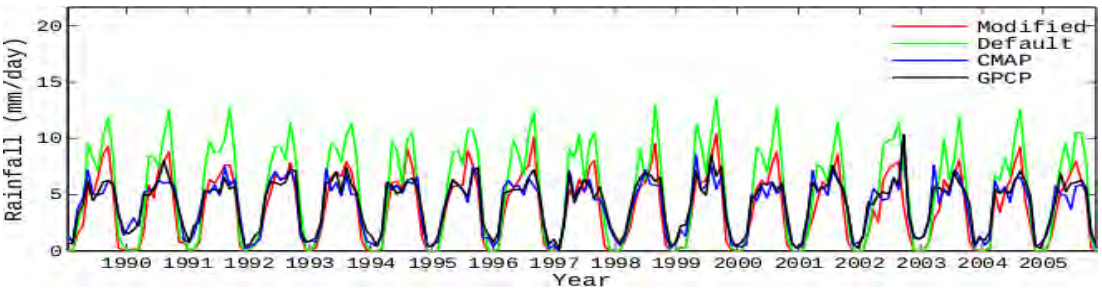
(a) Cluster 9



(b) Cluster 10



(c) Cluster 11



(d) Cluster 12

Figure 6.12: The time series of areal mean of simulated and observed monthly precipitations for clusters 9-12.

Chapter 7

Conclusions

We have examined the precipitation modeling of the RegCM4 for the Horn of Africa. Qualitative evaluation of the simulated precipitation was made for the four alternative convective schemes; the modified Anthes-Kuo scheme, the Grell scheme with Arkawa-Schubert closure, the Grell scheme with Fritsch-Chappell closure, and the Emanuel scheme.

Comparison of the modeled precipitation with CMAP and GPCP observed datasets revealed that Grell-FC is the best scheme for the Horn of Africa. Grell-FC is far better than the other Schemes in reproducing the observed 20-yrs mean climatology. Results obtained by performing cluster analysis also indicated that Grell-FC captured well the precipitation distribution over most of the clustered region, where Emanuel/Grell-As and Kuo schemes overestimated/underestimated the observed precipitation. In spite of high correlation by Emanuel scheme, Grell-FC scheme captured better the annual cycle with minimum positive bias. It is also noted from areal mean monthly time series analysis that simulation using Grell-FC is superior to the other schemes in reproducing the inter-annual variability of precipitation over the Horn of Africa.

After determining the best convective scheme, several sensitivity experiments were done to improve the performance of RegCM4 using the selected Grell-FC scheme particularly to increase the correlation of the simulated precipitation with the observational datasets and to reduce the positive bias over most of the clusters. The sensitivity experiments involved adjusting minimum and maximum precipitation efficiency, minimum and maximum shear effect on precipitation efficiency, minimum and maximum convective heating, and ABE removal timescale.

It has been found that adjusting one of the above parameters alter the correlation and the bias of the simulated precipitation. But unfortunately, none of the parameters can be adjusted to improve both the correlation and bias at the same time. The adjustment of minimum and maximum shear effect on precipitation efficiency to 0.10 and 0.30 respectively provided an improvement on correlation over all of the clusters except cluster 3. Setting minimum and maximum convective heating to -150 and 400 respectively was effective in improving the bias except over clusters 6 and 12.

Finally, evaluation of simulated precipitation over 1989-2005 was performed to observe the effect of adjusting minimum and maximum shear effect on precipitation efficiency to 0.10 and 0.30 respectively, on the performance of the RegCM4 in representing long period precipitation and in capturing inter-annual variability.

It has been found that the modified Grell-FC scheme improve the correlation over all of the 12 delineated regions and for all months. It also reduced the positive bias of September, October and November. The modification improved (reduced) the overestimation by Grell-FC scheme over most of the clusters. The inter-annual variability, the pattern and amount of precipitation, has been captured very well over clusters 6, 11 and 12 throughout the entire 17 years period.

Appendix
Sensetivity Experiments

Table 1: Correlation coefficient with respect to CMAP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	0.95	0.75	0.66	0.70	0.72	0.93	0.66	0.68	0.72	0.74	0.99	0.88
E I run 1	0.94	0.83	0.64	0.67	0.79	0.94	0.79	0.74	0.69	0.73	0.96	0.88
E I run 2	0.95	0.69	0.67	0.67	0.79	0.94	0.72	0.67	0.69	0.69	0.98	0.88
E I run 3	0.91	0.36	0.47	0.64	0.73	0.92	0.65	0.64	0.70	0.71	0.94	0.83
E I run 4	0.95	0.76	0.41	0.69	0.81	0.95	0.84	0.67	0.75	0.71	0.98	0.88
E I run 5	0.95	0.81	0.49	0.69	0.80	0.96	0.81	0.71	0.77	0.75	0.99	0.89
E I run 6	0.95	0.77	0.52	0.67	0.79	0.95	0.87	0.67	0.78	0.71	0.97	0.89
E II run 1	0.94	0.78	0.68	0.76	0.75	0.91	0.71	0.72	0.66	0.78	1.00	0.86
E II run 2	0.95	0.69	0.64	0.72	0.79	0.93	0.76	0.63	0.77	0.74	0.99	0.89
E II run 3	0.94	0.60	0.61	0.69	0.71	0.93	0.59	0.61	0.55	0.70	0.97	0.89
E II run 4	0.94	0.82	0.66	0.72	0.80	0.94	0.75	0.74	0.84	0.72	0.98	0.88
E II run 5	0.95	0.88	0.62	0.68	0.82	0.94	0.79	0.67	0.79	0.69	0.99	0.89
E II run 6	0.95	0.83	0.65	0.71	0.79	0.93	0.86	0.62	0.77	0.72	0.99	0.89
E II run 7	0.94	0.78	0.69	0.72	0.76	0.92	0.73	0.57	0.70	0.72	0.99	0.88
E II run 8	0.93	0.57	0.68	0.69	0.68	0.91	0.52	0.68	0.35	0.73	0.98	0.86
E II run 9	0.94	0.80	0.60	0.78	0.80	0.93	0.80	0.74	0.80	0.73	0.99	0.90
E III run 1	0.90	0.83	0.71	0.56	0.74	0.92	0.62	0.79	0.71	0.69	0.95	0.83
E III run 2	0.94	0.69	0.65	0.64	0.72	0.91	0.49	0.66	0.70	0.67	0.98	0.87
E III run 3	0.95	0.72	0.61	0.71	0.79	0.92	0.79	0.59	0.65	0.74	0.99	0.89
E III run 4	0.95	0.74	0.68	0.67	0.79	0.93	0.70	0.69	0.64	0.66	0.99	0.88
E III run 5	0.94	0.80	0.62	0.63	0.79	0.92	0.72	0.62	0.66	0.74	0.99	0.87
E III run 6	0.94	0.75	0.69	0.71	0.74	0.93	0.74	0.68	0.66	0.74	0.99	0.86
E III run 7	0.91	0.73	0.76	0.61	0.71	0.92	0.62	0.79	0.65	0.72	0.97	0.79
E IV run 1	0.95	0.69	0.65	0.70	0.83	0.93	0.81	0.53	0.76	0.72	0.99	0.89
E IV run 2	0.96	0.85	0.55	0.69	0.79	0.94	0.81	0.60	0.70	0.76	0.99	0.90
E IV run 3	0.95	0.82	0.44	0.67	0.82	0.94	0.89	0.61	0.68	0.74	0.98	0.88
E V	0.93	0.75	0.54	0.70	0.81	0.96	0.72	0.74	0.81	0.69	0.96	0.88
E VI	0.94	0.74	0.55	0.74	0.80	0.95	0.75	0.67	0.84	0.73	0.96	0.87
E VII	0.93	0.63	0.52	0.70	0.77	0.94	0.59	0.72	0.73	0.73	0.97	0.88
E VIII	0.92	0.75	0.47	0.53	0.78	0.93	0.71	0.59	0.80	0.67	0.94	0.79
E IX	0.93	0.64	0.45	0.69	0.73	0.94	0.75	0.73	0.86	0.69	0.96	0.89

Table 2: Significance level with respect to CMAP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	1.00	0.99	0.98	0.99	0.99	1.00	0.98	0.99	0.99	0.99	1.00	1.00
E I run 1	1.00	1.00	0.98	0.98	1.00	1.00	1.00	0.99	0.99	0.99	1.00	1.00
E I run 2	1.00	0.99	0.98	0.98	1.00	1.00	0.99	0.98	0.99	0.99	1.00	1.00
E I run 3	1.00	0.74	0.87	0.97	0.99	1.00	0.98	0.97	0.99	0.99	1.00	1.00
E I run 4	1.00	1.00	0.81	0.99	1.00	1.00	1.00	0.98	1.00	0.99	1.00	1.00
E I run 5	1.00	1.00	0.89	0.99	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00
E I run 6	1.00	1.00	0.92	0.98	1.00	1.00	1.00	0.98	1.00	0.99	1.00	1.00
E II run 1	1.00	1.00	0.99	1.00	0.99	1.00	0.99	0.99	0.98	1.00	1.00	1.00
E II run 2	1.00	0.99	0.98	0.99	1.00	1.00	1.00	0.97	1.00	0.99	1.00	1.00
E II run 3	1.00	0.96	0.97	0.99	0.99	1.00	0.96	0.96	0.94	0.99	1.00	1.00
E II run 4	1.00	1.00	0.98	0.99	1.00	1.00	0.99	0.99	1.00	0.99	1.00	1.00
E II run 5	1.00	1.00	0.97	0.98	1.00	1.00	1.00	0.98	1.00	0.99	1.00	1.00
E II run 6	1.00	1.00	0.98	0.99	1.00	1.00	1.00	0.97	1.00	0.99	1.00	1.00
E II run 7	1.00	1.00	0.99	0.99	1.00	1.00	0.99	0.95	0.99	0.99	1.00	1.00
E II run 8	1.00	0.95	0.99	0.99	0.98	1.00	0.92	0.98	0.74	0.99	1.00	1.00
E II run 9	1.00	1.00	0.96	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00
E III run 1	1.00	1.00	0.99	0.94	0.99	1.00	0.97	1.00	0.99	0.99	1.00	1.00
E III run 2	1.00	0.99	0.98	0.98	0.99	1.00	0.89	0.98	0.99	0.98	1.00	1.00
E III run 3	1.00	0.99	0.96	0.99	1.00	1.00	1.00	0.96	0.98	0.99	1.00	1.00
E III run 4	1.00	0.99	0.99	0.98	1.00	1.00	0.99	0.99	0.98	0.98	1.00	1.00
E III run 5	1.00	1.00	0.97	0.97	1.00	1.00	0.99	0.97	0.98	0.99	1.00	1.00
E III run 6	1.00	0.99	0.99	0.99	0.99	1.00	0.99	0.99	0.98	0.99	1.00	1.00
E III run 7	1.00	0.99	1.00	0.97	0.99	1.00	0.97	1.00	0.98	0.99	1.00	1.00
E IV run 1	1.00	0.99	0.98	0.99	1.00	1.00	1.00	0.92	1.00	0.99	1.00	1.00
E IV run 2	1.00	1.00	0.94	0.99	1.00	1.00	1.00	0.96	0.99	1.00	1.00	1.00
E IV run 3	1.00	1.00	0.85	0.98	1.00	1.00	1.00	0.97	0.99	0.99	1.00	1.00
E V	1.00	1.00	0.93	0.99	1.00	1.00	0.99	0.99	1.00	0.99	1.00	1.00
E VI	1.00	0.99	0.94	0.99	1.00	1.00	1.00	0.98	1.00	0.99	1.00	1.00
E VII	1.00	0.97	0.92	0.99	1.00	1.00	0.96	0.99	0.99	0.99	1.00	1.00
E VIII	1.00	0.99	0.88	0.92	1.00	1.00	0.99	0.96	1.00	0.98	1.00	1.00
E IX	1.00	0.97	0.85	0.99	0.99	1.00	1.00	0.99	1.00	0.99	1.00	1.00

Table 3: RMSE with respect to CMAP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E I run 1	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E I run 2	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E I run 3	0.01	0.07	0.01	0.03	0.01	0.03	0.00	0.02	0.05	0.09	0.01	0.04
E I run 4	0.01	0.06	0.01	0.03	0.01	0.02	0.00	0.02	0.04	0.07	0.01	0.03
E I run 5	0.01	0.08	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.06	0.01	0.03
E I run 6	0.01	0.06	0.01	0.03	0.01	0.02	0.00	0.02	0.04	0.07	0.01	0.03
E II run 1	0.01	0.08	0.01	0.04	0.01	0.03	0.00	0.02	0.06	0.07	0.02	0.04
E II run 2	0.01	0.07	0.01	0.04	0.01	0.03	0.00	0.02	0.05	0.08	0.02	0.04
E II run 3	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E II run 4	0.01	0.08	0.01	0.05	0.01	0.03	0.00	0.02	0.05	0.08	0.02	0.05
E II run 5	0.01	0.07	0.01	0.05	0.01	0.03	0.00	0.02	0.05	0.07	0.02	0.05
E II run 6	0.01	0.07	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.02	0.04
E II run 7	0.01	0.07	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.02	0.04
E II run 8	0.01	0.08	0.01	0.03	0.01	0.03	0.00	0.02	0.05	0.08	0.01	0.04
E II run 9	0.01	0.09	0.01	0.06	0.01	0.03	0.00	0.03	0.06	0.09	0.02	0.06
E III run 1	0.01	0.05	0.01	0.03	0.01	0.03	0.00	0.02	0.03	0.11	0.02	0.06
E III run 2	0.01	0.07	0.01	0.03	0.01	0.03	0.00	0.02	0.05	0.09	0.02	0.04
E III run 3	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.06	0.07	0.02	0.04
E III run 4	0.01	0.09	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E III run 5	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.06	0.08	0.02	0.04
E III run 6	0.01	0.08	0.01	0.04	0.01	0.03	0.00	0.02	0.05	0.07	0.01	0.04
E III run 7	0.01	0.05	0.01	0.03	0.01	0.03	0.00	0.02	0.03	0.11	0.02	0.06
E IV run 1	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E IV run 2	0.01	0.06	0.01	0.04	0.01	0.02	0.00	0.02	0.04	0.06	0.01	0.03
E IV run 3	0.01	0.07	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.06	0.01	0.03
E V	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.02	0.05
E VI	0.01	0.08	0.01	0.04	0.01	0.03	0.00	0.02	0.06	0.08	0.02	0.05
E VII	0.01	0.06	0.01	0.04	0.01	0.03	0.00	0.02	0.04	0.08	0.02	0.05
E VIII	0.01	0.05	0.01	0.03	0.01	0.03	0.00	0.02	0.03	0.10	0.02	0.04
E IX	0.01	0.06	0.01	0.04	0.01	0.03	0.00	0.02	0.04	0.08	0.02	0.05

Table 4: Bias with respect to CMAP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E I run 1	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E I run 2	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E I run 3	-0.00	0.01	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.00
E I run 4	-0.00	0.01	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E I run 5	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E I run 6	-0.00	0.01	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E II run 1	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.02	-0.01	0.00	0.00
E II run 2	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E II run 3	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E II run 4	0.00	0.02	0.00	0.01	-0.00	0.01	-0.00	0.00	0.01	-0.00	0.00	0.01
E II run 5	0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.01
E II run 6	0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00
E II run 7	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00
E II run 8	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E II run 9	0.00	0.02	0.00	0.01	-0.00	0.01	-0.00	0.01	0.01	0.00	0.01	0.01
E III run 1	-0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.02	0.00	-0.00
E III run 2	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E III run 3	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00
E III run 4	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E III run 5	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E III run 6	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.02	0.00	0.00
E III run 7	-0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.03	0.00	-0.00
E IV run 1	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E IV run 2	-0.00	0.01	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E IV run 3	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E V	0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.01
E VI	0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.01
E VII	-0.00	0.01	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.01
E VIII	-0.00	0.01	0.00	-0.00	-0.00	-0.01	-0.00	0.00	0.01	-0.03	-0.00	-0.01
E IX	-0.00	0.01	0.00	0.00	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00

Table 5: Correlation coefficient with respect to GPCP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	0.96	0.71	0.60	0.80	0.75	0.96	0.44	0.78	0.77	0.71	0.99	0.87
E I run 1	0.95	0.80	0.66	0.78	0.78	0.96	0.69	0.78	0.70	0.71	0.97	0.88
E I run 2	0.95	0.62	0.72	0.77	0.79	0.96	0.53	0.76	0.73	0.68	0.98	0.88
E I run 3	0.92	0.28	0.55	0.73	0.70	0.95	0.65	0.71	0.71	0.68	0.94	0.83
E I run 4	0.96	0.62	0.52	0.79	0.80	0.98	0.72	0.75	0.71	0.71	0.98	0.90
E I run 5	0.96	0.69	0.57	0.80	0.81	0.98	0.75	0.79	0.78	0.76	0.99	0.90
E I run 6	0.96	0.66	0.63	0.78	0.78	0.97	0.82	0.72	0.68	0.74	0.97	0.90
E II run 1	0.95	0.67	0.67	0.84	0.76	0.95	0.50	0.80	0.75	0.77	1.00	0.85
E II run 2	0.96	0.68	0.60	0.82	0.80	0.96	0.70	0.74	0.78	0.70	0.99	0.89
E II run 3	0.95	0.46	0.68	0.79	0.73	0.95	0.35	0.73	0.65	0.66	0.97	0.90
E II run 4	0.96	0.67	0.71	0.81	0.78	0.96	0.73	0.82	0.82	0.73	0.98	0.88
E II run 5	0.96	0.74	0.64	0.79	0.81	0.96	0.75	0.78	0.77	0.69	0.99	0.90
E II run 6	0.96	0.75	0.68	0.82	0.78	0.96	0.77	0.75	0.81	0.74	0.99	0.90
E II run 7	0.95	0.70	0.67	0.82	0.76	0.95	0.68	0.71	0.71	0.73	0.99	0.88
E II run 8	0.95	0.62	0.69	0.78	0.71	0.94	0.26	0.75	0.46	0.69	0.98	0.86
E II run 9	0.96	0.65	0.67	0.87	0.78	0.96	0.79	0.81	0.79	0.74	0.99	0.90
E III run 1	0.90	0.79	0.75	0.66	0.72	0.95	0.57	0.79	0.68	0.65	0.96	0.84
E III run 2	0.95	0.61	0.65	0.75	0.74	0.94	0.25	0.76	0.79	0.65	0.99	0.87
E III run 3	0.96	0.57	0.64	0.82	0.80	0.96	0.69	0.73	0.71	0.73	0.99	0.90
E III run 4	0.96	0.72	0.63	0.77	0.80	0.95	0.49	0.77	0.74	0.65	0.99	0.88
E III run 5	0.95	0.73	0.64	0.74	0.79	0.94	0.52	0.75	0.74	0.72	0.99	0.86
E III run 6	0.96	0.68	0.65	0.81	0.75	0.96	0.56	0.78	0.73	0.72	0.98	0.86
E III run 7	0.91	0.68	0.75	0.72	0.72	0.94	0.57	0.80	0.69	0.71	0.98	0.79
E IV run 1	0.96	0.54	0.63	0.81	0.81	0.96	0.75	0.67	0.75	0.72	0.99	0.90
E IV run 2	0.97	0.74	0.64	0.79	0.80	0.97	0.63	0.70	0.66	0.77	0.99	0.91
E IV run 3	0.96	0.75	0.58	0.78	0.82	0.97	0.79	0.74	0.63	0.74	0.98	0.89
E V	0.94	0.61	0.67	0.80	0.79	0.97	0.71	0.77	0.74	0.73	0.96	0.87
E VI	0.94	0.58	0.67	0.83	0.79	0.97	0.74	0.70	0.76	0.77	0.97	0.87
E VII	0.95	0.46	0.58	0.80	0.76	0.96	0.44	0.76	0.75	0.75	0.97	0.88
E VIII	0.92	0.64	0.55	0.61	0.76	0.95	0.68	0.67	0.76	0.66	0.95	0.81
E IX	0.94	0.54	0.53	0.78	0.74	0.96	0.63	0.77	0.88	0.69	0.97	0.90

Table 6: Significance level with respecyt to GPCP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	1.00	0.99	0.96	1.00	0.99	1.00	0.85	1.00	1.00	0.99	1.00	1.00
E I run 1	1.00	1.00	0.98	1.00	1.00	1.00	0.99	1.00	0.99	0.99	1.00	1.00
E I run 2	1.00	0.97	0.99	1.00	1.00	1.00	0.93	1.00	0.99	0.98	1.00	1.00
E I run 3	1.00	0.63	0.93	0.99	0.99	1.00	0.98	0.99	0.99	0.98	1.00	1.00
E I run 4	1.00	0.97	0.92	1.00	1.00	1.00	0.99	1.00	0.99	0.99	1.00	1.00
E I run 5	1.00	0.99	0.95	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
E I run 6	1.00	0.98	0.97	1.00	1.00	1.00	1.00	0.99	0.99	0.99	1.00	1.00
E II run 1	1.00	0.98	0.98	1.00	1.00	1.00	0.90	1.00	1.00	1.00	1.00	1.00
E II run 2	1.00	0.98	0.96	1.00	1.00	1.00	0.99	0.99	1.00	0.99	1.00	1.00
E II run 3	1.00	0.86	0.98	1.00	0.99	1.00	0.73	0.99	0.98	0.98	1.00	1.00
E II run 4	1.00	0.98	0.99	1.00	1.00	1.00	0.99	1.00	1.00	0.99	1.00	1.00
E II run 5	1.00	0.99	0.98	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
E II run 6	1.00	1.00	0.98	1.00	1.00	1.00	1.00	0.99	1.00	0.99	1.00	1.00
E II run 7	1.00	0.99	0.98	1.00	1.00	1.00	0.98	0.99	0.99	0.99	1.00	1.00
E II run 8	1.00	0.97	0.99	1.00	0.99	1.00	0.59	1.00	0.87	0.99	1.00	1.00
E II run 9	1.00	0.98	0.98	1.00	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00
E III run 1	1.00	1.00	1.00	0.98	0.99	1.00	0.94	1.00	0.99	0.98	1.00	1.00
E III run 2	1.00	0.97	0.98	1.00	0.99	1.00	0.56	1.00	1.00	0.98	1.00	1.00
E III run 3	1.00	0.95	0.98	1.00	1.00	1.00	0.99	0.99	0.99	0.99	1.00	1.00
E III run 4	1.00	0.99	0.97	1.00	1.00	1.00	0.89	1.00	0.99	0.98	1.00	1.00
E III run 5	1.00	0.99	0.98	0.99	1.00	1.00	0.92	0.99	0.99	0.99	1.00	1.00
E III run 6	1.00	0.98	0.98	1.00	0.99	1.00	0.94	1.00	0.99	0.99	1.00	1.00
E III run 7	1.00	0.98	1.00	0.99	0.99	1.00	0.95	1.00	0.99	0.99	1.00	1.00
E IV run 1	1.00	0.93	0.97	1.00	1.00	1.00	0.99	0.98	0.99	0.99	1.00	1.00
E IV run 2	1.00	0.99	0.98	1.00	1.00	1.00	0.97	0.99	0.98	1.00	1.00	1.00
E IV run 3	1.00	0.99	0.95	1.00	1.00	1.00	1.00	0.99	0.97	0.99	1.00	1.00
E V	1.00	0.97	0.98	1.00	1.00	1.00	0.99	1.00	0.99	0.99	1.00	1.00
E VI	1.00	0.95	0.98	1.00	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.00
E VII	1.00	0.87	0.95	1.00	1.00	1.00	0.85	1.00	1.00	0.99	1.00	1.00
E VIII	1.00	0.97	0.94	0.97	1.00	1.00	0.99	0.98	1.00	0.98	1.00	1.00
E IX	1.00	0.93	0.92	1.00	0.99	1.00	0.97	1.00	1.00	0.99	1.00	1.00

Table 7: RMSE with respec to GPCP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	0.01	0.09	0.01	0.04	0.01	0.02	0.00	0.02	0.06	0.07	0.01	0.04
E I run 1	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E I run 2	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.04
E I run 3	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.08	0.01	0.05
E I run 4	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.06	0.01	0.03
E I run 5	0.01	0.08	0.01	0.03	0.01	0.02	0.00	0.02	0.06	0.06	0.01	0.03
E I run 6	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.04	0.07	0.01	0.03
E II run 1	0.01	0.08	0.01	0.04	0.01	0.03	0.00	0.03	0.06	0.07	0.02	0.04
E II run 2	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E II run 3	0.01	0.08	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E II run 4	0.01	0.09	0.02	0.04	0.01	0.03	0.00	0.03	0.05	0.08	0.02	0.05
E II run 5	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.03	0.05	0.07	0.02	0.05
E II run 6	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E II run 7	0.01	0.08	0.01	0.04	0.01	0.02	0.00	0.02	0.05	0.07	0.02	0.04
E II run 8	0.01	0.09	0.01	0.03	0.01	0.03	0.00	0.02	0.05	0.08	0.01	0.05
E II run 9	0.01	0.10	0.02	0.05	0.01	0.03	0.00	0.03	0.06	0.09	0.02	0.05
E III run 1	0.01	0.05	0.01	0.04	0.01	0.03	0.00	0.02	0.03	0.11	0.02	0.06
E III run 2	0.01	0.08	0.01	0.03	0.01	0.03	0.00	0.02	0.05	0.09	0.02	0.05
E III run 3	0.01	0.09	0.01	0.04	0.01	0.02	0.00	0.02	0.06	0.07	0.02	0.04
E III run 4	0.01	0.09	0.01	0.03	0.01	0.02	0.00	0.02	0.06	0.07	0.01	0.04
E III run 5	0.01	0.09	0.01	0.04	0.01	0.02	0.00	0.02	0.06	0.07	0.01	0.04
E III run 6	0.01	0.08	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E III run 7	0.01	0.05	0.01	0.04	0.02	0.03	0.00	0.02	0.02	0.10	0.02	0.06
E IV run 1	0.01	0.08	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.07	0.01	0.04
E IV run 2	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.06	0.01	0.03
E IV run 3	0.01	0.07	0.01	0.03	0.01	0.02	0.00	0.02	0.05	0.06	0.01	0.03
E V	0.01	0.09	0.01	0.03	0.01	0.03	0.00	0.02	0.06	0.07	0.02	0.05
E VI	0.01	0.09	0.01	0.04	0.01	0.03	0.00	0.02	0.06	0.07	0.02	0.05
E VII	0.01	0.06	0.01	0.03	0.01	0.03	0.00	0.02	0.04	0.08	0.02	0.05
E VIII	0.01	0.06	0.01	0.03	0.01	0.03	0.00	0.02	0.03	0.09	0.02	0.05
E IX	0.01	0.06	0.01	0.03	0.01	0.03	0.00	0.02	0.04	0.08	0.02	0.05

Table 8: Bias with respect to GPCP

Cluster	1	2	3	4	5	6	7	8	9	10	11	12
Default	-0.00	0.02	0.00	0.00	-0.00	0.00	-0.00	0.00	0.02	-0.02	0.00	-0.00
E I run 1	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.00
E I run 2	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.00
E I run 3	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.01
E I run 4	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.01
E I run 5	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E I run 6	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.01
E II run 1	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.02	-0.01	0.00	-0.00
E II run 2	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E II run 3	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.00
E II run 4	-0.00	0.02	0.00	0.01	-0.00	0.01	-0.00	0.00	0.01	0.00	0.00	0.01
E II run 5	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.00	0.00	0.01
E II run 6	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00
E II run 7	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E II run 8	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.00
E II run 9	-0.00	0.03	0.00	0.01	-0.00	0.01	-0.00	0.01	0.02	0.00	0.00	0.01
E III run 1	-0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	-0.00	-0.01
E III run 2	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.02	0.00	-0.00
E III run 3	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.02	-0.01	0.00	-0.00
E III run 4	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.02	-0.01	0.00	-0.00
E III run 5	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.02	-0.01	0.00	-0.00
E III run 6	-0.00	0.02	0.00	0.00	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E III run 7	-0.00	0.01	0.00	-0.00	-0.00	-0.00	-0.00	0.00	0.00	-0.02	-0.00	-0.01
E IV run 1	-0.00	0.02	0.00	0.00	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E IV run 2	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E IV run 3	-0.00	0.02	0.00	0.01	-0.00	-0.00	-0.00	0.00	0.01	-0.01	0.00	-0.00
E V	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.01	-0.00	0.00	0.01
E VI	-0.00	0.02	0.00	0.01	-0.00	0.00	-0.00	0.00	0.02	-0.00	0.00	0.00
E VII	-0.00	0.01	0.00	0.00	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00
E VIII	-0.00	0.01	0.00	-0.00	-0.00	-0.01	-0.00	0.00	0.01	-0.02	-0.00	-0.01
E IX	-0.00	0.02	0.00	0.00	-0.00	0.00	-0.00	0.00	0.01	-0.01	0.00	0.00

Bibliography

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, and E. Nelkin (2003) *The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present).* J. Hydrometeor., 4:1147-1167.

Anthes R. (1977) *A cumulus parameterization scheme utilizing a one dimensional cloud model.* Mon. Wea. Rev., 105:270-286.

Anyah, R. (2005) *Modeling the variability of the climate system over lake victoria basin.* Ph.D. dissertation, North Carolina State University.

Anyah, R. O., and F. H. M. Semazzi (2007) *Variability of East African rainfall based on multiyear RegCM3 simulations.* Int.J. Climatol., 27:357-371.

Arakawa A, and Schubert WH. (1974) *Interaction of a cumulus cloud ensemble with the large scale environment. Part I.* J. Atmos. Sci., 31:674-701.

Bowden, J. H., and Semazzi, H. M (2007). *Empirical Analysis of Intraseasonal Climate Variability over the Greater Horn of Africa.* Amer. Meteor. Soc.,20:5715-5731.

Chih-Ping Wei, Yen-Hsien Lee, and Che-Ming Hsu (2000) *Empirical Comparison of Fast Clustering Algorithms for Large Data Sets.* Proc. the 33rd Hawaii International Conference on System Sciences.

Davis, N., Bowden, J., Semazzi, F., and Xie, L. (2009) *Customization of RegCM3 Regional Climate Model for Eastern Africa and a Tropical Indian Ocean Domain*. Amer. Meteor. Soc., 22:3595-3616.

Dickinson, R. E., P. J. Kennedy, and A. Henderson-Sellers (1993) *Biosphere-Atmosphere Transfer Scheme (BATS) version 1e as coupled to the NCAR Community Climate Model*, Tech. Rep., National Center for Atmospheric Research.

Diro G. T., Grimes D. I. F., Black E., O'Neill A., and Pardo-Iguzquiza E. (2008) *Evaluation of reanalysis rainfall estimates over Ethiopia*. Int. J. Climatol, Published online in Wiley InterScience.

Elguindi N., Bi X., Giorgi F., Nagarajan B., Pal J., Solmon F., Rauscher S., and Zakey A. (2010) *RegCM Version 4.0 User's Guide*.Trieste, Italy.

Emanuel KA. (1991) *A scheme for representing cumulus convection in large-scale models*. J. Atmo. Sci., 48:2313-2335.

Emanuel KA, and Zivkovic-Rothman M. (1999) *Development and evaluation of a convection scheme for use in climate models*. J. Atmos. Sci., 56:1756-1782.

Fairall, C. W., E. F. Bradley, J. S. Godfrey, G. A. Wick, J. B. Edson and G. S. Young (1996) *Cool-skin and warm layer effects on sea surface temperature*. J. Geophys. Res., 101:1295-1308.

Fritsch, J. M., and C. F. Chappell (1980) *Numerical prediction of convectively driven mesoscale pressure systems.part I: Convective parameterization*, J. Atmos. Sci., 37: 1722-1733.

Giorgi, F. and X. Q. Bi and Y. Qian (2003) *Indirect vs. direct effects of anthropogenic sulfate on the climate of East Asia as simulated with a regional coupled climate-chemistry/aerosol model*. J. Climatic Change, 58:345-376.

Gochis David J., W. James Shuttleworth, Zong-Liang Yang (2001) *Sensitivity of the Modeled North American Monsoon Regional Climate to Convective Parameterization*. Mon. Wea. Rev., 130:1282-1298.

Grell, G. A. (1993) *Prognostic evaluation of assumptions used by cumulus parameterizations*, Mon. Wea. Rev., 121:764-787.

Grell, G. A., Dudhia, J. and Stauffer, D. R. (1994) *A description of the fifth-generation penn state/NCAR Mesoscale Model (MM5)*. Tech. Rep., NCAR/TN-398+STR, National Center for Atmospheric Research.

Holton, James R. (2004) *An Introduction to dynamic meteorology, fourth edition*. Elsevier academic press.

Indeje, M., Semazzi, F. H. M., and Ogallo, L. J. (2000) *ENSO signals in East African rainfall seasons*. Int. J. Climatol., 20:19-46.

Jackobson, Mark Z. (2005) *Fundamentals of Atmospheric modelling, second edition*. Cambridge university press.

Kaufman, L., and Rousueeuw, P. J. (1990) *Finding Groups in Data: an Introduction to Cluster Analysis*. John Wiley and Sons.

Kiehl, J. T., J. J. Hack, G. B. Bonan, B. A. Boville, B. P. Breigleb, D. Williamson, and P. Rasch (1996) *Description of the NCAR community climate model (CCM3)*, Tech. Rep., NCAR/TN-420+STR, National Center for Atmospheric Research.

Krichak, S. O (2008) *Reional Climate Model Simulation of Present-day Regional Climate over the Eropian part of Russia with RegCM3*. Russian Meteorology and Hydrology, 33(1):20-26.

Ng, R. and J. Han (1994) *Efficient and Effective Clustering Methods for Spatial Data Mining*, Proc. 20th International Conference on Very Large Databases, Santiago, Chile.

Oleson, K. W. and G. Y. Niu and Z. L. Yang and D. M. Lawrence and P. E. Thornton and P. J. Lawrence and R. Stockli and R. E. Dickinson and G. B. Bonan and S. Levis and A. Dai and T. Qian (2008) *Improvements to the Community Land Model and their impact on the hydrological cycle*. J. Geophys. Res., 113:G01021.

Pal J. S., Small E. E., Eltahir A. B. (2000) *Simulation of regional-scale water and energy budgets: Representation of subgrid cloud and precipitation processes within RegCM*. J. Geophys. Res., 105(D24):29579-29594.

Pal J. S., Giorgi F., Bi X., Elguindi N., Solomon F., Gao X., Rauscher S. A., Francisco R., Zakey A., Winter J., Ashfaq M., Syed F. S., Bell J. L., Diffenbaugh N. S., Karmacharya J., Konare A., Martinez D., da Rocha R. P., Sloan L. C., Steiner A. L. (2007) *Regional Climate Modeling for the Developing World: The ICTP RegCM3 and RegCNET*. Bull. Amer. Meteor. Soc., 88:1395-1409.

Saha, Kshudiram (2008) *The Earth's Atmosphere Its Physics and Dynamics*. Springer.

Salzmann, Nadine Denise (2006) *The Use of Results from Regional Climate Models for Local-scale Permafrost Modelling in Complex High-mountain Topography Possibilities, Limitations and Challenges for the Future*. Ph.D. dissertation, University of Zurich.

Segele, Z., Leslie, L., Lamb, P. (2008) *Evaluation and adaptation of a regional climate model for the Horn of Africa: rainfall climatology and interannual variability*. Int. J. Climatol., Published online in Wiley InterScience.

Wang Y, Sen O. L., and Wang B. (2003) *A Highly Resolved Regional Climate Model (IPRC-RegCM) and Its Simulation of the 1998 Severe Precipitation Event over China. Part I: Model Description and Verification of Simulation*. J. Climate, 16:1721-1736.

Xie P., and P. A. Arkin (1997) *A Global precipitation: a 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs.* Bull. Amer. Meteor. Soc., 78:2539-2558.

Zeng, X.,M. Zhao, and R. E. Dickinson (1998) *Intercomparison of bulk aerodynamic algoriths for the computation of sea surface fluxes using toga coare and tao data.* J. Climate, 11:2628-2644.

Zhijie Xu, Laisheng Wang, Jiancheng Luo,Jianqin Zhang (2005) *A Modified Clustering Algorithm for Data Mining.* Sponsored by National Natural Science Foundation of China and Innovation Project of IGSNRR.

Declaration

This thesis is my original work, has not been presented for a degree in any other University and that all the sources of material used for the thesis have been dully acknowledged.

Name: Temesgen Gebremariam

Signature:_____

Place and time of submission: Addis Ababa University, June 2011

This thesis has been submitted for examination with my approval as University adviser.

Name: Dr. Gizaw Mengistu

Signature:_____