ECHAM4.5 Global Circulation Model as a seasonal forecasting system for southern Africa: *coupled vs. uncoupled*

by
Asmerom Fissehatsion Beraki

Submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy (meteorology)

in the
Faculty of Natural and Agricultural Sciences
University of Pretoria

December 2015

ECHAM4.5 Global Circulation Model as a seasonal forecasting system for southern Africa: *coupled vs. uncoupled*

Student: Asmerom Fissehatsion Beraki
Promoter: Prof. Willem Landman
Co-promoter: Dr. David DeWitt
Department: Department of Geography, Geoinformatics and Meteorology
Faculty: Faculty of Natural and Agricultural Sciences
University: University of Pretoria
Degree: Doctor of Philosophy

Abstract

The predictive skill of seasonal forecast arises from the slowly evolving climate processes where the signature, that noticeably influence the mean state of weather conditions, mainly resides in the ocean. The interaction of the ocean and atmosphere is therefore the minimum level of complexity required for seasonal timescale. The practice of contemporary seasonal prediction is presumably achievable with the use of two distinct GCM (Global Climate Model) configurations commonly referred to as one- and two-tiered forecasting systems based on the manner in which the atmosphere and ocean exchange information. One-tiered forecasting systems (Coupled climate models) are placed at the highest hierarchy in the science of numerical modelling in terms of complexity. They are hypothesized to represent the state of the art of seasonal forecasting which inherently renders them to be convenient for seasonal climate prediction purposes. Notwithstanding, it may be important to appraise whether or not two-tier forecasting systems (uncoupled models) offer comparable levels of skill that are currently attainable by state-of-the-art coupled climate models under a constrained computational resources environment. Such a restrictive environment is commonly found in developing countries such as South Africa. With this in mind, the study attempts to test the notion under a perfect model framework where the atmospheric global climate model is forced with the best estimate of predicted sea-surface temperature (SST), while the two systems are kept similar in all other aspects. The framework

eliminates differences between the two forecasting systems due to model biases and in fact enables the discrimination of the role of coupling on seasonal forecast skill.

Due to the enormous computational resources required to develop and run an operational forecast system based on coupled models, their engagement for real-time forecasts has been negligible in South Africa. However, motivated by the recent advances in computing infrastructures in South Africa due to the establishment and maintenance of the Centre for High Performance Computing (CHPC) as well as international collaboration, the study pioneered in Africa the emergence of the South African Weather Service Coupled Model (SCM) also referred to as the ECHAM4.5-MOM3-SA. The model couples the ECHAM4.5 atmospheric general circulation model (AGCM) and Modular Ocean Model version 3 (MOM3) using the multiple program multiple data (MPMD) coupler paradigm. The model employs an atmospheric initialization strategy that is different from other models that couple the same atmosphere and ocean models. The study reveals that the South African coupled model has skill levels for ENSO (El Niño Southern Oscillation) forecasts comparable with other coupled models currently administered by international centres. Furthermore the model is also found to be skilful in predicting upper air dynamics, surface air temperature, rainfall and equatorial Indian Ocean Dipole (IOD).

In the two-tiered experiment, the AGCM is constrained by the lower boundary conditions derived from predicted SST anomalies of two ocean-atmosphere coupled general circulation models (CGCMs) combined through equal weighting. In addition, the SST uncertainty amplitude (lower and upper bounds) defined from this combination is also considered as separate forcing fields. As with the CGCM, the AGCM is initialized with the realistic state of the atmosphere and soil moisture. Results from hindcasts show that this optimized forecasting system demonstrates large-scale consistent skill improvements for surface temperature and rainfall totals relative to forcing the AGCM with persisted SST anomalies and the AMIP-2 (Atmospheric Model Intercomparison Project) type simulations. Model evaluation further reveals that the AGCM is able to forecast anomalous upper air atmospheric dynamics (circulation) over the tropics up to several months ahead. In addition, the contribution of the predicted SST, which is based on a multi-model approach, is shown to be of significant importance for optimized AGCM results. However, the AGCM appears to be weakly sensitive to soil moisture initialization which may suggest an internal weakness of the model. The study has addressed some optimization issues for atmospheric models

and proposed an optimal AGCM configuration that can serve as baseline against which more advanced models can be tested.

Finally, the comparative experiments reveal that the GCM configurations widely differ in their performances and the superiority of one model over the other is mostly dependent on the ability to *a priori* determine an optimal global SST field for forcing the AGCM. In fact, the AGCM offers comparable predictive capabilities with the CGCM when the CGCMs' skilful predicted SST evolution can in turn be used to force the AGCM. This finding supports the notion that the further enhancement of seasonal forecasting practices favours the use and further improvement of CGCMs (should computational resources be permissible) since the potential for further improvement of AGCM-based forecasts heavily depends on the improvement of CGCMs.

Declaration

I, Asmerom F. Beraki, herewith declare that this thesis, which I submit for the degree PhD in Meteorology at the University of Pretoria, is my own work and has not previously been submitted for a degree at this or any other tertiary institution.

Signature: 

Date: 12 August 2015

Acknowledgments

I owe a special gratitude to my supervisor Prof. Willem Landman for his inspirational thoughts, direction and expert advice throughout the project and co-supervisor Dr. David DeWitt for his ingenious technical support and skill transfer pertaining to the coupling of the ocean and atmospheric models as well as his expert advice. My appreciation extends to my colleague Mr. Cobus Olivier for his technical assistance in and writing computer codes needed for data manipulation and pre-processing model inputs. Furthermore, I am thankful to my wife Tiegist Geberslassie and my parents for the encouragement and support. The study was sponsored by the Applied Centre for Climate & Earth Systems Science (ACCESS) and also supported by the Water Research Commission (WRC) through project K5/1913. The study was impossible without the Centre for High Performance Computing (CHPC) computational support. Furthermore, the blade server and computer storage devices obtained through the SATREPS (Science and Technology Research Partnership for Sustainable Development) a collaborative project between Japan and South Africa was essential for processing massive model output data. The Max-Planck-Institut für Meteorologie (MPI) has kindly provided the ECHAM4.5 AGCM code. The work was also impossible without the atmosphere and ocean data assimilation products respectively of NCEP (National Centers for Environmental Prediction) and GFDL (Geophysical Fluid Dynamics Laboratory) and the GFDL ocean model code. Last but not least, the SAWS' fully-fledged support plays a positive role toward the success of the study. The valuable comments of the examiners are also appreciated.

Contents

List of acronyms

AC	Anomaly Correlation
ACCESS	Applied Centre for Climate and Earth Systems Science
AGCM(s)	Atmospheric Global Circulation Model(s)
AMIP	Atmospheric Model Intercomparison Project
CBS	Commission for Basic Systems
CCA	Canonical Correlation Analysis
CFS	Climate Forecasting System
CGCM(s)	Coupled General Circulation Models (or Global climate Models)
CHPC	Centre for High Performance Computing
CMAP	Merged Analysis of Precipitation
CPC	Climate Prediction Center
CRU	Climate Research Unit
CSIR	Council for Scientific and industrial Research
DEMETER	Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction)
DMI	Dipole Mode Index
DOE	Department of Energy
ECHAM4.5	Hamburg, Germany model
ECMWF	European Centre for Medium Range Weather Forecasts
ENSO	El Niño-Southern Oscillation
EOF	Empirical Orthogonal Function
GCM(s)	General Circulation Model(s) / Global climate Model(s)
GFDL	Geophysical Fluid Dynamics Laboratory

GH	Geopotential Heights
GPC	Global Producing Centre
HIRLAM	High Resolution Limited Area Model
IC	Initial Condition
IOD	Indian Ocean Dipole
IRI	International Research Institute for Climate and Society
ITCZ	Inter-Tropical Convergence Zone
LRF	Long-Range Forecast(ing)
OAGCM(s)	Ocean-Atmosphere coupled General Circulation Model(s)
ODA	Ocean Data Assimilation
OI	Optimum Interpolation
MAE	Mean Absolute Error
MF	Météo-France
MMS	Multi Model System
MOM3	Modular Oceanic Model version 3
MPI	Message Passing Interface
MPMD	Multiple Program Multiple Data
MSE	Mean Square Error
MSSS	Mean Square Skill Score
NCEP	National Centers for Environmental Prediction
NH	Northern Hemisphere
NWP	Numerical Weather Prediction
PCA	Principal Component Analysis
RMSE	Root Mean Square Error

ROC	Relative Operating Characteristic
SADC	Southern African Development Community
SAM	Southern Annular Mode
SASAS	South African Society for Atmospheric Sciences
SATREPS	Science and Technology Research Partnership for Sustainable Development
SAWS	South African Weather Service
SCM	SAWS Coupled Model
SDT	Signal Detection Theory
SH	Southern Hemisphere
SOI	Southern Oscillation Index
SPW	South Pacific Wave
SST(s)	Sea-Surface Temperature(s)
T42	Triangular truncation at wave number 42
T42L19	T42 horizontal resolution with 19 vertical hybrid sigma layers
UKMO	United Kingdom Met Office
WMO	World Meteorological Organisation
XBT	expendable bathythermograph

1. Introduction

Whilst significant progress has been made to understand and model the Earth's climate system, many aspects of these complex natural processes may still not be fully understood (Henderson-Seller and McGuffeie, 2001). Improving the predictability of the mean state of the atmosphere, to the large extent, is expected to arise from the improvement of numerical model's formulations (i.e., dynamical and physical processes; Staniforth and Wood, 2008), data assimilation (initial conditions; Derber and Rosati, 1989; Moore and Anderson, 1989; Balmaseda et al., 2007; Balmaseda and Anderson, 2009) and the representation of interactions between climate system components (see Goddard *et al.*, 2012).

Many studies (e.g., Barnston *et al.*, 1999; Mason, *et al.*, 1999) succeeded to reveal the physical foundations that make the predictability of the mean state of the atmosphere at longer timescales possible. At the seasonal timescale, the predictability mainly arises from the slowly evolving boundary forcings which include, *inter alia*, sea surface temperature (SST), sea ice, snow cover and land surface conditions notably soil moisture (Brankovic *et al.*, 1994; Palmer and Anderson 1994; Barnston *et al.*, 1999; Goddard and Mason, 2002). In essence, seasonal-to-interannual prediction is a boundary-value problem despite that numerical studies (Shukla, 1981; Shukla and Gutzler, 1983) indicated that some of the climate predictability is also explained by the low frequency component of the atmospheric initial conditions (planetary waves) notably at the sub-seasonal timescale. Initial conditions, therefore, have a telling influence at the initial phase of the simulation notwithstanding the gradual degradation of their importance as the simulation progresses in time (Goddard *et al.*, 2001). Nonetheless, the role of atmospheric conditions in this study is allowed to evolve spontaneously using an atmospheric initialization strategy that exploits the use of realistic atmospheric states (Balmaseda and Anderson, 2009).

The need to understand the contribution of various climate forcings and their interactions within the ocean-atmosphere-land continuum particularly from a practical point of view may be important to improve the predictive skill of climate models. Most of the climate predictive signals are believed to originate from the ocean and thus the interaction of the ocean and atmosphere is

fundamentally important (Goddard *et al.*, 2001) because various interannual components of the general circulation which together account for its low-frequency variability are associated with SST anomalies. These anomalies are themselves a result of coherent atmosphere-ocean interactions. The large thermal inertia of the oceanic surface mixed layer results in these anomalies to have timescales longer than those associated with sub-seasonal variations in the atmosphere (Holton, 2004). Therefore, estimation of the evolution of SST anomalies, which are often relatively predictable, and subsequently employing them in atmospheric general circulation models (AGCMs), potentially provides means of generating forecasts of seasonal-average weather (Graham *et al.*, 2000). Moreover, in many regions of the globe the largest climate signal is associated with the SST variations of El Niño and La Niña events. These events influence the trade winds associated with large Darwin-Tahiti pressure differences, and the entire system of atmospheric and oceanic variations is referred to as the El Niño-Southern Oscillation (ENSO). In fact, ENSO is a dramatic example of interannual climate variability associated with ocean-atmosphere coupling. But maybe more important for this prediction study is that ENSO events have the potential to be predictable on a seasonal timescale and with several months lead-time (e.g., Zebiak and Cane, 1987; Stockdale *et al.*, 1998; Landman and Mason, 2001). Predicting the interannual variability of ocean areas other than the tropical Pacific has been enjoying attention only recently and in particular with the advent of (ocean-atmosphere) coupled global climate models (CGCMs; Stockdale *et al.*, 1998; Saha *et al.*, 2006). In addition to SST evolution, soil moisture anomalies can persist long enough to also trigger local and remote responses (Dirmeyer *et al.*, 2003) and its importance stirs many modellers' interest (e.g., Walker and Rowntree, 1977; Koster *et al*, 2004; Conil *et al*, 2009; Douville, 2010) despite that soil moisture anomalies are not as readily observable or predictable as SST anomalies. Nonetheless, the contribution of soil moisture is not suppressed in this work meaning that the forecast models (coupled and uncoupled) considered here are initialized with realistic soil moisture.

It is common practice to either interactively couple the ocean-atmosphere in the case of a CGCM or prescribe the ocean state (SST) for an AGCM. Although the spontaneous two-way feedback mechanism between the ocean and atmosphere provides CGCMs a distinctive advantage over AGCMs (Palmer and Anderson, 1994; Graham *et al.*, 2005), it remains still an open debate whether

these GCMs are really different in their performance even at longer lead-time. In fact several numerical studies indicate the advantage of an interactive evolution of SST which is only supported in ocean-atmosphere coupled models (e.g., Yu and Mechoso, 1999; Fu *et al.*, 2002; Graham *et al.*, 2005; Kug *et al.*, 2008; Chaudhari *et al.*, 2013). However, similar other studies reported the existence of only marginal differences between coupled and uncoupled models (e.g. Boville and Hurrell, 1998; Jha and Kumar, 2009; Colfescu *et al.*, 2013). Mainly motivated by the recently steadily growing trend in the use of CGCMs outside South Africa and local computing power enhancement, the thesis, therefore, embarks not only on the establishment of a suitable modelling framework but also elucidates how various modelling questions and challenges are addressed by placing reasonable emphasis on the Southern African climate regime within the context of the global picture.

Research problem

The advantage of a multi-model ensemble approach has been reported in many forecasting studies over recent years (e.g., Krishnamurti *et al.,* 2000; Palmer *et al.,* 2004; Doblas-Reyes *et al.,* 2005; Landman and Beraki, 2012). This advantage is due to the fact that GCMs differ in their parameterizations and therefore differ in their performance under different conditions (Krishnamarti *et al.,* 2000). Using a suite of several GCMs not only increases the effective ensemble size, it also leads to probabilistic simulations that are skilful over a greater portion of the region and a greater portion of the time series. Multi-model ensembles are nearly always better than any of the individual ensembles (Doblas-Reyes et al. 2000, Krishnamurti *et al*. 2000). The benefits from combining ensembles are a result of the inclusion of complimentary predictive information since a combination scheme should be able to extract useful information from the results of individual models from local regions where their skill is higher (Krishnamurti et al. 2000). In fact, the most striking benefit obtained from multimodel ensembles is the skill-filtering property in regions or seasons when the performance of the individual models varies widely (Graham *et al.,* 2000). However, there is the need to similarly explore the role of a multi-model SST forecasts and their effects on the predictive skill of an AGCM. In addition, the need to test whether an optimal two-tiered approach (prescribed SST forcing an AGCM) is achievable through generating a single forcing field by combining SST forecasts, or whether it is more desirable to exploit uncertainties associated with the forcing field by collectively considering various SST forecasts. This testing should be done in a manner that optimizes the representation of the uncertainties that arise both from boundary and initial conditions and at the same time taking into account computational implications.

Historically, two-tier forecasting systems were the first to appear in the scene as physically based seasonal forecasting tools and are still practiced globally (e.g. Kirtman *et al.,* 1997; Graham *et al.,* 2000; Tennant, 2003). Before that statistical forecasting methods were used for operational prediction, also in South Africa (Landman, 2014). In spite of the enormous cost implications and complexity, one-tier forecasting systems appear to gain preferences to two-tier forecasting systems over recent years and their use by operational centers is steadily growing (e.g. Stockdale *et al.,* 1998; Palmer *et al.,* 2004; Graham et al., 2005; Saha *et al.,* 2006; Molteni *et al.,* 2007). More

importantly, it has been demonstrated that southern African midsummer rainfall variability is sufficiently predictable with coupled models as demonstrated through the use of outputs obtained from projects such as DEMETER (Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction) project (Palmer *et al.*, 2004), especially during El Niño and La Niña seasons (Landman *et al.*, 2012; Landman and Beraki, 2012). Moreover, coupled models may also produce improved forecasts for South African seasonal rainfall and temperatures at lead-times of several months (Landman *et al.*, 2012; Lazenby *et al.*, 2014). In fact, it is commonly believed that coupled climate models are placed at the highest hierarchy in the science of numerical modelling in terms of complexity (Stockdale *et al.*, 1998; Palmer *et al.*, 2004). Furthermore, they are hypothesized, in theory, to represent the state of the art in seasonal forecasting which inherently renders them to be convenient for operational seasonal climate prediction purposes. Notwithstanding, due to the enormous computational resources required to develop and run coupled models, their engagement for seasonal forecasts in South Africa has not previously been considered feasible. The recent advances in computing infrastructures in South Africa due to the Centre for High Performance for Computing (CHPC) compels the need to develop and use coupled models in South Africa, at least in operational forecasting research despite that these computational resources are shared among a large range of disciplines. Notwithstanding improved computational capabilities there is still a limitation on how models to be tested can be configured and to what extent model development can happen in the region. In addition, it may also be important to consider whether two-tiered forecasting systems offer comparable levels of skills that are currently attainable by state-of-the-art coupled models (Troccoli *et al.*, 2008) on one hand, and the inhibiting factor of the computational requirement to operate such coupled systems on the other hand in a real-time forecasting environment. The latter consideration may be of particular importance in developing countries with advanced capabilities, and especially at operational centers within these countries tasked to produce real-time seasonal forecast output such as the SAWS. Moreover, although both model configurations are used at a number of operational centers, as noted above, their comparison on seasonal prediction in an operational environment is largely under explored. It is worth emphasizing that under a constrained resources scenario that it may be beneficial to objectively assess the relative merit or limitations of these forecasting systems.

Objectives and aims

Given the research problems stated above, the aim of the study is to investigate the role of coupling on predictive skill by establishing a suitable modelling framework that comprises of an interactive ocean-atmosphere coupled GCM and atmospheric GCM.

Thus the specific study objects are:

1. To introduce an optimally configured coupled GCM initialized with the best possible initialization strategy, in order to produce hindcasts that mimic a truly operational configuration at lead-time of several months. (This objective is addressed in chapter 2)
2. To introduce an optimally configured atmospheric-only GCM forced with realistic initial atmospheric states and the best available description of the surface boundary conditions as reflected in projected global SST, in order to produce hindcasts that mimic a truly operational configuration at lead-time of several months. (This objective is addressed in chapter 3)
3. To investigate the predictive potential of each forecasting system to represent key synoptic (regional) climatic systems and important diagnostic variables through the use of appropriate measures of skills; (This objective is addressed in chapters 2 and 3)
4. To identify deficiencies and sensitivities of the two systems in terms of representing climate processes in a manner that may promote further understanding of the coupled climate system and the subsequent lead to the improvement of the models. (This objective is addressed in chapters 2 and 3)
5. To conduct performance comparison of coupled and uncoupled climate forecasting systems through standard verification procedure. (This objective is addressed in chapter 4)

Thesis outline

Chapter 2 presents, in the form of an already *published peer-reviewed* journal paper, the SAWS coupled model description and evaluation. The SAWS optimized two-tiered global ensemble prediction system is given in Chapter 3, in the form of a journal paper *accepted*. The paper comprehensively describes, among others, the SAWS modelling experience from a historical perspective and the issue of optimization with the use of AGCMs in seasonal forecasting. Chapter 4 elucidates the skill comparison between one- and two-tried forecasting systems in the form of an already *published peer-reviewed* journal paper. Chapter 5 presents summary and general conclusions along with those areas that need further attention and future modelling directions.

Given that all figures and tables are specific to published and accepted papers, a list is not provided in the contents section. Furthermore, each paper (Chapters 2 – 4) is followed by the reference list specific to that paper and journal requirements. The reference list for the rest of Chapters is provided at the end of Chapter 5.

2. Dynamical seasonal climate prediction using an ocean-atmosphere coupled climate model: description and evaluation

Preface

This chapter consists of published journal paper and is cited as follows:

Beraki, A.F., D. G. DeWitt, W.A. Landman, and C. Olivier (2014): Dynamical seasonal climate prediction using an ocean-atmosphere coupled climate model developed in partnership between South Africa and the IRI, *Journal of Climate* **27:**1719-1741.

The study introduces an optimally configured seasonal forecasting system based on a CGCM and creates a comprehensive hindcast dataset that mimics a truly operational setup. Furthermore, the CGCM uses an initialization strategy that is different from previous versions of the model that coupled the same atmosphere and ocean models. The paper therefore addresses the objective 1 of the study while it conducts a multifaceted statistical analysis to assess the strength and weakness of the CGCM toward fulfilling objectives 3 and 4 relevant to the coupled model. Suitable model output is subsequently produced to be used in the chapters dealing with SST forcing and model comparison.

The paper was co-authored with David G. DeWitt, Willem A. Landman and Cobus Olivier. The conceptualisation of the paper, most of the modelling work, data analysis and the actual article writing were done by me.

Dynamical seasonal climate prediction using an ocean-atmosphere coupled climate model developed in partnership between South Africa and the IRI

Asmerom F. Beraki

South African Weather Service, Pretoria, South Africa and

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

David G. DeWitt

International Research Institute for Climate and Society, Columbia University, Palisades, New York

Current affiliation NOAA/NWS

Willem A. Landman

Council for Scientific and Industrial Research, Natural Resources and the Environment, Pretoria, South Africa

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

Cobus Olivier

South African Weather Service, Pretoria, South Africa and

Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

Abstract

The recent increase in availability of high-performance computing (HPC) resources in South Africa allowed the development of an Ocean-Atmosphere coupled general circulation model (OAGCM). The South African Weather Service (SAWS) OAGCM that coupled the ECHAM4.5 AGCM and Modular Oceanic Model version 3 (MOM3) is the first OAGCM to be developed in Africa for seasonal climate prediction. This model employs an initialization strategy that is

different from previous versions of the model that coupled the same atmosphere and ocean models. Evaluation of hindcasts performed with the model revealed that the OAGCM is successful in capturing the development and maturity of El-Niño and La-Niña episodes up to 8 months ahead. A model intercomparison also indicated that the ECHAM4.5-MOM3-SA has skill levels for the Niño-3.4 region SST comparable with other coupled models administered by international centres. Further analysis of the coupled model revealed that La-Niña events are more skillfully discriminated than El-Niño events. However, as is typical for OAGCM the model skill was generally found to decay faster during the spring barrier.

The analysis also showed that the coupled model has useful skill up to several months lead-time when predicting the equatorial Indian Ocean Dipole (IOD) during the period spanning between the mid austral spring and the start of the summer seasons which reaches its peak in November. The weakness of the model in other seasons was mainly caused by the western segment of the dipole which eventually contaminates the Dipole Mode Index (DMI). The model is also able to forecast the anomalous upper air circulations, particularly in the equatorial belt, and surface air temperature in the southern African region as opposed to precipitation.

1. Introduction

The most physically realistic and computationally expensive method of modelling the climate system is to model all components of the system believed to be relevant at the timescales of interest. At the seasonal lead-time for instance, the minimum level of complexity required is a model which coupled the atmosphere and the ocean (e.g., Stockdale et al. 1998; Palmer et al. 2004; DeWitt 2005; Graham et al. 2005; Guérémy et al. 2005; Saha et al. 2006).

The South African modelling community has over the past decade or so dedicated a large amount of resources to utilize Atmospheric General Circulation Models (AGCMs) as operational seasonal forecast tools (Landman et al. 2012). These models have all been developed outside of South Africa, but have been used extensively for operational seasonal forecast production as well as for research by many institutions including, inter alia, the South African Weather Service (COLA T30 – Kirtman et al. 1997; ECHAM4.5 – Roeckner et al. 1996), the Universities of Cape Town (HadAM3P – Pope et al. 2000) and the Council for Scientific and Industrial Research (CCAM – McGregor 1996). Due to the enormous computational resources required to develop

and run an operational forecast system based on coupled models, their engagement for real-time forecasts in South Africa has not previously been tractable. In fact, only a few institutions which are designated as global producing centres by the World Meteorological Organization (WMO) for long-range forecasts using coupled models for operational seasonal forecasting (Stockdale et al. 2009).

More recently, however, the ECHAM4.5 AGCM (Roeckner et al. 1996) has been coupled with the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Oceanic Model version 3 (MOM3; Pacanowski and Griffes 1998) at the South African Weather Service (SAWS) hereafter referred to as the "ECHAM4.5-MOM3-SA" Ocean-Atmosphere coupled General Circulation Model (OAGCM). In addition, this coupled model employs an initialization strategy that capitalizes on best available information (Balmaseda and Anderson 2009). The use of real-time atmospheric states for initialization becomes possible with an atmospheric initial condition interface introduced in the model configuration. This interface is based on the vertical interpolation scheme originally suggested by Majewski (1985) that employs the integration of the hydrostatic equation but with numerical adjustment (I. Kirchner 2001, unpublished manuscript) coded in a software package referred to as Interpolation of European Centre for Medium-Range Weather Forecasts (ECMWF) Reanalysis Data (INTERA; available online at http://wekuw.met.fu-berlin.de/;IngoKirchner/nudging/nudging/software/index.html). We used this software to develop the interface that makes the OAGCM's configuration uniquely different from previous systems involving the ECHAM4.5 AGCM coupled with the MOM3 OGCM (e.g., DeWitt 2005; hereafter referred to as "D05"). Our motivation for this work is twofold. First, it has been demonstrated that southern African midsummer rainfall variability has been shown to be sufficiently predictable by using the coupled model outputs such as from DEMETER (Development of a European Multimodel Ensemble system for seasonal to inTERannual prediction) project (Palmer et al. 2004) and the IRI, especially during El Niño and La Niña seasons (Landman et al. 2012; Landman and Beraki 2012). As noted above, coupled models are largely assumed or hypothesized to represent the state of the art of seasonal forecasting. In fact, it has been conclusively shown through the DEMETER project that coupled forecasting systems can predict both the evolution of SSTs and atmospheric conditions at enhanced levels of skill. This fact, indeed, stimulates the need to use coupled models in South Africa and renders them ideal candidates for seasonal climate prediction.

Second, with the inception of the Centre for High Performance for Computing (CHPC), the computational resources in South Africa has grown exponentially, consequently creating an environment for computationally intensive modelling research locally which would have been impossible otherwise. This recent advances in computing infrastructures compounded with the support from international institutions such as the International Research Institute for Climate and Society (IRI) in developing the coupled model described have paved the way for utilising and for further development of such state-of-the-art coupled models for seasonal forecast production and research. The aim of this paper is therefore to describe and evaluate the ECHAM4.5-MOM3-SA Ocean-Atmosphere Coupled Model (OAGCM) developed in partnership between South Africa and IRI.

The remainder of the paper is organized as follows. In sections 2 we describe the coupled model. The methodology of generating the hindcasts along with the initialization strategy is explained in section 3. In section 4 we evaluate the performance of the coupled model as a seasonal forecasting tool. A summary and conclusions are given in section 5.

2. Coupled model description

The ECHAM4.5 AGCM (Roeckner et al. 1996) is coupled with the GFDL MOM3 (Pacanowski and Griffes 1998) using the Multiple Program and Multiple Data (MPMD) fully parallelized coupler paradigm (Komori et al. 2008). Essentially, this means that the atmosphere and ocean models are the same as standalone versions except for changes needed to handle the passing of data in between. Each model is treated as an independent set of Message Passing Interface (MPI) parallel processes. In contrast, D05 employed the Ocean Atmosphere Sea Ice Soil (OASIS) coupling software (Terray et al. 1999) produced by the European Centre for Research and Advanced Training in Scientific Computation (CERFACS)" to couple the models despite that the principle on which the exchange of information between the AGCM and OGCM remains similar. The atmosphere and ocean models along with the coupling scheme are described next.

a. Atmospheric model

The AGCM is originally evolved from the spectral weather forecast model of the European Centre for Medium Range Weather Forecasts (ECMWF; Simmons et al. 1989). Numerous modifications (in dynamics and physics) have been applied to this model at the Max Planck Institute for

Meteorology (MPI) to make it suitable for climate predictions and it is the fourth generation in a series. This has been shown to have promising seasonal predictive capability for the southern Africa region (Landman et al. 2009).

The prognostic variables are represented by truncated series of spherical harmonics with triangular truncation at wave number 42 (T42) except for the moisture and trace substances. Vertically 19 unevenly spaced hybrid sigma layers are used. The model employs the vertical coordinate system of Simmons and Burridge (1981) and a semi-Lagrangian transport scheme of Williamson and Rasch, (1994) for water vapour, cloud water and trace substances. It uses the Longwave radiation of Fouquart and Bonnel (1980) and shortwave radiation of Morcrette et al. (1986). Cumulus convection is parameterized using the mass flux scheme of Tiedtke (1989) but incorporates the modifications introduced by Nordeng (1994). The turbulent surface fluxes are calculated from Monin–Obukhov similarity theory (Louis 1979), but different from its predecessors, a higher-order closure scheme (Brinkop and Roeckner 1995) is used to simulate the vertical diffusion of heat, momentum, moisture and cloud water. Horizontal diffusion is computed using the Laursen and Eliasen (1989) scheme. The orographic gravity waves are represented by the wave drag parameterization due to Miller et al. (1989). We refer the reader to Roeckner et al. (1996) for a complete model description.

b. Ocean model

The Ocean model MOM3 is a finite-difference treatment of the primitive equations of motion using the Boussinesq and hydrostatic approximations in spherical coordinates. Spatially it covers the global ocean ranges between 74^{o} South and 65^{o} North. The coastline and bottom topography are realistic but the minimum and maximum ocean depths are assumed 100 and 6000m respectively. The artificial high-latitude meridional boundaries are impermeable and insulating. The model has a 0.5^{o} uniform zonal resolution, variable meridional resolution with a 0.5^{o} between 30^{o}S and 10^{o}N, gradually increasing to 1.5^{o} at 30^{o}N and fixed at 1.5^{o} in the extratropics. There are 25 layers in the vertical with 17 layers in the upper levels between 7.5m and 450 m. The vertical mixing scheme is the nonlocal K-profile parameterization (KPP) scheme of Large et al. (1994). The horizontal mixing of tracers and momentum is Laplacian. The momentum mixing uses the space-time-dependent scheme of Smagorinsky (1963) and the tracer mixing uses Redi (1982) diffusion along with Gent and McWilliams (1990) quasi-adiabatic stirring.

c. Coupling procedure

The two GCMs exchange information once per simulation day. The AGCM feeds the OGCM with heat, momentum, freshwater, and surface solar flux. The OGCM, in turn, feeds the AGCM sea-surface temperature (SST) information. The coupling strategy used in this configuration is anomaly coupling on the AGCM side and full-field coupling on the OGCM side meaning that the anomalous atmospheric fluxes are super-imposed on the observed climatology. This procedure is the same as followed by Ji et al. (1998). The climatological AGCM fluxes are computed using a long-term climatology obtained from the uncoupled AGCM forced with observed SST. The climatological AGCM fluxes are subtracted from the fluxes computed by the AGCM component model in the coupled model to form anomalies. In addition, since the ocean model lacks a sea-ice model, the OGCM SST is relaxed toward the observed climatology in high latitudes to suppress the generation of spurious ice.

3. Retroactive forecasts design

The OAGCM uses initial states of the atmosphere, land surface and ocean. While the use of the ocean and land surface states is straightforward, the atmospheric state needs cautious treatment prior to initializing the coupled model. In this process, the model is initialized with the National Centers for Environmental Prediction (NCEP) daily atmospheric initial states, interpolated into the AGCM's vertical and horizontal resolution in a manner that respects numerical stability as explained above. The atmospheric initial conditions in D05, however, were taken from simulations made with the (ECHAM4.5) AGCM forced by the temperature from the uppermost layer of the ODA product, which is equivalent to the (MOM3) OGCM SST. Despite that the atmospheric initial conditions become less important as the lead-time increases (Goddard et al. 2001), it is worth emphasizing that the fast development of both computational technology and observational network (particularly with the advent of meteorological satellite information) has an immense contribution on the improvement in forecast quality. Theoretically, improving the predictability of the mean state of the atmosphere, to a large extent, is expected to arise from the improvement of, apart from dynamical and physical processes, optimal estimate of the state of the climate system (Balmaseda and Anderson 2009; Doblas-Reyes et al. 2013). The use of realistic atmospheric and land surface (soil moisture) states in the ECHAM4.5-MOM3-SA configuration

is, therefore, viewed from this perspective. The contribution of this initialization strategy to the overall forecast quality improvement is underway using different simulations of the ECHAM4.5 AGCM only.

The OAGCM is initialized using slightly different atmospheric initial states to build an ensemble prediction system. The technique is however applied only to the atmospheric state (section 3a) meaning that the OAGCM is constrained with a fixed ocean state for all ensemble members which fall within the proximity of the forecast date (in our case the 4th of each month). The uncertainties which arise from the initial conditions are accounted for by taking 10 consecutive daily atmospheric states back from the forecast date in each month and year. For the November hindcasts for example the atmospheric initial conditions cover the period from October 26 to November 4 for 28 years starting from 1982 to 2009. This approach is slightly different from the Climate Forecasting System (CFS) of NCEP (Saha et al. 2006) which considers pentad initial conditions. Each retrospective forecast is of 9 months length. The procedures of generating the various initial states are described below.

a. Atmospheric initial states

The atmospheric initial conditions are obtained from the National Centers for Environmental Prediction, Department of Energy (NCEP/DOE) Atmospheric Model Intercomparison Project (AMIP) II Reanalysis (R2) dataset (Kanamitsu et al. 2002) except that the lower layer atmospheric temperature is assimilated from the upper layer of the GFDL ocean data assimilation (ODA) system in order to minimize the imbalance between the (near-equatorial) upper-ocean mass field and wind stress (D05). The NCEP/DOE atmospheric states are transformed to the horizontal and vertical resolution (T42L19) of the ECHAM4.5 AGCM as noted in section 2(a). In general, the process involves: 1) conversion of pressure to a hybrid-sigma coordinate system (Simmons and Burridge, 1981), 2) computation of vorticity and divergence from meridional and zonal wind components and 3) transformation of grid to spectral space. The latter component is applied on prognostic variables, i.e., temperature, vorticity and divergence only as the specific humidity needs to remain in the corresponding Gaussian grid resolution.

The vertical coordinate system transformation requires careful treatment to ensure that an initial state is produced that is numerically and gravitationally stable. The ECHAM4.5 AGCM was

found to be sensitive to be numerically unstable when a linear or non-linear interpolation scheme without adjustment was employed. The difficulties associated with the vertical interpolation were noted in various previous studies (e.g. Majewski 1985; Shen et al. 1986; Gaertner and Castro 1996). The horizontal truncation may also potentially introduce imbalances presumably due to normal mode variations between the NCEP/DOE and the model. To minimize the problem, the vertical interpolation was conducted in a manner that preserves the structure of the vertical stratification of the atmosphere. The scheme is based on the vertical integration of the hydrostatic equation with adjustment (I. Kirchner 2001, unpublished manuscript). The interpolation scheme is originally suggested by Majewski (1985) and is also widely used for conversion between models fields of different resolution in the HIRLAM (High Resolution Limited Area Model; Undén et al. 2002) community.

b. Preparation of land surface state

The AGCM land surface model is initialized with observed soil moisture states. The soil moisture is taken from the Climate Prediction Center (CPC) monthly mean dataset (Fan and van den Dool 2004). The CPC product is interpolated to the AGCM resolution using a bi-linear interpolation procedure. The AGCM uses the simple biosphere model (Sellers et al. 1986) and soil hydrology parameterization scheme suggested by Dümenil and Todini (1992). Many studies highlighted the role of soil moisture initialization on the skill of climate models (e.g., Walker and Rowntree 1977; Koster et al. 2004; Conil et al. 2009; Douville 2010). However, it is beyond the scope of this work to assess the sensitivity of the OAGCM to soil moisture initialization alone. The goal is rather to optimize the forecasting system for predictive skill in an operational context.

c. Preparation of ocean state

The ocean initial conditions are taken from ODA system produced at the GFDL that employs an optimum interpolation scheme (Derber and Rosati 1989). The ODA uses expendable bathythermograph (XBT) data for the subsurface and relaxes the SST to observed values with a 5-day time scale. The use of the product is done by the horizontal and vertical interpolation procedure described by D05. The procedure reportedly leads to a reasonable balanced ocean initial state for use in making SST forecasts.

4. Performance statistics

The OAGCM's deterministic and probabilistic skill has been explored for different months and seasons along with several lead-times.

The verification is based on 3360 (12 months x 28 years x 10 ensemble members) hindcasts each consisting of 9-month integrations. The model runs are grouped according to the forecast date (if they were issued) to a set of hindcasts with 10 ensemble members. Each ensemble set mimics a set of operational forecasts issued on the 4th of each month starting from 1982 to 2009. The model bias in the mean annual cycle was removed from the model forecasts prior to comparing the statistics, that is, computing the anomalies of the model about its own drifted climatology as a function of different initialization time and lead-months (Wang et al. 2002; Schneider et al. 2003; D05).

The model surface and upper air data were compared against observed data compiled from different sources. For the surface variables, rainfall and air temperatures, the observed data sets used for comparison were the CPC Merged Analysis of Precipitation (CMAP; Xie and Arkin 1997) and Climate Research Unit (CRU; New et al. 2000). The SST forecasts were compared against optimum interpolation SST (OISST) version 2 (Reynolds et al. 2002). For the upper air analysis, the NCEP/DOE R2 (Kanamitsu et al. 2002) was used as a proxy for observation.

a. Deterministic forecast verification

Although operational seasonal forecasts are commonly issued probabilistically, it is also often informative to investigate their deterministic forecast performance. It is worth noting first that no cross-validation is conducted on the OAGCM SST hindcasts meaning that all the verification scores are computed directly from hindcasts as in D05. The most commonly used measures of skill in predicting the SST of the Equatorial Pacific Ocean are anomaly correlation (AC) and root-mean-square error (RMSE). Fig.1 shows the AC of the Nino3.4 basin (5^{o}S-5^{o}N, 170^o-120^{o}W) for 12 initial condition (IC) months and 9 months lead-time integrations. The model is successful in predicting the Nino3.4 SSTs well ahead of time and in most instances the ACs exceed 0.6 up to an 8 month lead-time for the ICs considered here. An AC 0.5 to 0.6 is commonly used as an indicator for the skilful forecast of the equatorial Pacific SST in the forecasting literature (e.g., Kirtman 2003; Schneider et al. 2003; D05). The ECHAM4.5-MOM3-SA struggles to maintain the defined

skill threshold at higher lead-times (> 6 month lead-time) for May, June and July target months. This result is consistent with the CFS of NCEP (Saha et al. 2006). This sudden decay in skill near April is commonly referred to as *the spring barrier* in the literature and as Saha et al. (2006) suggested the spring barrier renders the austral winter months, most notably July forecasts to be more difficult.

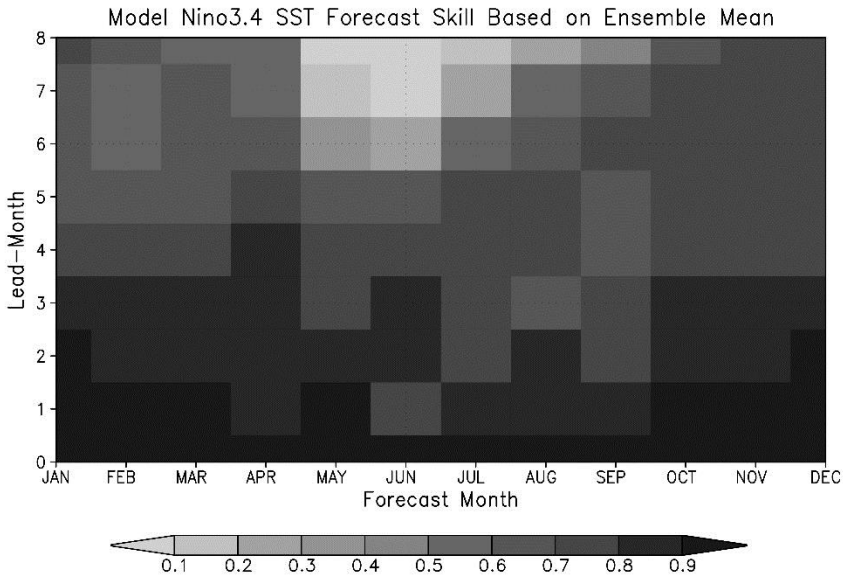

FIG. 1. OAGCM Skill for the SST forecasts in the Niño3.4 as measured with AC as a function of lead-time (vertical) and target (horizontal) months based on the monthly mean SST over the period 1982-2009.

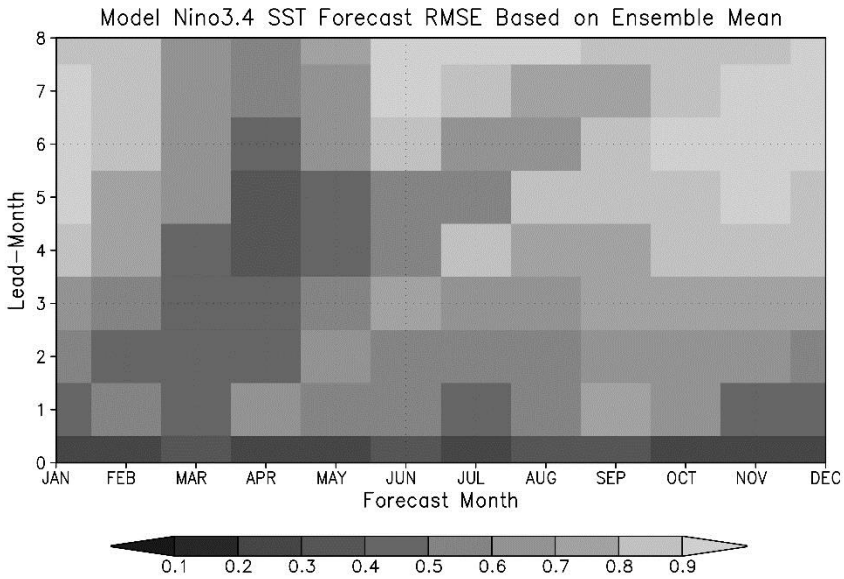

FIG. 2. OAGCM overall accuracy predicting the Niño3.4 SST. The RMSE is computed based on the monthly mean SST over the period 1982-2009

The overall accuracy of model SST forecasts for the Nino3.4 region is also assessed using the RMSE. Our model has very low error concentrations for nearly all IC cases considered here, but with errors increasing as a function of lead-time (Fig. 2). The model error is confined within the range of 0.1 and 0.5 ^{o}C. At increased lead-times (at about 5 or more months lead-times), such as for October, November, December and January ICs, the model has relatively large biases.

The time evolution of observed vs. model simulated the El Niño-Southern Oscillation (ENSO) phenomenon is shown in Fig. 3.The SST indices are area average anomalies for the Nino3.4 region similar to the index used for the computation of the AC (Fig. 1) or RMSE (Fig. 2) but for seasons instead of months. According to Fig. 3 the model agrees very well with observation for different lead-times for the austral summer (December to February; DJF) and austral fall (September to November; SON). The other seasons are relatively less skilful and uncertain specifically during the austral winter (June to August; JJA) as expected due to the spring barrier noted earlier. The skill enhancement shown in Fig. 1, to the large extent, is the contribution of the model's ability to capture the amplitude of the El Niño phases (Fig. 3) accurately except during

the early 1990s. However, the model seems to overestimate the annual amplitude or interannual variability of La Niña phases. This tendency is, however, not as striking when assessing the probabilistic skill of the model when the skill is decomposed in to various categories (see section 3(b)).

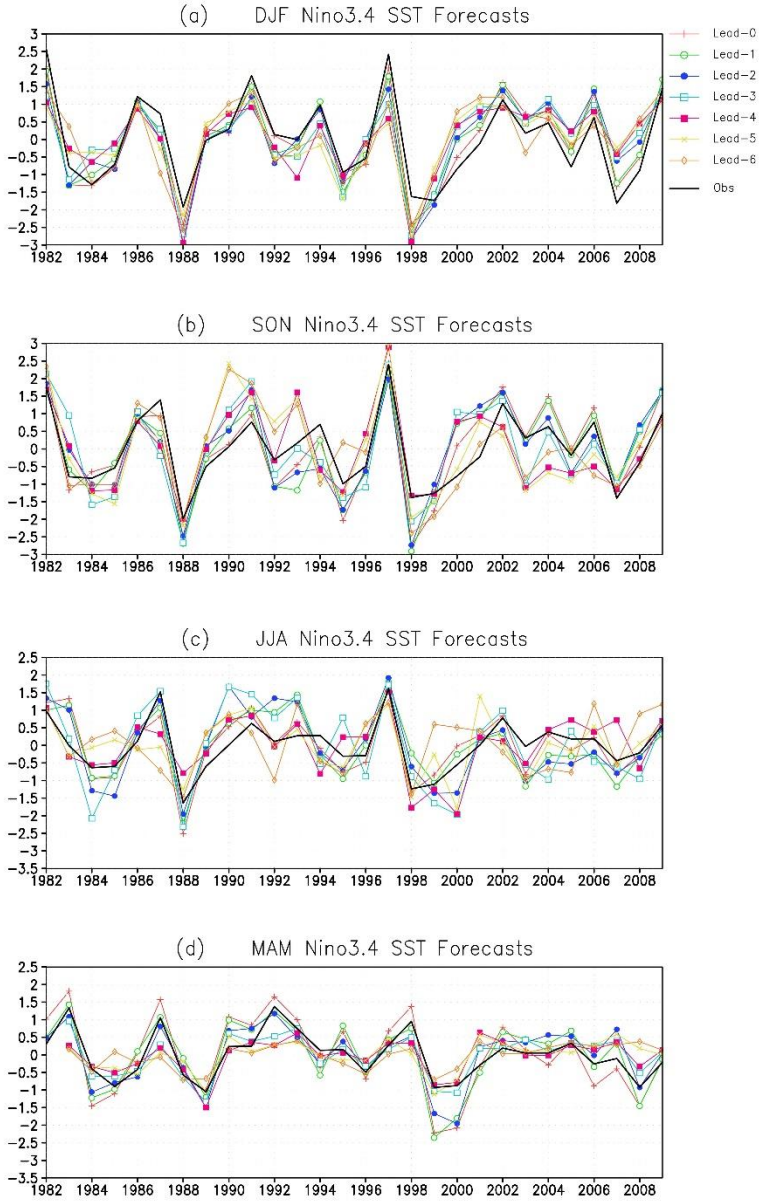

FIG. 3. The time evolution of El Niño Southern Oscillation (ENSO) as simulated by the ECHAM4.5-MOM3-SA OAGCM.

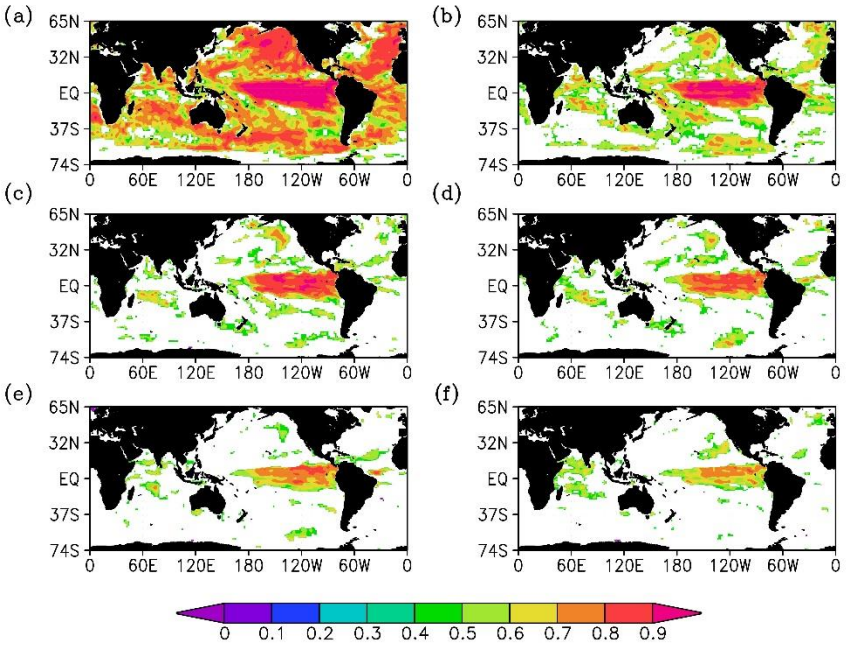

FIG. 4. Near global SST skill (AC) of the OAGCM during the start of the austral summer (December) for 6 months lead-time (0-5). Only statistical significant at 95% are shown.

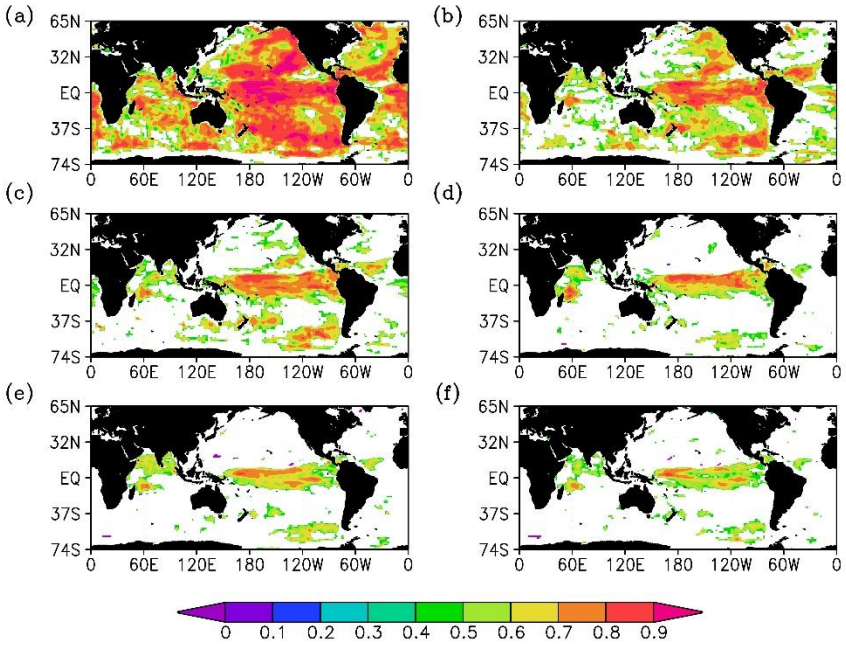

FIG. 5. Same as Fig. 4 but for the austral winter (June).

The global skill distribution of the ECHAM4.5-MOM3-SA for the start of the austral summer (December) at 0 to 5 months lead-time is shown in Fig. 4. The central and eastern equatorial Pacific region remains the area of highest predictability and is associated with coherent spatial skills statistically significant at the 95% level except at zero month lead-time (Fig. 4a) when a similar skill is also found in other ocean basins. During austral winter, the model forecast skill is similar during austral summer except that the highest skill area is also expanded towards the western part of the equatorial Pacific region. The magnitude of the skill is, however, relatively weakened toward the west as a function of lead-time (Fig. 5). The austral autumn and spring seasons (not shown) reveal a great deal of similarities with the austral summer and winter respectively.

Model intercomparison is also a useful tool and commonly practiced in the area of model evaluation (e.g., Landman and Mason 2001; Schneider et al. 2003; Alves et al. 2004; D05; Saha et al. 2006). The SST prediction for the Nino3.4 area is usually used as a benchmarking for this type of comparison since ENSO is the most predictable component of the climate system (Fig. 9). For this purpose, we used the DEMETER models (Palmer et al. 2004), the CFS coupled model (Saha et al. 2006), D05 and a statistical model [multimodel system (MMS); Beraki et al. 2012] to investigate whether our model has a reasonable skill level compared to other similar models. The data for the DEMETER models are only available for 1981-2001 whereas the other models hindcasts presented here are from 1982 to 2009. These differences may pose some difficulties in making objective or fair judgement. For this reason, we restrict the hindcast period to 1982-2001 for this model intercomparison purpose. In addition, no interpolation was performed on the individual models rather the observed SST is interpolated to each model's resolution to minimize the noise that might be introduced as a result. The DEMETER OAGCMs considered in this comparison comprise of the UKMO (United Kingdom Met Office; Pope et al. 2000), MF (Météo-France; Déqué 2001) and ECMWF (European Centres for Medium-Range Weather Forecast; Gregory et al. 2001).

The skill and accuracy of the different models computed from their hindcasts initialized in November and February are illustrated in Figs. 6 and 7, respectively. The choice of the November

and February ICs were dictated by the availability of the DEMETER models' hindcast data. In the November IC, our coupled model demonstrates a competitive skill with most of the models considered here. The ECHAM4.5-MOM3-SA (denoted as SCM) maintains its AC at or above 0.8 across all lead-times; the skill for ECMWF, CFS and MMS starts decaying faster at lead-time 5 (Fig. 6a). In the February IC (Fig. 6b), all models tend to show similar tendency at all lead-times to that of the November IC. The exception is that the skill level in February initialized runs is generally low. The ECMWF, CFS and MMS decay faster relative to the other models at longer lead-times.

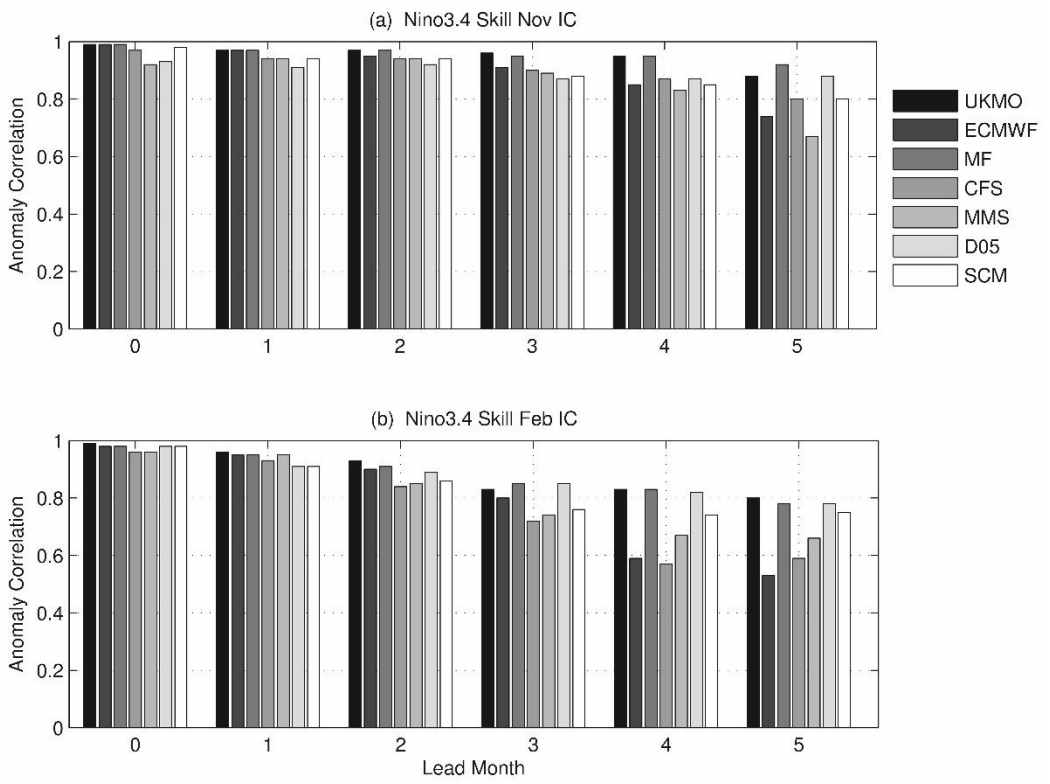

FIG. 6. Anomaly correlation by various prediction methods of monthly for the Niño-3.4 forecasts as a function of different lead-month (horizontal). The skill scores are base on the November (a) and February (b) initialized hindcasts. The ECHAM4.5-MOM3-SA is denoted by SCM; The MMS refers to CCA based statistical Multi-Model ENSO prediction system.

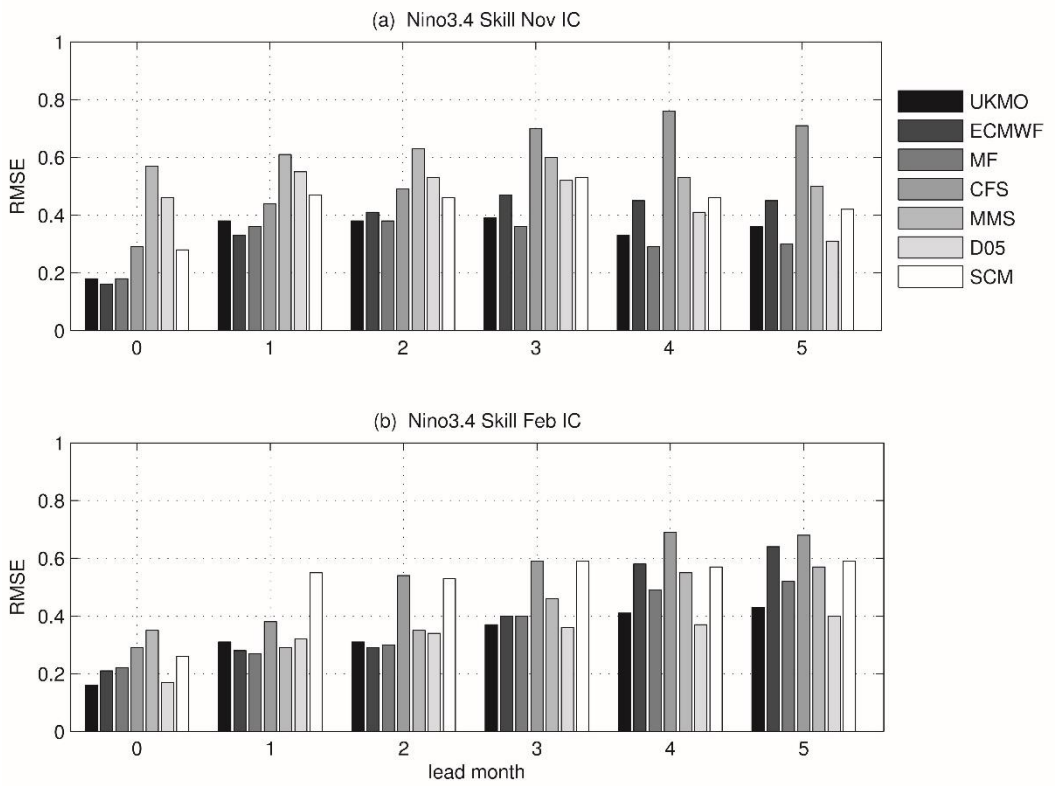

FIG. 7. RMSE by various prediction methods of monthly mean for the Niño-3.4 forecasts as a function of different lead-month (horizontal). The level of bias in each model computed using the November (a) and February (b) initialized hindcasts. The ECHAM4.5-MOM3-SA (as in Fig. 6) compared with the local CCA based empirical model, DEMETER coupled models (UKMO, ECMWF and MF), D05 and CFS of NCEP coupled model.

It is imperative to accompany the AC with a measure of accuracy or bias because the AC is not sensitive to bias since a biased forecasting system can still produce good AC. The RMSE computed for each model forecast set against the OI-SST is shown in Fig. 7. In the November IC (Fig. 7a), the ECHAM4.5-MOM3-SA achieves a comparable level of accuracy relative to the DEMETER models where MF is found to have the highest skill. The MMS (specifically at the start) and CFS (as the lead-time increases) show a gradual growth in error. In the February IC case (Fig. 7b), errors grow the fastest with increasing lead-time for the ECHAM4.5-MOM3-SA (1 to 3 months lead).

The MMS shows a tendency of greater error growth in the context of ENSO forecasts at the start of the austral summer even though it demonstrates a robust performance in the February initialized ENSO forecasts. ECHAM4.5-MOM3-SA performance is nearly comparable with the OAGCMs which are performing best in the cases we considered except for the relative error growth noted earlier.

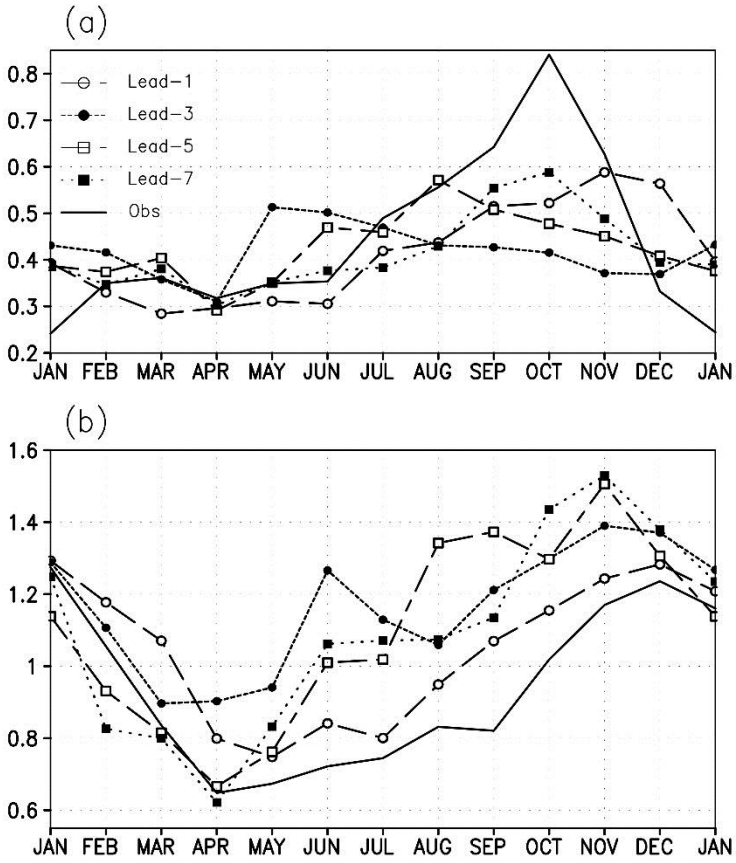

FIG. 8. Seasonal cycle of the standard deviation of anomalies of (a) the DMI and (b) the Niño3.4 index at various lead-months (as shown in the inset). Anomalies are computed formed by removing the respective climatological seasonal cycle each lead-time and observations.

The skill of the ECHAM4.5-MOM3-SA OAGCM in predicting ENSO during the austral summer is an improvement over D05 for the first four lead-months (0 to 3; Figs. 6a and 7a) and then tends to decay faster after that. In addition, our coupled model simulated the amplitude of ENSO more realistically than D05 (Fig. 9) when the amplitude of the seasonal variation of ENSO peaks (Fig. 8). Notwithstanding, in the February initialized hindcasts (Figs. 6b and 7b), D05 did

well compared to our model. The result suggested that the impact of the initialization strategy apparently becomes noticeable when ENSO becomes active (Fig. 8). Generally, the two models, however, demonstrated comparable skill levels particularly in the context of rainfall and Indian Ocean Dipole (IOD) forecasts.

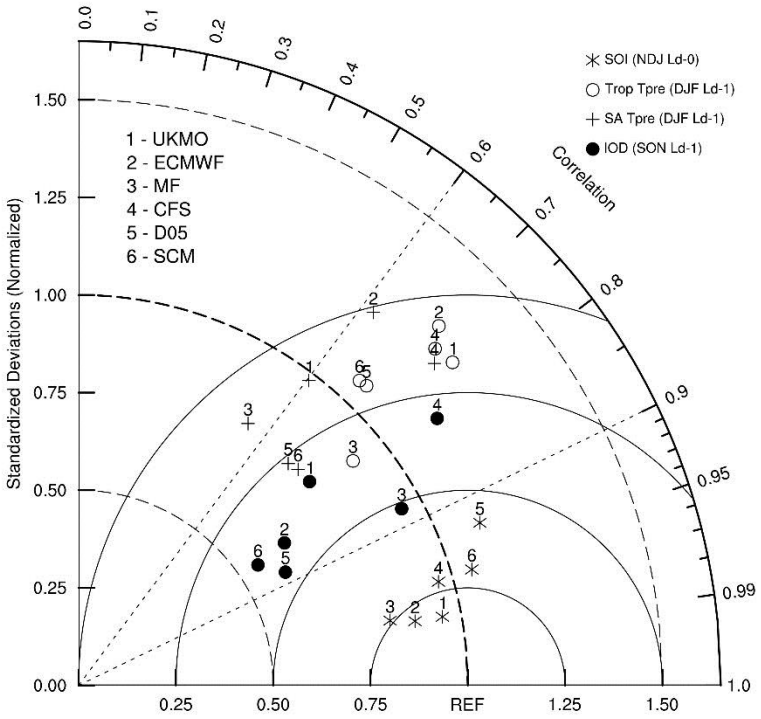

FIG. 9. Taylor diagram (Taylor 2001) by various prediction methods (as shown in the inset) based on the ensemble mean for Southern Oscillation index (SOI; asterisks), IOD (solid circles), and rainfall totals for the tropical region between 20°S and 20°N (open circles) and Southern Africa south of the equator (crosses). The standard deviation is normalized by the respective observation (see text). ECHAM4.5-MOM3-SA is denoted by ‘‘SCM.’’

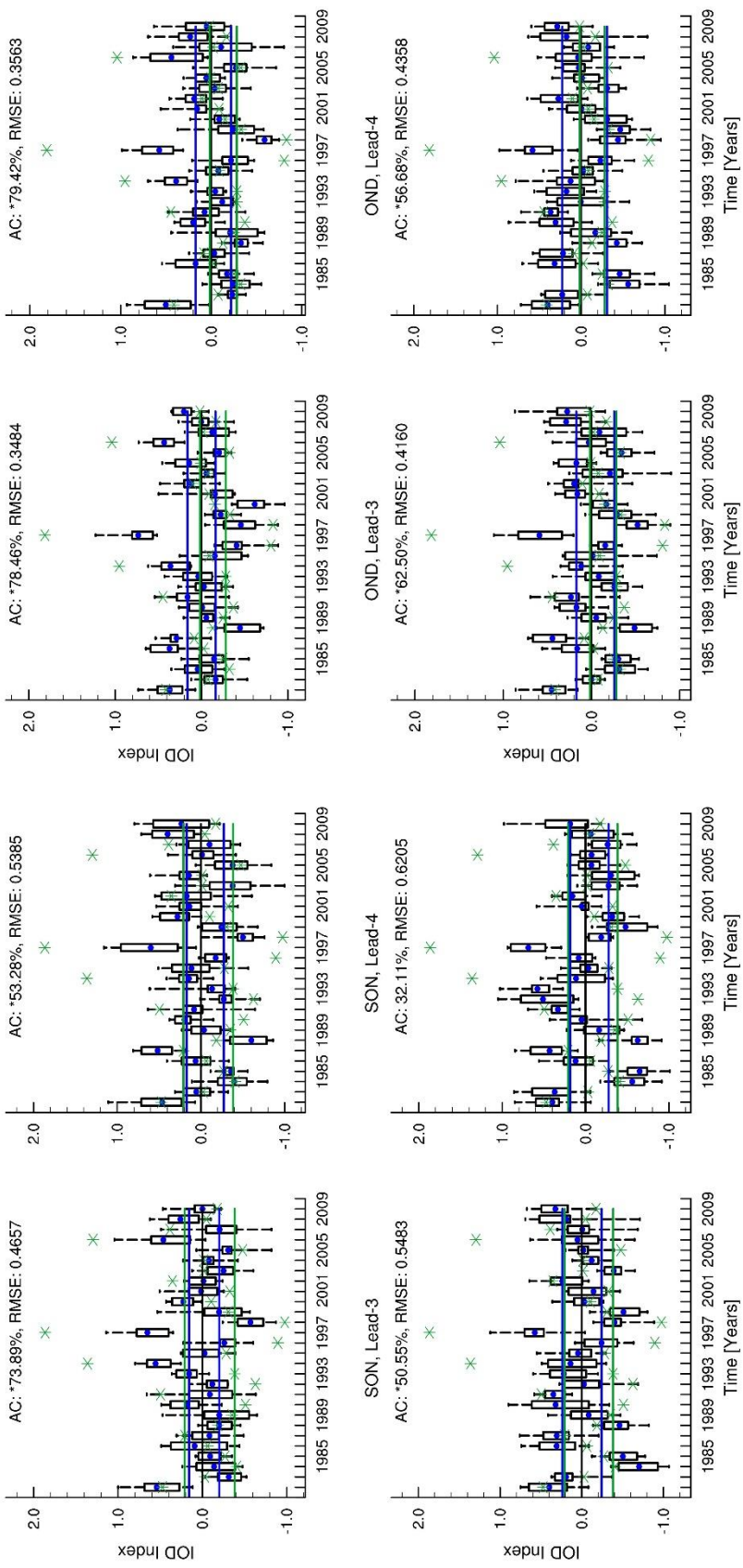

FIG. 10. Time series of IOD index from the mid austral spring to the start of summer at 2,5 and 8 months lead for the period of 1982-2009. The OAGCM ensemble spread is shown by the box-whisker representation with 25% and 75% terciles of the ensemble members. The blue dots and green stars represent the ensemble mean and OI SST respectively. The blue (ensemble mean) and green (OI SST) lines around zero black line also depict the historical quartiles based on 28 years of the index.

Previous studies highlighted the role of the coupling of the equatorial Indian Ocean basin with southern African rainfall variability (e.g., Reason 1999, 2002; Reason and Mulenga 1999; Washington and Preston 2006). It was hypothesized that this coupling phenomenon is found to drive the southern African extratropical climate system through the influence of large-scale rainfall bearing systems such as the relative annual position of the Inter-Tropical Convergence Zone (ITCZ), the South Atlantic anticyclone, and the midlatitude westerlies (Reason, et al. 2006). However, the SST prediction over the equatorial Indian Ocean has been found to be more complex and challenging. In fact, state-of-the-art coupled climate models are most often unable to replicate the observed evolving SST patterns over this ocean region (Collins et al. 2004; Landman et al. 2009). It is therefore not surprising that our model also had difficulty in simulating the observed SST patterns over the equatorial Indian Ocean sub-domain. The model shows some skill in simulating the SSTs over the western equatorial Indian Ocean off the coast of the African sub-continent except during austral autumn. However, the model manages to capture the eastern part of the equatorial Indian Ocean SST patterns near the coast of north-western Australia starting from the mid austral spring toward the beginning of the austral summer season. To substantiate this notion, we conducted model intercomparison to investigate the models' ability to simulate the equatorial Indian Ocean Dipole (IOD) using the Dipole Mode Index (DMI; Saji et al. 1999). The DMI is the SST anomaly difference between western (50^{o}E–70^{o}E, 10^{o}S–10^{o}N) and eastern (90^{o}E–110^{o}E, 10^{o}S-Eq) tropical Indian Ocean and commonly used to measure the strength and phases of the IOD (Saji et al. 1999). The model intercomparison analysis conducted during active period of IOD and ENSO (Fig., 8; Zhao and Hendon 2009) suggested that all the coupled models considered demonstrated marginal skill relative to ENSO despite that IOD is more predictable than rainfall (Fig. 9). Most of the coupled models overestimated or underestimated the amplitude of the IOD except for the MF coupled model. Notwithstanding, all models showed comparable level of skill, in the range of 0.8 and 0.9 AC, in predicting IOD for austral spring (September to November; SON) at 0-month lead.

The seasonal variation of the IOD fully develops during the austral spring (SON; Fig. 8; Saji et al. 1999; Zhao and Hendon 2009). The model generally underestimated the amplitude of the seasonal variation of IOD particularly for the first few lead-months as opposed to the model's tendency to overestimate the amplitude of the seasonal variation of ENSO. Notwithstanding, the

best performance of the model closely followed the observed peak of the seasonal cycle of the standard deviation of the IOD. Fig. 10 shows the predicted time evolution of the IOD during SON and OND (October to December) at different lead-times (1-4 months) over the verification period of 1982-2009 using a box-whisker representation. The model was found to be skilful during the SON and OND seasons. For most cases the observations (green asterisks) are dressed with the ensemble spread and tend to cluster within the same categories as delineated by the historical 25% and 75% percentiles of both the observation (green line) and ensemble mean (blue line). However, there are cases where observations lie outside the ensemble spread specifically during the1990s. These outliers might be significantly contributed to the model's tendency to damp the amplitude of seasonal variation of the IOD (Fig. 8). This suggests that there is still room for further improvement by simply increasing the ensemble size of the OAGCM integrations. In addition, the model performance during the same period but for individual months measured using the AC and RMSE is also shown in Table 1. The model demonstrates good skill (statistically significant at 5% confidence level) up to 5 months lead-time which attains its peak during November. Nonetheless, the model tends to show the smallest error growth during December presumably due to the subsidence of the IOD maturity or variation.

TABLE 1. OACGM skill and error growth in predicting the IOD for different lead-months during late autumn and the beginning of summer seasons as measured by AC and RMSE respectively. The skill scores were computed against observed DMI computed from the OI SST. The * represents that AC is statistically significant at 95% level.

	AC			RMSE		
Lead	Oct	Nov	Dec	Oct	Nov	Dec
0	*94.51	*92.75	*72.70	0.3645	0.2225	0.3472
1	*77.20	*83.21	*61.63	0.5526	0.4102	0.2692
2	*74.27	*77.29	*44.95	0.6036	0.4054	0.3054
3	*50.23	*77.61	*54.35	0.6894	0.4179	0.2844
4	*50.95	*59.99	*47.09	0.6852	0.4767	0.3214
5	31.36	*56.76	*47.84	0.7651	0.4887	0.3134
6	14.77	*45.63	*44.60	0.8535	0.5325	0.3255
7	10.35	38.19	28.26	0.8844	0.5712	0.3733
8	12.40	38.22	38.53	0.8485	0.6142	0.3518

We also extended our analysis on the upper air fields of the model using the mean square skill score (MSSS; Murphy 1988). This score can easily be computed using the Mean Absolute Error (MAE) or Mean Square Error (MSE) where the latter is employed here. Usually the reference (control) forecasts are provided either by the climatology or persistence of the variable of interest (Wilks 2006). The skill score therefore represents improvements in the forecast skill relative to the reference. The MSSS has a value of one for perfect forecasts. The MSSS could be positive (negative) when the accuracy of the forecast is superior (inferior) to the accuracy of the reference forecast. When the MSE of the reference and forecast are equal, the MSSS becomes zero which implies no improvement in the forecast system relative to the reference forecast. The spatial distribution of global actual-skills (MSSS) of the OAGCM during the austral summer for geopotential height (GH) is shown in Fig. 11. The skill score is computed from the ensemble mean of the model against the NCEP/DOE. On a synoptic scale, it is evident that the model, initialized in November, performs well at simulating the 850 and 500 hPa GH over the equatorial region, specifically over the latter. Of particular interest is that the OAGCM outscores the forecast of climatology further south, particularly on those key ocean basins which are recognized modes of atmospheric variability such as the South Pacific Wave (SPW) train (Mo and Ghil 1987) and the Southern Annular Mode (SAM).

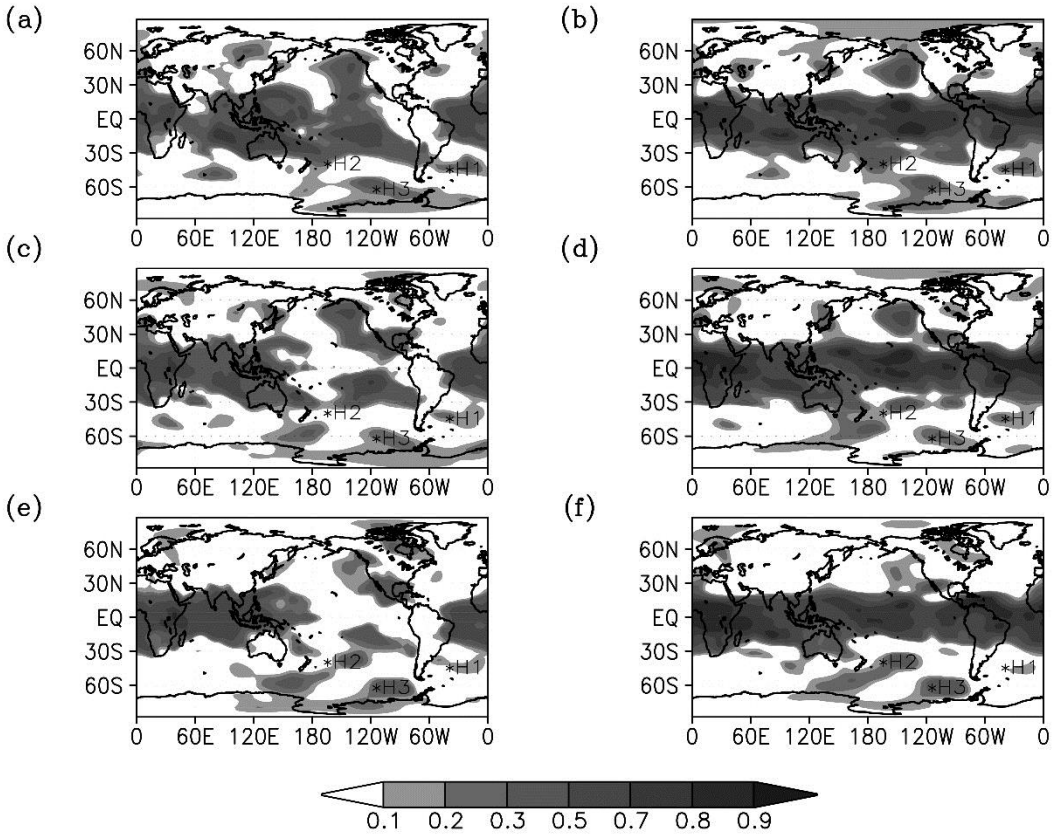

FIG. 11. Actual skill of November initialized OAGCM integrations both for 850 hPa (left panel) and 500hPa (right panel) geopotential heights. (a,b) NDJ (0-month lead-time), (c,d) DJF (1-month lead-time) and (e,f) JFM (2-month lead-time). The MSSS is computed against the NCEP/DOE upper air climate data as a proxy for observation and climatological forecast as a reference. The PSA indicated with * on the three locations (H1, H2 and H3; Yuan and Li 2008).

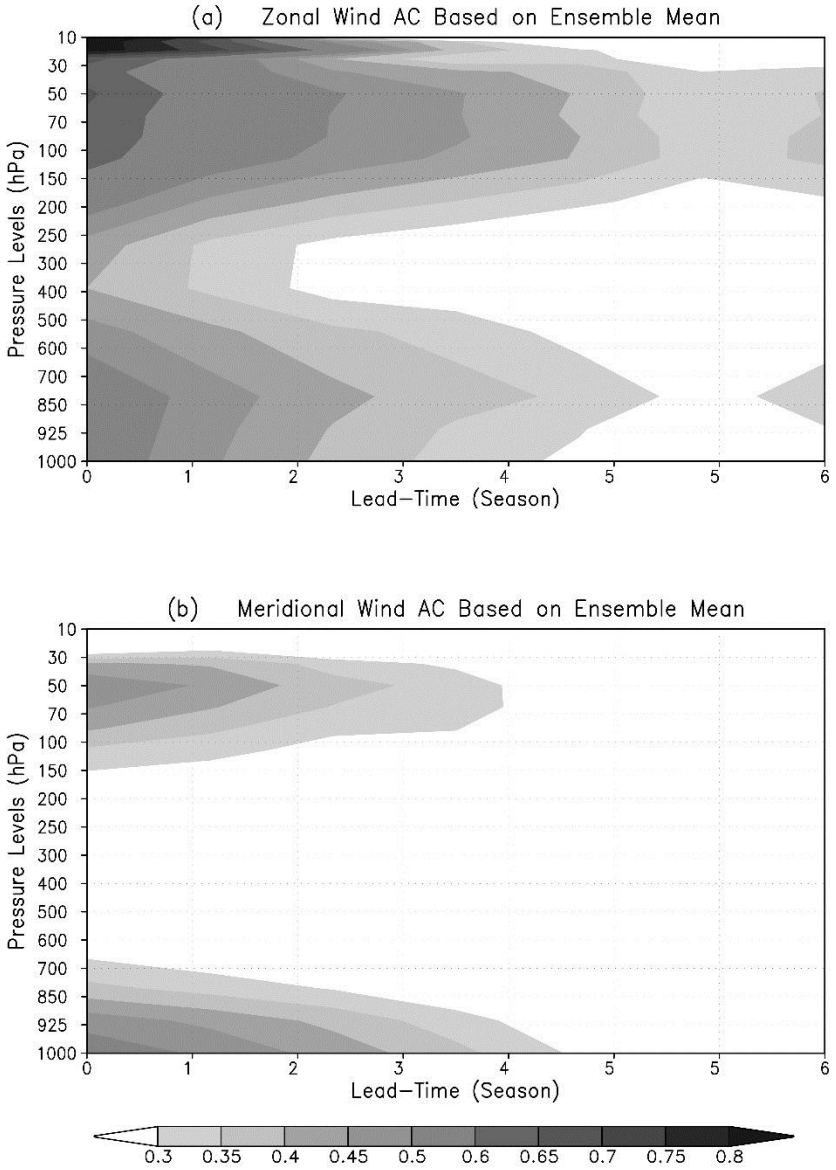

FIG. 12. Anomaly correlation of zonal (a) and meridional (b) mean wind anomalies of the equatorial region (20oSo to 20oN) at various lead-time as a function of pressure levels.

The OAGCM's performance in predicting wind components was also evaluated based on the ensemble mean integrations. Fig. 12 shows the equatorial zonal and meridional wind anomaly (20^{o}S to 20^{o}N) skill during the austral summer (DJF) at various forecast lead-times (seasons) as a

function of pressure levels computed against the NCEP/DOE. The result suggested that the model showed some skill on the lower tropospheric and upper stratospheric tropical flow as opposed to the extratropical flow. The analysis further revealed that the zonal wind appears to be more predicable than the meridional flow which might be attributed to ENSO forcing. Saha et al. (2006) demonstrated similar skill on the stratospheric zonal flow based on the CFS coupled model. Our coupled model is, however, struggling to predict the upper air flow between the upper troposphere and lower stratosphere. Mathole et al. (2014) recently identified similar deficiency in the ECHAM4.5-MOM3-SA OAGCM. They indicated that the OAGCM was unable to simulate the observed pole ward migration of the eddy driven southern extratropical jet stream and lower stratospheric cooling which is presumably attributed to the lack of proper stratospheric ozone prescription, anthropogenic forcings and coarse vertical resolution.

b. Probabilistic forecast verification

Evaluating the model's ability to predict ENSO phases probabilistically provides additional insight into the models ability to capture important modes of variability. The model testing is done here in a setting that mimics a true operational forecasting approach. First we present typical examples of forecast plumes for the 1982 and 1988 El-Niño and La-Niña events respectively as illustrated in Fig. 13, and it can be seen that the coupled model successfully captures the development and maturity of these two typical ENSO episodes.

In a probabilistic verification framework for seasonal forecasting, the observed and predicted fields are often separated into three categories of above-normal, near-normal and below-normal conditions based on pre-defined thresholds emanated from model history (climatology). Despite that ENSO (anomalous and neutral) is a relatively more predictable component of the climate system, results from the near-normal category are omitted here owing to the low skill generally associated with this category in other variables such as surface temperature and rainfall (Van den Dool and Toth 1991). In addition, the signature of neutral ENSO is not as influential as anomalous ENSO when used as a predictor in a statistical remapping framework (e.g. Landman and Beraki, 2012). This categorization, results in a 2 x 2 contingency table. The contingency tables are subsequently used to compute the reliability diagrams, relative operating characteristics (ROC) curves, area underneath of the ROC curve and other commonly used measures of probabilistic skill such as the Brier (Skill) Score.

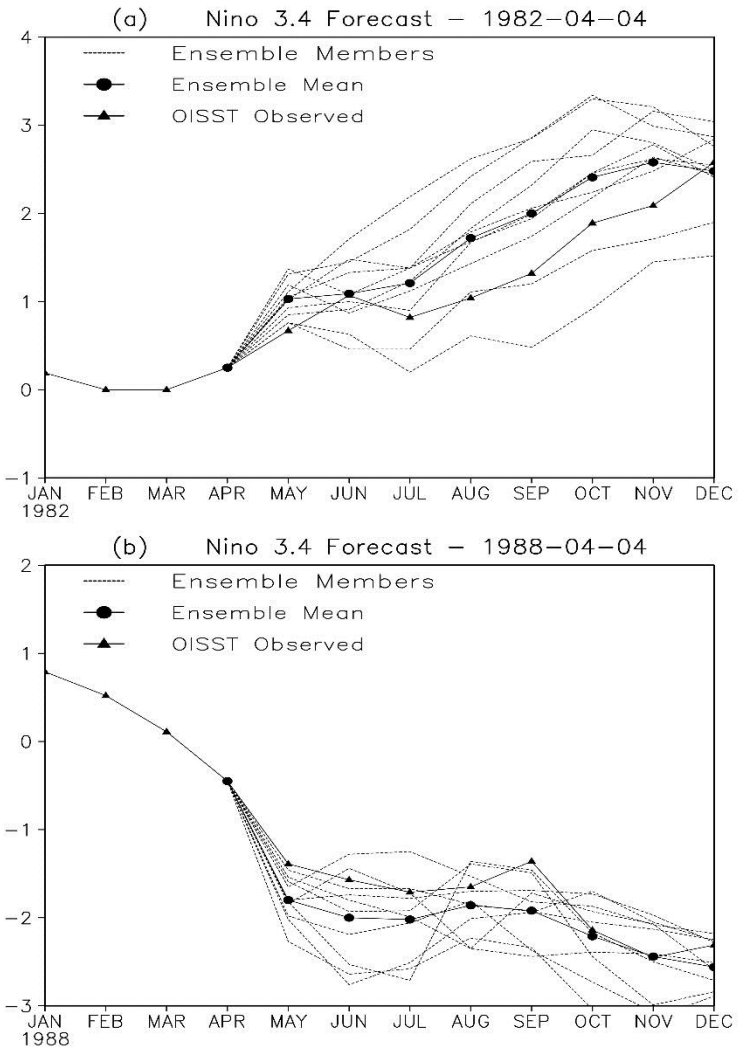

FIG. 13. ECHAM4.5-MOM3-SA forecast plume of Niño-3.4 SST anomalies (K) initialized from the 10 NCEP/DOE atmospheric initial states on the 4th of April 1982 (top panel) and 1988 (lower panel). All members are shown in dot lines, the ensemble mean is solid line marked with closed circle, and the observation is in black line marked with triangle as shown in the legend.

The ROC is a highly flexible method for representing the quality of dichotomous, categorical and probabilistic forecasts (Mason & Graham, 1999). It is derived from Signal Detection Theory (SDT) which was first introduced to the Meteorological community by Mason (1982). The ROC curve (Swets 1973; Mason 1982; Harvey et al. 1992) is derived from a contingency table (Wilks 2006) in which the hit rate and the false-alarm rate are compared. In

probabilistic forecasting system, a warning can be issued when the forecast for a predefined event exceeds some threshold (Mason 1979). Optimally, the ROC curve is desired to lie toward the upper most left corner of a ROC diagram. The diagonal line represents no skill and a curve lays below the no skill line implies that the forecasting system is not better than guess work. The area under the ROC curve is computed numerically and normalized to constitute what is referred to as a ROC score. The ROC score of a skillful forecasting system always exceeds the 0.5 limit.

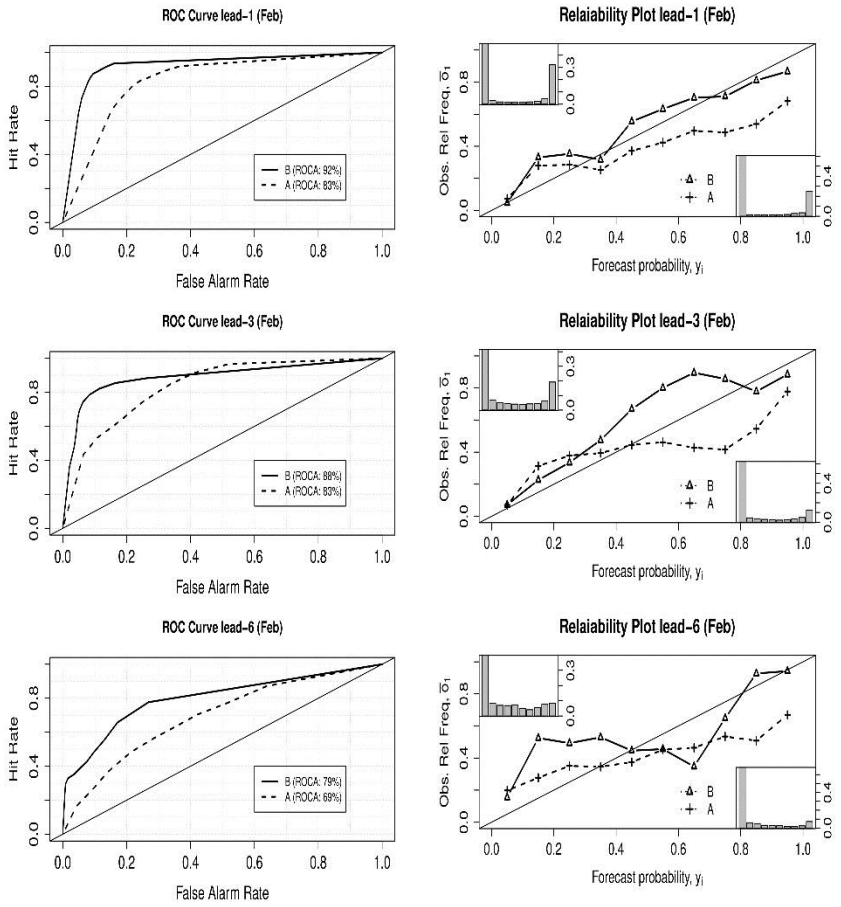

FIG. 14. ROC curve (left panel) and reliability diagrams (right panel) of ECHAM4.5-MOM3-SA probabilistic forecasts that show the warming and cooling phases of ENSO for different lead-times of February as shown on the title of each plot. "B" and "A" in the legend denote La Niña and El-Niño respectively. The histograms on the topleft (cold) and bottom right (warm) corners each reliability diagrams plots imply the frequency of forecast usage in different bins.

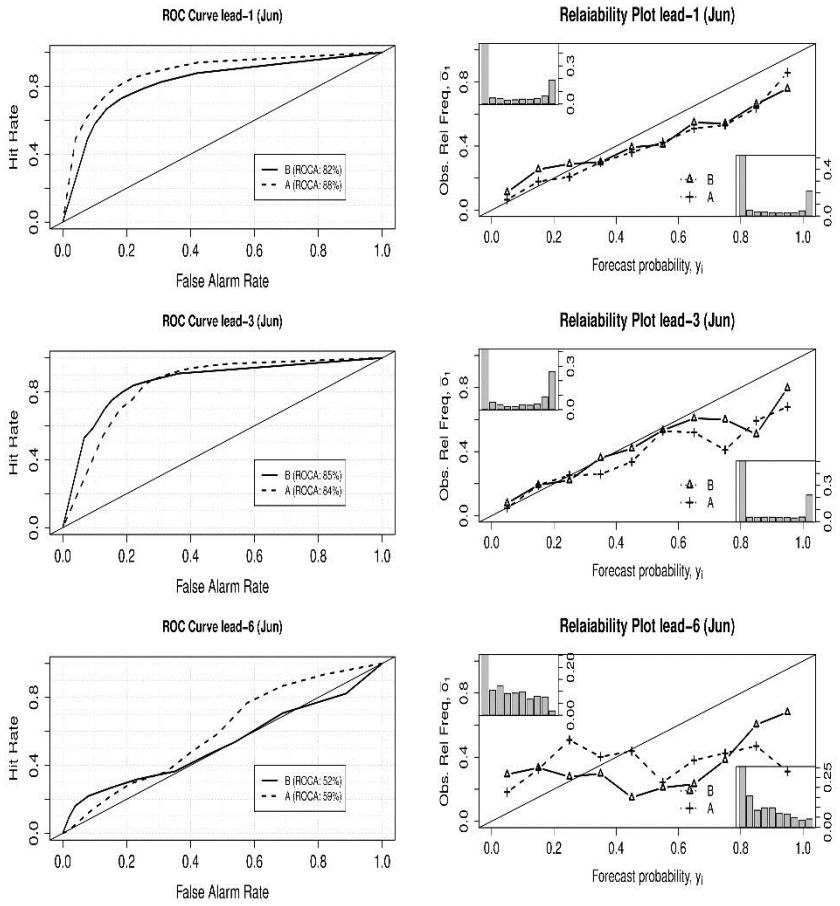

FIG. 15. as in Fig. 14 but June as the target month.

It is worthwhile mentioning, however, that the ROC is not sensitive to biases (systematic or nonsystematic; Murphy 1988) that may be embedded in the forecast system. This implies that a biased forecast can still produce a good ROC curve. It is useful to view the ROC as measure of potential skill and is often accompanied by a corresponding reliability diagrams. Reliability diagram is a type of conditional distribution which shows given each forecast probability interval, how frequently observation actually ended up in one or another category (Hartmann et al. 2002). The reliability diagram is constructed from the computation of the hit rate for the set of forecasts for individual probability bins separately and then plotted against the corresponding forecast probabilities. The most reliable forecasting systems have curves in close proximity of the diagonal line of perfect reliability.

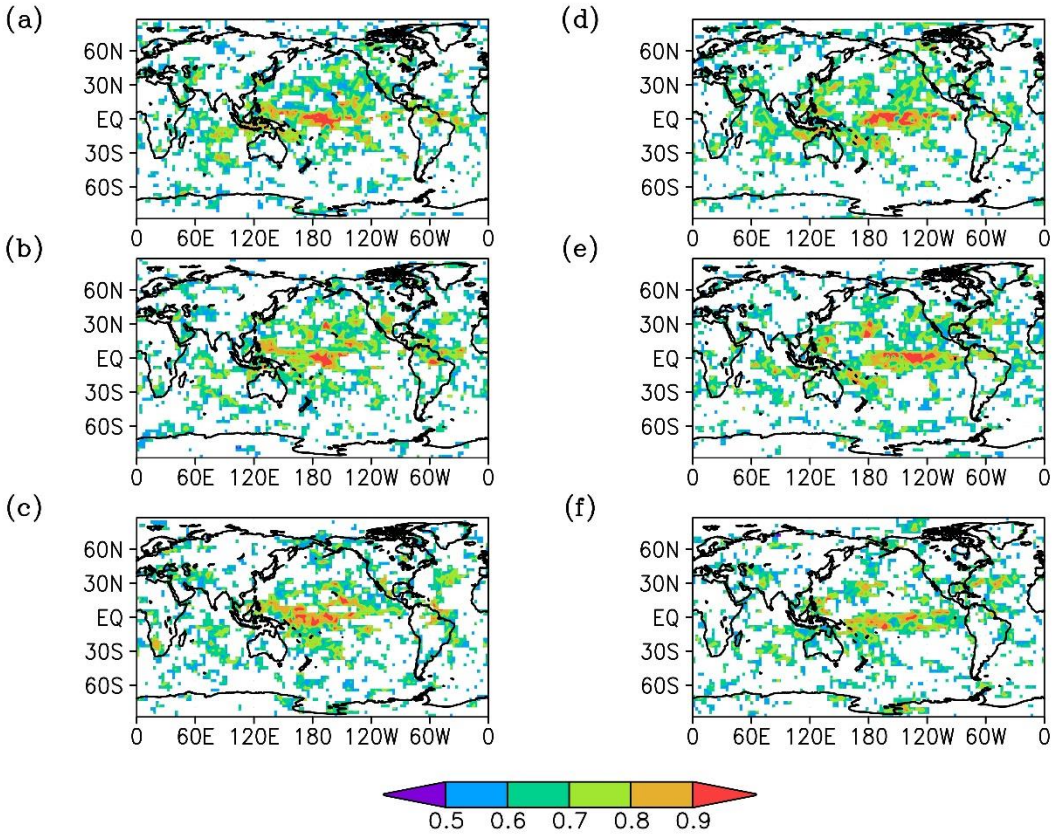

FIG. 16. Global Distribution of ROC area for seasonal rainfall totals (mm) skill of the OAGCM during the austral summer from NDJ (lead-0) to JFM (lead-2) both for below- (a-c) and above-normal (d-f) categories. Only statistically significant values at the 95% level shades are shown.

The ROC and reliability diagrams curves were calculated for each forecast lead-time. Fig. 14 shows diagrams for three lead-months (1-,3- and 6- month lead-times) to describe the model performance. The corresponding frequency histograms showing the relative frequency of use of the forecast bins which are also referred to as “sharpness diagrams“ both for below- and above-normal are shown on the top-left and bottom-right corners of each plot respectively. These histograms reveal how strongly and frequently the issued forecast probabilities depart from the climatological probabilities. At a1-month lead-time, the model exhibits good reliability to predict both the cold and warm phases of ENSO during late austral summer (December) although it shows a tendency of over forecast relative to the latter. It suggests that the quality of forecast manifested

in the ROC curves is attainable as the forecasting system is unbiased where the strength is more robust for the cold phase category. At a 3-month lead-time the model reveals fairly high reliability to predict both cold and warm phases at lower probability bins, but gradually diverges to be over and under forecast for warmer and colder categories at higher probabilities respectively. The model still has good reliability at a 6 month lead-time in spite of both categories being overpredicted. During the start of the austral winter (Fig. 15), the model exhibits high reliability to predict both the cold and warm phases of ENSO at 1- and at 3-month lead-times. Notwithstanding, at a 6-month lead-time the reliability of the model becomes weak. The deterioration of skill at this lead-time and longer is also captured in the AC (Fig. 1) and is presumably attributed as suggested earlier to the spring barrier. It is more evident from both Fig. 14 and Fig. 15 that in the Niño-3.4 region, cold events are more skilfully predicted compared to warm events. These results are similar to that found in previous ENSO predictability studies (Kirtman 2003; D05).

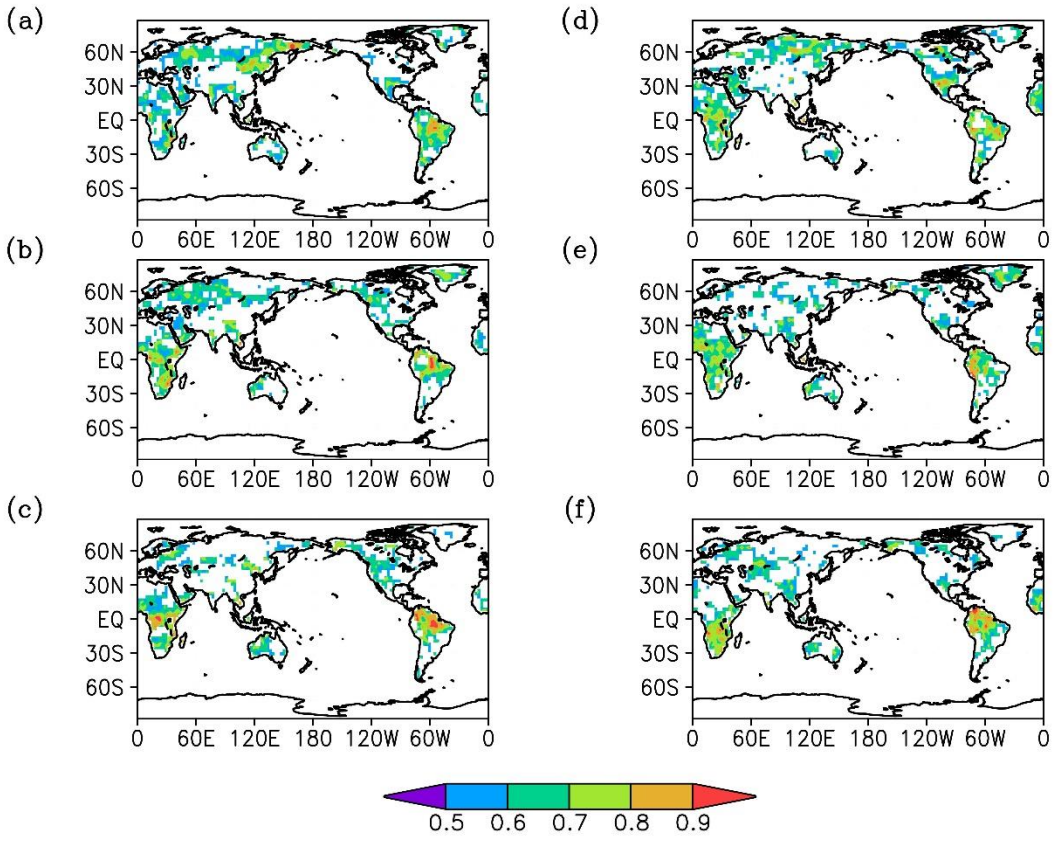

FIG. 17. As Fig. 16 but for 2m surface temperatures (K).

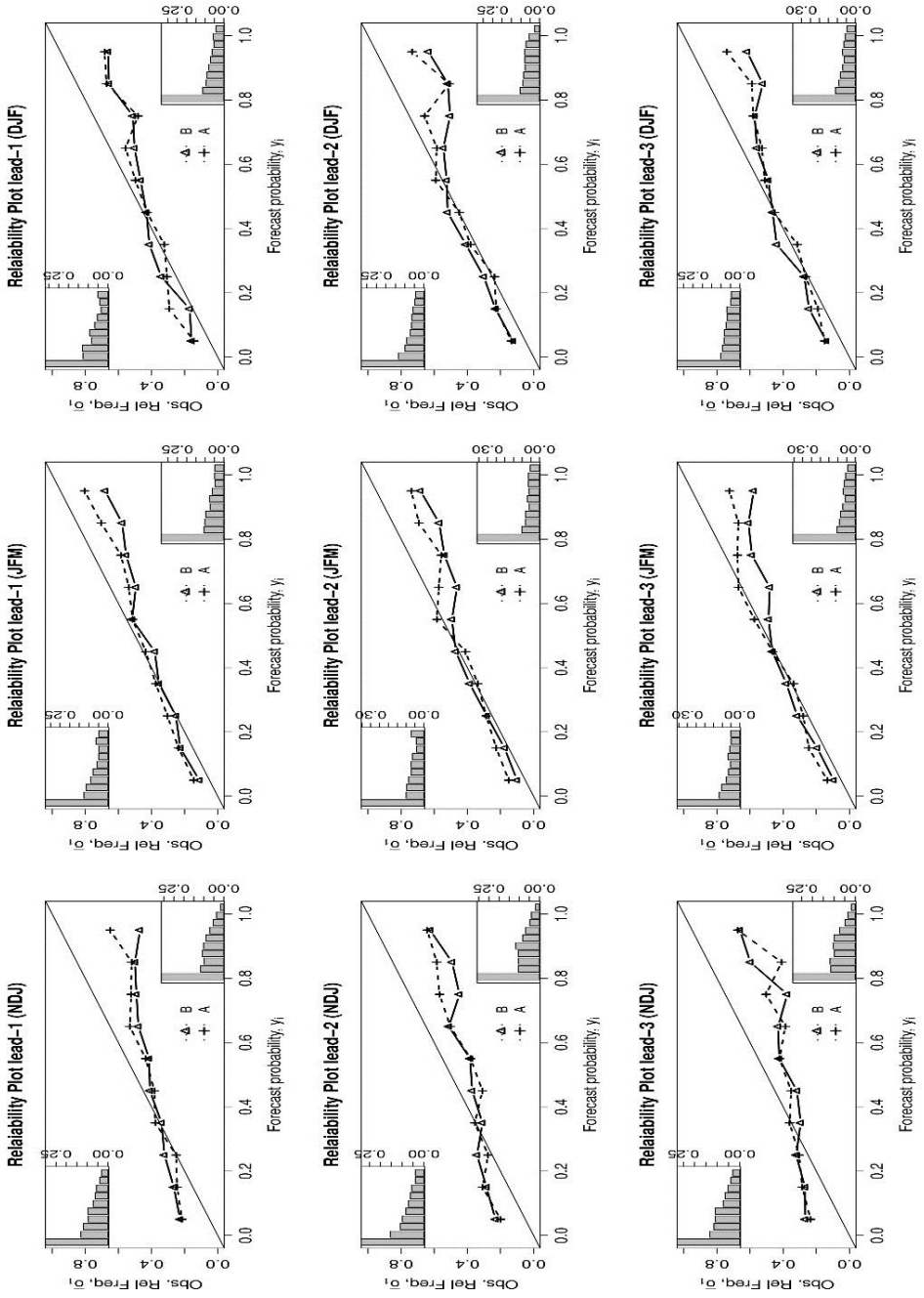

FIG. 18. Reliability diagrams of the OAGCM in predicting below- and above-normal surface air temperature conditions during the three rolling seasons centred around the austral summer season for the southern African region (South of the Equator). "B" and "A" in the legend denote cold and warm events respectively. The frequency of utilization the different probability bins for both below- and above-normal categories are also shown on the left top-upper and bottom-left corners of each diagram

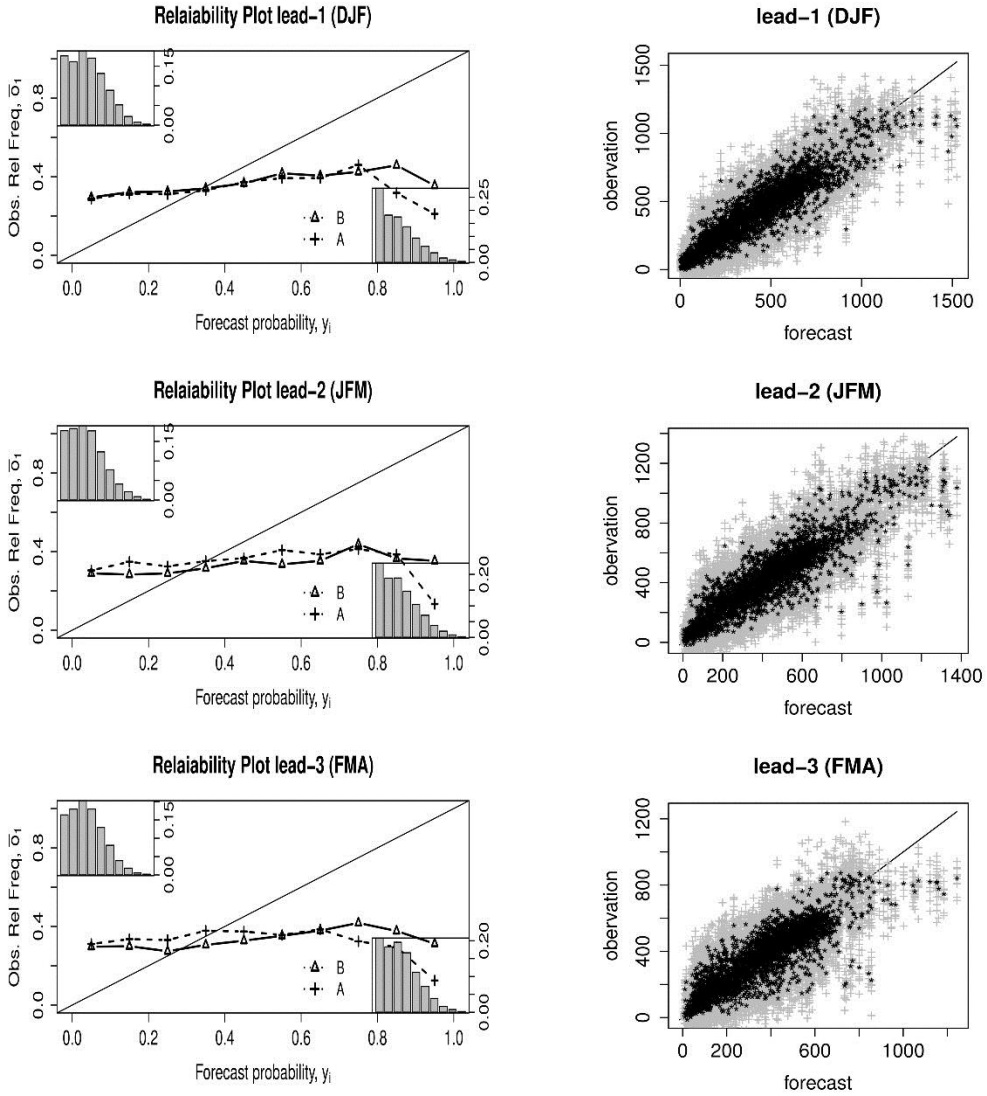

FIG. 19. (left) Reliability diagrams as in Fig 18 but for November initialized hindcasts rainfall totals. The scatter diagrams (right) used the same inputs as in the reliability diagrams; the ensembles members are shown in grey (+) and the ensemble mean is in black (*). No spatial average applied on the fields meaning that each grid point is contributed in each plot.

The global distributions of ROC scores demonstrated by the OAGCM during the austral summer based on the November hindcasts predicting years of wet and dry conditions are shown in Fig. 16. Only those scores which are statistically significant at 95% are shown. The significance

test is conducted using a variant of the Mann-Whitney non-parametric procedure that explicitly accounts for variance adjustment caused by incidents of ties (Mason and Graham 2002; Wilks 2006). It is apparent that the OAGCM is successful in discriminating below- and above-normal rainfall conditions over the larger part of the globe. Maximum skill is obtained on the equatorial Pacific region across all lead-times. Similar skill patterns are demonstrated for the larger part of southern Africa ranging from ROC sores of predominantly 0.6 to patches of 0.8. Similarly, the global surface temperature ROC score distribution of the coupled model is further demonstrated in Fig. 17. This verification result suggests that the model is able to significantly discriminate cold and warm episodes over the larger part of the globe. The performance of the model is more consistent and stronger in predicting surface temperatures than rainfall probabilistically, a result also found in other models (Barnston and Smith 1996; Colman and Davey 1999). These global results also support what has been discussed above with respect to the reliability diagrams for the southern Africa region in the sense that the model is more reliable in providing cold and warm events as opposed to dry and wet.

Fig. 18 shows the reliability diagrams obtained from the different initialized model hindcasts for unusually warm and cold events in the vicinity of the austral summer (DJF) at 1-3 months lead-times. The sharpness diagram both for below- and above-normal are also shown on the top-left and bottom-right corners of each plot respectively. The model is reliably discriminating warm and cold episodes with virtually no skill deterioration as a function of lead-time. At higher probabilities, however, the model exhibits a slight tendency of overconfidence despite the model being increasingly conservative in providing warnings at higher probability bins as shown in the sharpness diagrams. The model also demonstrates similar skill levels during the January to March and the November to January seasons. Notwithstanding, the OAGCM has generally shown a tendency of issuing warnings of certain events while such events (dry or wet condition; Fig. 19) are less frequently observed in the southern African region during the austral summer. The skill of the OAGCM in predicting surface temperature probabilistically, as one may expect, is by far more reliable than the rainfall forecasts where the model generally suffers from overconfidence. Nevertheless, the weakness is presumably attributed to the fact that the model is not equally successful across the whole of southern Africa as shown in Fig. 16. Besides, the overconfidence bias is apparently caused by rainfall conditions of higher seasonal totals (right tail of the scatter

diagram; Fig. 19). The reliability and scatter analysis used identical inputs and both considered the contribution of each grid point and each ensemble member (no spatial average was applied). Previous studies (e.g., Landman et al. 2012; Landman and Beraki 2012) have similarly suggested that the most common slope of the reliability curves found for seasonal rainfall forecasts for the region are shallower than the diagonal line.

5. Summary and conclusions

Coupled climate models represent the state of the art of seasonal forecasting which inherently renders them to be exceptionally convenient for seasonal climate prediction purposes. Notwithstanding, owing to the enormous computational needs of and complexity associated with OAGCMs, their engagement for seasonal forecasts in South Africa was initially not considered feasible particularly in an operational environment.

The substantial augmentation of the computational resources in South Africa due to the resent CHPC intervention brought new hope to South African climate modellers. Founded mainly on this motivation, we attempted to explore the advantage of coupled climate models in the area of research and seasonal forecast production. The emergence of the ECHAM4.5-MOM3-SA OAGCM in South Africa is the first ever locally developed coupled climate model which is configured for seasonal forecasts production. Moreover, it employs an initialization strategy that capitalizes on the best available atmospheric information, thusly making the forecasting system uniquely different from previous coupled models using the same atmosphere and ocean models.

The model evaluation in the context of ENSO forecast showed that the OAGCM was plausibly skillful in most instances in capturing the development and maturity of El-Niño and La-Niña episodes up to an 8 moth lead-time. The result was also complemented by low error concentrations confined within the range of 0.1 to 0.5 RMSE. In a probabilistic sense, the analysis revealed that La-Niña events are more skillfully discriminated than El-Niño events by the model. However, the model skill was generally found to decay faster during the spring barrier.

The model intercomparison revealed that the ECHAM4.5-MOM3-SA OAGCM demonstrated comparable level of skill for the Niño-3.4 region SST forecast with state-of-the-art coupled models administered by other international centres such as the UKMO, MF, ECMWF and CFS-NCEP, IRI and locally developed CCA based statistical model (MMS). The initialization

strategy introduced in the ECHAM4.5-MOM3-SA configuration found to be beneficial when the seasonal variation of ENSO attains its peak as opposed to the D05 version. This result is rather encouraging and further implies that the proposed forecasting system is robust.

Further verification analysis confirmed that the coupled model demonstrated remarkable skill up to several month lead-times in predicting the equatorial IOD during the period spanning between the mid austral spring and the start of the main summer seasons which reaches its peak in November. This may suggest that IOD is more predictable when its seasonal variation becomes strong. The investigation also unveiled that the weakness of the model in other seasons was mainly caused by the western segment of the dipole which eventually contaminates the DMI although the cause of the deficiency is not clear. The complexity of the equatorial IOD prediction reportedly challenges coupled climate models even though observational and theoretical studies conclusively demonstrated the role of the dipole structure in modulating southern Africa and Australian rainfall variability at the seasonal timescale.

The ECHAM4.5-MOM3-SA was also found to be successful in simulating the observed upper air circulation as represented by the 850 and 500 hPa GH in the equatorial belt with a pronounced skill on the latter. Further south, the model was fairly skilful on those key ocean basins such as SPW and SAM despite that the model was mostly unable to outscore a climatological forecast. In addition, the model is fairly skillful in simulating the lower tropospheric and upper stratospheric equatorial flow during the austral summer. Notwithstanding the zonal wind appeared to be more predicable than the meridional wind that might be attributed to ENSO forcing.

The OAGCM probabilistic forecast for the austral summer season for rainfall totals and surface air temperatures was found to be informative and fairly useful. The model was evidently successful in discriminating below- and above-normal rainfall conditions over the larger part of the globe where the signal is more pronounced on the equatorial Pacific region. Similarly, the verification result indicated that the model was able to discriminate cold and warm episodes. Nonetheless, as one may expect, the performance of the model was more consistent and more skilful in predicting surface air temperatures than rainfall totals probabilistically. The findings is further supported, at least for the southern African window, by the fact that the model is more reliable in issuing forecasts of cold or warm seasons as opposed to dry or wet. Probabilistic rainfall forecasts are biased toward overconfidence.

The advent of fully coupled ocean–atmosphere models (e.g., Stockdale et al. 1998) promised improved seasonal forecasts. However, in spite of the promise of enhanced seasonal forecast skill, coupled models have not been administered in South Africa for operational seasonal forecast production because these models effectively require double the computing resources of their atmosphere-only counterparts. Recent advances in computing infrastructures in South Africa and the support from international institutions such as the IRI in developing the coupled model described here have paved the way for utilising and for further development of such state-of-the-art coupled models for seasonal forecast production and research.

Acknowledgments. The work is financially supported by the Water Research Commission (WRC) and Applied Centre for Climate & Earth Systems Science (ACCESS). The Authors are also gratefully appreciative for the CHPC computational support. Uwe Schulzweida of the Max-Planck-Institut für Meteorologie (MPI) has kindly provided the ECHAM4.5 AGCM code. This work has benefitted from comments from Nico Kroese and anonymous reviewers. The work was also impossible without the atmosphere and ocean data assimilation products respectively of NCEP and GFDL and the GFDL ocean model.

REFERENCES

Alves, J. O. S., M.A. Balmaseda, D. L. T. Anderson, and T. N. Stockdale, 2004: Sensitivity of dynamical seasonal forecasts to ocean initial conditions. *Quart. J. Roy. Meteor.* Soc., **130,** 647–668.

Balmaseda, M., and D. Anderson, 2009: Impact of initialization strategies and observations on seasonal forecast skill. *Geophys. Res. Lett.*, **36,** L01701, doi:10.1029/2008GL035561.

Barnston, A. G., and T. M. Smith, 1996: Specification and Prediction of Global Surface Temperature and Precipitation from Global SST Using CCA. *J. Climate*, **9**, 2660–2697.

Beraki, A.F., W.A. Landman, D.G. DeWitt, C. Olivier, K. Mathole, T. Ndarana, 2012: Modelled Sea-Surface Temperature Scenario Considerations and Southern African Seasonal Rainfall

and Temperature Predictability. *In press*.Water Research Commission Rep. xxxx/x/xx, 127 pp.

Brinkop. S., and E. Roeckner, 1995: Sensitivity of a general circulation model to parameterizations of cloud-turbulence interactions in the atmospheric boundary layer. *Tellus*,**47A,** 197-220.

Colman, A. and M. Davey, 1999: Prediction of summer temperature, rainfall and pressure in Europe from preceding winter North Atlantic Ocean temperature. *Int. J. Climatol*., **19,** 513–536. doi: 10.1002/(SICI)1097-0088(199904)19:5<513::AID-JOC370>3.0.CO;2-D

Collins, D.C., C.J. Reason, and F. Tangang, 2004: Predictability of Indian Ocean sea surface temperature using canonical correlation analysis. *Climate Dyn.,* **22**, 481-497.

Conil, S., H. Douville and S. Tyteca, 2009: Contribution of realistic soil moisture initial conditions to boreal summer climate predictability. *Climate Dyn.,* **32,** 75-93.

Derber J., and A. Rosati, 1989: A global oceanic data assimilation system. *J. Phys. Oceanogr*., **19,** 1333–1347.

Déqué, M., 2001: Seasonal predictability of tropical rainfall: Probabilistic formulation and validation. *Tellus*,**53A,** 500–512.

DeWitt, D. G. 2005: Retrospective forecasts of interannual Sea Surface Temperature anomalies from 1982 to present using a directly coupled Atmosphere-Ocean General Circulation Model. *Mon. Wea. Rev*., **133,** 2972–2995.

Doblas-Reyes, F.J., J. Garcia-Serrano, F. Lienert, A.P. Biescas and L.R.L. Rodrigues, 2013: Seasonal climate predictability and forecasting: status and prospects. *WIREs Clim Change* 2013, **4,** 245–268. doi: 10.1002/wcc.217.

Douville, H., 2010: Relative contribution of soil moisture and snow mass to seasonal climate predictability: A pilot study. *Climate Dyn.,* **34,** 797–818.

Dümenil, L., and E. Todini, 1992: A rainfall-runoff scheme for use in the Hamburg climate model. *Advances in theoretical hydrology*. J. P. O'Kane, Ed.,Elsevier, 129-157.

Fan Y., and H. van den Dool, 2004: Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present. *J. Geophys. Res*., **109,** D10102, doi:10.1029/2003JD004345.

Fouquart, Y., and B. Bonnel, 1980: Computations of solar heating of the earth's atmosphere: A new parameterization. *Beitr. Phys. Atmos.*, **53,** 35–62.

Ji, Ming, David W. Behringer, Ants Leetmaa, 1998: An Improved Coupled Model for ENSO Prediction and Implications for Ocean Initialization. Part II: The Coupled Model. *Mon. Wea. Rev.*, **126**, 1022–1034.

Gaertner, M.A., and M. Castro, 1996: A new method for vertical interpolation of the mass field. *Mon. Wea. Rev.*, **124,** 1596-1603.

Gent, P.R., and J. C. McWilliams, 1990: Isopycnal mixing in ocean circulation models. *J. Phys. Oceanogr*., **20,** 150–155.

Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R, Basher, and M. A. Cane, 2001: Current approaches to seasonal-to-interannual climate predictions. *Int. J.Climatol.*, **21,** 1111–1152.

Graham, R. J., M. Gordon, P. J. McLean, S. Ineson, M. R. Huddleston, M. K. Davey, A. Brookshaw, and R. T. H. Barnes, 2005: A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus*, **57A,** 320–319.

Gregory, D., J. J. Morcrette, C. Jakob, A. C. M. Beljaars, and T. Stockdale, 2000: Revision of convection, radiation and cloud schemes in the ECMWF Integrated Forecasting System. *Quart. J. Roy. Meteor. Soc.*, **126,**1685–1710.

Guérémy. J. F., M. Déqué, A. Braun, and J. P. Piedelièvre, 2005: Actual and potential skill of seasonal predictions using the CNRM contribution to DEMETER: coupled versus uncoupled model. *Tellus*, **57A,** 308–319.

Hartmann, H. C., T. C. Pagano, S. Sorooshian, R. Bales, 2002: Confidence builders: Evaluating seasonal climate forecasts for user perspectives. *Bull. Amer. Meteor. Soc.*, **83,** 683–698.

Harvey, L. O., K. R. Hammond, C. M. Lusk, and E. F. Mross, 1992: The application of signal detection theory to weather forecasting behavior. *Mon. Wea. Rev.*, **120;** 863–883.

Kanamitsu, M., A. Kumar, J. K. Schemm, H. M. H. Juang, W. Wang, F. Yang, S. Y. Hong, P. Peng, W. Chen, and M. Ji, 2002: NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteor. Soc*., **83**: 1019-1337.

Kirtman, B. P., J. Shukla, B. Huang, Z. Zhu, E. K. Schneider, 1997: Multiseasonal Predictions with a coupled tropical ocean-global atmosphere system. *Mon. Wea. Rev*., **125,** 789-808.

——, B. P., 2003: The COLA anomaly coupled model: Ensemble ENSO prediction. *Mon. Wea. Rev.*, **131,** 2324–2341.

Komori, N., A. Kuwano-Yoshida, T. Enomoto, H. Sasaki, and W. Ohfuchi, 2008: High-Resolution Simulation of the Global Coupled Atmosphere-Ocean System: Description and Preliminary Outcomes of CFES (CGCM for the Earth Simulator). *High Resolution Numerical Modelling of the Atmosphere and Ocean*. K. Hamilton W. Ohfuchi, Ed., Springer, 241-260.

Koster, R. D., M. J. Suarez, P. Liu, U. Jambor, A. Berg, M. Kistler, R. Reichle, M. Rodell, and J. Famiglietti, 2004: Realistic initialization of land surface states: Impacts on subseasonal forecast skill. *J. Hydrometeor.*, ***5,*** 1049-1063.

Landman, W. A., and S. J. Mason, 2001: Forecasts of near-global sea surface temperatures using canonical correlation analysis. *J. Climate*, **14**: 3819-3833.

——, F. Engelbrecht, A. Beraki, C. Engelbrecht, M. Mbedzi, T. Gill, and L. Ntsangwane, 2009: Model output statistics applied to multi-model ensemble long-range forecasts over South Africa. Water Research Commission Rep. 1492/1/08, 56 pp.

——, and A. Beraki, 2012: Multi-model forecast skill for mid-summer rainfall over southern Africa. *Int. J. Climatol.*, **32,** 303–314.

——, D. DeWitt, D-E. Lee, A. Beraki, and D. Lötter, 2012: Seasonal rainfall prediction skill over South Africa: 1- vs. 2-tiered forecasting systems. *Wea. Forecasting*, **27,** 489-501.

Large, W. G., J. C. McWilliams, and S. C. Doney, 1994: Oceanic vertical mixing: A review and a model with a nonlocal boundary layer parameterization. *Rev. Geophys.*, **32,** 363–403.

Laursen L, and E. Eliasen, 1989: On the effects of the damping mechanisms in an atmospheric general circulation model. *Tellus*, **41A,** 385-400.

Louis, J. F., 1979: A parametric model of vertical eddy fluxes in the atmosphere. *Bound.-Layer Meteor.*, **17,** 187–202.

Majewski, D. 1985: Balanced initial and boundary values for a limited area model. Beitr. *Phys. Atmosph.*, **58,** 147-159.

Mason, I. 1979: On reducing probability forecasts to yes/no forecasts. *Mon. Wea. Rev.*, **107,** 207–211.

——, 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30,** 291–303.

Mason, S.J., and N.E. Graham, 2002: Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: statistical significance and interpretation. *Quart. J. Roy. Meteoro. Soc.*, **128**, 2145–2166.

Mathole, K., T. Ndarana, A.F. Beraki and W. A. Landman, 2014: Assessing the Importance of Lower Stratospheric Processes on the Predictability of Summer Rainfall over South Africa. *S. Afr. J. Sci.*, in press.

McGregor, J. L., 1996: Semi-Lagrangian advection on conformal-cubic grids. *Mon. Wea. Rev.*, **124,** 1311-1322.

Miller, M.J., T. N. Palmer, and R. Swinbank, 1989: Parameterization and influence of sub-grid scale orography in general circulation and numerical weather prediction models. *Meteor. Atmos. Phys.*, **40,** 84-109.

Mo, K.C.n and M. Ghil, 1987: Statistics and dynamics of persistent anomalies. *J. Atmos. Sci*., **44,** 877-901.

Morcrette J-J, L. Smith and emperature dependence of the absorption in longwave radiation parameterizations. *Beitr. Phys. Atmosph*., **59,** 455-469.

Murphy, A.H., 1988: Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Mon. Wea. Rev*., **116,** 2417–2424.

New, M., M. Hulme, and P. D. Jones, 2000: Representing twentieth century space–time climate variability. Part 2: development of 1901–96 monthly grids of terrestrial surface climate. *J. Climate*, **13,** 2217–2238.

Nordeng, T. E., 1994: Extended versions of the convective parameterization scheme at ECMWF and their impact on the mean and transient activity of the model in the Tropics. ECMWF Research Department Tech. Memo. 206, European Centre for Medium-Range Weather Forecasts, 41 pp.

Pacanowski, R. C., and S. M. Griffes, 1998: MOM 3.0 manual. NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, 608 pp.

Palmer, T. N., and Coauthors, 2004: Development of a European Multimodel Ensemble System for Seasonal-to-Interannual Prediction (DEMETER). *Bull. Amer. Meteor. Soc.*, **85,** 853–872.

Pope, V. D., M. Gallani, P. R. Rowntree, and R. A. Stratton, 2000: The impact of new physical parametrizations in the Hadley Centre climate model—HadAM3. *Climate Dyn.*, **16,** 123–146.

Reason, C. J. C., 1999: Interannual warm and cool events in the subtropical/mid-latitude South Indian Ocean region. *Geophys. Res. Lett..*, **26,** 215–218.

——, C.J.C., W. Landman, and W. Tennant, 2006: Seasonal to Decadal Prediction of Southern African Climate and Its Links with Variability of the Atlantic Ocean. Bull. *Amer. Meteor. Soc.*, **87**, 941–955.

——, and H.M. Mulenga, 1999: Relationships between South African rainfall and SST anomalies in the South West Indian Ocean. *Int. J. Climatol*., **19,** 1651–1673.

——, 2002: Sensitivity of the southern African circulation to dipole SST patterns in the South Indian Ocean. *Int. J. Climatol*., **22,** 377–393.

Redi, M. H., 1982: Oceanic isopycnal mixing by coordinate rotation. *J. Phys. Oceanogr*., **12,** 1155–1158.

Reynolds, R. W., N. A. Rayner, T. M. Smith, and D. C. Stokes, 2002: An improved in situ and satellite SST analysis for climate. *J. Climate*, **15,** 1609-1625.

Roeckner, E., and Coauthors, 1996: Simulation of present-day climate with the ECHAM4 model: Impact of model physics and resolution. Max Planck Institute for Meteorology Rep. 93, Hamburg, Germany, 171 pp.

Saha, S., and Coauthors, 2006: The NCEP Climate Forecast System. *J. Climate*, **19,** 3483–3517.

Saji, N. H, B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, 1999: A dipole mode in the tropical Indian Ocean. *Nature* **401,** 360-363.

Schneider, E. K., D. G. DeWitt, A. Rosati, B. P. Kirtman, L. Ji, and J. J Tribbia. 2003: Retrospective ENSO forecasts: Sensitivity to atmospheric model and ocean resolution. *Mon. Wea. Rev*., **131,** 3038–3060.

Sellers, P. J., Y. Mintz, Y. C. Sud, and A. Dalcher. 1986: A Simplified Biosphere Model (SiB) for use within general circulation model. *J. Atmos. Sci*., **43,** 505 – 531.

Shen, R., E. R. Reiter, and J. F. Bresch, 1986: A simplified hydrodynamic mesoscale model suitable for use over high plateau regions. *Meteor. Atmos. Phys.,* **34,** 251-296

Simmons, A. J., and D. M. Burridge, 1981: An energy and angular-momentum conserving vertical finite difference scheme and hybrid vertical coordinates. *Mon. Wea. Rev*., **109,** 758–766.

Simmons, A. J., D. M. Burridge, M. Jarraud, C. Girard, and W. Wergen, 1989: The ECMWF medium-range prediction models: Development of the numerical formulations and the impact of increased resolution. *Meteor. Atmos. Phys*. **40,** 28-60.

Smagorinsky, J., 1963: General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Wea. Rev.*, **91,** 99–164.

Stockdale, T. N., D. L. T. Anderson, J. O. S. Alves, and M. Balmaseda, 1998: Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature*, **392,** 370–373.

——, O. Alves, G. Boer, M. Deque, Y. Ding, A. Kumar, K. Kumar, W. Landman, S. Mason, P. Nobre, A. Scaife, O. Tomoaki, and W.-T. Yun, 2009: Understanding and Predicting Seasonal to Interannual Climate Variability - the producer perspective. World Climate Conference 3, Geneva, Switzerland, 31 August – 4 September 2009.) .

Swets, J. A., 1973: The relative operating characteristic in psychology. *Science,* **182,** 990–1000.

Taylor, K.E., 2001: Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res*., **106**, 7183-7192.

Terray, L., A. Piacentini, and S. Valcke, 1999: OASIS 2.3, Ocean Atmosphere Sea Ice Soil: User's guide. CERFACS Tech. Rep. TR/CMGC/99/37, Toulouse, France, 82 pp. [Available online at http://www.cerfacs.fr/3-25801-Technical-Reports.php.]

Tiedtke, M., 1989: A comprehensive mass flux scheme for cumulus parameterization in largescale models. *Mon. Wea. Rev.,* **117,** 1779-1800.

Undén, P., and Coauthors, 2002: HIRLAM-5 Scientific Documentation, Sveriges meteorologiska och hydrologiska institut (SMHI), 144 pp.

Van den Dool, H., and Z. Toth, 1991: Why do forecasts for ''near normal'' often fail? *Wea. Forecasting*, **6,** 76–85.

Walker, J., and P. R. Rowntree, 1977: The effect of soil moisture on circulation and rainfall in a tropical model. *Quart. .J. Roy. Meteor. Soc*., **103,** 29–46.

Wang, G., R. Kleeman, N. Smith, and F. Tseitkin, 2002: The BMRC coupled general circulation model ENSO forecast system. *Mon. Wea. Rev.*, **130,** 975–991.

Washington, R., and A. Preston, 2006: Extreme wet years over southern Africa: Role of the Indian Ocean sea surface temperatures. *J. Geophys. Res*., **111,** D15104, doi:10.1029/2005JD006724.

Williamson, D. L., and P.J. Rasch, 1994: Water vapor transport in the NCAR CCM2. *Tellus*, **46A**, 34-51

Wilks, D. S., 2006: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 627 pp.

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

Yuan, X., and C. Li, 2008: Climate modes in southern high latitudes and their impacts on Antarctic sea ice, *J. Geophys. Res.*, **113**, C06S91, doi:10.1029/2006JC004067.

Zhao, M., and H.H. Hendon, 2009: Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model, *Quart. J. Roy. Meteor. Soc.*, **135**: 337–352.

Synopsis

The study introduced an ocean-atmosphere coupled climate model which has been suitably configured for seasonal climate forecasting. The development of the CGCM and its operational configuration directly addresses objective 1 of the thesis. The work has demonstrated robustness of the forecasting system through thorough statistical analysis that encompasses both deterministic and probabilistic verification approaches. In addition, the CGCM was compared against other CGCMs by considering climate drivers such as ENSO and IOD as benchmarking. This comparison is an important step towards understanding the relative strength and weakness of the CGCM from an operational point of view which satisfy objectives 3 and 4 relevant to the coupled model. Moreover, modelling configuration is an important step towards achieving the final objective of the project that deals with the comparison between one- and two-tiered forecasting systems. The comparison can be done once the required modelling framework is fully attained with the optimization of a seasonal forecasting system based on an atmospheric-only model as described in the next chapter.

3. The Impact of Optimization on the Predictive Skill of a two-tiered Seasonal Forecasting System

Preface

This chapter consists of journal paper under review and is cited as follows:

Beraki A.F., W. Landman, D. DeWitt and C. Olivier (2015): Global Dynamical Forecasting System Conditioned to Robust Initial and boundary forcings: Seasonal Context. *International Journal of Climatology* (accepted).

Towards addressing the objective 2 of the thesis, the next part of the study explores ways to best optimize an AGCM in order to maximize its seasonal predictive skills from an operational perspective. The analysis further provides an insight into the strengths and weaknesses of the AGCM in an effort towards addressing the objectives 3 and 4 which are relevant to the atmospheric model. The study also endeavours to establish a baseline skill level against which the coupled model, which is reported on in the proceeding chapter, is comprehensively compared in the following chapter. While the paper is addressing the impact of an optimization on the overall predictive skill of the AGCM, the role of predicted SST based on a multi-model approach and the sensitivity of the AGCM to soil moisture initialization are also investigated though pairwise analyses.

The paper was co-authored with Willem A. Landman, David DeWitt and Cobus Olivier. The conceptualisation of the paper, most of the modelling work, data analysis and the actual article writing were done by me.

Global Dynamical Forecasting System Conditioned to Robust Initial and Boundary Forcings: Seasonal Context

Asmerom F. Beraki,[a,c*] Willem A. Landman,[b,c] David DeWitt [d] and Cobus Olivier[a]

[a] South African Weather Service, Pretoria, South Africa Private Bag X097, Pretoria, 0001, South Africa

[b] Council for Scientific and Industrial Research, Natural Resources and the Environment, Pretoria, South Africa

[c] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

[d] National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland, USA

Abstract

We propose how seasonal climate prediction with the use of an atmospheric general circulation model (AGCM) can be optimized. The AGCM predictive skill is extensively examined under various forecast strategies that mimic truly operational prediction. It is shown that the AGCM predictive skill is found to produce superior results given a suitable sea surface temperatures (SSTs) as forcing and is subject to an initialization strategy that uses realistic atmosphere and soil moisture states. Evaluation of hindcasts performed with the model further revealed that the AGCM is able to forecast anomalous upper air atmospheric dynamics (circulation) over the tropics up to several months ahead. The AGCM probabilistic forecasts for rainfall and surface air temperatures during the austral summer season are also found to be informative and useful. The contribution of the predicted sea-surface temperature, which is based on a multi-model approach, is shown to be of significant importance for best AGCM results. The AGCM may also benefit from the initial condition interface in the AGCM's configuration which is implicitly considered in the analysis. Notwithstanding, the AGCM's predictive skill does not vary much whether the AGCM is initialized with realistic or climatological soil moisture which is presumably suggestive of the AGCM's internal weakness..

KEY WORDS: seasonal forecasting, model evaluation, model sensitivity, model initialization, multi-model SST

1. Introduction

The fundamental process driving the global climate system is heating by incoming short-wave solar radiation and the subsequent cooling by outgoing long-wave infrared radiation into space (e.g. Paltridge and Platt, 1976; Goody and Yung, 1996). This heat budget exchange keeps the global temperature of the Earth close to constant (assuming that anthropogenic forcing is neglected), and to a great extend modulates the global atmospheric circulation, which in turn, influences complex smaller scale flow patterns found in the atmosphere. A climate model simulation is therefore an attempt to mimic these complex atmospheric processes that govern Earth's climate with a great deal of simplification. Many aspects of the climate system are not yet well understood, and a significant number of the uncertainties are still directly related to the lack of knowledge of the Earth system (Henderson-Sellers and McGuffeie, 2001).

Numerical Weather Prediction (NWP) may be categorized into different types based on the physical basis they rely on and the timescale they pertain to. In the broader sense, they may cascade on the range of weather forecasting to climate change (Kalnay *et al*., 1998). Weather forecast, at its nascent, is realized from the physical processes inherited to the atmosphere and, to the large extend, it is assumed to be an atmospheric initial condition problem (AMS, 2001). Climate change, on the other hand, refers to the possible shift of the climate system envelope in the future when the atmospheric composition is altered substantially as a result of external forcings notably of anthropogenic origin (Viner, *et al*., 1995).

Numerical Weather Prediction in the seasonal time-scale is of our particular interest here. Seasonal forecasting continues to be a rapidly developing field with considerable effort devoted to developing state-of-the-art global climate models with a great deal of sophistication both internationally (e.g., Stockdale *et al*., 1998; Palmer *et al*., 2004; Graham *et al., 2*005; Saha *et al., 2*006; 2014; Molteni *et al*., 2007) as well as nationally (e.g. Beraki *et al.,* 2014). Historically all these developments stem from the growing scientific evidence on the availability of sufficient physical basis for predicting the mean state of the atmosphere on longer time-scales (Shukla, 1981; 1983; Shukla and Gutzler, 1983; Mason, *et al*., 1999; Barnston *et al.*, 1999). With the emergence

of ensemble methods (e.g. Houtekamer and Lefaivre, 1996; Hansen *et al*., 1997; Vitart *et al*., 1997; Tett *et al*., 1999), climate prediction becomes more appealing and marked the inception of probabilistic dynamical climate forecasts (e.g. Tracton and Kalnay 1993; Toth and Kalnay, 1993; Palmer *et al*., 1993; Molteni *et al*., 1996). Notwithstanding, the science of seasonal forecast practice becomes more noticeable when the slowly evolving boundary conditions notably Sea Surface Temperature (SST) anomalies are discovered to influence the mean state of weather conditions (Palmer and Anderson 1994; Barnston *et al*., 1999). The ability of the coupled climate models in predicting the evolution of (specifically equatorial Pacific Ocean) SSTs with elevated levels of skill up to several months ahead is also an added value (Palmer *et al*., 2004).

In the context of seasonal forecasting, the forecast period, lead-time and persistence issues have a significant importance as far as the quality of a particular forecast assessment is concerned (WMO, 2010). Theoretically, improving the predictability of the mean state of the atmosphere, to the large extent, is expected to arise from the improvement of numerical model's formulations (i.e., dynamical and physical processes; Staniforth and Wood, 2008) and data assimilation (initial conditions; Derber and Rosati, 1989; Moore and Anderson, 1989; Balmaseda *et al*., 2007; Balmaseda and Anderson, 2009). The fast development of both computational technology and the global observational network (particularly with the advent of meteorological satellite information) has an immense contribution to the forecast quality improvements despite that the skill of the forecast deteriorates with the increase of forecast lead-time due to the proportional growth of uncertainties originated mainly from climate forcing imperfections (Lorenz, 1963, Reynolds *et al.*, *1*994; Mason, *et al*., 1999; Goddard *et al*., 2001).

Despite that many leading operational institutions including, inter alia, the Met Office (Arribas, *et al*., 2011), NCEP (National Centers for Environmental Prediction; Saha *et al*., 2014), ECWMF (European Centre for Medium-Range Weather Forecasts; Molteni *et al*., 2011) and BoM (Cottrill *et al*., 2013) use state-of-the-art coupled (ocean-atmosphere) general circulation models (CGCMs) for their seasonal predictions, atmospheric general circulation models (AGCMs) may also provide a feasible alternative presumably with comparable level of skills attainable by the CGCMs (e.g. Boville and Hurrell, 1998; Jha and Kumar, 2009). The study is, therefore, primarily focus on how an AGCM's predictive skill may optimally be maximized under a constrained computational resources environment, a situation commonly found in developing countries

including South Africa and also establish a baseline against which more advanced models can be tested.

To achieve our goal, we use the ECHAM4.5 AGCM (Roeckner *et al.*, 1996) and the study exploits the novel idea of a multi-model approach in establishing an SST forcing field to constrain the AGCM. The advantage of a multi-model approach has been reported in many forecasting studies over recent years (e.g., Krishnamurti *et al.*, 2000; Palmer et al. 2004; Doblas-Reyes *et al.*, 2005; Hagedorn *et al*., 2005; see also Figure 1). Here we make use of that advantage in order to create a "best field" of SST to force an AGCM. In addition, the AGCM configuration employs an initialization strategy that capitalizes on best available information (Balmaseda and Anderson 2009) where daily realistic atmospheric states are used to account for the uncertainties and means of building the ensembles (Beraki *et al*., 2014). For convenience, the term "*data assimilation*" is used here to imply the use of reanalysis products in the model's configuration. Many forecasting studies also use a similar strategy for initialization (e.g., Saha et al., 2014; Arribas, *et al*., 2011; Cottrill *et al*., 2013). Other variants are also commonly applied as a means of deriving perturbed atmospheric initial states to build ensemble prediction systems such as EOF (empirical orthogonal function) based perturbation (e.g., Zhang and Krishnamurti, 1999; Mandonça and Bonatti, 2009), breeding of growing modes (Toth and Kalnay, 1993), singular vectors (Molteni *et al.*, 1996; 2011) and time-lagged average (Hoffman and Kalnay, 1983).

The work is primarily motivated by the fact that the forecasting system contributes to the SAWS Multi-Model System (MMS; Landman and Beraki, 2012) which is being utilized as the vehicle through which SAWS routinely issues the official seasonal forecasts for South Africa, that the ECHAM4.5 is the model through which SAWS maintains its status as one of WMO's Global Producing Centers, and that this AGCM has been proven useful for applications forecast development at SAWS since 2007. In addition the model has been shown to have promising seasonal predictive capability for the Southern Africa region (Landman *et al.*, 2009) and also elsewhere (S. Mason 2015, personal communication).

The paper is structured as follows. In sections 2 we briefly discuss the SAWS climate prediction system from its historical context. The methodology of generating the hindcasts is explained in section 3. In section 4, we evaluate the performance of the forecasting system. The

AGCM's sensitivity to climate forcings is elucidated in section 5. A summary and conclusions are given in section 6.

2. Overview of the SAWS climate prediction system

Locally, seasonal forecast practice started in the early 1990s (e.g., Landman and Mason, 1999; Mason, 1998). The monthly to seasonal predictability of extratropical atmosphere is relatively lower than the tropics owing to strong hydrodynamical instabilities associated with baroclinicity that exits in the middle latitudes (Holton, 2004). In spite of the southern Africa subcontinent being strongly affected by extratropical instabilities, observational and modelling studies succeeded to show the existence of potential predictability (e.g. Klopper *et al*., 1997; Landman and Goddard, 2002; Reason and Rouault, 2005, Tennant and Hewitson, 2002). All these modelling studies are mainly founded on the knowledge of the tropical heat modulation on the mid-latitude circulation particularly during the austral mid-summer when its signature becomes noticeable (Shukla 1981; Mason *et al.,* 1996).

In South Africa, the use of AGCMs in operational seasonal forecasts became more visible when South African based institutions such as the University of Cape Town, University of Pretoria, the Council for Scientific and Industrial Research (CSIR) and the South African Weather Service (SAWS) began to display their forecasts on the website of the Global Forecasting Centre for Southern Africa (www.GFCSA.net) in 2003. See Landman (2014) for a description of some of the seasonal forecasting efforts in South Africa.

Nonetheless, the SAWS experience with regard to the use of AGCMs date back to 1995. In this historical context, the Center for Ocean–Land–Atmosphere Studies global spectral model (COLA; Kirtman *et al*., 1997) was the first model to be operationally introduced on the SAWS first Super Computer (SV1 Cray; Tennant, 2003). After more than a decade of service, the COLA AGCM was concurrently replaced by the ECHAM4.5 AGCM (Roeckner *et al.,* 1996) in 2007 when the Cray Super computer was substituted by the NEC SX8 Super Computer. Since then one of the major advancements in the area of seasonal forecast development in the region includes the SAWS's acquisition of Global Producing Centre for Long-Range Forecasts status from the World Meteorological Organisation (WMO).

With the inception of the Centre for High Performance for Computing (CHPC) in South Africa, the computational infrastructure in this country has grown exponentially. This drastic

computational development also stimulated the emergence of ocean-atmosphere global circulation models (OAGCMs) locally (Beraki *et al*., 2014). In fact, the expansion and optimization of the SAWS global climate prediction system would have been impossible without this computational support. The CHPC was initiated by the Department of Science and Technology and its aim is to provide High Performance Computing (HPC) facilities and expertise for research in South Africa.

3. Methodology

3.1. Model Description

The ECHAM4.5 AGCM is the fourth generation of the Max Planck Institute for Meteorology, Hamburg, Germany. In this experiment, T42 (triangular truncation at wave number 42) horizontal resolution and 19 unevenly spaced hybrid sigma layers are used. The prognostic variables are represented by truncated series of spherical harmonics except for the moisture and trace substances. The model employs the hybrid sigma vertical coordinate system of Simmons and Burridge (1981). The three-dimensional transport of water vapour, cloud water and trace substances is computed with a semi-Lagrangian transports scheme of Williamson and Rasch (1994). The AGCM uses respectively the longwave and shortwave radiation of Fouquart and Bonnel (1980) and Morcrette *et al*. (1986). Cumulus convection is parameterized using the mass flux scheme suggested by Tiedtke (1989) but incorporates the modifications introduced by Nordeng (1994). The turbulent surface fluxes are calculated from Monin–Obukhov similarity theory (Louis 1979), but different from its predecessors, a higher-order closure scheme (Brinkop and Roeckner, 1995) is used to simulate the vertical diffusion of heat, momentum, moisture and cloud water. Horizontal diffusion is computed using the Laursen and Eliasen (1989) scheme. The orographic gravity waves are represented by the wave drag parameterization due to Miller *et al*. (1989). The AGCM uses the simple biosphere model (Sellers *et al., 1*986) and the soil hydrology parameterization scheme suggested by Dümenil and Todini (1992). We refer the reader to Roeckner *et al*. (1996) for a complete model description.

3.2. Retroactive forecasts design

The optimization of the ECHAM4.5 forecasting system is mainly brought about by forcing the model with SST fields retroactively produced through a multi-model approach. In addition, it exploits the initialization strategy that maximizes the use of realistic atmospheric states as noted

earlier. The initial and boundary forcings that collectively explore the uncertainly envelope are used together as a means of building sets of hindcast ensembles. In this process, the AGCM is initialized and forced with slightly different atmospheric initial conditions (ICs) and boundary conditions (BCs) respectively. The uncertainties that arise from the initial conditions are accounted for by taking 10 consecutive daily realistic atmospheric states back from the forecast date in each month and year. For the November hindcasts for instance the atmospheric initial conditions cover the period from October 26 to November 4 for 28 years starting from 1982 to 2009. This combination gives rise to thirty ensemble integrations each consists of 9 months length.

Furthermore, various independent retroactive simulations of the ECHAM4.5 are also undertaken as a function of possible combinations of ICs and BCs. These model simulations are used for comparison and sensitivity analyses (see section 4). In this comparative framework, the SAWS optimized forecasting system, referred to as "ScA" for convenience, is taken as a control and the rest (ScB – ScC; see Table 1 for definitions) may be assumed as perturbations. Notwithstanding, some of these experiments share the same of ICs or BCs. For this reason, it is more appropriate to describe the procedure of generating the different BCs and ICs that form the bases of the different experiments.

3.2.1. Preparation of the boundary conditions

Multi-model ensemble mean SST anomalies are used as input to derive the lower boundary forcing for the AGCM. Hindcasts of SST anomalies obtained from two CGCMs are objectively combined (equal weights applied to ensemble averaged hindcasts) to produce a multi-model ensemble (MME) of global SST anomalies. The CGCMs are the SAWS coupled model (SCM; Beraki *et al*., 2014) and NCEP CFS v2 (Climate Forecasting System Version 2; Saha *et al*., 2014). The advantage of using MME mainly arises from the fact that CGCMs differ in their performance under different conditions and are presumably nearly always better than any of the individual ensembles (Doblas-Reyes *et al.,* 2000; Krishnamurti *et al.,* 2000). In fact, the most striking benefit obtained from MME is the skill-filtering property in regions or seasons when the performance of the individual models varies widely (Graham *et al.,* 2000). Coupled model SST forecasts are used since it has been shown that using such forecasts in order to force the ECHAM4.5 AGCM is an improvement over using statistical SST forecasts to force the model (Landman *et al*., 2014). Figure 1 shows that forecast quality of ENSO is enhanced as a result of model combination, particularly

at the extended lead-times over several months. The latter supports the notion of using MME SST forcing in the AGCM's configuration.

Forcing the AGCM with all ensemble realizations of the participating models and multi-model ensemble mean is computationally inhibiting. First of all, each boundary condition forcing causes a double fold increase in the AGCM integrations by the size of atmospheric ICs as it is also important to account for uncertainties arising from the atmospheric states. Second, it is hardly possible to maintain consistency between a comprehensive set of hindcasts and real-time model forecasts because the number of participating models in the MMS may change as models are either added to or withdrawn from the system. Another drawback with such an initial and boundary value configuration is that deficiencies in the true representation of the uncertainty amplitude can result because these models sometimes tend to reveal a great deal of resemblance among themselves while at other times they are vastly differ from one another. Hence the prescription of the SST scenarios in a manner that optimizes the representation of the uncertainty envelope as well as taking into account computational limitations is vital. Besides, this strategy minimizes the risk of redoing costly hindcast runs because the reconstruction of the retroactive forecasts of the AGCM is only needed when there is a significant shift in the uncertainty envelope of the SSTs themselves rather than changes in the participating models *per se*. To achieve our goal, the AGCM (ScA and ScB; Table 1) is forced with only three SST scenarios which comprise the multi-model ensemble-mean and its amplitude band estimated from the uncertainty term (S. Mason 2010, personal communication).

The uncertainty term was identified independently for all lead-months (lead-0 to lead-8) from the historical standard error. In this context, a slight perturbation was applied on the SST retroactive forecasts using empirical orthogonal function analysis (EOF; Lorenz, 1956; North, 1984). The EOF analysis searches optimum directions that explain the maximum variation that deviates from the reference. In this case, the first normalized EOF mode is retained to describe the uncertainty term assuming that the variance is best explained by the most dominant EOF mode. A similar variant of SST prescription is also used to derive one of the operational AGCMs at the International Research Institute for climate and society (IRI; T. Barnston 2007, personal communication).

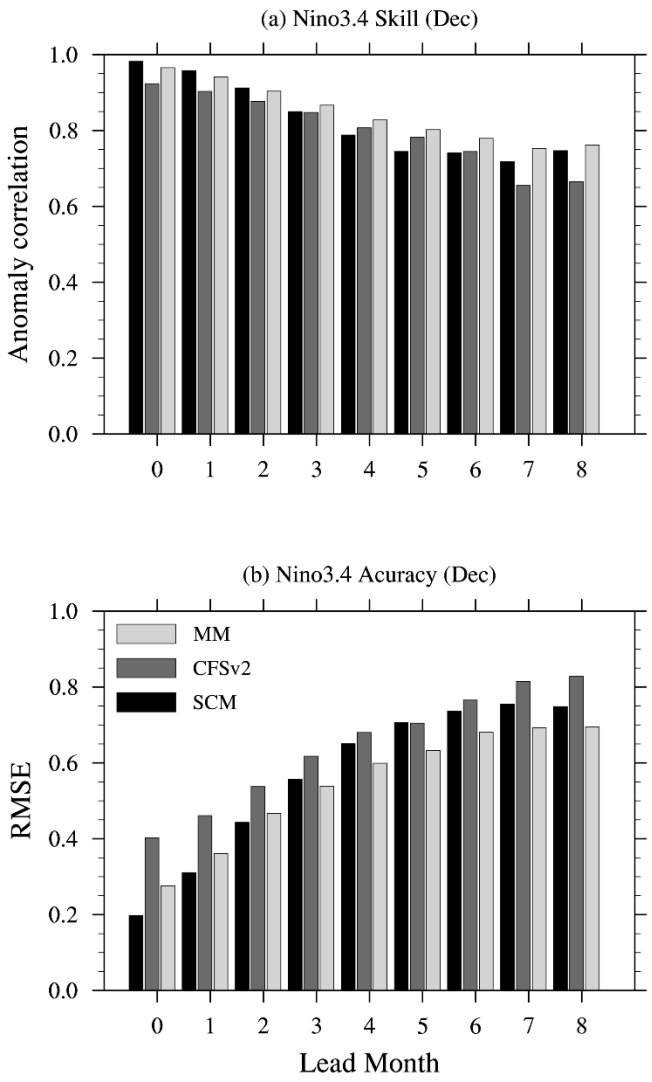

Figure 1. Anomaly correlation (a) and root mean square error (RMSE; b) by various prediction methods of monthly mean for Niño-3.4 forecats as a function of different lead months (horizaontal). The skill scores and level of accuracy are for the beginning of the austral summer (December) when the seasonal variation of ENSO fully attains its climax. The comparion shows how the multi-model (MM) mean SST, used to force the AGCM, improves (degrades) the quality of Niño-3.4 forecats relative to the the two coupled models (see text) used in the combination.

Table 1. ECHAM 4.5 AGCM configuration as a function of ICs and BCs.

Model configuration scenario	Atmospheric State (Initialization)	Land-surface State	Ocean State	Description	Ensemble size
ScA	Real-time NCEP/DOE reanalysis dataset (Kanamitsu *et al*., 2002)	Real-time CPC monthly mean soil moisture (Fan and van den Dool 2004)	Multi-Model SST	With data-assimilation; SAWS new global prediction system	30
ScB	Real-time NCEP/DOE reanalysis dataset (Kanamitsu *et al*., 2002)	CPC Climatology (Fan and van den Dool 2004)	Multi-Model SST	With data assimilation	30
ScC	Real-time NCEP/DOE reanalysis dataset (Kanamitsu *et al*., 2002)	Real-time CPC monthly mean(Fan and van den Dool 2004)	Persistence (OI; Reynolds, *et al., 2*002)	With data assimilation	10
ScD	AGCM atmospheric state	Climatology (Claussen *et al*., 1994)	Persistence (OI; Reynolds, *et al., 2*002)	Without data assimilation; The SAWS phased out seasonal prediction system	6
ScE	AGCM atmospheric state	Climatology (Claussen *et al*., 1994)	AMIP-2 (Taylor *et al.*, 2000)	Without data assimilation; Baseline integration	6

To discriminate the contribution of the multi-model SST forcing to the overall predictive skill of the AGCM, a different model simulation is performed with persisting observed SST anomalies taken from Optimum Interpolation version 2 (OI; Reynolds, *et al., 2*002) as lower

boundary conditions to the AGCM. In this regard, experiments ScA and ScC are forced respectively with the MME and persisting observed SSTs while the AGCM uses the same atmospheric and soil moisture initialization strategy in both experiments. In addition, the previous forecasting system of SAWS (ScD) is similarly constrained with persisting observed SST anomalies.

The baseline skill of the model is delineated using the AMIP-2 (Gates, *et al.*, 1999) type simulations where the AGCM is forced by the lower boundary conditions generated from a high resolution AMIP-2 SST and Sea-Ice (Taylor *et al.*, 2000) informed by the AGCM's atmospheric state (ScE). However, all the other model configurations noted earlier use AMIP-2 Sea-Ice climatology.

3.2.2. Preparation of the initial conditions

The atmospheric initial conditions are acquired from the NCEP, Department of Energy (DOE) Atmospheric Model intercomparison Project (AMIP) II Reanalysis (R2) dataset (Kanamitsu *et al.*, 2002). The NCEP/DOE atmospheric states are transformed to the horizontal and vertical resolution (T42L19) of the ECHAM4.5 AGCM in a manner that maintains numerical and gravitational stability as explained in Beraki *et al.* (2014). Experiments ScA, ScB and ScC (Table 1) use the NCEP/DOE atmospheric states except that the lower layer atmospheric temperature over the ocean is assimilated from the respective SSTs as described in section 3.2.1.

In order to quantify the benefit arising from the use of realistic atmospheric states to initialize the model, the AGCM is also constrained with slightly different atmospheric states taken from the AGCM's simulations (in experiments ScD and ScE) performed using a time-lagged average initialization method (Hoffman and Kalnay, 1983) forced with persisted SST anomalies which essentially means that the model restarts from its own atmospheric state (without data assimilation). The latter configuration may have the advantage of having stable model simulations as the risk of initialization shock is significantly minimized (Balmaseda and Anderson, 2009). Despite that the atmospheric initial conditions become less important as the lead time increases (Goddard *et al.*, 2001), as noted earlier, the fast development of both computational technology and the global observational network (particularly with the advent of meteorological satellite information) may lead to a significant improvement in forecast quality because the predictability of the mean state of the atmosphere is also expected to benefit from the improvement of the

optimal estimate of the state of the climate system (Balmaseda and Anderson, 2009; Doblas-Reyes *et al., 2*013).

The AGCM land surface model is initialized with observed soil moisture states since many studies suggested the role of soil moisture initialization on the skill of climate models (e.g., Walker and Rowntree, 1977; Koster *et al*, 2004; Conil *et al*, 2009; Douville, 2010). In spite of the main motivation of this modelling study being the optimization of the forecasting system for predictive skill in an operational context, the sensitivity of the AGCM to soil moisture initialization is also contrasted using two independent model hindcasts (ScA and ScB; Table 1). The hindcasts are reproduced using the same information of atmospheric state and SST forcings while the model was forced with observed climatological and real-time soil moistures. The soil moisture is obtained from the Climate Prediction Center (CPC) monthly mean dataset (Fan and van den Dool, 2004). Experiments ScD and ScE, however, use climatological soil moisture described by Claussen *et al.* (1994).

3.3. Verifying data

The model surface and upper air data are compared against the gridded data derived from different sources of observations. For the surface variables, seasonal or monthly rainfall totals and air temperatures were acquired respectively from CMAP (CPC Merged Analysis of Precipitation; Xie and Arkin, 1997) and the Climatic Research Unit (CRU; New *et al.*, 2000). For the upper air data analysis, the NCEP/DOE (Kanamitsu *et al.*, 2002) is used as a proxy for observation.

4. Retroactive forecast skill

The AGCM's performance is explored for different seasons along with several lead-times. The verification is based on 10080 (12 months x 28 years x 30 ensemble members) hindcasts each consisting of 9-month integrations in order to establish forecast lead-times of up to eight months. The model runs are grouped according to their proximity to the forecast date (as if they were issued in real-time operational forecasts) to a set of hindcasts with 30 ensemble members. Each ensemble set mimics a set of operational forecasts issued on the 4th of each month starting from 1982 to 2009 as described in Beraki *et al*. (2014). The model bias in the mean annual cycle was removed from the model forecasts prior to comparing the statistics, that is, computing the anomalies of the model

about its own drifted climatology as a function of different initialization time and lead-months (Wang *et al.*, 2002; Schneider *et al.*, 2003; DeWitt, 2005).

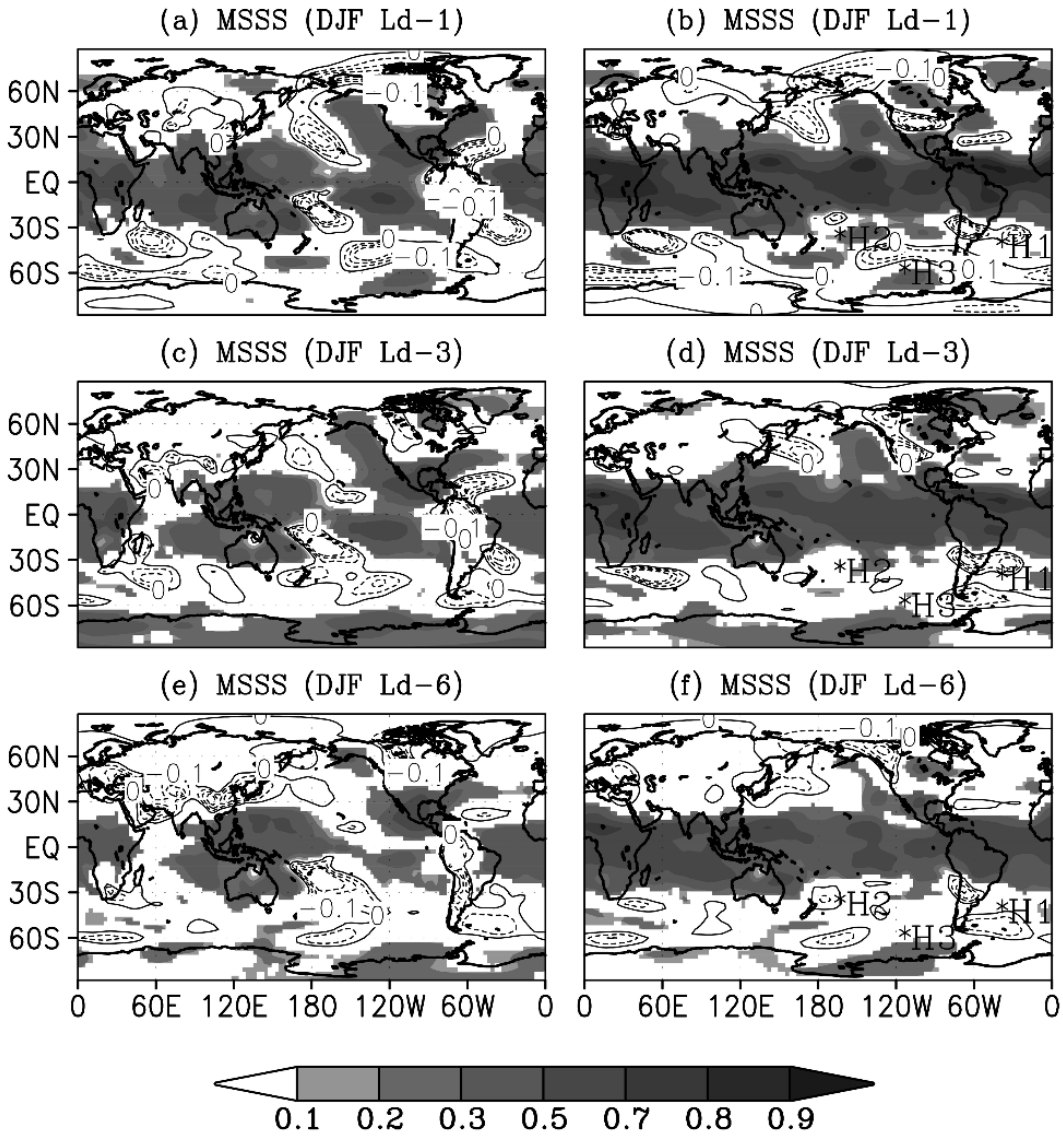

Figure 2. Actual skill of the AGCM during the austral summer (December-January-February; DJF) both for 850 hPa (left panel) and 500hPa (right panel) geopotential heights. (a,b) 1-month lead-time, (c,d) (3-month lead-time) and (e,f) 6-month lead-time. The MSSS is computed against the NCEP/DOE upper air climate data. Shades show statistical significance at the 95% level. The sginficance test is performed with a bootstrap non-parametric procedure (Wilks, 2006). The asterisks indicate the three centre of actions of PSA (H1, H2 and H3; Yuan and Li, 2008).

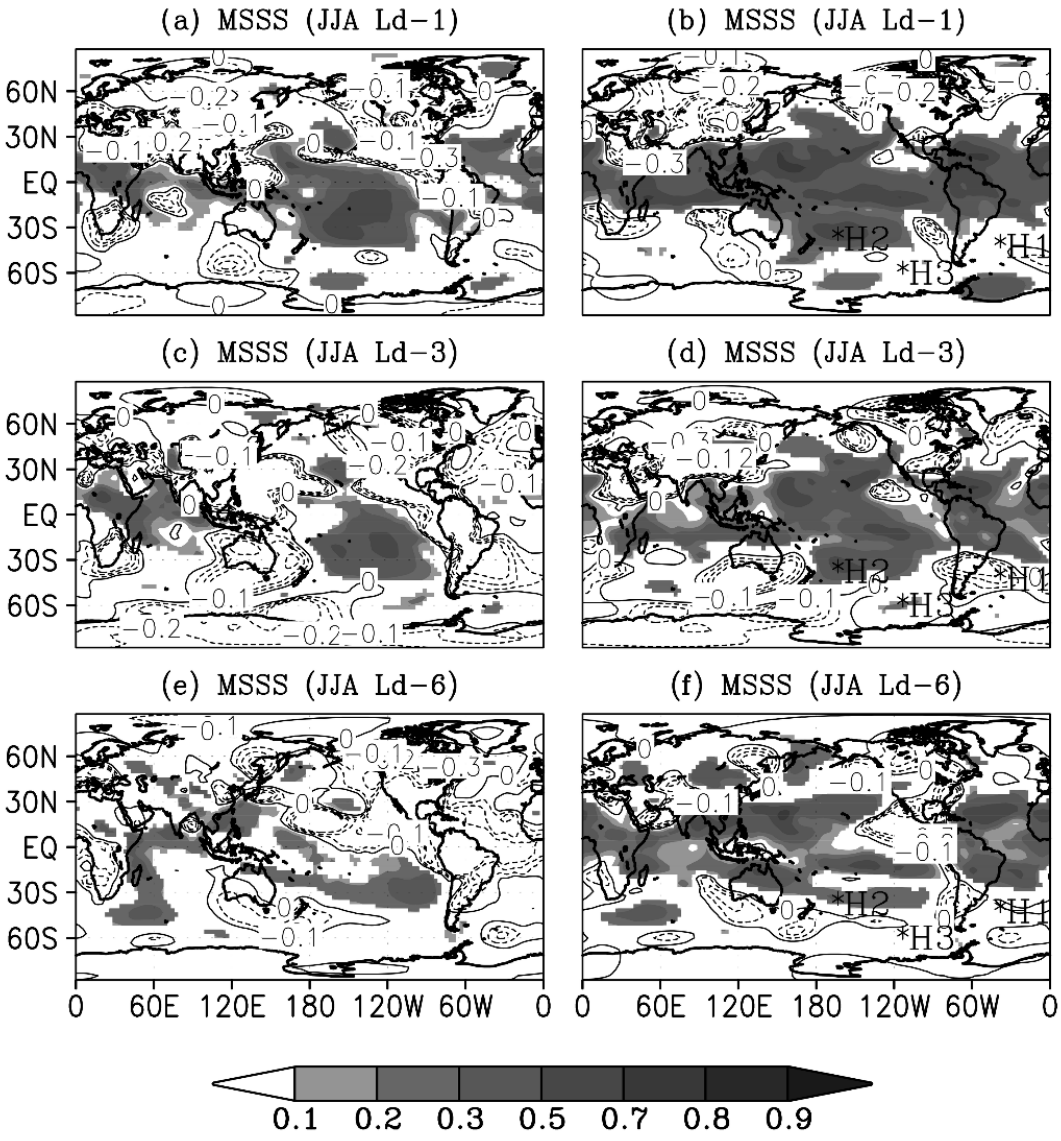

Figure 3. As in Figure 2 but austral winter (June-July-August; JJA).

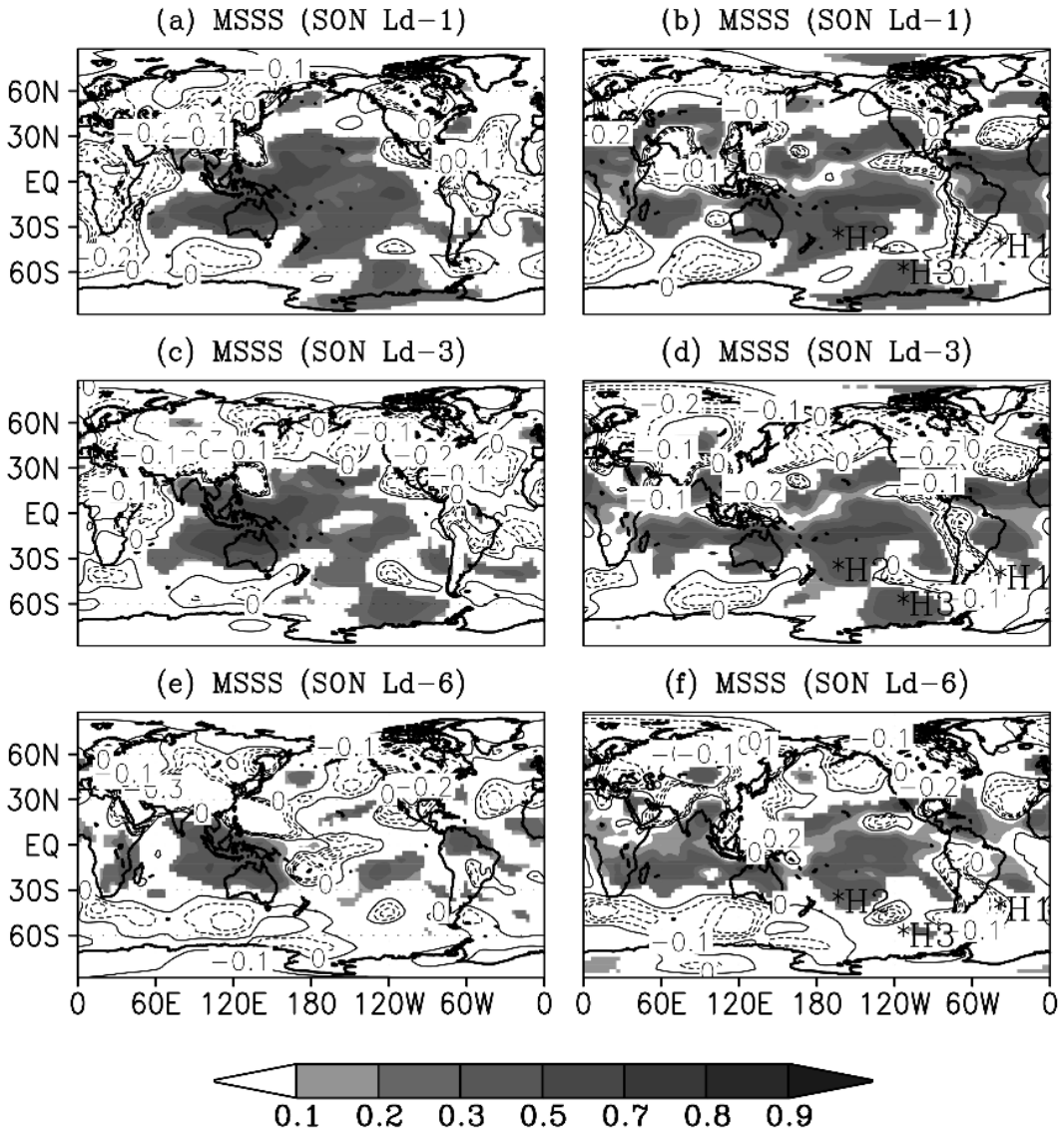

Figure 4. As in Figure 2 but austral spring (September-October-November; SON).

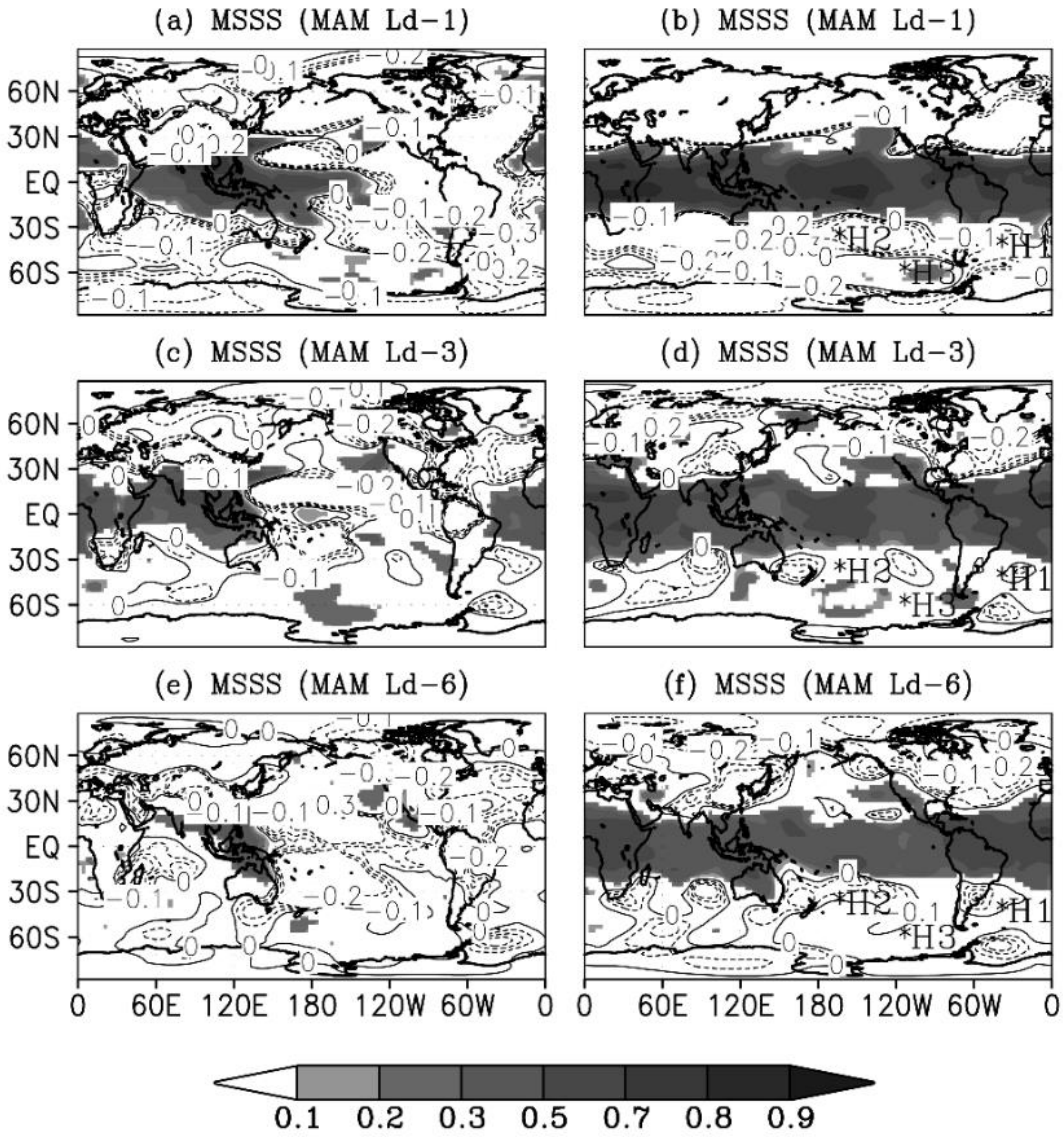

Figure 5. As in Figure 2 but austral autumn (March-April-May; MAM).

We first analyze upper air fields of the model using a flexible and appealing verification score referred to as the mean square skill score (MSSS; Murphy, 1988). The MSSS may be computed using the mean square error [i.e., MSSS = 1-MSE_f/MSE_r], where the reference (MSE_r) could be provided either by the climatology or a different or previous forecasting system. The skill score therefore represents improvement (degradation) in the forecast skill relative to the reference forecast. In this context, the MSSS approaches one if the actual forecast perfectly outperforms the reference forecast; likewise, the positive (negative) gradient between the upper and lower bounds

implies the extent to which the forecasting system is superior (inferior) relative to the reference forecast. The MSSS approaches zero when there is marginal improvement in the forecast system. Figure 2 shows the global skill distribution of the AGCM during the austral summer for the geopotential height fields (GH) at 1, 3 and 6 months lead-time. The skill score is computed from the ensemble mean of the model against the equivalent NCEP/DOE data. On a synoptic scale, the model appears to be skilful at simulating the 850 and 500 hPa GH particularly over the equatorial region between about 30°S and 30°N. The model also shows some degree of predictive skill in eastern Indian Ocean and Pacific region during the austral winter (Figure 3) and spring (Figure 4) for 850 hPa GH. In the austral autumn, the skill is more restricted to the equatorial Indian Ocean off the eastern African coast (Figure 5). In Addition, the skill of the AGCM is rigorous and consistent up to 6 months lead-time for the 500hPa GH for all main austral seasons which attains its peak during the DJF (December-January-February). However, the model's 850hPa GH predictive skill was found to decay faster as a function of lead-time. Over the mid-and high-latitude regions, the AGCM's predictive skill is marginal though it reveals statistically significant (at 95%) skills on those key ocean basins presumably associated with those known modes of climate variability such as the Pacific South America (PSA; Mo and Ghil, 1987). The statistical significance test has been performed using a bootstrap nonparametric procedure (resampling with replacement for 1000x; Wilks, 2006). The degradation of skill might be attributed to the extratropical instabilities (such as fronts and subsynoptic-scale storms; Holton, 2004) and is found to limit the skill of seasonal forecast in southern Africa region (e.g., Klopper *et al*., 1997; Landman and Goddard, 2002; Reason and Rouault, 2005, Tennant and Hewitson, 2002).

Furthermore, Figure 6 shows the global skill distribution of the AGCM during both the austral summer and winter for the 200hPa velocity potential at 1, 3 and 6 months lead-time. The AGCM appears to perform better than climatological forecasts notably over the western and eastern Pacific region in both seasons. The AGCM is also able to retain its predictive skill as the lead-time increases although the skill decays at 6 month lead-time during the JJA (Figure 6(f)). Notwithstanding, the AGCM's fidelity is reduced to the climatological forecast mostly over the mid-and high-latitude regions. The result is consistent with the GH analyses (Figures 2 &3). The 200hPa velocity potential contains information concerning the tropical circulation (such as Hadley, Walker and monsoon) driven by different dynamical causes (Tanaka *et al.*, 2004). The tendency

of the AGCM to perform better in the equatorial Pacific region may attributed to ENSO signal and enhance the model's fidelity in predicting tropical rainfall (Lee *et al.*, 2011).

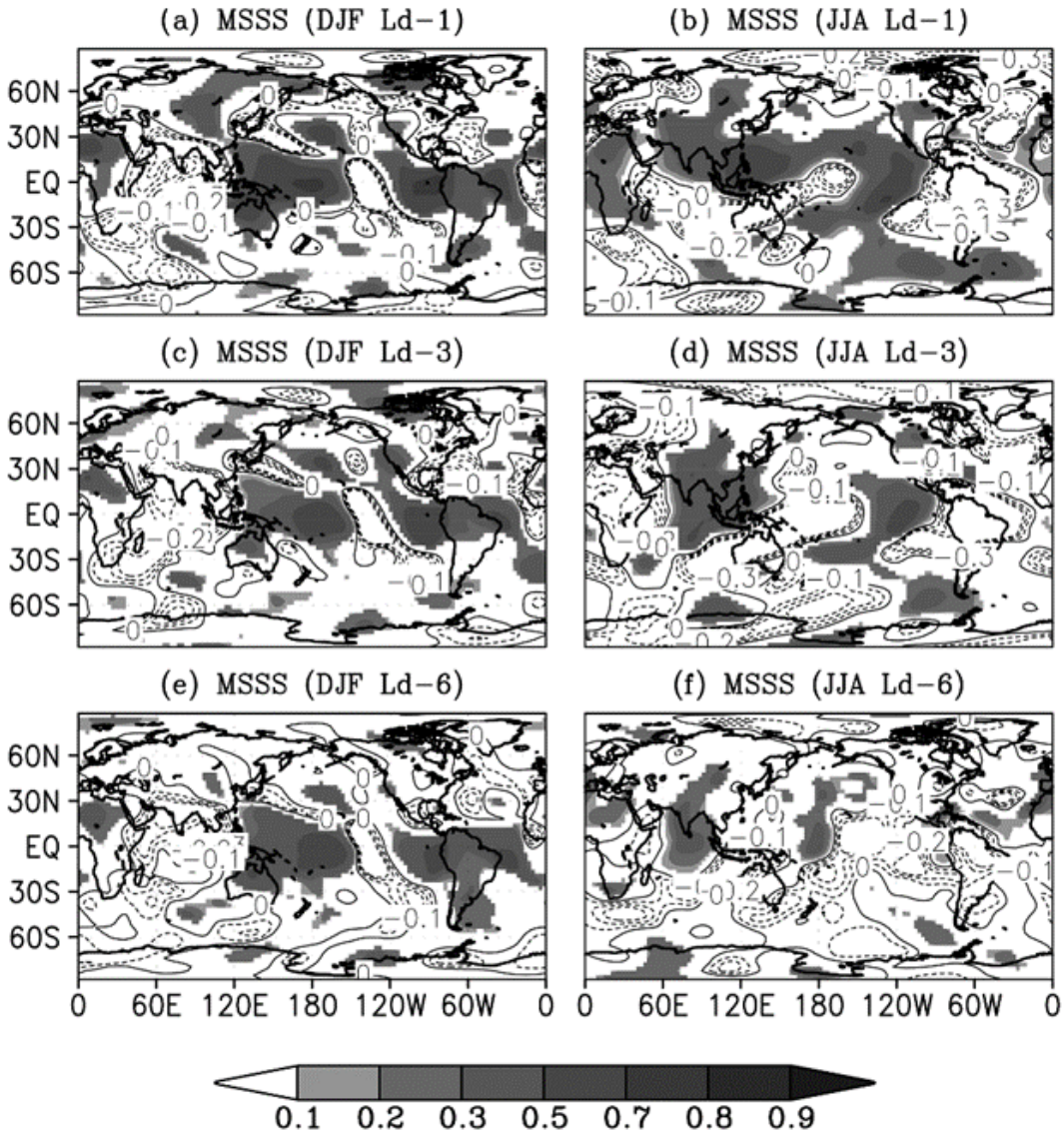

Figure 6. Actual skill of the AGCM in predicting velocity 200 hPa potential during the austral summer (DJF; left panel) and austral winter (JJA; right panel). (a,b) 1-month lead-time, (c,d) (3-month lead-time) and (e,f) 6-month lead-time. The MSSS is computed against the corresponding NCEP/DOE field. Shades show statistical significance at the 95% level.

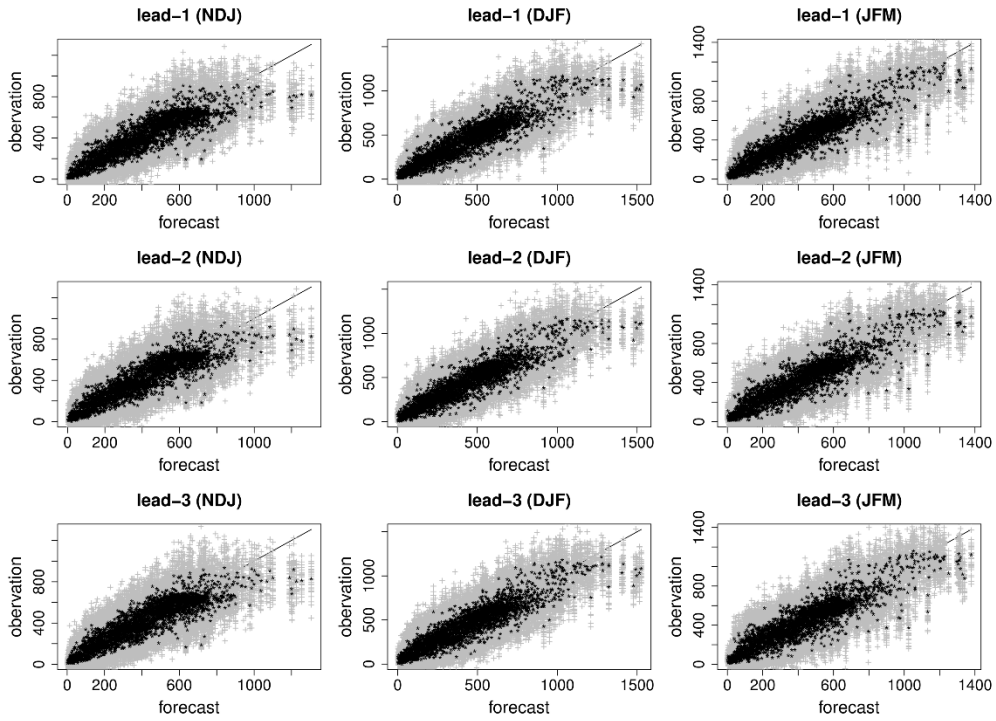

Figure 7. Seasonal rainfall totals in the vicinity of the austral summer for the Southern Africa region modelled versus observed scatter plots. The model forecasts are from 1-3 month lead-time as shown on the title of each scatter plot. The AGCM uses 30 ensemble size respectively; the ensembles member are shown in grey (+) and the ensemble mean is in black (*).

Evaluating the model's ability in predicting rainfall and temperature probabilistically is required since seasonal climate prediction is inherently probabilistic. The model evaluation is performed on model hindcasts generated in a manner that mimics a true operational forecasting configuration as noted earlier. This approach provides a better insight into the model's ability or weakness in an operational context. As a starting point, we examine the southern Africa (south of the Equator) seasonal rainfall distribution in the vicinity of the austral summer (rolling seasons from November to March) using an exploratory method to visually inspect whether the forecasting system is biased. Figure 7 shows the scatter plots the AGCM's simulations plotted against the corresponding observed (CMAP) data. The scatter plots represent each grid point, each time step and all ensemble members in order to suppress any generalization. All ensemble members and

ensemble mean are shown in grey (+) and black (*) respectively. The rainfall scatter plots yield good agreement with the CMAP which implies that the forecasting system is mostly unbiased. The uncertainty band proportionally grows as a function of magnitude (mm) and the cluster deviates gradually from the diagonal line toward the extreme right tail suggesting that the forecast may be biased for rainfall extremes or on those regions which receive seasonal rainfall totals above 800mm. The AGCM forecast shares a great deal of similarity in extent and shape of the distribution as shown in the caption of each plot among the three seasons considered (Figure 7). For an unbiased system, each pair of predicted and observed values tends to cluster around the perfect forecast (diagonal) line.

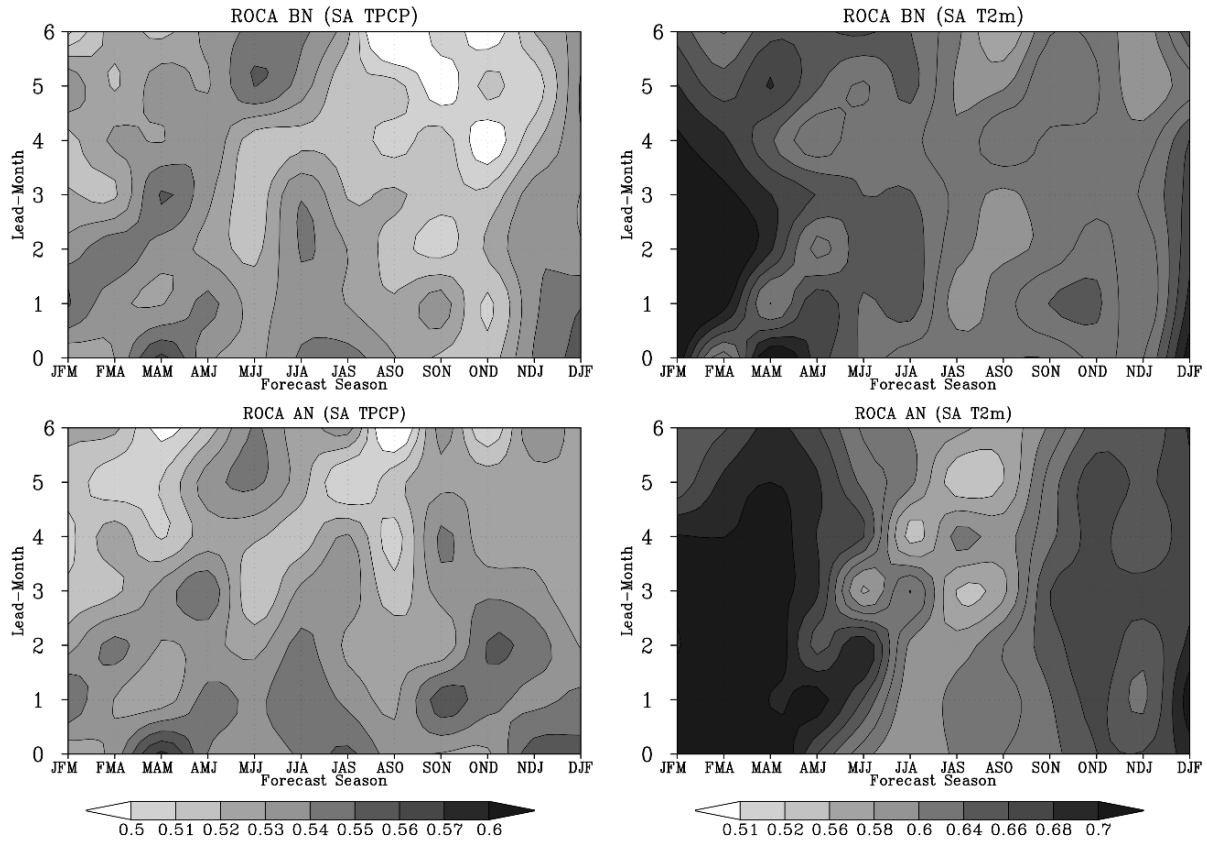

Figure 8. AGCM probabilistic skill for Southern Africa seasonal rainfall totals (left panel) and mean surface air temperatures (right panel) as measured with the ROC area as a function of lead time (vertical) and target (horizontal) over the period 1982–2009.

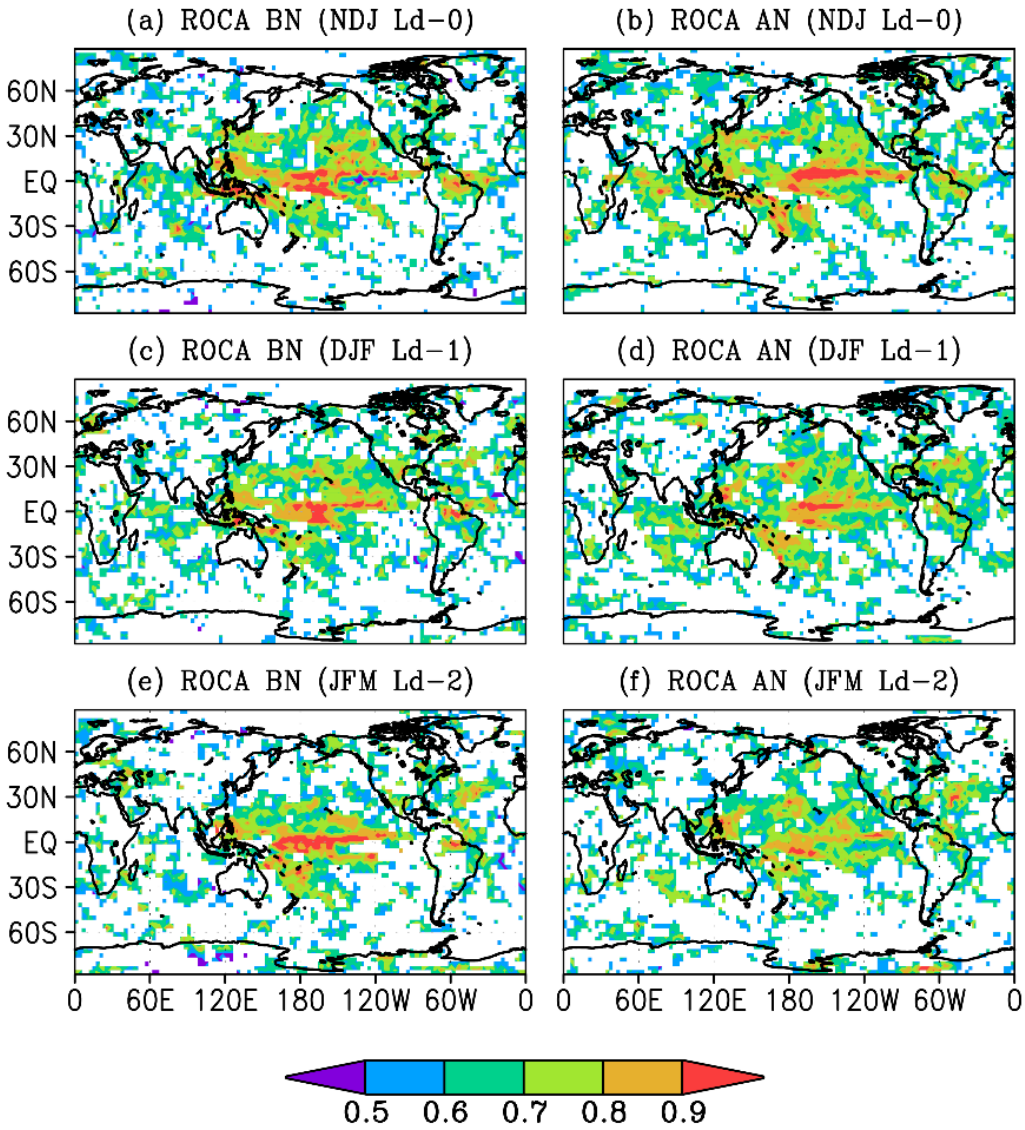

Figure 9. Global potential probabilistic rainfall skill of the AGCM during the austral summer for NDJ (November-December-January; lead-0), DJF (lead-1) and JFM (January-February-March; lead-2) for both below- (a-c) and above-normal (d-f) categories. Only statistically significant at the 95% level shading is shown.

One of the main attributes of interest for probabilistic forecasts is discrimination (are the forecasts discernibly different given different outcomes?) which is tested here through calculating the relative operating characteristic areas (ROC; Mason and Graham, 1999). ROC applied to

probabilistic forecasts indicates whether the forecast probability was higher when an event such as a flood or drought season occurred compared to when it did not occur. ROC scores for the rainfall categories for example represent the respective areas beneath the ROC curve that is produced by plotting the forecast hit rates against the false alarm rates. If the area would be ≤ 0.5 the forecasts have no skill, and for a maximum ROC score of 1.0, perfect discrimination has been obtained. Figure 8 shows the model's ability to discriminate wet (warm) and dry (cold) episodes over the southern Africa region using 12 rolling seasons as a function of lead-time by taking into consideration the contribution of each ensemble member and each grid point as in the scatter diagrams (Figure 7). The model is reasonably skilful for predicting rainfall conditions of the outer terciles particularly in the vicinity of the austral summer. Notwithstanding, the model skill is marginal for below-normal rainfall conditions during the austral spring. For surface air temperatures, maximum skill of the model is concentrated on the austral summer and autumn seasons for below- and above-normal surface temperature conditions. The model skill during the winter season particularly for above-normal is relatively low. It has been shown in previous studies that most of the predictability is found during the mid-summer in the southern Africa region when tropical influences start to dominate the atmospheric circulation across the region, with almost no predictability evident during austral winter and spring when the seasonal rainfall of South Africa is mostly influenced by transient weather systems (e.g. Landman *et al*., 2012).

Further, the global probabilistic skill distribution of the AGCM is assessed for both rainfall and surface air temperatures. The skill scores presented are based on the November initialized hindcasts for the austral summer from 1982-2009. Only those scores which are statistically significant at the 95% are retained. As in Beraki *et al*. (2014), the significance test is conducted using a variant of the Mann-Whitney non-parametric procedure that explicitly accounts for variance adjustment caused by incidents of ties (Mason and Graham, 2002; Wilks, 2006). It is noticeable that the AGCM is skilful in discriminating below- and above-normal rainfall conditions particularly on the equatorial Pacific region consistently across all lead-times (Figure 9). Similarly, the global surface temperature ROC score distribution of the AGCM is demonstrated in Figure 10. The analysis reveals that the model is able to discriminate significantly cold and warm episodes over the larger part of the globe during the austral summer.

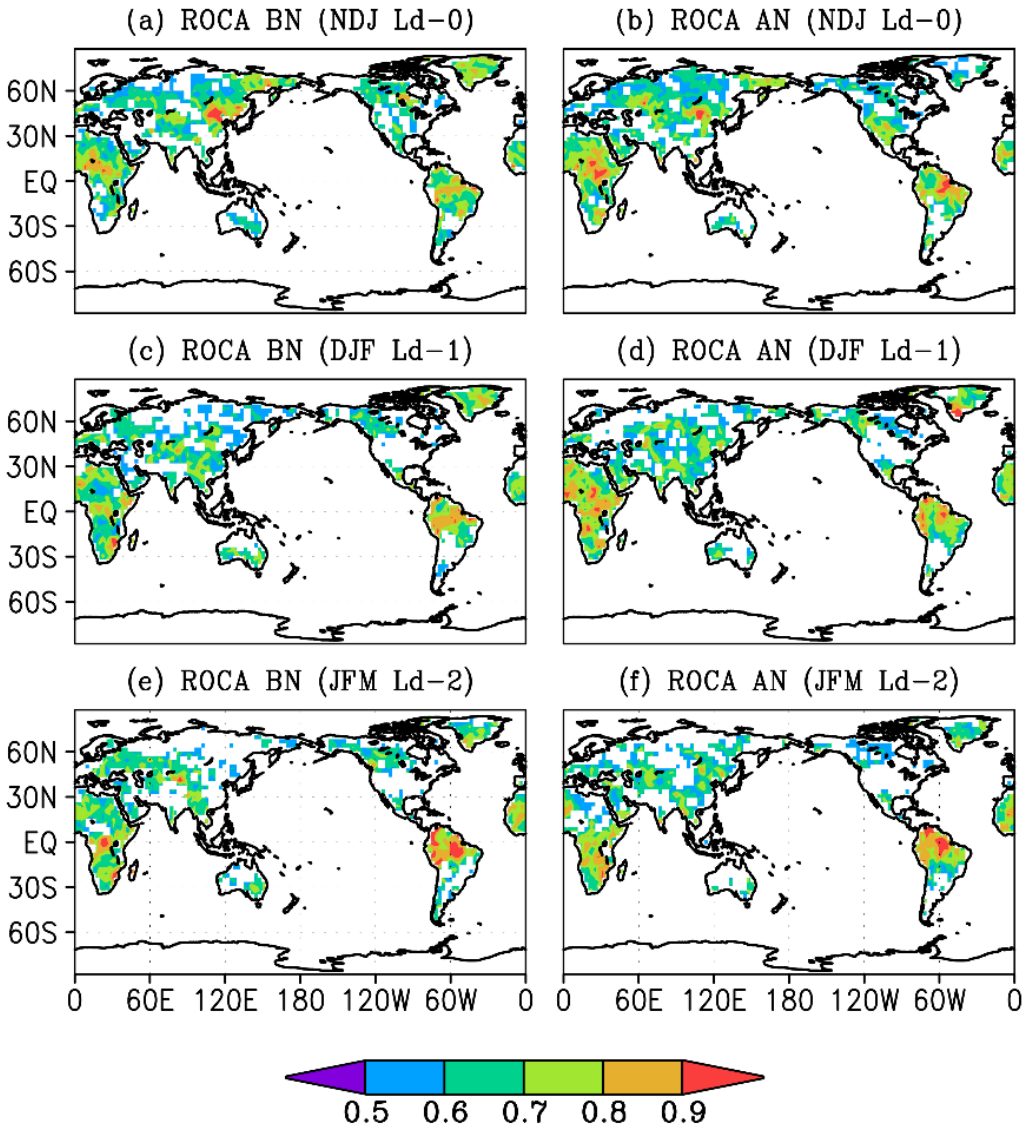

Figure 10. Global potential probabilistic seasonal average 2m temperature (oK) skill of the AGCM during the austral summer from NDJ (lead-0) to JFM (lead-2) both for below- (a-c) and above-normal (d-f) categories. Only statistically significant values at the 95% level are shaded. The analysis over ocean is excluded.

5. Model Sensitivity and performance comparisons

The forecasting system evaluated in section 4 has also been compared with different configurations of the same model including the previous operational forecasting system of the SAWS (2007-2013; ScD) and AMIP-2 (baseline skill; ScE). The comparison is mainly aimed at investigating

how the proposed forecasting system has positioned itself against other possible model configurations of the same model. The comparison aspect with other CGCMs is beyond the scope of the current work as it deserves in depth analysis and we defer it for future work.To achieve our goal, we have performed different retroactive simulations that span a range of possible combinations as a function of ICs and BCs. These combinations range from persisted SST without data-assimilation to multi-model predicted SST forcings with data assimilation options as summarized in Table 1. Meanwhile, pairwise comparisons of these various model simulations allow the characterization of the AGCM's sensitivity and the discrimination of the respective contributions.

Figure 11 shows the extent to which the AGCM's predictive skill has improved (degraded) in predicting 850 and 500hPa GHs, and 200hPa velocity potential attributed to the multi-model SST anomaly forcing and implies that the benefit that could have been missed should the AGCM be forced with persisted SST anomalies. The model uses the same atmospheric and soil moisture initialization procedures in order to isolate the role of the multi-model SST in the predictive skill of the AGCM. According to Figure 11, the model predictive skill appears to significantly (95%) strengthen across the equatorial region for all lead-months considered particularly for the 500hPa GH. However, its benefit is more confined to the equatorial Indian, western Pacific Oceans and Australia for 850hPa GH, and western and eastern Pacific Ocean (with a gradual tendency to extend over the western equatorial Atlantic Ocean as the lead-time increases) for 200hpa velocity potential.

The contribution of soil moisture on the skill of the AGCM in predicting upper level dynamics (circulation) is shown in Figure 12. This actual skill concentration is computed from the two independent AGCM's retroactive simulations constrained with a real-time (realistic) and climatological soil moisture used as forecast and reference. In both cases, the AGCM uses the NCEP/DOE atmospheric states and multi-model SST anomalies as a lower boundary condition to facilitate the sensitivity analysis. The contribution of soil moisture initialization on the general improvement of the AGCM's predictive skill is not significant with the exception of a few scattered patches. Similar pairwise analyses performed on surface air temperature and total precipitations have also produced an indistinguishable signal (not shown). The result suggests that the response of the AGCM to soil moisture initialization is not clearly manifested as reported in

similar modelling studies elsewhere (e.g., Koster *et al.*, 2004; Seneviratne *et al.*, 2010). This AGCM's internal weakness is presumably attributed to the framework used to couple the land surface and atmosphere or the land surface scheme itself and therefore needs additional attention at a later stage.

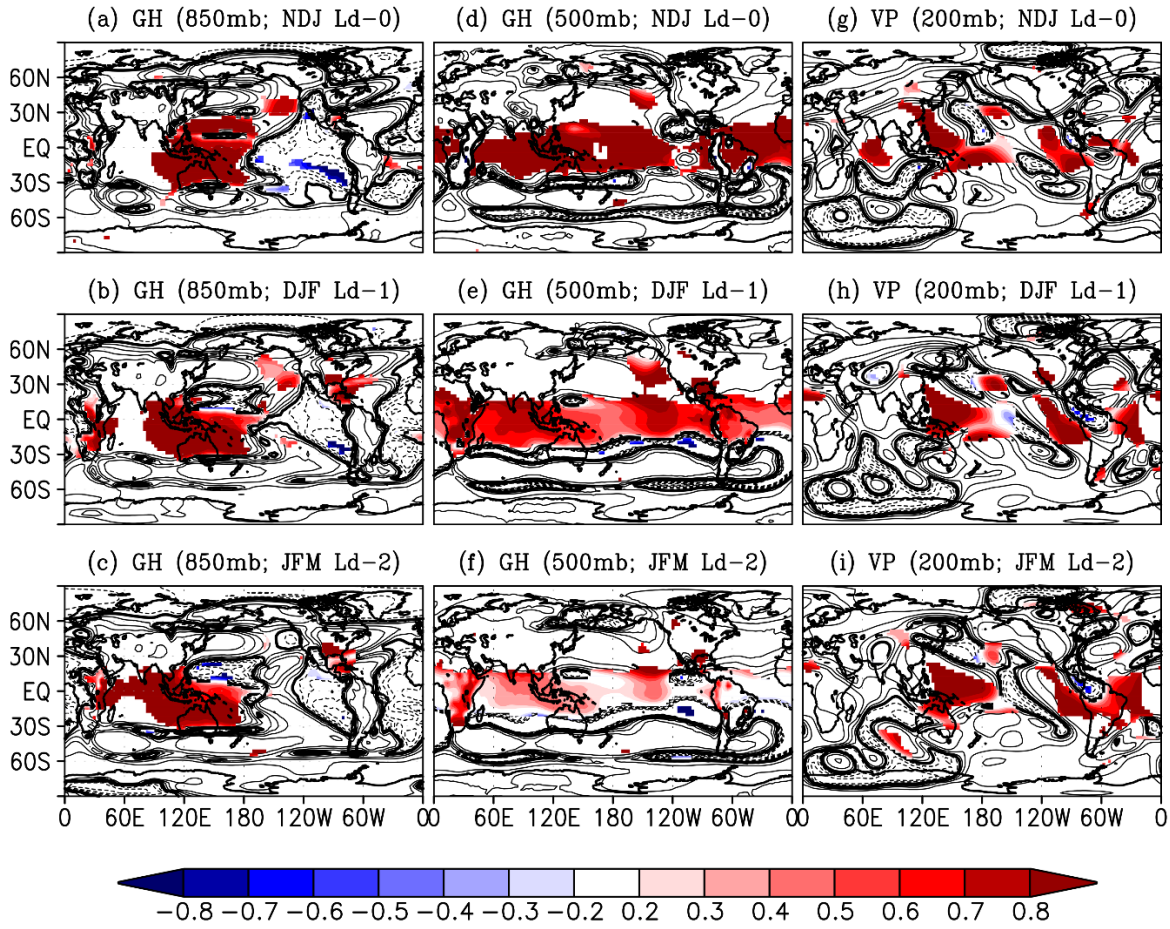

Figure 11. Contribution of multi-model SST anomalies on the predictive skill of the AGCM in predicting both for 850 hPa (left panel) and 500hPa GHs. This skill concentration computed from the AGCM's retroactive integrations respectively forced with a multi-model forecast and persisting observed SST anomalies (the reference forecast). The +ve (-ve) shades imply that the model largely benefits when it was forced with multi-model predicted (persisting observed) SST anomalies while maintaining the two integrations similar in every other respect.

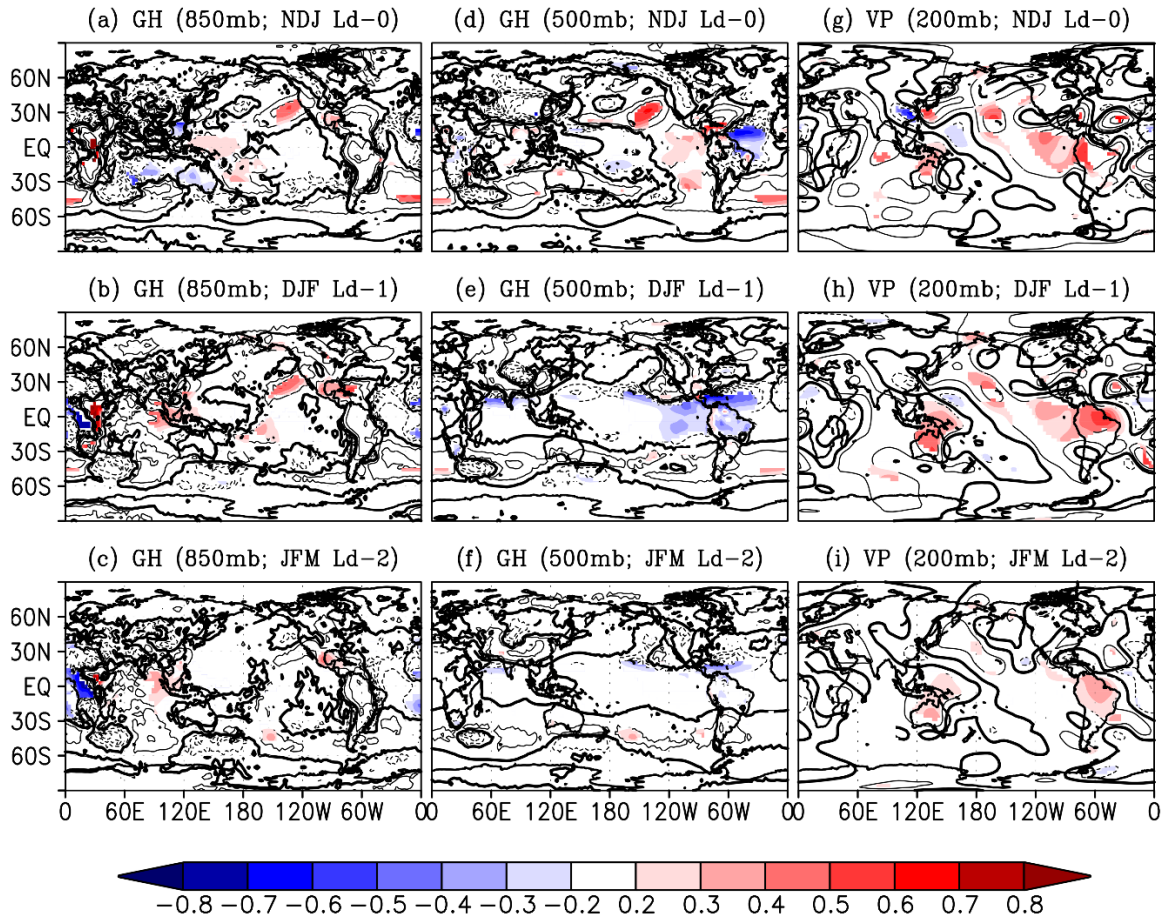

Figure 12. Contribution of soil moisture on the predictive skill of the AGCM in predicting upper level circulation. The AGCM's retroactive integrations respectively forced with a real-time (realistic) and climatological soil moisture forcings were used as forecast and reference respectively. The +ve (-ve) shades imply that the model largely benefits when it was forced with a real-time (climatological) soil moisture while maintain the two integrations similar in every other respect.

The overall skill and accuracy of those various forecast methods outlined in Table 1 in predicting both t2m temperatures and rainfall for the tropical and southern Africa regions at several month lead-times are also computed and placed in the correlation and standard deviation space as depicted using a Taylor Diagram (Taylor, 2001; Figure 13). The result reveals that, in most instances, the optimized forecasting system is found to outperform most of the other model configurations for the variables and regions considered in the analysis consistently across all lead-

times. It is worth emphasising that the two forecast scenarios, denoted by "ScA" and ScB" in Figure 13, remain in close proximity in the correlation and standard deviation space which attests again the AGCM's insensitivity to soil moisture initialization noted earlier.

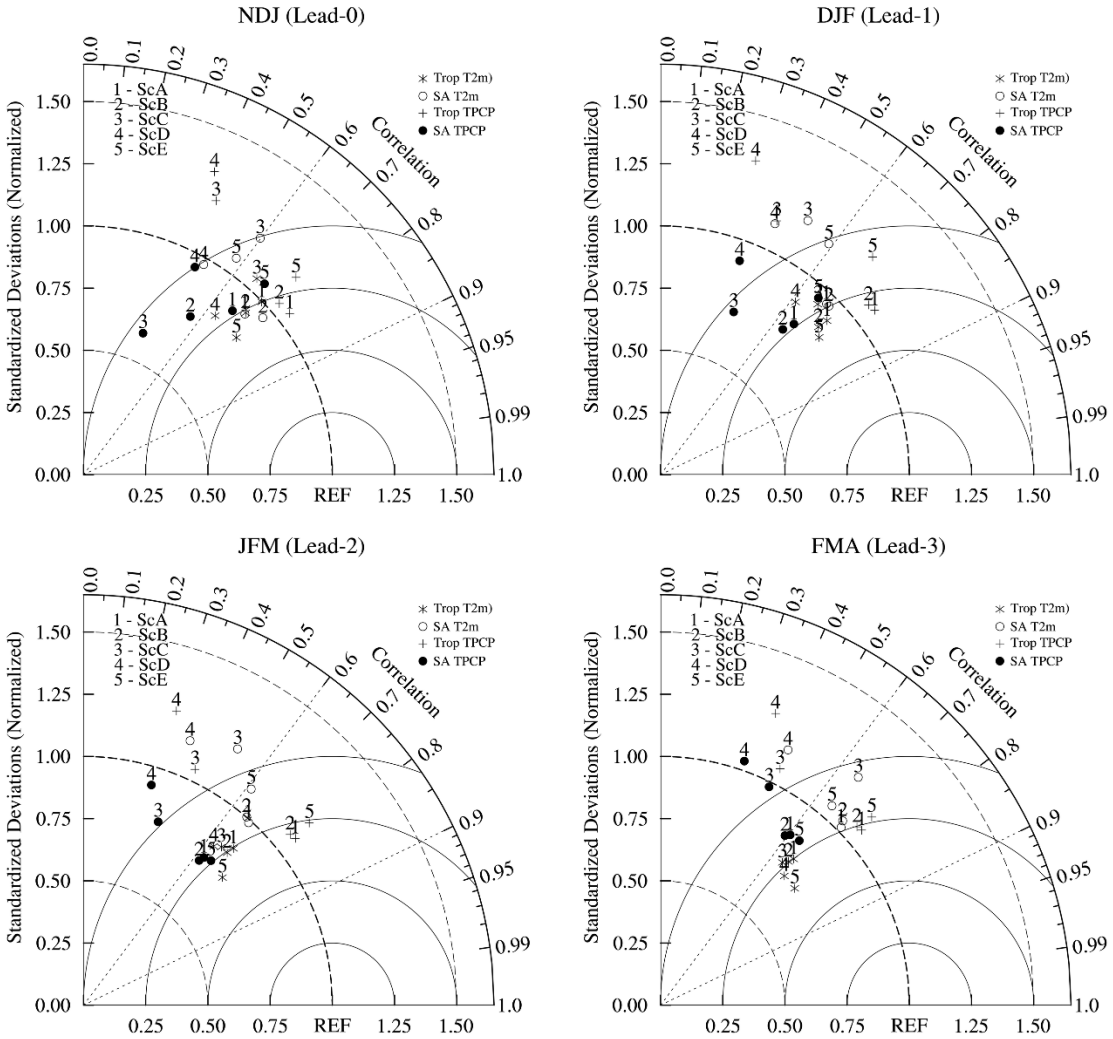

Figure 13. Taylor diagram (Taylor 2001) by different model configuration scenarios (as shown in the inset) and lead-months based on the ensemble mean from the November initialized hindcasts for tropical region between 20oS and 20oN and southern Africa south of the Equator for both seasonal rainfall totals (mm) and surface air temperatures (oK). The standard deviation is normalized by the respective observations to facilitate the comparison. The lead-time is not applicable for the baseline skill (denoted by "ScE").

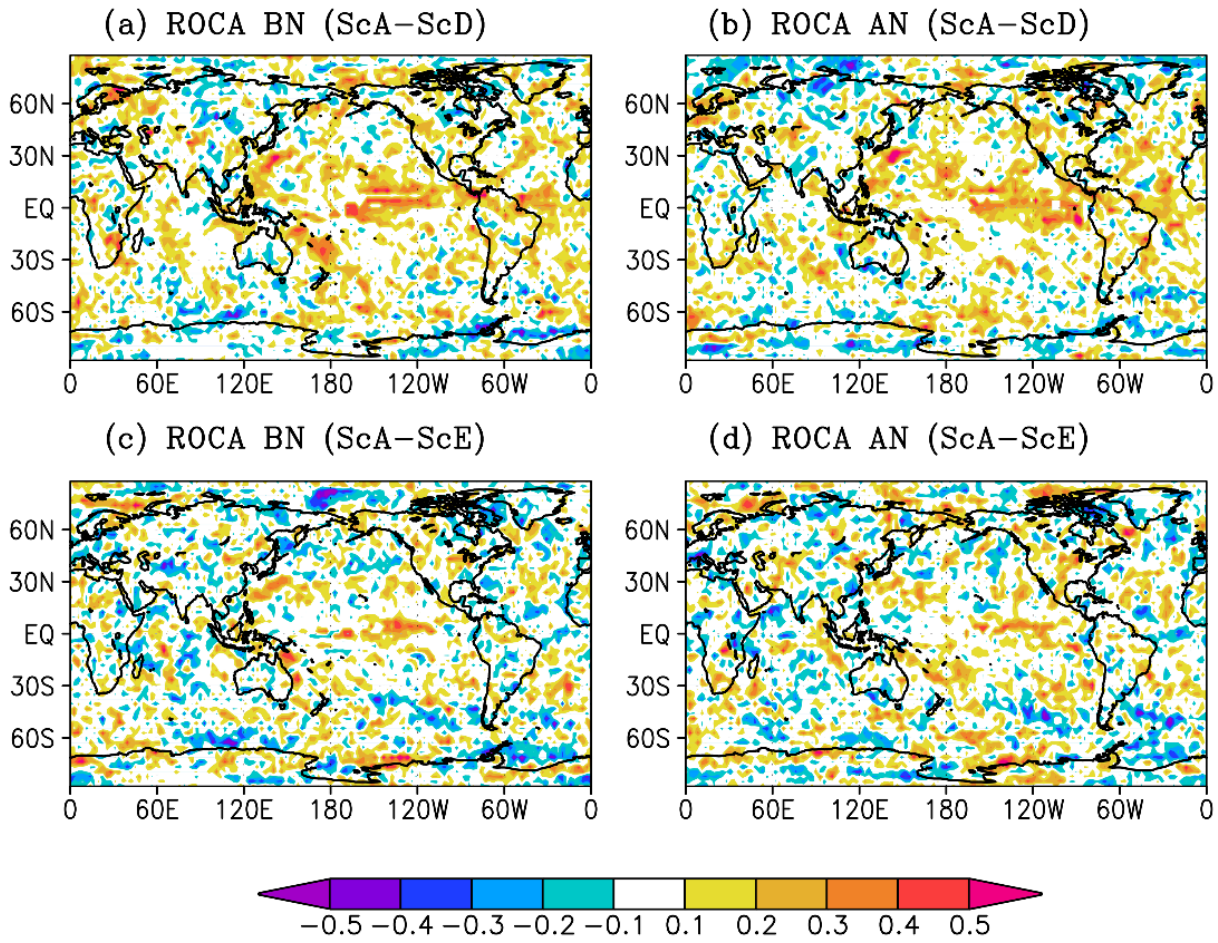

Figure 14. The maps show the global probabilistic skill (ROC areas) improvements of the optimized forecasting system in simulating the austral summer (DJF; lead-1) rainfall totals relative to the old operational (upper panel) and AMIP2 type simulations of the same model (lower panel). +ve (-ve) values indicate relative skill improvement (degradation) for below-normal (a, c) and above-normal (b, d) categories.

Further we compared the probabilistic skills of the forecasting system in predicting rainfall and temperatures against persistence and the baseline skill of the same model. The baseline skill is obtained from the AMIP-2 (Gates, *et al.*, 1999) type simulations of the AGCM. The persistence and AMIP-2 model configurations are lacking the initial condition interface as opposed to the new forecasting system reported in this article. These two independent integrations of the AGCM consist of six ensemble members each. The global skill of the model in simulating the austral

summer rainfall totals is measured using the ROC area. ROC scores are first computed for each system independently against the Merged Analysis of Precipitation of Climate Prediction Centre (CMAP-CPC; Xie and Arkin, 1997). Figure 14 shows the improvement (degradation) of skill of the optimized forecasting system relative to persistence (upper panel) and AMIP-2 (lower panel) both for below- and above-normal rainfall. In both cases, the positive split largely dominates the global skill distribution both for wet and dry DJF seasons. This result indicates that the optimized system is mostly successful in discriminating events from nonevents comparing to persistence or baseline simulation though it has the advantage due to the disproportionate ensemble size over the latter model configurations. The improvement is more evident over the tropics. The ability of the model to discriminate the categories has also improved over southern Africa region.

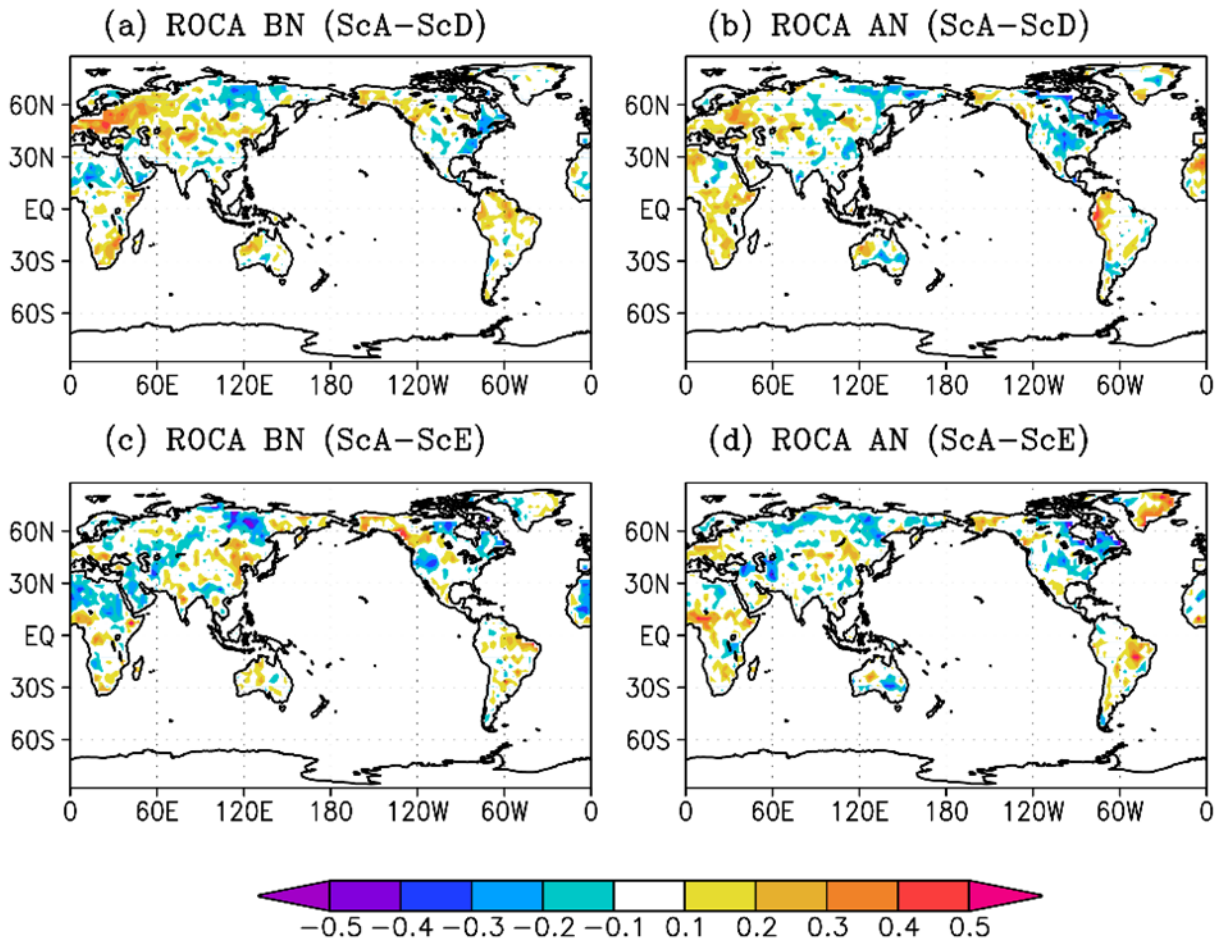

Figure 15. The same as Figure 13 but for seasonal average surface temperatures.

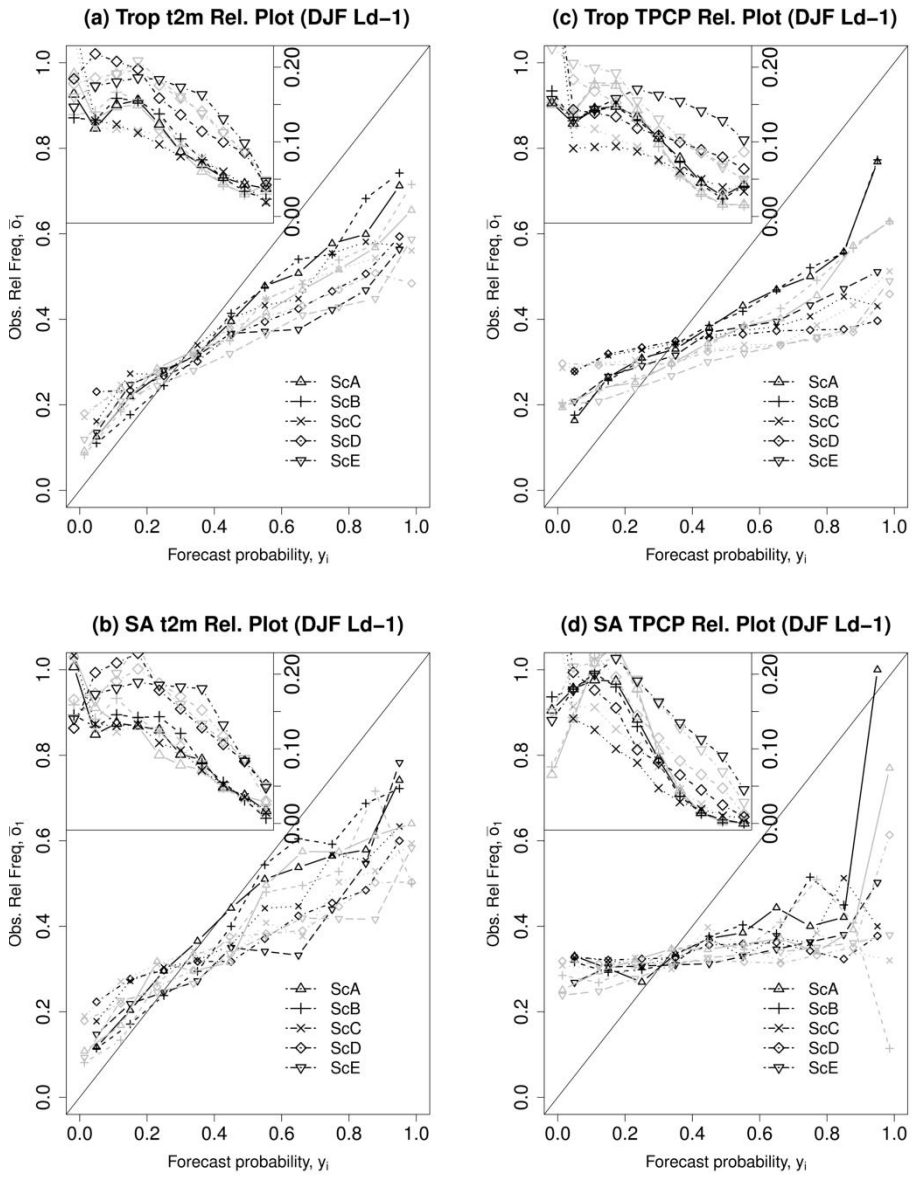

Figure 16. Reliability diagrams by different model configuration scenarios (as shown in the inset) in predicting below- and above-normal surface air temperature and rainfall conditions during the austral summer seasons (DJF) at one month lead-time for the tropical region between 20oS and 20oN (left raw) and the Southern African region (south of the equator; bottom raw). The frequency of utilization the different probability bins for both below- and above-normal categories are also shown on the top-left corners of each diagram. The grey and black lines represent cold (dry) and warm (wet) events, respectively.

Similarly, the global skill difference distribution during the austral summer for surface temperatures is also shown in Figure 15. The scores are independently calculated for each system against observed temperature acquired from the Climatic Research Unit (CRU; New *et al* 2000). The optimized forecasting system mostly outscores the persistence and baseline models in terms of discriminating cold and warm seasons from the rest of the seasons. The level of skill improvement is found to be stronger in predicting surface temperature than rainfall during mid-summer.

Figure 16 shows the reliability diagrams for the southern Africa and tropical regions obtained from various model configurations by verifying hindcasts for unusually warm (wet) and cold (dry) events during the austral summer (DJF) at one month lead-time. The reliability curves are for the 33rd percentile (dry/cold) and 67th percentile (wet/hot) thresholds. Also shown is the relative frequency of use of the forecast bins, which commonly referred to as the "sharpness diagrams" on the left top corner of each plot both for below- and above-normal conditions. Reliability diagram is a type of conditional distribution which shows given each forecast probability interval, how frequently observation actually ended up in one or another tercile (Hamill, 1997; Hartmann, *et al*, 2002). The reliability diagram is constructed from the computation of the hit rate for the set of forecasts for individual probability bins separately and then plotted against the corresponding forecast probabilities (see Wilks, 2006). The forecast utility with a closer proximity to the diagonal line is more reliable. The curves above (below) the diagonal imply that observed wet or hot (dry or cold) seasons tend to occur more (less) frequently than predicted. According to Figure 16, the optimized forecasting system (ScA) demonstrates better reliability and manages to reduce the over-confidence problem comparing to the other model configurations particularly the persistence (ScD) and baseline skill (ScE). The skill improvement is more pronounced on surface air temperatures than rainfall for the cases being considered. Notwithstanding, ScA and ScB exhibit similar reliability levels which consistently confirms the AGCM's problem with soil moisture initialization presented earlier. It is worth mentioning, however, that all the forecasting systems show a tendency to give warnings of certain events while such events (notably dry or wet condition) are less frequently observed. The poor reliability of rainfall forecast by the AGCM is presumably attributed to the fact that all grid points (skilful and

non-skilful) were contributed in the analysis apart from the internal weakness of the model (Figure 9). The other possible contributor is that the model is biased toward higher seasonal rainfall totals (Figure 7). Previous modelling studies (e.g., Landman *et al.*, 2012; Landman and Beraki, 2012) reported that the most common slope of the reliability curves found for southern Africa seasonal rainfall forecasts was shallower than the diagonal line which is also the case for tropical rainfall as presented here. In addition, the different forecasting methods tend to fall mostly on the climatological probabilities particularly for rainfall and more conservative to issue warning at higher probabilities despite that the optimized forecasting system is relatively flatter or exhibits better sharpness. The versification result generally suggests that the skill of the model is more encouraging in predicting surface temperatures than rainfall probabilistically, a result also found with other models (Barnston and Smith, 1996; Colman and Davey, 1999; Beraki *et al*., 2014). The fact that the optimized forecasting system outscores the baseline skill in predicting surface air temperature and rainfall outer percentiles is an encouraging improvement from the previous forecasting system.

6. Summary and conclusions

In spite of a number of modelling institutions in the developed world which have developed and administer state-of-the art CGCMs and use these models routinely in their seasonal predictions, AGCMs should be able to continue to provide useful forecast information and be a feasible alternative to coupled models particularly under a constrained computational resources environment, a situation commonly found in developing countries. In this paper we have demonstrated how an AGCM, the ECHAM4.5, can be a strong competitor given suitable SSTs as forcing and is subject to an initialization strategy that uses realistic atmosphere and soil moisture states.

In this AGCM optimization study, the model is initialized with the NCEP/DOE daily atmospheric initial states assimilated in a manner that respects numerical stability in the model vertical atmospheric stratification. In addition, the AGCM is constrained by lower boundary conditions derived from multi-model ensemble SST forecasts and an associated predefined uncertainty envelope. This optimized forecasting system, demonstrates large-scale consistent skill improvements for surface temperature and rainfall totals relative to the SAWS previous

operational system. The fact that the new configuration outscores the baseline skill and persistence of the same model manifests the robustness of the proposed forecast system.

Evaluation of hindcasts performed with the model further revealed that the AGCM was able to forecast upper air atmospheric dynamics (circulation) over the tropics up to several months ahead. However, the model demonstrated marginal skill over the extra-tropical regions.

The AGCM probabilistic forecasts for the austral summer season for rainfall totals and surface air temperatures were found to be informative and useful. For example, the model was able to discriminate significantly wet (warm) and dry (cold) episodes over the larger part of the globe where the signal for rainfall is more pronounced over the equatorial Pacific region.

In addition, the pairwise analysis suggested that the contribution of the multi-model SST forcing in enhancing the AGCM's predictive skill was found to be significantly important. The AGCM might also implicitly benefit from the atmospheric data assimilation and the initialization strategy included in the analysis since the major contribution is presumably expected to arise from the sub-seasonal timescale which is the subject of future work. Notwithstanding, the model appeared to be insensitive to soil moisture initialization suggestive of the AGCM's internal weakens presumably on the framework used to couple the land surface and atmosphere or the land surface scheme itself.

The modelling work presented here reports on some of the seasonal forecasting efforts using global climate models currently happening in South Africa. The work is not about model development or model versions per se, but is rather about how an AGCM could possibly be optimized to improve its seasonal predictive skills operationally in a modelling environment faced with challenges unique to a developing country such as South Africa that is at the same time also technically advanced. Notwithstanding the computing resources available in South Africa that made this work possible, these resources are shared among a large range of users other than climate modellers, therefore still placing a limitation on how models to be tested can be configured and to what extent model development can happen in the region. There is an effort in South Africa to further develop fully coupled models (e.g. Beraki *et al.,* 2014) as predecessors to useful earth systems models (Smith *et al.*, 2014), and to do modelling similar to what other developing nations like Brazil has achieved (e.g. Nobre *et al.,* 2012). In order to achieve the objective of making South Africa an internationally recognized competitor in the field of earth system model development, its modellers need to at first improve their understanding of the various components of an earth

system model and how to optimize their configurations to suit available computing capabilities. The paper has addressed some optimization issues for atmospheric models and at the same time proposed an optimal AGCM that can serve as baseline against which more advanced models can be tested.

Acknowledgments

The work is financially supported by the Water Research Commission (WRC) and Applied Centre for Climate & Earth Systems Science (ACCESS). The authors are also gratefully appreciative for the CHPC computational support. The Max-Planck-Institut für Meteorologie (MPI) has kindly provided the ECHAM4.5 AGCM code. The work was also impossible without the NCEP reanalyses product. The authors also wish to thank Dr Simon Mason and anonymous reviewers for their valuable comments.

References

AMS Council. 2001. Statement on seasonal to interannual climate prediction. *Bulletin of the American Meteorological Society* **82,** 701–703.

Arribas A, Glover M, Maidens A, Peterson K, Gordon M, MacLachlan C, Graham R, Fereday D, Camp J, Scaife AA, Xavier P, McLean P, Colman A, Cusack S. 2011. The GloSea4 ensemble prediction system for seasonal forecasting. *Monthly Weather Review* **139**, 1891-1910.

Balmaseda M, Anderson D, Vidard A. 2007. Impact of Argo on analyses of the global ocean. *Geophysical Research Letters* **34**: L16605. DOI: 10.1029/2007GL030452.

Balmaseda M, Anderson D. 2009. Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters* **36**: L01701. DOI: 10.1029/2008GL035561.

Barnston AG, Smith TM. 1996. Specification and prediction of global surface temperature and precipitation from global SST using CCA. *Journal of Climate* **9**: 2660-2697.

Barnston AG, Leetmaa A, Kousky V, Livezey R, O'Lenic E, Van den Dool H, Wagner AJ, Unger D. 1999. NCEP Forecasts of the El Niño of 1997-98 and Its U.S. Impacts. *Bulletin of the American Meteorological Society* **80**: 1829-1852.

Beraki AF, DeWitt DG, Landman WA, Olivier C. 2014. Dynamical Seasonal Climate Prediction Using an Ocean–Atmosphere Coupled Climate Model Developed in Partnership between South Africa and the IRI. *Journal of Climate* **27**: 1719-1741.

Boville, BA, Hurrell JW. 1998. Comparison of the Atmospheric Circulations Simulated by the CCM3 and CSM1. *Journal of Climate*, **11**, 1327-1340.

Brinkop S, Roeckner E. 1995. Sensitivity of a general circulation model to parameterizations of cloud-turbulence interactions in the atmospheric boundary layer. *Tellus* **47**: 197-220.

Colman A, Davey M. 1999. Prediction of summer temperature, rainfall and pressure in Europe from preceding winter North Atlantic Ocean temperature. *International journal of climatology* **19**: 513-536.

Conil S, Douville H, Tyteca S. 2009. Contribution of realistic soil moisture initial conditions to boreal summer climate predictability. *Climate Dynamics* **32**: 75–93.

Cottrill, A., Hendon H H, Lim E-P, Langford S, Shelton K, McClymont D, Jones D, Kuleshov Y. 2013. Seasonal forecasting in the Pacific using the coupled model POAMA-2. *Weather and Forecasting* **28**, 668–680, doi:10.1175/WAF-D-12-00072.1.

Derber J, Rosati A. 1989. A global oceanic data assimilation system. *Journal of Physical Oceanography* **19**: 1333–1347.

DeWitt DG. 2005. Retrospective forecasts of interannual Sea Surface Temperature anomalies from 1982 to present using a directly coupled Atmosphere-Ocean General Circulation Model. *Monthly Weather Review* **133**: 2972–2995.

Doblas-Reyes FJ, Andreu-Burillo I, Chikamoto Y, García-Serrano J, Guemas V, Kimoto M, Mochizuki T, Rodrigues LRL, Van Oldenborgh GJ. 2013. Initialized near-term regional climate change prediction. *Nature communications* **4**: 1715. doi:10.1038/ncomms2704.

Doblas-Reyes FJ, Hagedorn R, Palmer TN. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus* **57A**: 234–252.

Douville H. 2010. Relative contribution of soil moisture and snow mass to seasonal climate predictability: A pilot study. *Climate Dynamics* **34**,797–818.

Dümenil L, Todini E. 1992. A rainfall-runoff scheme for use in the Hamburg climate model. In Theoretical Hydrology, O'Kane, JP (ed.). Elsevier: Amsterdam; 129-157.

Fan Y, van den Dool H. 2004. Climate Prediction Center global monthly soil moisture data set at 0.5° resolution for 1948 to present. *Journal of Geophysical Research* **109**: D10102. DOI: 10.1029/2003JD004345.

Fouquart Y, Bonnel B. 1980. Computations of solar heating of the earth's atmosphere: A new parameterization. *Beitraege zur Physik der Atmosphaere* **53**: 35–62.

Gates WL, Boyle JS, Covey C, Dease CG, Doutriaux CM, Drach RS, Fiorino M, Gleckler PJ, Hnilo JJ, Marlais SM, Phillips TJ, Potter GL, Santer BD, Sperber KR, Taylor KE, Williams DN. 1999. An Overview of the Results of the Atmospheric Model Intercomparison Project (AMIP I). *Bulletin of the American Meteorological Society* **80**: 29–55.

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R, Cane MA. 2001. Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology* **21**: 1111–1152.

Goody RM, Yung YL. 1996. Atmospheric radiation: theoretical basis. 2nd edn, Oxford University Press: Oxford.

Graham RJ, Gordon M, McLean PJ, Ineson S, Huddleston MR, Davey MK, Brookshaw A, Barnes RTH. 2005. A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus* **57**: 320–319.

Hamill TM. 1997. Reliability Diagrams for Multicategory Probabilistic Forecasts. *Weather and Forecasting* **12**: 736–741.

Hansen JM, Sato M, Ruedy R, Lacis A, Asamoah K, Beckford K, Borenstein S, Brown E, Cairns B, Carlson B, Curran B, de Castro S, Druyan L, Etwarrow P, Ferede T, Fox M, Gaffen D, Glascoe J, Gordon H, Hollandsworth S, Jiang X, Johnson C, Lawrence N, Lean J, Lerner J, Lo K, Logan J, Luckett A, McCormick MP, McPeters R, Miller R, Minnis P, Ramberran I, Russell G, Russell P, Stone P, Tegen I, Thomas S, Thomason L, Thompson A, Wilder J, Willson R, Zawodny J. 1997. Forcings and chaos in interannual to decadal climate change. *Journal of Geophysical Research* **102**: 25679–25720.

Hartmann HC, Pagano TC, Sorooshian S, Bales R. 2002. Confidence builders: Evaluating seasonal climate forecasts for user perspectives. *Bulletin of the American Meteorological Society* **83**: 683–698.

Henderson-Seller A, McGuffeie K. 2001. Forty years of numerical climate modelling. *International Journal of Climatology* **21**:1067-1109.

Hoffman RN, Kalnay E. 1983. Lagged average forecasting, an alternative to Monte Carlo forecasting. *Tellus* **35**: 100-118.

Holton JR. 2004. An Introduction to Dynamic Meteorology. Elsevier Academic Press: San Diego, CA.

Houtekamer PL, Lefaivre L. 1997. Using ensemble forecasts for model validation. *Monthly Weather Review* **125**: 2416–2426.

Jha B, Kumar A. 2009. Comparison of the atmospheric response to ENSO in coupled and uncoupled model simulations. *Journal of Climate* **137**, 479-487.

Kalnay E, Stephen JL, Ronald DM. 1998. Maturity of operational numerical weather prediction: Medium range. *Bulletin of the American Meteorological Society* **79**: 2753-2769.

Kanamitsu M, Ebisuzaki W, Woollen J, Yang SK, Hnilo JJ, Fiorino M, Potter GL. 2002. NCEP-DOE AMIP-II reanalysis(R-2). *Bulletin of the American Meteorological Society* **83**: 1631 – 1643.

Kirtman BP, Shukla J, Huang B, Zhu Z, Schneider EK. 1997. Multiseasonal Predictions with a coupled tropical ocean-global atmosphere system. *Monthly Weather Review* **125**: 789-808.

Klopper E, Landman WA, van Heerden J. 1998. The predictability of seasonal maximum temperature in South Africa. *International Journal of Climatology* **18**: 741-758.

Koster RD, Dirmeyer PA, Guo Z, Bonan G, Chan E, Cox P, Gordon CT, et al. 2004. Regions of strong coupling between soil moisture and precipitation. *Science* **305**: 1138-1140.

Krishnamurti TN, Kishtawal CM, Zang Z, LaRow T, Bachiochi D, Williford E, Gadgil S, Surendran S. 2000. Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate* **13**: 4196–4216.

Landman WA. 2014. How the International Research Institute for Climate and Society has contributed towards seasonal climate forecast modelling and operations in South Africa. *Earth Perspectives*, **1**, 1-13.

Landman WA, Beraki A. 2012. Multi-model forecast skill for mid-summer rainfall over southern Africa. *International Journal of Climatology* **32**: 303–314. DOI: 10.1002/joc.2273.

Landman, WA, Beraki A. DeWitt D. Lötter D. 2014. SST prediction methodologies and verification considerations for dynamical mid-summer rainfall forecasts for South Africa, *Water SA* **40**, 615-622.

Landman WA, DeWitt D, Lee D-E, Beraki A, Lötter, D. 2012. Seasonal rainfall prediction skill over South Africa: 1- vs. 2-tiered forecasting systems. *Weather and Forecasting* **27**: 489-501. DOI: 10.1175/WAF-D-11-00078.1.

Landman WA, Goddard L. 2002. Statistical recalibration of GCM forecasts over southern Africa using model output statistics. *Journal of Climate* **15**: 2038-2055.

Landman, WA, Kgatuke M, Mbedzi M, Beraki A, Bartman A, du Piesanie A. 2009. Performance comparison of some dynamical and empirical downscaling methods for South Africa from a seasonal climate modelling perspective. *International Journal of Climatology* **29:** 1535-1549. DOI: 10.1002/joc.1766.

Landman W, Mason SJ. 1999. Operational long-lead prediction of South African rainfall using canonical correlation analysis. *International Journal of Climatology* **19**: 1073-1090.

Laursen L, Eliasen E. 1989. On the effects of the damping mechanisms in an atmospheric general circulation model. *Tellus*, **41**: 385-400.

Lee, D. Y., Ashok K. Ahn J-B. 2011., Toward enhancement of prediction skills of multimodel ensemble seasonal prediction: A climate filter concept, *Journal of Geophysical Research*, **116**, D06116, doi:10.1029/2010JD014610.

Lorenz EN, 1956. Empirical orthogonal functions and statistical weather prediction.
Science Report 1, Statistical Forecasting Project, Department of Meteorology, MIT
(NTIS AD 110268), 49 pp.

Lorenz EN. 1963. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences* **20**: 130-141.

Louis J-F. 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Bound-Layer Meteorology* **17**: 187–202.

Mason SJ. 1998. Seasonal forecasting of South African rainfall using a non-linear discriminant analysis model. *International Journal of Climatology* **18**: 147-164.

Mason SJ, Goddard L, Graham NE, Yulaeva E, Sun L, Arkin PA. 1999. The IRI seasonal climate prediction system and the 1997/98 El Ni˜no event. *Bulletin of the American Meteorological Society* **80**:

1853–1873

Mason SJ, Graham NE. 1999. Conditional Probabilities, Relative Operating Characteristics, and Relative Operating Levels. *Weather and Forecasting* **14**: 713–725.

Mason SJ, Graham NE. 2002. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. *Quarterly Journal of the Royal Meteorological Society* **128**: 2145-2166.

Mason SJ, Joubert AM, Cosijn C, Crimp SJ. 1996. Review of seasonal forecasting techniques and their applicability of southern Africa. *Water SA,* **22,** 203–209.

Mendonça AM, Bonatti J. 2009. Experiments with EOF-Based Perturbation Methods and Their Impact on the CPTEC/INPE Ensemble Prediction System. *Monthly Weather Review* **137**: 1438-1459.

Miller MJ, Palmer TN, Swinbank R. 1989. Parameterization and influence of sub-grid scale orography in general circulation and numerical weather prediction models. *Meteorology and Atmospheric Physics* **40**: 84-109.

Mo KC, Ghil M. 1987. Statistics and dynamics of persistent anomalies. *Journal of Atmospheric Sciences* **44**: 877-901.

Molteni F, Buizza R, Palmer TN, Petroliagis T. 1996. The ECMWF Ensemble Prediction System: Methodology and validation. *Quarterly Journal of the Royal Meteorological Society* **122**: 73–119.

Molteni F, Ferranti L, Balmaseda M, Stockdale T, Vitart F. 2007. ECMWF seasonal forecast system 3. *CLIVAR Exchange* **43**: 7-9.

Molteni F, Stockdale T, Balmaseda M, Balsamo G, Buizza R, Ferranti L, Magnunson L, Mogensen K, Palmer T, Vitart F. 2011. The new ECMWF seasonal forecast system (System 4), *ECMWF Technical Memorandum* **656**: 49.

Moore AM, Anderson DLT. 1989. The assimilation of XBT data into a layer model of the tropical Pacific Ocean. *Dynamics of Atmospheres and Oceans* **13**: 441-464.

Morcrette J-J, Smith L, Fouquart Y. 1986. Pressure and temperature dependence of the absorption in longwave radiation parameterizations. *Beitraege zur Physik der Atmosphaere* **59**: 455-469.

Murphy AH. 1988. Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. Monthly Weather Review **116**: 2417–2424.

Nobre P, De Almeida RA, Malagutti M, Giarolla E. 2012. Coupled Ocean–Atmosphere Variations over the South Atlantic Ocean. *Journal of Climate*, **25**, 6349–6358.

Nordeng TE. 1994. Extended versions of the convective parametrization scheme at ECMWF and their impact on the mean and transient activity of the model in the tropics. *Technical Memorandum* **206**, ECMWF, Shinfield Park, Reading, United Kingdom, 41 pp.

North GR. 1984. Empirical orthogonal functions and normal modes. *Journal of Atmospheric Sciences* **41**: 879–887

Palmer R, Larsen M, Sheppard E, Fukao S, Yamamoto M, Tsuda T, Kato S. 1993. Poststatistic steering wind estimation in the troposphere and lower stratosphere. *Radio Science* **28:** 261-271.

Palmer TN, Anderson DLT. 1994. The prospects for seasonal forecasting—A review paper. *Quarterly Journal of the Royal Meteorological Society* **120**: 755–793.

Palmer TN, Alessandri A, Anderson U, Cantelaube P, Davey M, Délécluse P, Déqué M, Díez E, Doblas-Reyes FJ, Feddersen H, Graham R, Gualdi S, Guérémy J-F, Hagedorn R, Hoshen M, Keenlyside N, Latif M, Lazar A, Maisonnave E, Marletto V, Morse AP, Orfila B, Rogel P, Terres J-M, Thomson MC. 2004. Development of a European ensemble system for seasonal to inter-annual prediction (DEMETER). *Bulletin of the American Meteorological Society* **85**: 853–872.

Paltridge GW, Platt CMR.1976. Radiative processes in meteorology and climatology. Elsevier: Amsterdam.

Reason CJC, Rouault M. 2005. Links between the Antarctic Oscillation and winter rainfall over western South Africa. *Geophysical Research Letters* **32**, L07705, DOI: 10.1029/2005GL022419.

Reynolds CA, Webster PJ, Kalnay E. 1994. Random error growth in NMC's global forecasts. *Monthly Weather Review* **122**: 1281–1305.

Reynolds RW, Rayner NA, Smith TM, Stokes DC. 2002. An improved in situ and satellite SST analysis for climate. *Journal of Climate* **15**: 1609-1625.

Roeckner EKA, Bengtsson L, Christoph M, Claussen M, Dűmenil L, Esch M, Giorgetta M, Schulzwedia U. 1996. The atmospheric general circulation model ECHAM4: The atmospheric

general circulation model ECHAM4: Model description and simulation of present-day climate. MPI Technical Report 218. MPI. Hamburg, Germany.

Saha S, Moorthi S, Wu X, Wang J, Nadiga S, Tripp P, Behringer D, Hou Y-T, Chuang H, Iredell M, Ek M, Meng J, Yang R, Mendez MP, van den Dool H, Zhang Q, Wang W, Chen M, Becker E. 2014: The NCEP Climate Forecast System Version 2. *Journal of Climate* **27**: 2185-2208.

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, Van den Dool HM, Pan H-L, Moorthi S, Behringer D, Stokes D, Peña M, Lord S, White G, Ebisuzaki W, Peng P, Xie P. 2006. The NCEP Climate Forecast System. *Journal of Climate* **19**: 3483-3517.

Sellers PJ, Mintz Y, Sud YC, Dalcher A. 1986. A Simplified Biosphere Model (SiB) for use within general circulation model. *Journal of Atmospheric Sciences* **43**: 505 – 531.

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner L, Orlowsky B, Teuling AJ. 2010. Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews* **99**:125–161.

Shukla J. 1981. Dynamical predictability of monthly means. *Journal of Atmospheric Sciences* **38**: 2547–2572.

Shukla J. 1983. Comments on "Natural variability and predictability" *Monthly Weather Review* **111**: 581-585.

Shukla J, Gutzler DS. 1983. Interannual variability and predictability of 500-mb geopotential heights over the Northern Hemisphere. *Monthly Weather Review* **111**: 1273–1279.

Simmons AJ, Burridge DM. 1981. An energy and angular-momentum conserving vertical finite difference scheme and hybrid vertical coordinates. *Monthly Weather Review* **109**: 758–766.

Schneider EK, DeWitt DG, Rosati A, Kirtman BP, Ji L, Tribbia JJ. 2003. Retrospective ENSO forecasts: Sensitivity to atmospheric model and ocean resolution. *Monthly Weather Review* **131**: 3038–3060.

Staniforth A, Wood N. 2008. Aspects of the dynamical core of a nonhydrostatic deep-atmosphere, unified weather and climate-prediction model. *Journal of Computational Physics* **227**, 3445–3464.

Stockdale TN, Anderson DLT, Alves JOS, Balmaseda MA. 1998. Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. *Nature* **392**: 370-373.

Smith MJ, Palmer PI, Purves DW, Vanderwel MC, Lyutsarev V, Calderhead B, Joppa LN, Bishop CM, Emmott S. 2014. Changing How Earth System Modeling is Done to Provide More Useful Information for Decision Making, Science, and Society. *Bulletin of the American Meteorological Society* **95**, 1453–1464.

Tanaka, H L, Noriko I, Kitoh A. 2004. Trend and interannual variability of Walker, monsoon and Hadley circulations defined by velocity potential in the upper troposphere, *Tellus* **56A**, 250–269.

Taylor KE, Williamson D, Zwiers F. 2000. The sea surface temperature and sea-ice concentration boundary conditions for AMIP II simulations, PCMDI Report 60. PCMDI, Lawrence Livermore National Laboratory, Livermore, CA.

Tennant WJ, Hewitson BC. 2002. Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *International Journal of Climatology* **22**: 1033–1048.

Tennant WJ. 2003. An assessment of intraseasonal variability from 13-yr GCM simulations. *Monthly Weather Review* **131**: 1975–1991.

Tett SFB, Stott PA, Allen MR, Ingram WJ, Mitchell JFB. 1999. Causes of twentieth-century temperature change near the Earth's surface. *Nature* **399**: 569–572.

Tiedtke M. 1989. A comprehensive mass flux scheme for cumulus parameterization in largescale models. *Monthly Weather Review* **117**: 1779-1800.

Tracton MS, Kalnay E. 1993. Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Weather and Forecasting* **8**: 379-398.

Toth Z, Kalnay E. 1993. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society* **74**: 2317-2330.

Viner D, Hulme M, Raper SCB. 1995. Climate change scenarios for the assessments of the climate change on regional ecosystems. *Journal of thermal Biology* **20**: 175-190.

Vitart F, Anderson JL, Stern WF. 1997. Simulation of Interannual Variability of Tropical Storm Frequency in an Ensemble of GCM Integrations. *Journal of Climate* **10**: 745–760.

Walker J, Rowntree PR. 1977. The effect of soil moisture on circulation and rainfall in a tropical model. *Quarterly Journal of the Royal Meteorological Society* **103**: 29–46. DOI: 10.1002/qj.49710343503.

Wang J, Cole HL, Carlson DJ, Miller ER, Beierle K, Paukkunen A, Lane TK. 2002. Corrections of humidity measurement errors from the Vaisala RS80 radiosonde—Application to TOGA COARE data. *Journal of Atmospheric and Oceanic Technology* **19**: 981-1002.

Wilks DS. 2011. *Statistical methods in the atmospheric sciences*. Vol. 100. Academic press: Press: San Diego, CA.

Williamson DL, Rasch PJ. 1994. Water vapor transport in the NCAR CCM2. *Tellus* **46**: 34-51.

WMO. 2010. Manual on the Global Data-processing and Forecasting System. WMO: Geneva. (available from WMO web site: www.wmo.int/pages/prog/www/DPFS/Manual_GDPFS.html).

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

Zhang Z, Krishnamurti TN. 1999. A Perturbation Method for Hurricane Ensemble Predictions. *Monthly Weather Review* **127**: 447–469.

Synopsis

A suitable optimization strategy which is based on the emerging concept of multi-model ensemble SST forcing and an atmospheric initialization procedure that uses realistic atmosphere and soil moisture states has been described. This optimized two-tiered forecasting system configured in a manner that mimics an operational setup talks to objective 2 of the thesis. Furthermore, the impact of the optimization on the overall predictive skill of the AGCM and the sensitivity of the AGCM to the different boundary forcings and soil moisture initialization have been demonstrated. This pairwise sensitivity analysis contributes towards the understanding of the strength and weakness of the AGCM from an operational point of view which address objectives 3 and 4 relevant to the uncoupled model. Hence, the study has established a baseline skill against which more advanced climate models such as coupled ocean-atmosphere models could be compared. The thesis has up to now established a modelling framework that supports both the contrast of the role of oceanic evolution of air-sea interaction (only supported in the CGCM as demonstrated in the proceeding chapter) and the use of prescribed multi-model SST forcing (only supported in the AGCM) on the predictive skill of seasonal forecasts while both GCMs effectively use the same atmospheric and soil moisture initialization strategy. The contrast in terms of forecast performance will be dealt with in the next chapter which addresses the last objective of the study, namely to do a thorough verification comparison between the two systems.

4. On the comparison between seasonal predictive skill of global circulation models: coupled versus uncoupled

Preface

This chapter consists of published journal paper and is cited as follows:

Beraki A.F., W. Landman, and D. DeWitt. (2015): Comparison on the seasonal predictive skill of global circulation models: coupled versus uncoupled. *J. Geophys. Res. Atmos*., 120, doi:10.1002/2015JD023839.

Towards addressing the final objective (objective 5) of the thesis, the paper presented below employs a suitable framework which comprises of an interactive ocean-atmosphere coupled GCM and optimized AGCM to undertake a skill comparison between the two GCMs over an extended verification period of about three decades. In this framework, the GCMs share a great deal of similarities in their respective configurations except for the way in which the SST information is flows within the GCMs. The configurations enable the isolation of the role of coupling from model biases. Moreover, the comparison provides an insight toward understanding the condition under which an AGCM and CGCM may be able to produce similar or different levels of skill.

The paper was co-authored with Willem Landman and David DeWitt. The conceptualisation of the paper, data analysis and the actual article writing were done by me.

On the comparison between seasonal predictive skill of global circulation models: coupled versus uncoupled

Asmerom F. Beraki,[a,c*], Willem A. Landman,[b,c] and David DeWitt[d]

[a] South African Weather Service, Pretoria, South Africa Private Bag X097, Pretoria, 0001, South Africa

[b] Council for Scientific and Industrial Research, Natural Resources and the Environment, Pretoria, South Africa

[c] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, South Africa

[d] National Oceanic and Atmospheric Administration, National Weather Service, Silver Spring, Maryland, USA

Abstract

The study compares one- and two-tiered forecasting systems as represented by the South African Weather Service (SAWS) Coupled Model (SCM) and its atmosphere-only version. In this comparative framework, the main difference between these Global Climate Models (GCMs) resides in the manner in which the sea-surface temperature (SST) is represented. The models are effectively kept similar in all other aspects. This strategy may allow the role of coupling on the predictive skill differences to be better distinguished. The result reveals that the GCMs differ widely in their performances and the issue of superiority of one model over the other is mostly dependent on the ability to *a priori* determine an optimal global SST field for forcing the Atmospheric General Circulation Model (AGCM). Notwithstanding, the AGCM's fidelity is reasonably reduced when the AGCM is constrained with persisting SST anomalies to the extent to which the Coupled General Circulation Model (CGCM)'s superiority becomes noticeable. The result suggests that the boundary forcing coming from the optimal SST field plays a significant role in leveraging a reasonable equivalency in the predictive skill of the two GCM configurations.

Key words: model comparison, AGCM, CGCM, Seasonal forecast, multi-model SST, sea-air interaction

1. Introduction

The practice of contemporary seasonal climate prediction requires state-of-the-art Global Climate Models (GCMs). The predictive skill of seasonal predictions mainly arises from the slowly evolving components of the climate system which are found to significantly modulate the mean state weather conditions [*Charney and Shukla, 1981; Palmer and Anderson, 1994; Barnston et al., 1999*]. Most of the signature of these slowly evolving systems is believed to originate from the ocean and thus the interaction between the ocean and the atmosphere is of paramount importance in the context of seasonal forecasting [*Goddard et al., 2001*]. In fact, GCMs are classified into two distinct configurations, commonly referred to as one- and two-tiered forecasting systems. These configurations are based on the manner in which information flows between the ocean and the atmosphere. In the atmosphere-only configuration, the Atmospheric General Circulation Models (AGCMs) are forced with independently predicted or persisted SST (sea-surface temperature) anomalies [*Bengtsson et al., 1993; Graham et al., 2005; Kug et al., 2008*] with the assumption that the atmosphere responds to SST but does not in turn affect the ocean [*Copsey et al., 2006*]. On the other hand, in one-tiered forecasting systems, the way in which the ocean and atmosphere interact and evolve mimics processes found in nature [*Palmer and Anderson, 1994*]. Therefore, this spontaneous two-way feedback mechanism provides Coupled (ocean-atmosphere) General Circulation Models (CGCMs) a distinctive advantage over AGCMs [*Graham et al., 2005*].

Historically, two-tiered forecasting systems were the first to appear on the scene as seasonal forecasting tools and are still practiced globally [*e.g. Kirtman et al., 1997; Graham et al., 2000; Tennant and Hewitson, 2002; Beraki et al., 2015*]. Despite the enormous cost implications and complexity, one-tiered forecasting systems appear to have gained preference over two-tiered forecasting systems over recent years and their use by operational centers is steadily growing [e.*g. Stockdale et al., 1998; Palmer et al., 2004; Graham et al., 2005; Saha et al., 2006; Molteni et al., 2007; Beraki et al., 2014*]. This proliferation of interest is presumably stimulated by the fast development of computational technology complemented by the fact that many intercomparison studies suggest the superiority of CGCMs to AGCMs [e.*g. Yu and Mechoso 1999; Fu et al. 2002; Graham et al. 2005; Kug et al., 2008; Landman et al., 2012; Chaudhari et al., 2013; Zhu and Shukla, 2013; Shukla and Zhu, 2014*], even though similar studies report that only marginal

differences exit [*e.g. Boville and Hurrell, 1998; Jha and Kumar, 2009; Colfescu et al., 2013*]. Most of these numerical studies report the weakness of AGCMs in simulating the Asian monsoon during the austral winter, where air-sea interaction plays a significant role. In contrast, CGCMs are distinctively able to rectify the problem and to better represent air-sea coupling in the tropical Indian and western Pacific Oceans [*Fu et al., 2002; Kug et al., 2008*]. In addition, *Graham et al.* [*2005*] suggest that coupled models can provide substantial benefits for seasonal prediction not only in tropical regions, but also in the extratropics.

It is commonly believed that coupled climate models are placed at the highest hierarch in the science of numerical modelling in terms of complexity [*Stockdale et al., 1998; Palmer et al., 2004*]. In theory, they are largely hypothesized to represent state-of-the-art of seasonal forecasting which inherently renders them convenient for operational seasonal climate prediction purposes. Notwithstanding, it may also be important to consider whether two-tiered forecasting systems offer comparable levels of skills that are currently attainable by state-of-the-art coupled models [*Troccoli et al., 2008*] on one hand, and the inhibiting factor of the computational requirement to operate such coupled systems on the other hand. The latter consideration may be of particular importance in developing countries with less advanced capabilities, and especially at operational centers within these countries tasked to produce real-time seasonal forecast output. Moreover, although, as noted earlier, both model configurations are used at a number of operational centers, their comparison on seasonal prediction in an operational environment is less explored. It is worth emphasizing that it may be beneficial to objectively assess the relative merit or limitations of these forecasting systems under a constrained resources scenario. The aim of this paper is, therefore, to undertake a performance comparison of one- and two-tiered forecasting systems where the AGCM is constrained by the lower boundary conditions derived from predicted SST anomalies of two CGCMs' forecasts in contrast to persisted or empirically predicted SSTs [*e.g. Graham et al., 2005*] while the two systems share a great deal of similarities in other aspects. To achieve our goal, the South African Weather Service (SAWS) Coupled Model (SCM) [*Beraki et al., 2014*] and its atmospheric version [*Roeckner et al., 1996; Beraki et al., 2015*] are used. These two forecasting systems are currently running operationally at the SAWS as part of a multi-model system [*Landman and Beraki, 2012*]. The notion is also tested under a perfect model framework [*Colfescu et al. 2013*] and persistence (an AGCM forced with persisted SST anomalies) [*Graham et al.,*

2005]. The former configuration eliminates differences due to model bias between the CGCM and AGCM and enables the isolation of the role of coupling. In this framework, the AGCM is forced with the CGCM (SCM) retroactive SST simulations.

The paper is organized as follows: In section 2, the experimental design is described. Results from composite and time series analyses are presented on section 3. In section 4, differences in the predictive skill of the CGCM and the AGCM are elucidated. A summary and conclusions are given in section 5.

2. Experimental Design

2.1 Description of GCMs

The study compares the SCM and its atmosphere-only version as mentioned above. The SCM is described in depth in *Beraki et al.* [*2014*] while we only briefly describe the model here.

The SCM couples the ECHAM4.5 AGCM [*Roeckner et al., 1996*] and the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model version 3 (MOM3) [*Pacanowski and Griffes, 1998*] using the Multiple Program and Multiple Data (MPMD) fully parallelized coupler paradigm [*Komori et al., 2008*]. Under this coupling framework, the atmosphere and ocean models are treated as standalone versions apart from the interface that handles the exchange of information between the models.

While the AGCM, as in the two-tiered experiment, uses T42 (triangular truncation at wave number 42) horizontal resolution and 19 unevenly spaced hybrid sigma layers, the OGCM (Ocean General Circulation Model) has a 0.58^o uniform zonal resolution, with a variable meridional resolution of 0.5^o between 10^o S and 10^o N, gradually increasing to 1.5^o at 30^{o}S and 30^o N and fixed at 1.5^o in the extratropics. In the vertical, the OGCM uses 25 layers with 17 layers in the upper levels between 7.5m and 450m.

2.2 Retroactive forecasts

In this comparison experiment, the fundamental difference between the GCMs, as noted earlier, arises from the manner in which the ocean and atmosphere interact with each other. The two systems remain nearly identical in other respects. In the CGCM experiment, the AGCM and OGCM exchange information per simulation day. The AGCM feeds the OGCM with heat,

momentum, freshwater, and surface solar flux. The OGCM, in turn, feeds the AGCM SST information. The coupling strategy used in this configuration is an anomaly coupling on the AGCM side and full-field coupling on the OGCM side, meaning that the anomalous atmospheric fluxes are superimposed on the observed climatology as in *Ji et al.* [*1998*] and *DeWitt* [*2005*]. The ocean initial conditions are taken from the ODA (Ocean Data Assimilation) system produced at GFDL that employs an optimum interpolation scheme [*Derber and Rosati, 1989*]. However, it is worth mentioning that seasonal climate prediction skill may be dependent on the accuracy of ODA on which systems are initialized from [*Zhu et al., 2012*].

In the two-tiered experiment, however, the AGCM is constrained by the lower boundary conditions derived from the predicted SST of two CGCMs combined with equal weighting. The two CGCMs are the SCM [*Beraki et al., 2014*] and the NCEP CFS v2 (National Centers for Environmental Prediction, Climate Forecasting System Version 2) [*Saha et al., 2014*]. The benefit of the multi-model approach has been reported in many forecasting studies over recent years [*e.g. Krishnamurti et al., 2000; Palmer et al., 2004; Doblas-Reyes et al., 2005, Hagedorn et al., 2005; Kirtman et al., 2014*]. In addition, the SST uncertainty amplitude (lower and upper bounds) defined from this combination is also considered as separate forcing fields. To identify the uncertainty amplitude a slight perturbation was applied to the multi-model ensemble mean of the SST independently for all lead-months (lead-0 to lead-8) using empirical orthogonal function analysis (EOF) [*North, 1984*]. The first normalized EOF mode was retained to describe the uncertainty term assuming that the variance is best explained by the most dominant EOF mode and subsequently subtracted from or added to the multi-model SST [*Beraki et al., 2015*].

Although more emphasis is placed on the understanding of whether the AGCM is a viable option to the CGCM under a constrained computational resources scenario, it is worth emphasizing that the use of multi-model SST forcing into the AGCM's configuration deviates slightly from a perfect model framework [*Colfescu et al., 2013*]. For scientific purpose, this approach may not cleanly eliminate differences between the CGCM and AGCM due to model bias or isolate the role of coupling. However, from the perspective of operational forecasts, the multi-model SST forcing may offer a better optimization option. For interest of quantifying the extent of the disparity in biases and skill differences, the AGCM is also constrained by the predicted SST derived from the

CGCM only, as in *Colfescu et al.* [*2013*], and for convenience hereafter referred to as "AGCMc". Also included are model simulations performed with persisting observed SST anomalies taken from the Optimum Interpolation version 2 [*Reynolds, et al., 2002*] as a lower boundary condition to the AGCM (hereafter referred to as "AGCMp"). The rationale of including the AGCMp experiment is to gain additional insight into the AGCM's predictive skill relative to the CGCM when the AGCM is independently configured from the influence of the CGCM(s), as this may delineate the lower limit of the skill of the AGCM. Furthermore, this forecast strategy was used in similar comparative studies [*e.g. Boville and Hurrell 1998; Graham et al., 2005*].

The CGCM and the AGCM (also AGCMp and AGCMc) use the same atmospheric initialization strategy. The atmospheric initial conditions (ICs) are obtained from the NCEP/DOE (Department of Energy) Atmospheric Model Intercomparison Project (AMIP) II Reanalysis (R2) dataset [*Kanamitsu et al., 2002*]. The NCEP/DOE atmospheric states are transformed to the horizontal and vertical resolution (T42L19) of the ECHAM4.5 AGCM in a manner that maintains numerical and gravitational stability as explained in *Beraki et al.* [*2014*]. The only difference is that the lower layer atmospheric temperature over the ocean (atmosphere-ocean interface) is assimilated from the multi-model ensemble mean of the SST and from the GFDL-ODA ocean state for use in the AGCM and the CGCM respectively. This is done to minimize the imbalance between upper-ocean mass field and wind stress [*DeWitt 2005*]. The uncertainties that arise from the ICs are accounted for by taking 10 consecutive daily atmospheric states back from the forecast date in each month and year. The November hindcasts, for example, consider the ICs that extend over the 10-day period from October 26 to November 4 for 28 years starting from 1982 and ending in 2009. The combination of ocean state and atmospheric state gives rise 10 and 30 ensemble integrations of the CGCM and the AGCM respectively.

2.3 Observation data

The model surface and upper air data are compared against the respective observed data compiled from different sources. For the surface variables, rainfall and air temperatures are acquired from the Climate Prediction Center (CPC) Merged Analysis of Precipitation (CMAP) [*Xie and Arkin, 1997*] and the Climatic Research Unit (CRU) [Harris *et al., 2014*] respectively.

For pressure data analyses, the (NCEP/ DOE) [*Kanamitsu et al., 2002*] is used as a proxy for observation.

3. Climatological and temporal differences

The results presented in this section are taken from the coupled and uncoupled models' hindcast simulations for the 28 years from 1982 to 2009. In this study, the lead-time is defined from the starting month when the model is initialized. For example, hindcasts from November ICs for NDJ (November-December-January) are referred to as zero month lead-time hindcasts, while hindcasts for December-January-February (DJF), with the same initial conditions, are made at a one month lead-time, and so forth.

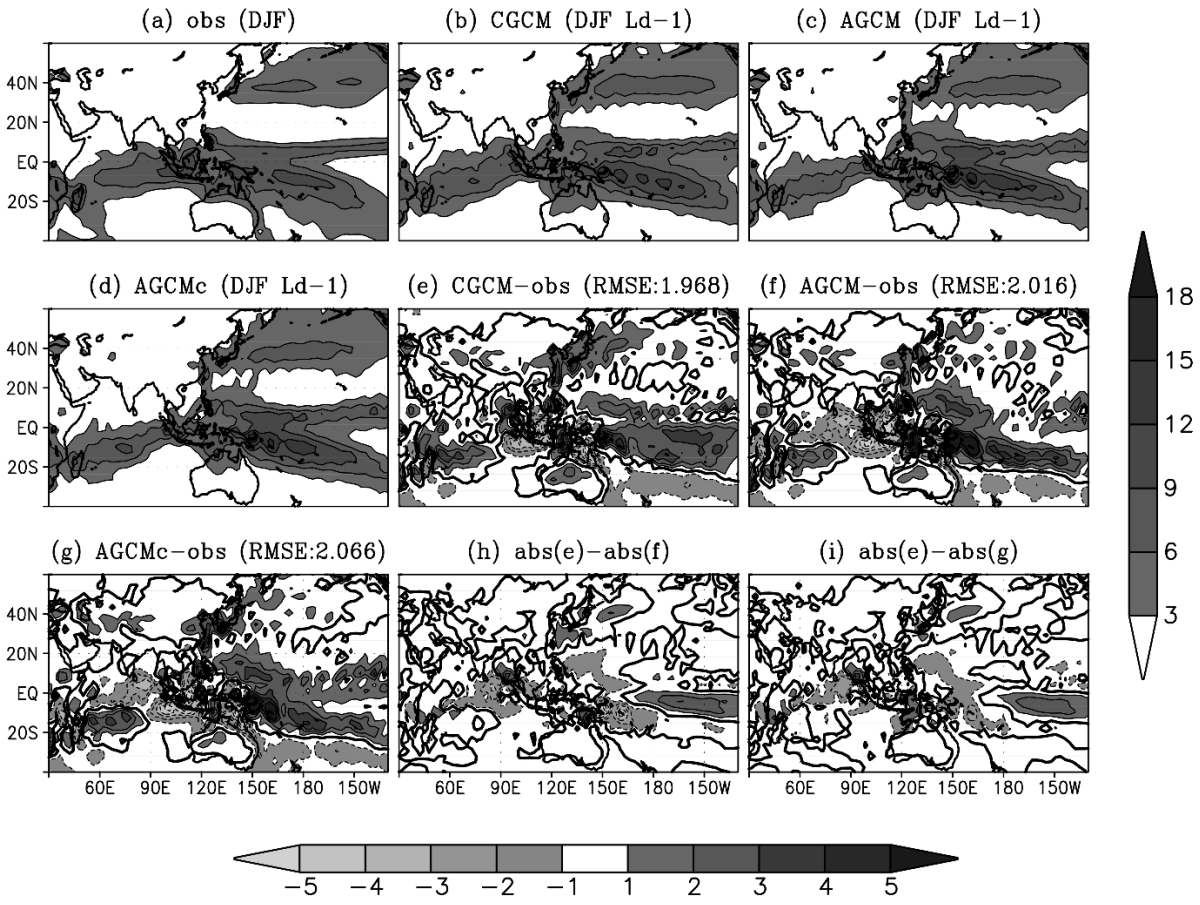

Figure 1. a Climatological representation of austral summer mean (DJF) precipitation (mm/day); (a) observation, (b) CGCM, (c) AGCM, (d) CGCM bias and (e) AGCM bias, (g) AGCMc (AGCM forced with the SCM predicted SST anomalies; see text) bias, (h) absolute value difference between CGCM bias and AGCM bias and (i) the same as (h) but for the AGCMc. The absolute bias differences enable to easily identify where exactly the GCMs differ, nonetheless the direction of the bias should be interpreted in conjunction to (e), (f) and (g). Also shown on the title of each bias plots is area averaged root mean square error (RMSE; e-g).

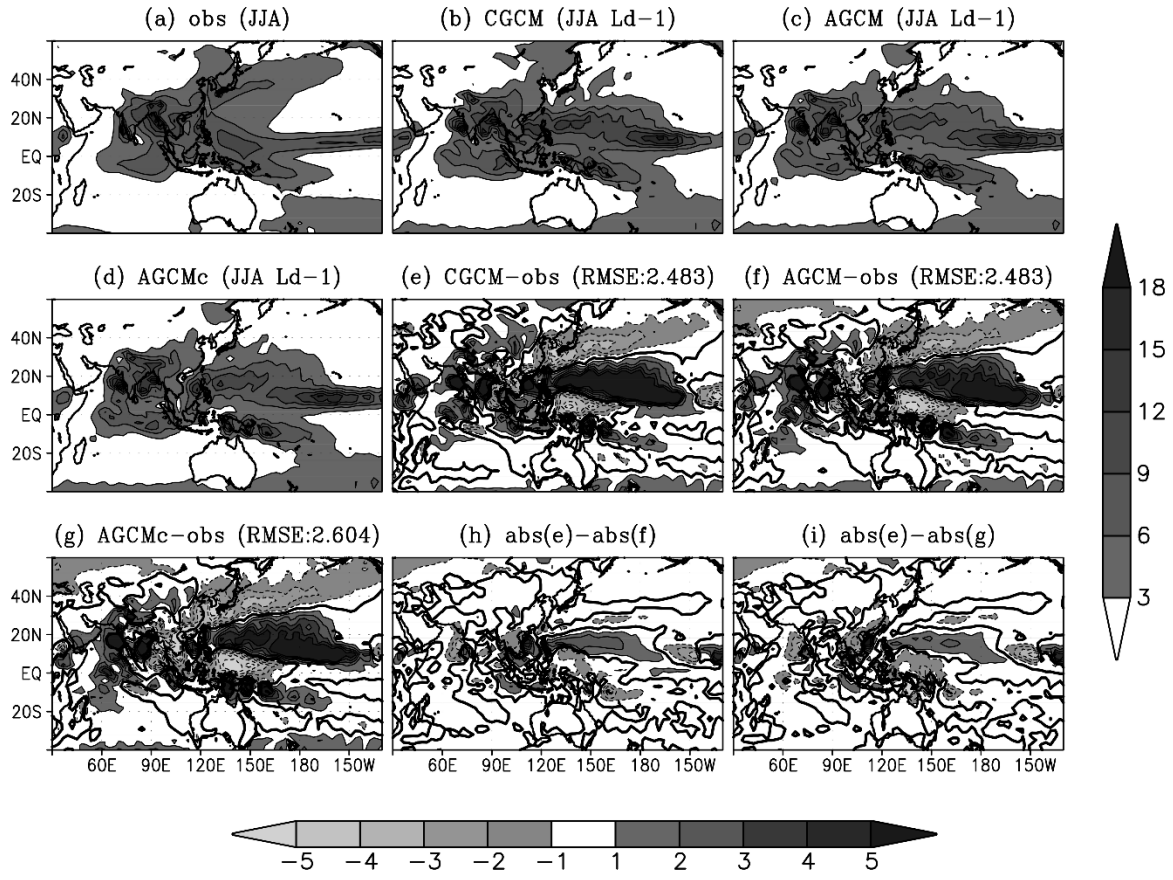

Figure 2. As Figure 1 but for austral winter.

First, we investigate the role of the oceanic evolution of sea-air interaction by zooming in on the equatorial Indo-Pacific (Asian monsoon) region. Since the region has become the subject of many similar numerical studies, as noted earlier, it may be used as a benchmark for comparative assessments here.

The composite analysis of rainfall during the austral summer (DJF) at a one-month lead-time for part of the global region that centers the equatorial Indian and Pacific Oceans is shown in Figure 1. The two models (CGCM and AGCM) capture the CMAP climatological distribution reasonably well. They also consistently manifest similar bias patterns, with the exception that the AGCM exhibits a greater dry bias over the western and central Indian Ocean and a wet bias over the western Pacific Ocean. The CGCM is more biased over the eastern Pacific region south of the Equator at about 120^{o}W. The AGCMc also exhibits similar bias patterns to the AGCM although the absolute bias difference is slightly shallower than the AGCM (Figure 1h, 1i). Similar composite analysis for the austral winter (June-July-August; JJA; Figure 2) suggests that both the CGCM and AGCM are able, once again, to represent the observed spatial patterns of rainfall reasonably well. The analysis also reveals that the CGCM appears to overestimate daily rainfall over the eastern equatorial Pacific region while the AGCM is more biased over the eastern Indian Ocean and western Pacific Ocean adjacent to Australia at about 10^{o}S. It is worth noting that the AGCM and the AGCMc hardly differ in terms of bias distributions as also shown in their area averaged root mean square error (rmse) differences (see Figures 1, 2) and all GCMs appear to be more biased during the austral winter than they are during the summer (Figure 1d, 1e).

The zonally averaged (over the zonal extent of Figures 1, 2) DJF and JJA rainfall (mm/day) depicted in Figure 3 show that the CGCM, AGCM and AGCMc forecasts are in good agreement with the observed rainfall. In the tropics, the symmetry and position of the ITCZ (intertropical convergence zone) are well represented in all simulations. The mid-latitude storm tracks are also adequately represented despite all forecast strategies overestimate the DJF and JJA rainfall over the Northern Hemisphere (NH) and Southern Hemisphere (SH) respectively. However, the difference among the forecasting systems seems to be that the CGCM shows a slight tendency to exaggerate the tropical peak during DJF and JJA comparing to the two AGCM configurations.

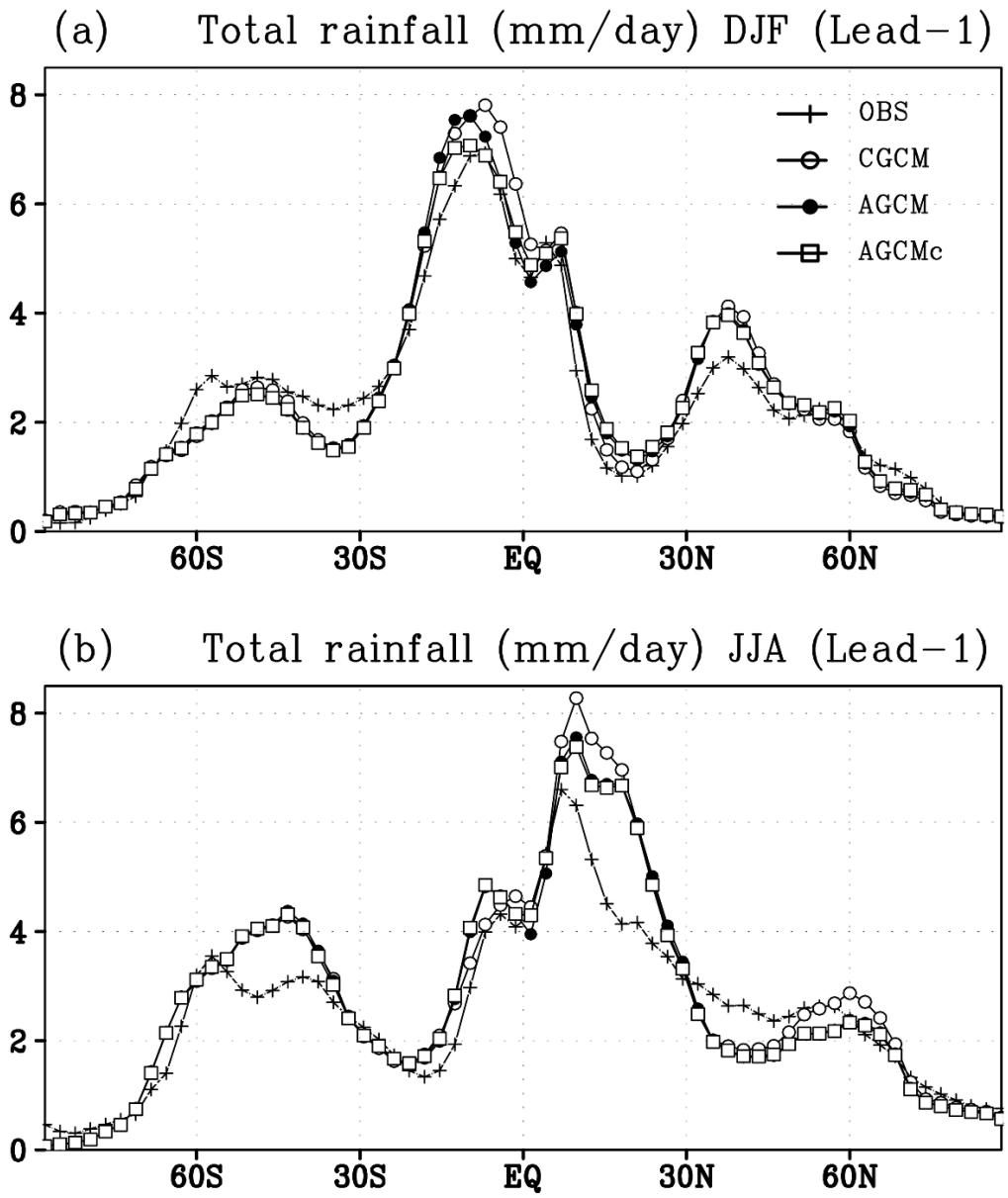

Figure 3. Zonally average total rainfall (mm/day) for DJF (a) and JJA (b) lead month 1 for the SCM, two different configurations of ECHAM 4.5 AGCM (as shown in the inset) and observation from CMAP [*Xie and Arkin, 1996*]. The temporal average is from 1982-2009 and the zonal extent is as in Figures 1, 2.

The time evolution of rainfall biases across the equatorial Indo-Pacific region is further demonstrated in Figures 4 to 5 through the use of Hovmöller diagrams. In the equatorial Indian

basin, during the JJA season, the GCMs are marginally biased with nearly indistinguishable differences in the Indian Ocean sector (Figure 4f) except that the AGCM is consistently underrepresenting the DJF CMAP estimates (Figure 4b). Notwithstanding, the two prediction systems widely differ in their temporal bias distribution over the equatorial Pacific Ocean. The CGCM overestimates the rainfall over the equatorial western Pacific between the equator and approximately 20^{o}S and the eastern Pacific confined within 15^{o}N and 25^{o}N during the austral summer (Figure 5a) and winter (Figure 6d) seasons. In contrast, the AGCM overestimates rainfall over the eastern and western Pacific Ocean during the austral winter and summer at about 15^{o}N (Figure 6e) and the region that lies between 15^{o}S and 15^{o}N (Figure 5d) respectively.

The time evolution of rainfall biases across the equatorial Indo-Pacific region is further demonstrated in Figures 4-6 through the use of Hovmöller diagrams. The absolute bias differences computed from the GCMs biases and ENSO (El-Niño Southern Oscillation) information are also included in the plots to enhance objective interpretation and to better characterize the meridional and temporal bias differences. The ENSO phases are represented with the Oceanic Niño Index (ONI) [*L'Heureux et al., 2012*]. However, it is worth noting that care should be exercised when identifying wet or dry biases because the absolute value differences only emphasize whether the AGCM or the CGCM is more biased. During JJA season, the GCMs exhibit nearly indistinguishable differences in the Indian Ocean sector (Figure 4i, g) whereas the AGCM and the AGCMc consistently underrepresent the DJF CMAP estimates (Figure 4b - e). Notwithstanding, the GCMs biases are more pronounced over the equatorial Pacific Ocean than over the equatorial Indian Ocean. According to Figure 5, the AGCM exhibits wetter bias than the CGCM or the AGCMc during the austral summer over the western Pacific region. During the austral winter, in contrast, both AGCM prediction strategies (though more strengthened in the case of the AGCMc; Figure 5i, g) are relatively biased in comparison to the CGCM in the vicinity of the equator while the CGCM is noticeably biased over the southern hemisphere around 15^{o}N. It is also noticeable that the CGCM overestimates the rainfall over the equatorial eastern Pacific during the austral summer (Figure 6d, e) while both AGCM configurations overestimate rainfall during the austral winter at about 15^{o}N (Figure 6i, g), with the exception that the CGCM is more biased than the AGCM in 1983, 1989-1992 and 1997.

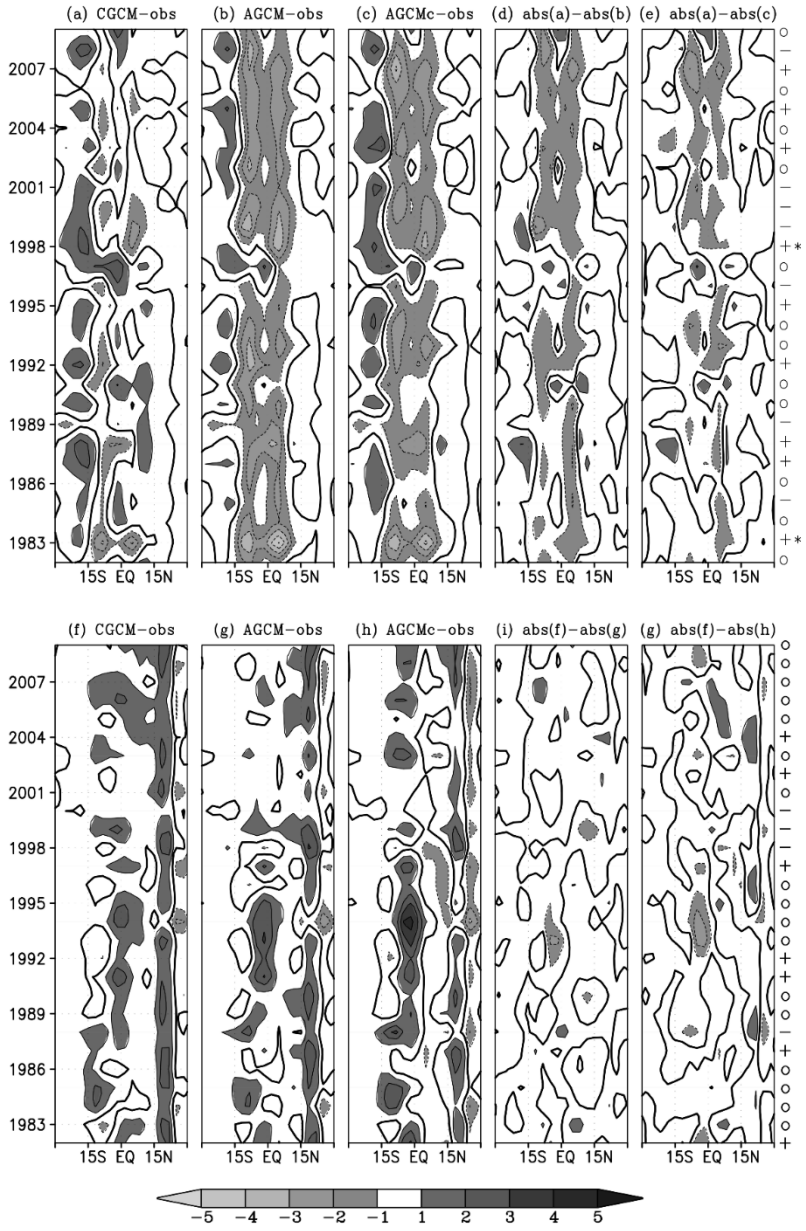

Figure 4. Hovmöller diagrams for the austral summer (top) and winter (bottom) rainfall at one month lead-time zonally averaged over the Equatorial Indian Ocean (50^{o}E - 110^{o}E). As shown in the title of each plot, rainfall biases are computed from the hindcast simulations of the GCMs against CMAP estimates and among the GCMs themselves. The absolute bias differences between the CGCM and AGCM (CGCMc) are also presented in (d), (e), (i) and (g) to highlight the meridoinal and temporal differences of the GCMs. The anomalous and neutral phases of ENSO, where - denotes La-Niña, + denotes El-Niño and o denotes neutral, are shown on the right side of

plots (e) and (g). The two strongest El-Nino episodes are also indicated with *.The ENSO phases are based on the Oceanic Niño Index (ONI) [*L'Heureux et al., 2012*] obtained from the National Oceanic and Atmospheric Administration (NOAA), Climate Prediction Centre (CPC; http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml).

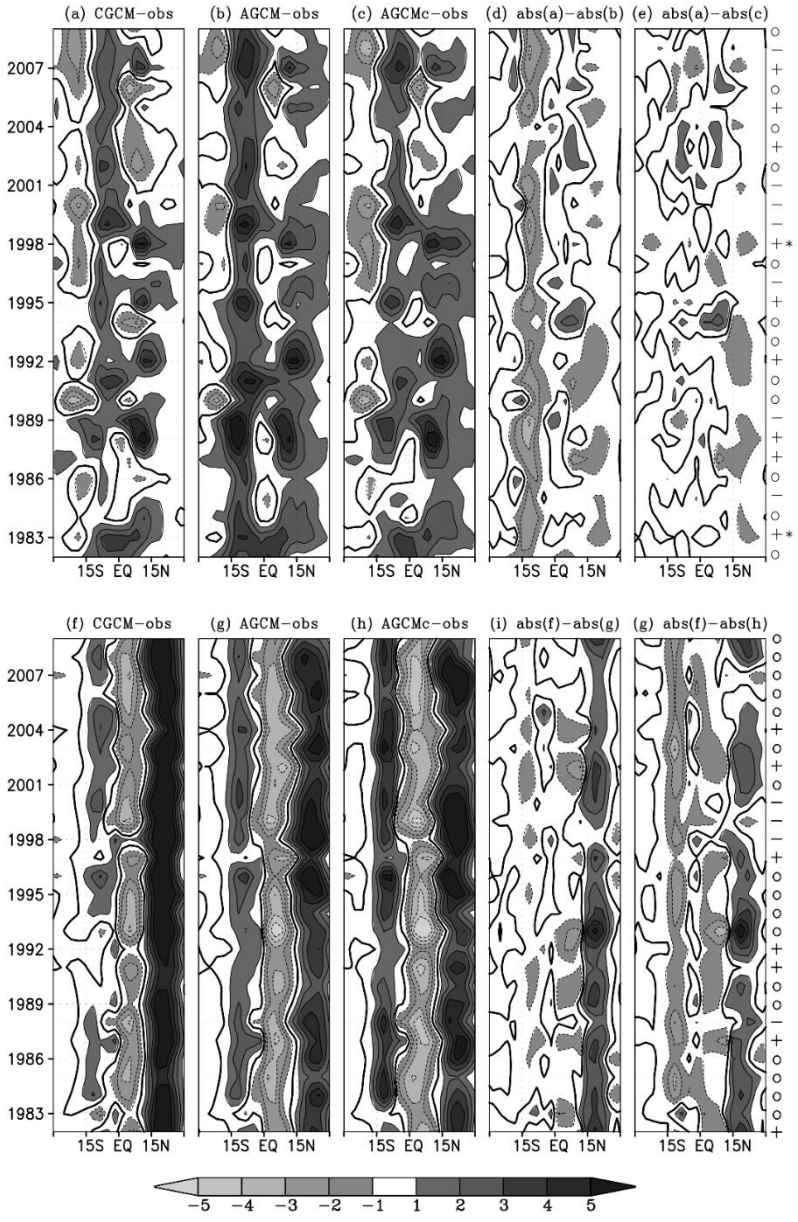

Figure 5. As in Figure 4 but for Western equatorial Pacific Sector (120°E - 170°E).

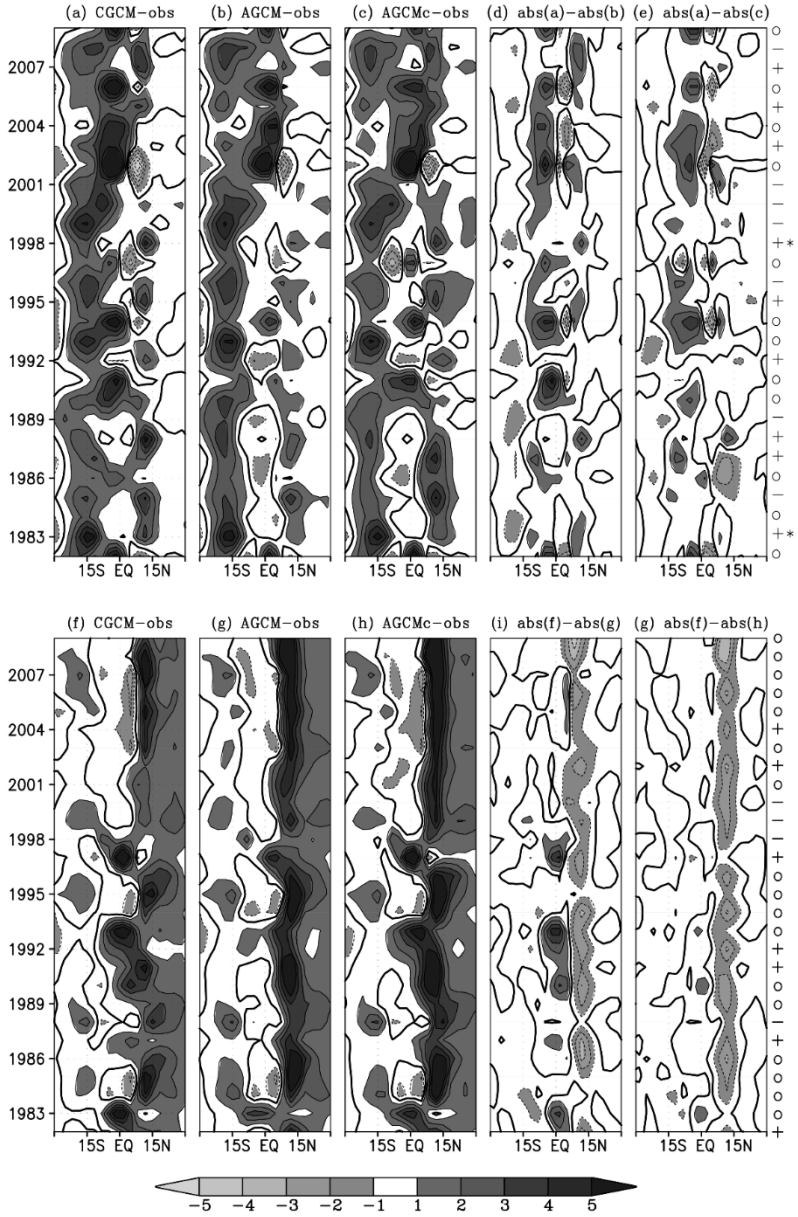

Figure 6. As in Figure 4 but for Eastern equatorial Pacific sector (170^{o}E - 60^{o}W).

The biggest bias differences between the AGCM (AGCMc) and the CGCM over the equatorial Pacific regions mostly coincide with neutral ENSO. In the western Pacific during the austral winter season (Figure 5i, g), the largest biases of the CGCM are found (for example in 1989, 1990, 2001, 2009) during neutral ENSO years over areas surrounding 15^{o}N; likewise, the AGCMc (largely reduced in the AGCM; Figure 5i) is noticeably biased (for example in 1984, 1985, 1988, 1993, 2003) over the area confined within the equator and 15^{o}S. During DJF over the

eastern equatorial Pacific, the CGCM biases are intensified in 1990, 1991, 2002 and 2009 (Figure 6d, e). Furthermore, the distribution of the model biases is more intense during the austral winter than during the summer, particularly for the western (Figure 4f-h) and eastern Pacific (Figure 5f-h). This result may suggest that the GCMs atmospheric response seems to depend on the accuracy of SST predictions. The dynamical ENSO predictive skill is minimized due to the NH spring barrier [*Saha et al., 2006; Beraki et al., 2014*]. The result further suggests that the CGCM's excessive rainfall in the tropics (Figure 3) may be attributed to the SST bias found within the time evolution of the air-seas coupling process which is slightly minimized in the AGCM (AGCMc) with the use of prescribed SST forcing. This may reinforce the conclusion that both GCMs are able to respond to the SST fluctuations equally and that the role of the evolution of sea-air interaction anticipated to favor the CGCM is not clearly established in the cases we have tested for the Asian monsoon regions, particularly during the austral winter despite that previous similar studies reported in favor of the CGCMs [e.g. *Fu et al., 2002; Kug et al., 2008*].

However, the result presented above does not consider whether those differences between the two forecasting systems are statistically significant. To approach the problem indirectly, we perform a statistical significance test using the Wilcoxon–Mann–Whitney non-parametric approach without involving observations [*Graham et al., 2005; Wilks, 2006*]. Figures 7 and 8 show the spatial extent and temporal frequencies when the CGCM and the AGCM rainfall fields are found to be different in their probability density functions (PDFs) with a statistical significance at the 95% level for the austral summer and winter seasons at a one month lead-time respectively. The PDFs are represented with 10 and 30 ensemble members of the CGCM and the AGCM respectively. According to this result, more than 80% of the time the source of variation between the coupled and uncoupled models arises mostly from the equatorial region. During the austral summer, more pronounced differences are found over the equatorial eastern Pacific Ocean, Brazil, Atlantic Ocean and the south eastern Australian coast (Figure 7a). Likewise, these differences are noticeable in the Pacific Ocean and Atlantic Ocean off the coast of West Africa during austral winter (Figure 8a). Most of the peaks in the areal extent differences between the two models are consistently concentrated over the Pacific region during JJA although the differences are confined over the eastern Pacific region during DJF. It is also noticeable that the peaks are mostly found during neutral ENSO conditions. As noted earlier, the GCMs also differ mostly over

the equatorial Pacific region in terms of biases (Figures 5 and 6). Similar analysis between the CGCM and the AGCMr (AGCMc), in which the GCMs use the same ensemble size, demonstrate that most of the source of differences similarly arise from the equatorial region. However, in both cases the differences in the areal extent over the tropical Pacific are reduced to some extent particularly between the CGCM and the AGCMc during JJA seasons (not shown). The result attests the realism of the bias differences discussed so far.

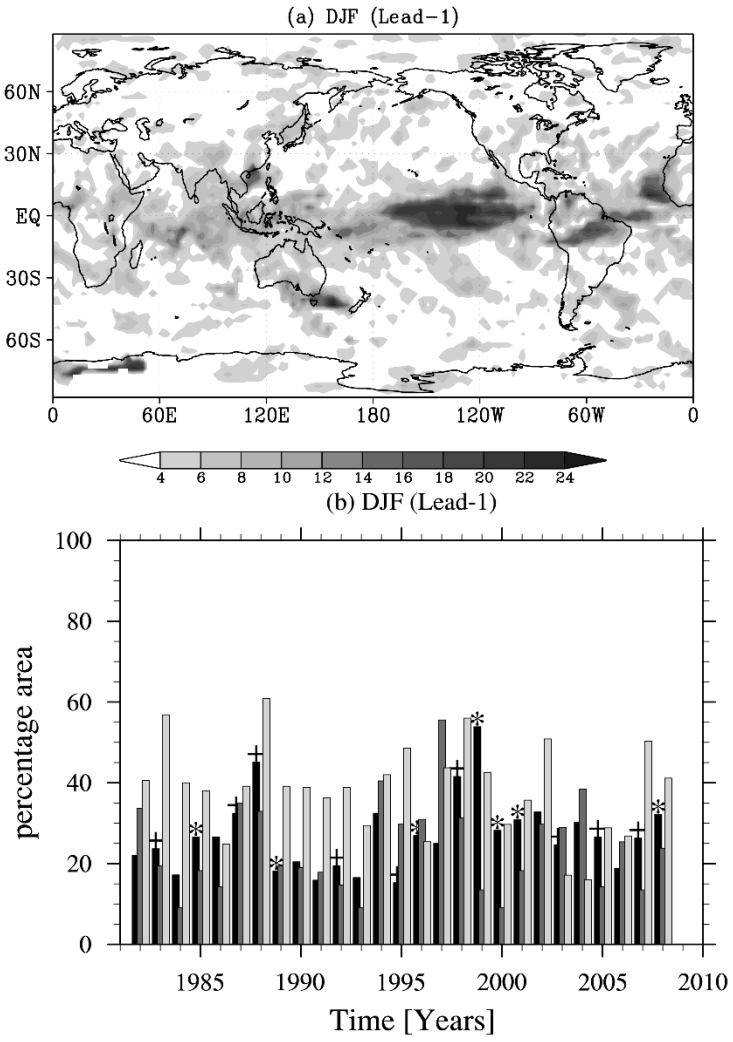

Figure 7. Number of years (out of 28, 1982–2009) when CGCM and AGCM ensemble distributions for DJF (lead-1) rainfall were found significantly different at the 95% level with a Wilcoxon–Mann–Whitney test performed at each grid point (a); time series of the percentage area where CGCM and AGCM ensemble distributions for austral summer (DJF Lead-1) rainfall are

found significantly different at the 95% level (b). Black, dark grey, and light grey bars denote equatorial Indian Ocean, western equatorial Pacific Ocean and equatorial eastern Pacific Ocean respectively. Annotation represent anomalous phases of ENSO where * denotes La-Niña and + denotes El-Niño. The ENSO phases are based on the Oceanic Niño Index (ONI) [*L'Heureux et al., 2012*].

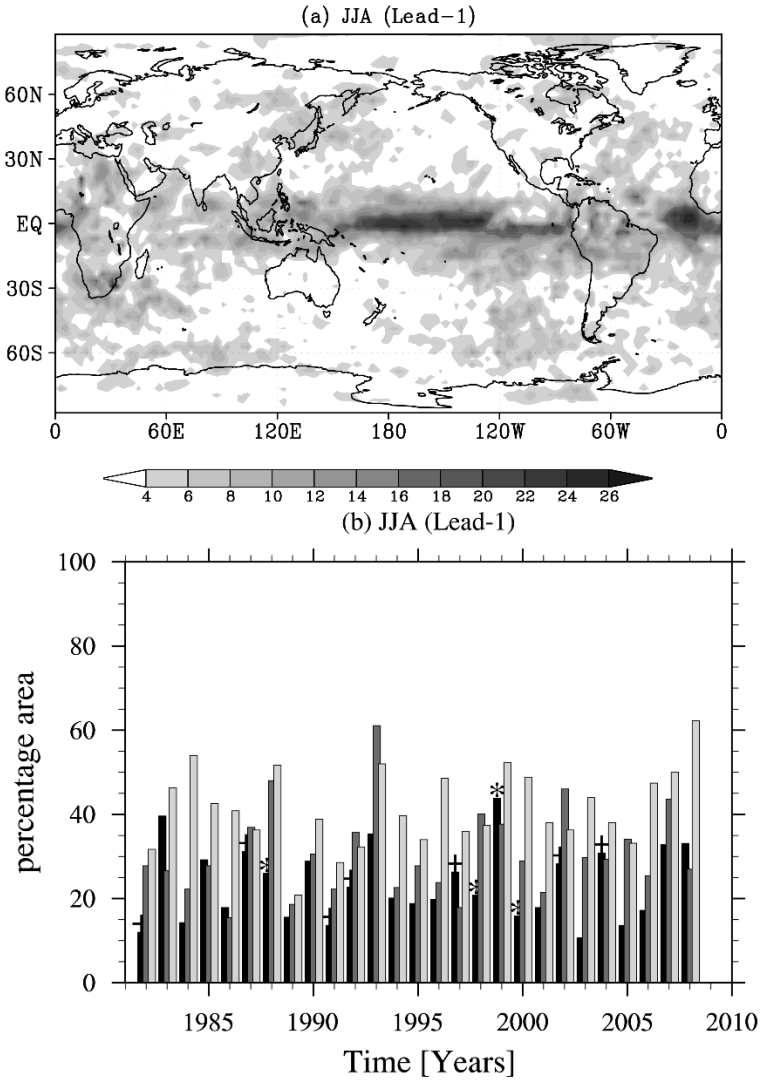

Figure 8. As in Figure 7 but for astral winter.

4. Differences in the Seasonal Predictive skills

The GCMs are investigated for their relative performance for different seasons and lead-times. This comparative analysis is based on the 28 years hindcast of the AGCM and the CGCM which consist of 30 and 10 ensemble members respectively. Furthermore, results from the AGCMc and the AGCMp, which each has the same ensemble size as the CGCM, are also presented. Each ensemble set mimics a set of operational forecasts because the sets were created in a manner similar to how operational forecasts at SAWS, a WMO-recognized Global Producing Centre, are being conducted. This approach offers a better insight into the relative enhancement or degradation of forecast quality in an operational environment.

4.1 Comparison based on the ensemble mean

The impact of the oceanic evolution of sea-air interaction (only supported in the CGCM) and the use of prescribed multi-model SST forcing (only supported in the AGCM) on the predictive skill of seasonal forecasts is compared by first evaluating the ensemble mean of the GCMs. The anomaly of each model is computed about its own drifted climatology before the statistics are applied in order to remove biases from the model forecasts as a function of lead months. We first concentrate on atmospheric pressure fields as the signatures of most climate drivers (ocean-atmosphere coupling phenomena) including, inter alia, ENSO [*Neelin et al., 1998; Wallace et al., 1998*], Indian Ocean Dipole (IOD) [*Saji et al., 1999*] and Pacific South America (PSA) [*Mo and Ghil, 1987*], which are represented in the mean sea level pressure (mslp) or geopotential height (GH) fields. To facilitate the comparison, we use the mean square skill score (MSSS) [*Murphy, 1988*]. The MSSS is computed using the mean square error (MSE). The MSE of each model is independently calculated first against the NCEP/DOE. In this comparison, the MSSS of the CGCM is identified by taking the AGCM as a reference forecast (i.e., $MSSS = 1\text{-}MSE_{CGCM}/MSE_{AGCM}$). The skill score therefore represents gains or losses in the forecast skill relative to the reference forecast. In this context, the MSSS approaches +1 (approaches $-\infty$) if the CGCM (AGCM) perfectly outperforms its counterpart and the gradient between +1 and $-\infty$ measures the degree of superiority (inferiority) of one model over the other. The CGCM and AGCM predictive skill equates when the ratio of MSE of one model to the other approaches 1.

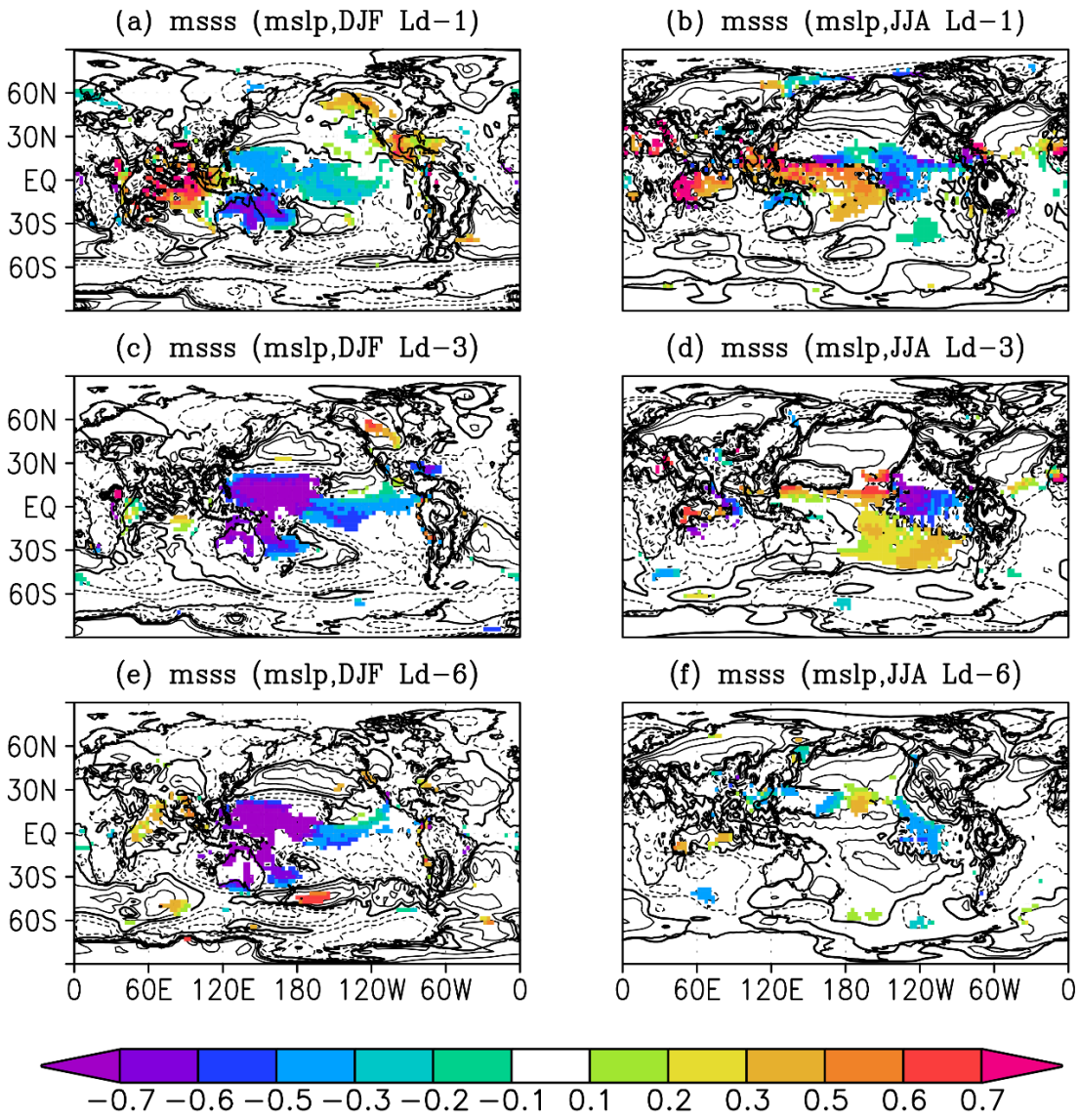

Figure 9. Skill improvement or degradation of the CGCM relative to the AGCM (reference) in predicting mslp (Pa) during the austral summer (December-January-February; DJF; left panel) and winter (June-July-August; JJA; right panel) for various months lead-time as shown in the title of each plot. The MSE of each model is first computed against the NCEP/DOE mslp that eventually returns the MSSS. The region with +ve (-ve) scores imply the superiority of the CGCM (AGCM) where those statistically significant at 95% level with reasonable differences are shaded. Also shown is contours with 0.2 interval. The significance test is performed with a bootstrap non-parametric procedure [*Wilks, 2006*].

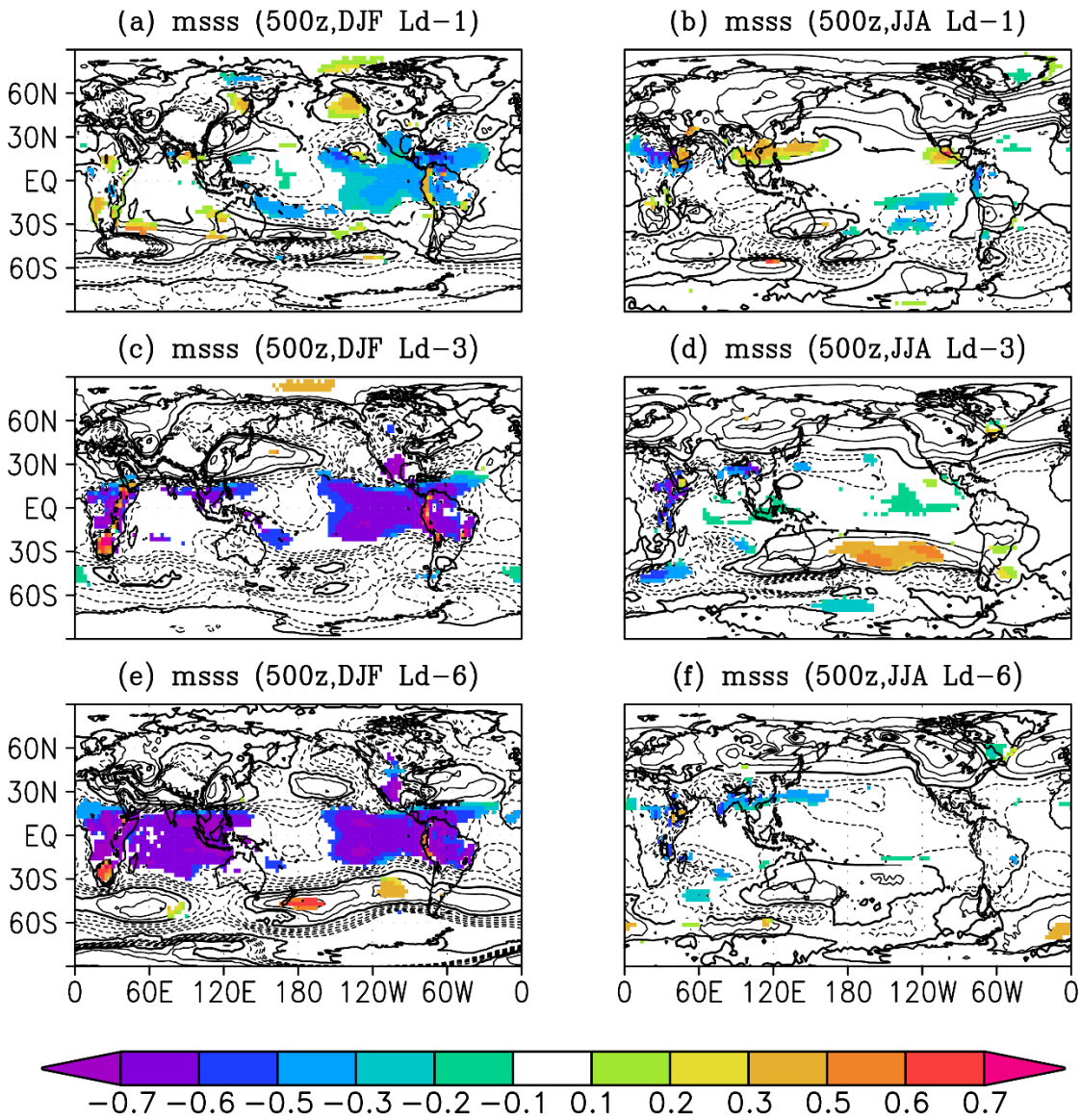

Figure 10. As Figure 9 but for 500hPa GH.

Figure 9 shows the extent to which the CGCM's predictive skill has improved (degraded) relative to the AGCM in predicting the mslp during DJF and JJA seasons with a 95% level of statistical significance. The significance level is identified using the bootstrap non-parametric procedure (sampling with replacement) [*Wilks, 2006*] where the analysis is repeated 1000 times. According to this result, the predictive skill of the CGCM during the austral summer appears to significantly strengthen across the equatorial Indian Ocean and in the vicinity of central and North America relative to the AGCM, but the benefit diminishes as a function of lead-time. The AGCM,

however, consistently outperform the CGCM on the equatorial Pacific region across all lead-times considered. Results from AGCMc (maps not shown) demonstrate a relative skill degradation predicting mslp. The finding may support the notion that the use of multi-model SST forcing in the AGCM configuration played a significant role for the best performance of the AGCM. *Beraki et al.* [*2015*] also showed that the skill and accuracy of the Nino3.4, derived from the multi-model ensemble SST used to force the AGCM, intensified during the austral summer as the lead-time increases.

During the austral winter, the coupled model gains significant advantage over the atmospheric model particularly at one and three month lead-times on the equatorial region, with the exception of the eastern Pacific sector. The strength of the CGCM skill over the equatorial Indian Ocean is also noticeable at short lead-times. However its skill is significantly reduced over the eastern part to the extent that the AGCM takes the lead as the lead-time increases (Figure 9d, 9f). In the prediction of the austral winter mslp over the Asian monsoon region, the CGCM outperforms the AGCM, although its superiority decays quickly as the lead-times increase.

In the 500hPa geopotential height (GH) comparative analysis, the AGCM's benefit, during the DJF season, is well manifested over the eastern part of the equatorial Pacific region and the northern South America sub-continent with a tendency of deepening as the lead-time increases (Figure 10c, 10e) despite the CGCM perform better over the southern Africa sub-continent. At enhanced lead-times (Figure 10c, 10e) the dominance of AGCM is also extended over the equatorial Africa and Indian Ocean. Notwithstanding, during the JJA season, the CGCM is found to outperform the AGCM over the Asian monsoon basin at a one month lead-time although the benefit is changed in favor of the AGCM at extended lead-times. The superiority of one model over the other during the winter season is mostly indistinguishable as opposed to the austral summer for the 500hPa GH. Previous evaluation studies [*Beraki et al., 2014; 2015*] conducted on these models using the same hindcasts showed independently that reasonable skill of the GCMs' in predicting pressure fields was found mostly over the equatorial region and the predictive skill presented here should therefore be viewed in relative terms.

So far, what has transpired from the comparative analyses of pressure fields is that the differences between the GCMs are a function of both space and time of the year (i.e., seasonality).

For example, the CGCM is superior over the Asian monsoon region during the austral winter, but skill deteriorates rapidly with increasing lead time or for a different season. Although both the CGCM and the AGCM are skillful but not necessarily over the same areas, seasons or even lead-times, by combining the forecast from these two models in a multi-model system may further improve on the forecasts.

In the 500hPa geopotential height (GH) comparative analysis, the AGCM's benefit, during the DJF season is well manifested over the eastern part of the equatorial Pacific region and the northern South America sub-continent with a tendency of deepening as the lead-time increases (Figure 10c, 10e), despite the CGCM performing better over the southern Africa sub-continent. At enhanced lead-times (Figure 10c, 10e) the dominance of the AGCM is also extended over equatorial Africa and the Indian Ocean. During the JJA season, the CGCM is found to outperform the AGCM over the Asian monsoon basin at a one month lead-time, although the benefit is changed in favor of the AGCM at extended lead-times. The superiority of one model over the other during the winter season is mostly indistinguishable as opposed to the austral summer for the 500hPa GH. Previous evaluation studies [*Beraki et al., 2014; 2015*] conducted on these models using the same hindcasts show independently that reasonable skill of the GCMs in predicting pressure fields taken climatological forecast as a reference is found mostly over the equatorial region and the predictive skill presented here should therefore be viewed in relative terms.

The finding is consistent with what has been discussed so far with regard to the prevalence of noticeable differences over the equatorial (notably Pacific) region in spite of the narrowing tendency in (bias and skill) differences under the perfect model framework. The other point worth mentioning is that the evolution of sea-air interaction expected to favor the CGCM is barely supported particularly at longer lead-time. This result may suggest that the differences are better explained by model biases which tend to be intensified during neutral ENSO episodes rather than the coupling issue, *per se*.

Furthermore, what has transpired from the comparative analyses of pressure fields is that the differences between the GCMs are a function of both space and time of the year (i.e., seasonality). For example, the CGCM is superior over the Asian monsoon region during the austral winter, but skill deteriorates rapidly with increasing lead-time or for a different season. Although both the CGCM and the AGCM are skillful but not necessarily over the same areas, seasons or

even lead-times, combining the forecast from these two models in a multi-model system may further improve on the forecasts.

Further, we examine the implication of variations in the pressure fields of the GCMs on ENSO and equatorial IOD (which are the main climate variability modes and most relevant at the seasonal timescale) and coupling responses (teleconnections). ENSO characteristics are represented using the Southern Oscillation Index (SOI), the mslp difference between Tahiti (17.5^{o}S, 149.5^{o}W) and Darwin (12.5^{o}S, 130.9^{o}E), using the method suggested by *Ropelewski and Jones* [*1987*]. Likewise, the IOD characteristics are measured using a pressure index with anomalous difference between the mslp in the east and western tropical Indian Ocean [*Saji et al., 1999*]. The indices of these climate drivers are deduced from mslp since previous observational studies show strong association between mslp and SST indices [e.g. *Philander, 1990; Behera and Yamagata, 2003*]. The comparison is based on the November initialized hindcasts since this month coincides with the onset of the seasonal peaks of ENSO although IOD is more active during the austral spring [*Beraki et al., 2014; Zhao and Hendon, 2009*]. By using a Taylor diagram [*Taylor, 2001*], Figure 11 presents skill comparisons by various forecasting methods in predicting ENSO and the equatorial IOD. The skill is represented in the correlation [*Wilks, 2006*] and standard deviation space of the Taylor diagram. The standard deviations are normalized with the corresponding observed standard deviation to facilitate the comparison. The result reveals that differences in the skill and interannual variability between the CGCM and the AGCM at zero-month (November-December-January; NDJ) and one month (DJF) lead-times in predicting ENSO is marginal despite the ability of the GCMs to predict the mslp varying significantly, mainly because the biases across the Pacific Ocean region (between east and west dissect) cancel each other. The CGCM, however, performs better than the AGCM in predicting the equatorial IOD. The result further shows that the CGCM's predictive skill is found to be consistently superior to the AGCMp's, a result also reported elsewhere [*e.g. Graham et al., 2005*]. Furthermore, the AGCMc's ability is slightly reduced in predicting both ENSO and IOD from the CGCM or AGCM, suggesting that the AGCM benefits from the multi-model SST forcing. At longer lead-times, the GCMs underestimate the observed variability with a sharp skill drop suggesting weakening of the atmospheric response to ocean variations. The predictability of these climate modes, notably ENSO, is much stronger up to several months lead-time when their strength and

evolution are measured using SST indices (e.g., Niño 3.4 index) [e.g. *Beraki et al., 2014*] as opposed to mslp derived indices (e.g., SOI) presented here.

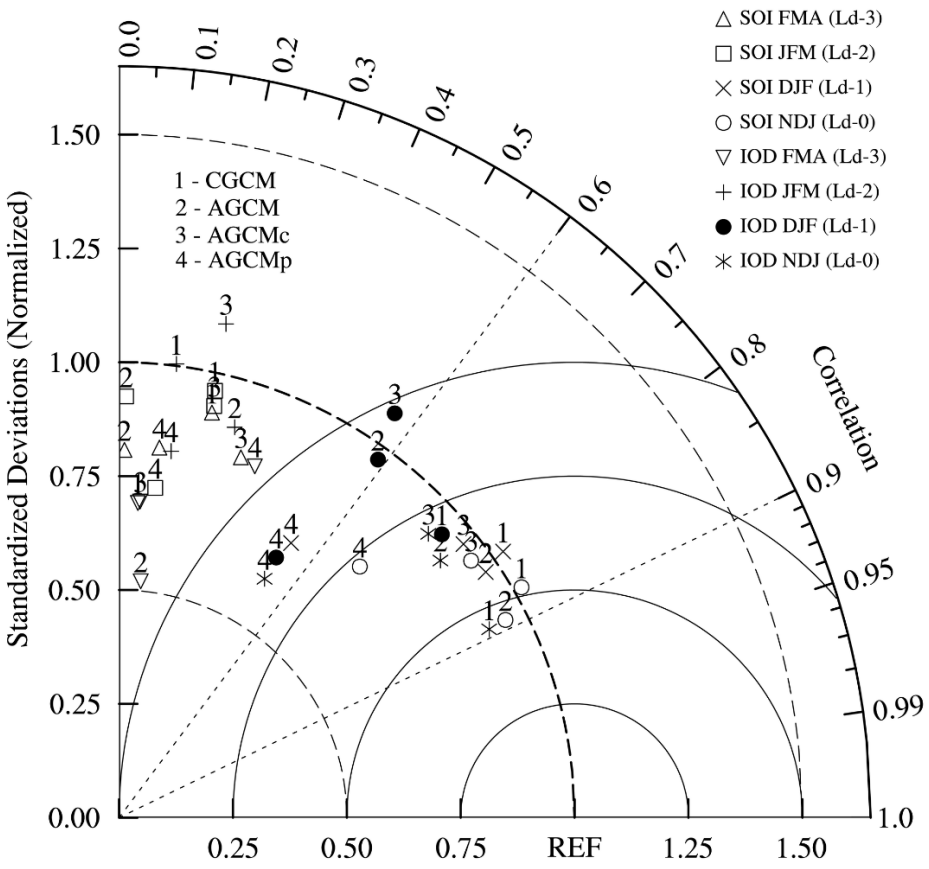

Figure 11. Taylor diagram by CGCM, AGCM, AGCMc and AGCMp (AGCM forced with persisted SST anomalies) based on the ensemble mean from the November initialized hindcasts for the equatorial Indian Ocean Dipole (IOD) and ENSO forecasts (SOI). The standard deviation is normalized by the respective NCEP/DOE. The indices are computed from the respective mslp fields (see text).

The rainfall analysis conducted over various ocean basins along the equatorial Indo-Ocean region and southern Africa sub-continent is presented in Figure 12. During the austral summer, all

prediction methods demonstrate nearly similar skills in predicting rainfall and tendencies representing its interannual variability for most regions considered. The exception is that the CGCM performs slightly better than the two AGCM configurations over the equatorial western Pacific at one month lead-time while both the AGCM and the AGCMc perform noticeably better than the CGCM over the equatorial eastern Pacific at three month lead-time. Furthermore, the AGCM simulations manifest a tendency to overestimate the interannual variability over the equatorial Indian Ocean and southern Africa region while the CGCM shows a similar tendency over the equatorial eastern Pacific basin as the lead-time increases. The other difference worth mentioning is that in most instances the AGCM skill is slightly better than the AGCMc and the improvement deepens over the equatorial Indian Ocean at three month lead-time.

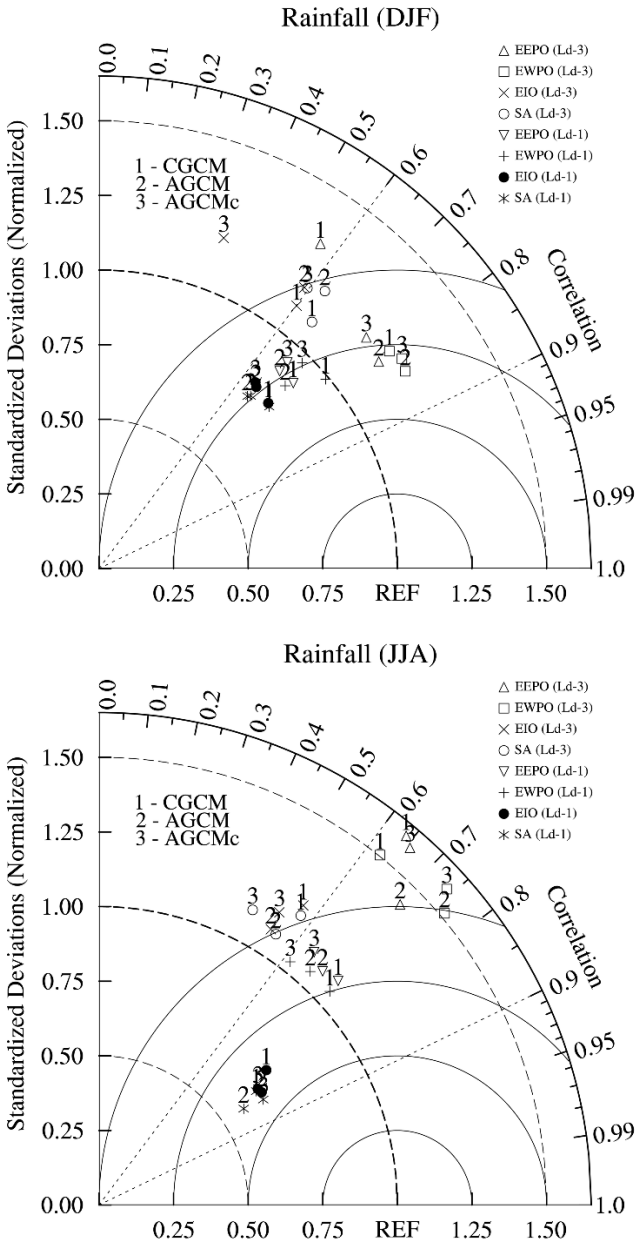

Figure 12. Taylor diagram by one- and two-tiered prediction methods (as shown in the inset) predicting spatially averaged rainfall based on their ensemble means for the southern Africa (SA; 35^{o}S-0 and 0-55^{o}E; masked over the ocean), the equatorial region (20^{o}S-20^{o}N) of Indian Ocean (EIO; 50^{o}E-110^{o}E), western Pacific Ocean (EWPO; 120^{o}E-170^{o}E), and eastern Pacific Ocean (EEPO; 170^{o}E - 60^{o}W). The verification is for the austral summer (top panel) and winter (lower panel) at one and three months lead-times. All standard deviations are normalized by CMAP for the respective basins.

The GCMs demonstrate similar levels of skills during the JJA season at one month lead-time as in the case for the DJF season. At three month lead-time, the biggest skill difference between the AGCM and CGCM is found over the eastern and western Pacific region in favor of the AGCM. The AGCM skill is also found to be better than the AGCMc noticeably over the eastern part of the basin. This skill improvement is presumably attributed to the multi-model SST forcing. By and large, rainfall variability over southern Africa and the equatorial Indian Ocean is severely underestimated in all prediction methods with a tendency to be slightly worse in the AGCM. In addition, the interannual rainfall variability is also over or underestimated more during the austral winter than summer season.

The result further indicates that (at least for the austral summer at a one-month lead-time) there is a noticeable similarity in the manner in which the GCMs vary in their skills in the prediction of ENSO (Figure 11), and of the rainfall of most regions (Figure 12). In both cases, the superiority of one model over the other is nearly indistinguishable which suggests ocean-wide atmospheric response to ENSO in both models. The contribution of the equatorial IOD is not clearly manifested on the rainfall predictability of most of the regions considered in the analysis in favor of the CGCM, despite the CGCM's superiority over the AGCM being noticeable in the prediction of IOD. This lack of teleconnection in the GCMs is not clear and is deferred for future work. Observational studies [*e.g. Yang et al. 2010*] report the strong association of IOD with the Asian monsoon during the peak season.

4.2 Comparison based on probabilistic forecasts

In this comparative experiment, the probabilistic scores are calculated from three equi-probable categories, defining below-normal, near-normal and above-normal. The categories are identified from the 33rd and 67th percentiles of the 28 years climatological record. The relative operating characteristic (ROC) area is commonly applied to probabilistic forecasts to measure the ability of a forecasting system to discriminate events such as flood or drought seasons from non-events [*Mason and Graham, 2002*]. Skillful probabilistic forecast, therefore, possesses higher frequencies of hit rates than false alarms in order to yield the area beneath the ROC curve to be

greater than 0.5. The global distributions of the ROC score differences between the CGCM and the AGCM during the austral summer based on the November initialized hindcasts in predicting years of wet and dry conditions are shown in Figure 13. The ROC scores are independently computed against the CMAP rainfall estimates first, and only those scores which are statistically significant at 95% are retained for the comparison, meaning that those probabilistic forecasts which are not better than guessing are omitted. The significance test is conducted using a variant of the Mann-Whitney non-parametric procedure that explicitly accounts for variance adjustment caused by incidents of ties [*Mason and Graham, 2002; Wilks, 2006*]. From a visual inspection, the CGCM is seemingly doing better over the southern African sub-continent, southern Indian Ocean and Pacific region near the equator. The AGCM, on the other hand, is more successful over the equatorial Indian Ocean off the coast of eastern Africa, over the central and eastern Pacific around 10^{o}N and southern Pacific region off the coast of north-eastern Australia. Broadly speaking, however, the two models are more or less similar in their ability to discriminate below- and above-normal rainfall conditions and most of the differences are as small as 0.1 or 0.15.

According to the global surface temperature skill comparison (Figure 14), the GCMs differ significantly in their ability to differentiate warm or cold episodes from non-events. The skill variations are, however, spatially and seasonally dependent. For instance, during the NDJ season at a zero month lead-time, the AGCM outperforms the CGCM over equatorial Africa for both below- and above-normal temperature conditions (Figure 14a, 14b) with a tendency to persist into the DJF season at a one month lead-time for the upper tercile (Figure 14d). But the condition is changed in favor of the CGCM as the lead-time increases for the lower tercile (Figure 14c, e). Elsewhere, there is an apparent equal distribution of the CGCM and the AGCM predominance.

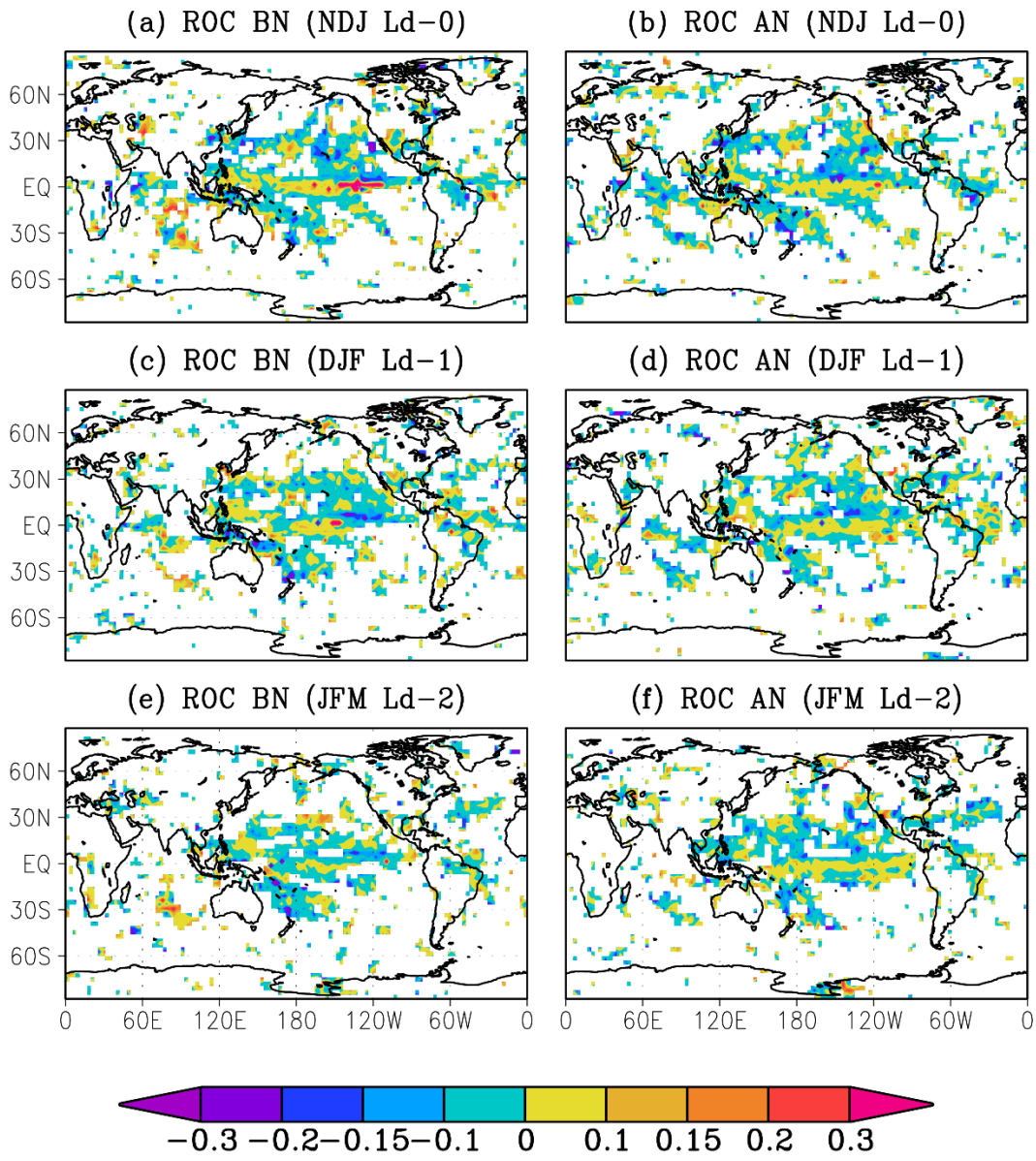

Figure 13. ROC area differences between the CGCM and AGCM for seasonal rainfall totals (mm). The +ve (-ve) scores imply the CGCM (AGCM) is better in discriminating dry (a,c,e) or wet (b,d,f) episodes than the AGCM (CGCM). These differences are computed using the November initialized integrations for various seasons and lead-times as shown in the title of each plot and the skills are independently computed first against the corresponding CMAP estimates. The differences are statistically significant at the 95% level.

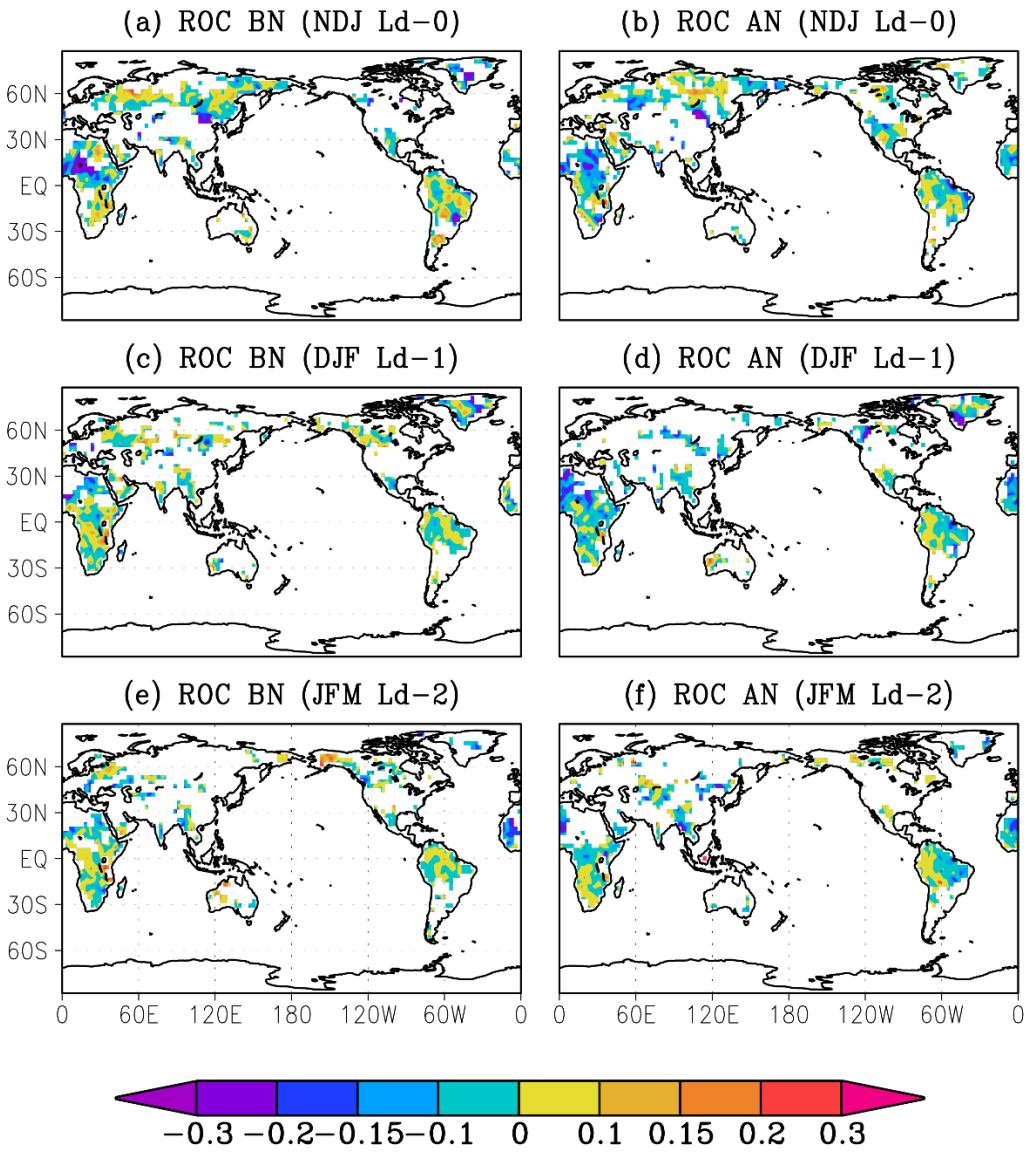

Figure 14. ROC area difference between the CGCM and AGCM for 2m temperatures. The +ve (-ve) scores imply the CGCM (AGCM) is better in discriminating cold (a,c,e) or hot (b,d,f) episodes than the AGCM (CGCM). These differences are computed using the November initialized integrations for various seasons and lead-times as shown in the title of each plot and the skills are independently computed first against the corresponding CRU estimates. The differences are statistically significant at the 95% level.

Table 1. Probabilistic skill of the GCMs in predicting cold and warm events as measured using the ROC area aggregated over three different regions[a].

	Tropics					Southern Extratropics					Southern Africa				
Lead	T1	T2	T2r	T2c	T2p	T1	T2	T2r	T2c	T2p	T1	T2	T2r	T2c	T2p
							Cold events								
0	0.66	0.67	0.65	0.63	0.68	0.59	0.56	0.55	0.52	0.56	0.62	0.60	0.59	0.59	0.60
1	0.69	0.68	0.67	0.65	0.65	0.60	0.60	0.60	0.56	0.57	0.69	0.66	0.63	0.66	0.60
2	0.70	0.70	0.69	0.68	0.64	0.59	0.61	0.60	0.59	0.54	0.77	0.73	0.72	0.73	0.62
3	0.73	0.72	0.71	0.70	0.62	0.60	0.62	0.62	0.63	0.52	0.77	0.74	0.72	0.74	0.59
							Warm events								
0	0.68	0.70	0.68	0.66	0.68	0.60	0.57	0.58	0.55	0.57	0.65	0.65	0.62	0.63	0.63
1	0.69	0.71	0.70	0.68	0.67	0.59	0.59	0.59	0.57	0.56	0.69	0.69	0.68	0.67	0.63
2	0.70	0.72	0.70	0.69	0.63	0.58	0.59	0.59	0.60	0.53	0.76	0.74	0.72	0.73	0.60
3	0.72	0.72	0.71	0.71	0.63	0.60	0.62	0.61	0.62	0.54	0.77	0.76	0.74	0.77	0.59

[a]The analysis is based on the November initialized hindcasts and 0 lead stands for NDJ, 1 for DJF etc. T1, T2, T2r, T2c and T2p represent respectively the CGCM, AGCM, AGCMr, AGCMc and AGCMp. The ensemble sizes used in the analysis for the various forecast strategies are as in Table 1. In this analysis, Tropics and southern Extratropics global zonal belt respectively 20^{o}S – 20^{o}N and 20^{o}S and 20^{o}S-90^{o}S. Sothern Africa (SA) as in Figure 12.

We extend the comparison further by including various model setup options of the AGCM. The analysis mainly focuses on the seasons surrounding the austral summer, since, as noted earlier, it is an active period of ENSO. Additionally, it is the period when maximum skill is mostly found, particularly at the southern Africa sub-continent at the seasonal timescale. Table 1 shows the ROC area analysis aggregated over different regions by various forecast methods in predicting below-

and above-normal surface air temperature conditions. In this comparison, the CGCM is compared with the AGCM with a full ensemble size, reduced ensemble size (AGCMr), the AGCM forced with the CGCM predicted SSTs (AGCMc) and persistence (AGCMp). The last three have the same ensemble size as that of the CGCM. The reduction in ensemble size is made by retaining the AGCM simulations that use multi-model ensemble mean SSTs as lower boundary conditions only. The idea is to explore how the AGCM forecast quality fluctuates by differing them both in terms of ensemble size and SST forcing. The result shows that the AGCM's fidelity in discriminating hot and cold events from non-events is reasonably reduced in the case of AGCMp and its superiority to the CGCM is noticeably lost in all regions and lead-times considered. The result is consistent to the deterministic skill presented above (Figure 11). The ability of the AGCM to distinguish events from non-events is also slightly reduced with the reduction of the ensemble size (AGCMr) and AGCMc. The latter improvement suggests that it is beneficial to use the multi-model approach to obtain the SST fields to force the AGCM.

In seasonal climate predictions, the forecast quality is better described by virtue of its reliability (calibration) and resolution (sharpness). These measures of skills are commonly practiced to compensate for the potential drawback of the ROC scores in the event where the system is not free of forecast biases. Consequently, we compare the GCMs using the Brier skill score [*Murphy, 1988*] and the reliability diagrams [*Hartmann et al. 2002*].

The Brier Score (BS) provides a handy measure of accuracy (bias) of probabilistic forecasts aggregated over all forecast probability bins. It has a negative orientation ranging between 0 and 1. In this context, the probabilistic forecast attains perfection when the BS approaches zero. The BS can be decomposed into three terms algebraically i.e., (BS = reliability - resolution + uncertainty) [*Murphy, 1973; Wilks, 2006*]. A skillful probabilistic forecast therefore attempts to have the lowest possible value and the largest possible value of reliability (B_{rel}) and resolution (B_{res}). Conversely, the uncertainty term (B_{unc}) is independent of the forecast itself and is determined by the inherent circumstance of the observed climatological frequency of the events [see *Wilks, 2006*].

In this comparative framework, the Brier Skill Score (B_{ss}) is used to measure the relative benefit of one model over the other, i.e., ($BSS = 1 - BS_{CGCM}/BS_{AGCM}$) similar to the MSSS (section 4.1). Likewise, the relative benefits of the reliability and resolution terms are assessed with the

same formula, except that the resolution term is normalized by the uncertainty term [*Graham et al., 2005*].

The CGCM and the AGCM comparison in terms of the BSS, B_{rel} and B_{res} for surface air temperature during the mid-austral summer at one month lead-time is presented in Table 2. The AGCM (with full ensemble size) exhibits a better performance than the CGCM in terms of BSS and B_{rel} for the three regions considered. The maximum benefit of reliability of the AGCM is found in the southern Africa region (73.1% and 69.7% for cold and warm events respectively) followed by the tropics and then the SH extra-tropics. However, in terms of B_{res}, the CGCM (AGCM) tends to outperform the AGCM (CGCM) in predicting cold (hot) events. The AGCMr also attains a better B_{rel} and B_{ss}, with the exception of the below-normal temperature over the southern Africa region, despite the fact that the B_{rel} drops by about 50%. It is noticeable that the CGCM mostly outscores the AGCMc in terms of B_{ss} and B_{res} while the AGCMc performs better that the CGCM in terms of B_{rel}. However, the CGCM has discernibly outperformed persistence (AGCMp) in all Brier terms for all regions.

Table 2. The CGCM's benefit relative to various AGCM configurations predicting 2m temperature[a].

	Tropics			Southern Extratropics			Southern Africa		
Tercile	Bss	Brel	Bres	Bss	Brel	Bres	Bss	Brel	Bres
					CGCM vs. AGCM				
Lower	-3.7	-43.1	**0.5**	-4.2	-41.1	**0.5**	**0.2**	-73.1	**3.5**
Upper	-7.9	-63.0	-2.7	-4.7	-41.0	-0.2	-7.0	-69.7	-2.0
					CGCM vs. AGCMr				
Lower	-0.8	-25.0	**1.7**	-3.0	-28.2	**0.2**	**7.6**	**16.4**	**6.9**
Upper	-4.3	-40.4	-1.1	-2.0	-17.9	-0.1	-1.0	-36.5	**1.6**
					CGCM vs. AGCMc				

Lower	**1.1**	-20.5	**2.9**	**2.0**	**4.8**	**1.7**	**1,2**	-57.5	**4.1**
Upper	**0.8**	-14.8	**2.0**	**0.8**	-0.7	**1.0**	-1.1	-37.6	**1.4**
					CGCM vs. AGCMp				
Lower	**3.0**	**1.0**	**3.2**	**2.9**	**12.1**	**1.6**	**13.4**	**40.5**	**10.1**
Upper	**3.0**	**8.0**	**2.6**	**3.9**	**22.4**	**1.2**	**8.1**	**19.5**	**7.2**

[a]The relative probabilistic skills are measured using the Brier skill score (Bss) and its algebraic decompositions, i.e., Brier reliability (Brel) and Brier resolution (Bres). Positive CGCM benefits are shown in bold type against various AGCM forecast strategies. The analysis is for DJF at one month lead-time. The CGCM, AGCMr, AGCMc and AGCMp configurations use 10 ensemble members while AGCM uses 30 ensemble members (see text).

To gain a deeper insight into their performance differences, the GCMs are further compared using reliability diagrams. The reliability diagram is a graphical tool that is constructed from the computation of the hit rate for the set of forecasts for individual probability bins separately (as opposed to the generalization in the case of the BSS and its decomposition terms), and then plotted against the corresponding forecast probabilities. [*Hartmann et al. 2002; Wilks, 2006*]. The most reliable forecasting system is determined by the extent of its proximity to the diagonal line (perfect reliability).

Figure 15 shows the reliability diagrams for the southern African and Tropical regions. The verification for unusually warm (wet) and cold (dry) events during the austral summer (DJF) at one month lead-time are for the CGCM and for the three cases of the AGCM hindcasts. Also shown is the relative frequency of the use of the forecast bins, which is commonly referred to as the "sharpness diagrams" on the left top corner of each plot both for below- and above-normal conditions. The result shows that the AGCM (forced by the multi-model SST forecasts) and the CGCM both demonstrate similar levels of skill in their ability to detect unusual conditions. Notwithstanding, the CGCM shows relatively more overconfidence than the AGCM at higher probability bins (particularly 0.8), which presumably clarifies the reason why the CGCM is heavily penalized in terms of the B_{ss} and B_{rel} (Table 1).

The result of Figure 15 further indicates that the CGCM and the AGCMc exhibit slightly better sharpness for probabilistic temperature and rainfall predictions respectively, as the sharpness diagrams are flattening when compared to all cases of the other forecast methods. Generally, however, these different forecasting methods tend to fall mostly in the lower or climatological probabilities particularly for rainfall suggesting that the GCMs are more reluctant to issue warnings with higher probabilities.

The reduction of the ensemble size has only caused a minor change in the reliability diagram's shape (i.e. a slight displacement of the curve towards overconfidence when compared to the use of the full ensemble), meaning that the skill drop is too small to alter the circumstance in favor of the CGCM. The AGCMc attains more or less a comparable reliability level to the AGCM (with full or reduced ensemble size), even though its reliability is slightly compromised for probabilistic temperature prediction during the austral summer. Notwithstanding, there is a substantial degradation of skill in favor of the CGCM when the AGCM is forced with persisted SST (AGCMp).

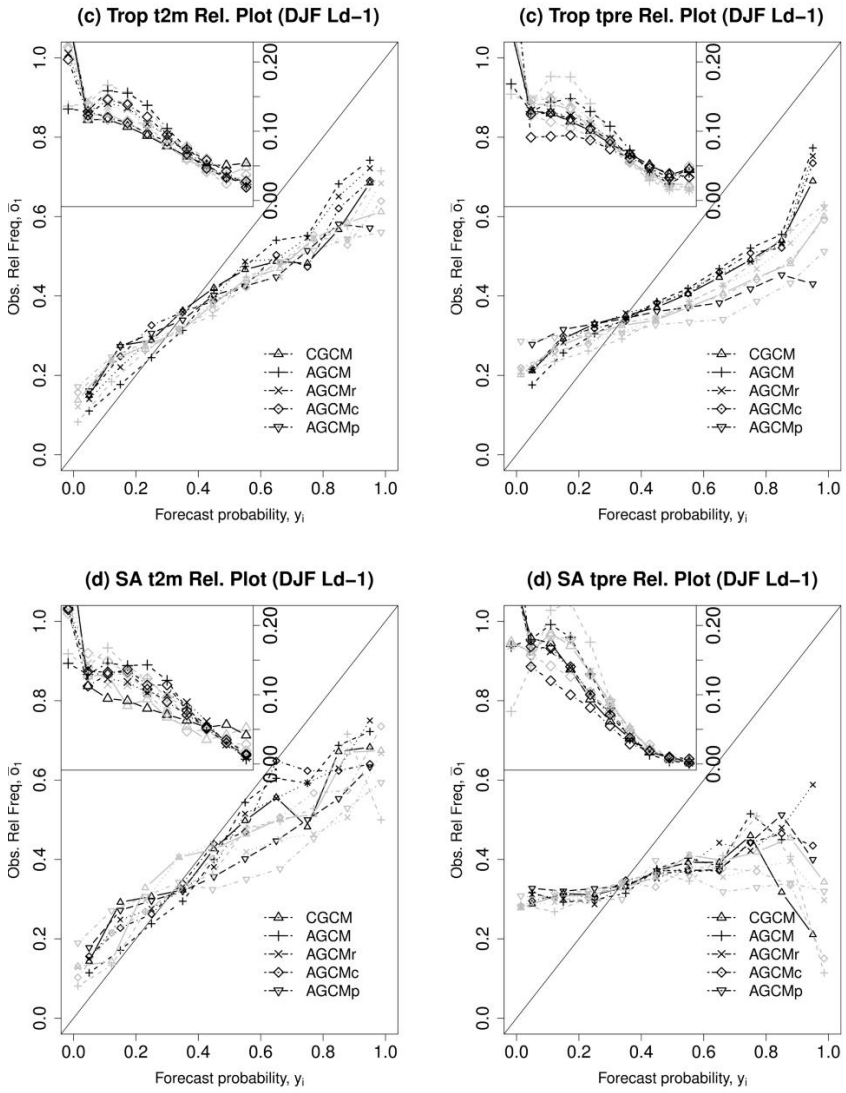

Figure 15. Reliability diagrams by the CGCM (10 ensemble size), AGCM (30 ensembles size), AGCMr (ensemble size reduced to 10), AGCMc (from the SCM predicted SST forced AGCM); ensemble size reduced to 10) and AGCMp (from persisted SST forced AGCM integrations; 10 ensemble size) in predicting below- and above-normal surface air temperature and rainfall conditions during the austral summer seasons (DJF) at one month lead-time for the Tropical region between 20°S and 20°N (top panel) and the Southern African region (bottom panel). The frequency of utilization the different probability bins for both below- and above-normal categories are also shown on the top-left corners of each diagram. The grey and black lines represent cold (dry) and warm (wet) events, respectively.

5. Summary and conclusions

The steady increase over recent years in the use of coupled models for seasonal forecasting has been at the expense of uncoupled models owing to the fast development of computational resources and the envisaged advantage of coupled models in representing state-of-the-art seasonal forecasts more realistically. Despite many numerical studies conclusively present evidence in favor of coupled models, a gap still exists whether these models are similar or differ widely in their predictive skill in an operational and hence practical environment. With this in mind, we revisit the subject under a practical model framework, where a multi-model SST anomalies and its uncertainty envelope are used to constrain the atmospheric model. This model comparison study uses the SCM and its atmosphere-only version, which run concurrently at the SAWS for seasonal forecast production in a multi-model environment. Furthermore, the two models are suitably configured in such a manner that the role of coupling on the predictive skill differences is better distinguished. In this experimental framework, the GCMs share a great deal of resemblance in their configuration except for the manner in which the SST information is communicated within the GCMs.

The analysis finds that the two models are able to represent the observed spatial patterns of rainfall and that climatologically, they do not differ strongly in terms of bias distribution both during the SH summer and winter seasons even though the models are somewhat more biased for the latter season. In addition, the comparative analysis reveals that the symmetry and position of the ITCZ and the mid-latitude storm tracks are well represented in both models with a tendency of the CGCM to overestimate the peak of the rainfall distribution in the tropics. This overestimation over the tropics is presumably attributed more to the SST bias than to the air-sea coupling process which is largely minimized in the AGCM with the use of multi-model SST forcing. There are two possible reasons that may substantiate the finding. Firstly, the intensity and distribution of biases are mostly found during the austral winter period with mostly marginal difference between the GCMs which tends to coincide with the poor predictive skill of ENSO (due to the NH spring barrier). Secondly, the biggest bias differences during the austral summer and winter seasons between the GCMs tend to coincide for the most part with neutral ENSO conditions.

Results from the predictive skill comparisons indicate that most of the differences in the skill of the GCMs arise over the tropical region. Outside the tropics, the superiority of one model over the other is mostly indistinguishable, and the skill levels are also generally lower than over the tropics. The result further indicates that there is a noticeable similarity in the manner in which the GCMs vary in their skills in the prediction of ENSO and rainfall over the equatorial Indo-Pacific basins and the southern Africa sub-continent. In both cases the superiority of one model over the other is mostly indistinguishable and suggests an ocean-wide atmospheric response to ENSO.

In addition, the AGCM's fidelity is drastically reduced in the case of AGCMp to the extent that where the superiority of the CGCM becomes noticeable. The benefit of the AGCM over the CGCM is also slightly reduced with decreasing ensemble size, but not to an extent that can lead to the shift in the superiority balance. Again, this result attests to the conclusion that the role of the multi-model SST forcing is paramount and is the reason why the AGCM and CGCM have comparable levels of forecast skill.

Generally, what has transpired from this comparative experiment is that the GCMs differ widely in their performances, and the issue of the superiority of one model over the other is mostly dependent on space and time (seasonality). One may conclude that the CGCM has the upper hand in the Asian monsoon region during the austral winter. However, the CGCM skill becomes weaker with the increase of lead-time or in a different season. The diversity in their predictive skill as function of space and time may be beneficial in complementing each other in some way.

The modelling work presented here suggests a circumstance under which AGCMs and CGCMs may be able to produce similar levels of skill, notwithstanding the fact that only two such models were considered. At the very least, the study has provided some guidance on how best to optimize an AGCM under circumstances in which limited computational resources only supports the use of AGCMs in an operational forecast environment, a situation commonly found in developing countries such as South Africa. An optimal AGCM configuration, however, depends heavily on skillful SST forecasts, here obtained through a multi-model SST forecast system. These predicted SSTs may be reproduced from a number of CGCMs and such ocean-atmosphere models are therefore essential for skillful seasonal climate predictions when AGCMs are used. A significant amount of work and investment has already gone into AGCM development, but the potential for further improvement of AGCM-based forecasts thus depends to a large extent on the

improvement of CGCMs. Nowadays many leading institutions make their seasonal forecast, including SST forecasts, freely available to national and regional centers under the auspices of the World Meteorological Organization (WMO) and so such SST forecasts can be assimilated into AGCM operational forecast systems.

Acknowledgment

The work was supported financially by the Water Research Commission (WRC) and Applied Centre for Climate & Earth Systems Science (ACCESS). The authors are also gratefully appreciative for the CHPC's computational support. Furthermore, computer hardware obtained through the SATREPS (Science and Technology Research Partnership for Sustainable Development) a collaborative project between Japan and South Africa was essential for processing massive model output data. The Max-Planck-Institut für Meteorologie (MPI) has kindly provided the ECHAM4.5 AGCM code. The work was also impossible without the NCEP reanalysis product. GCMs data used in the study are available for research purpose on request via the SAWS ftp server (ftp://ftp.saws.co.za; email: asmerom.beraki@weathersa.co.za).

References

Beraki A.F., W. Landman, D. DeWitt and C. Olivier 2015: Global dynamical forecasting system conditioned to robust initial and boundary forcings: Seasonal Context. *Int. J. Climatol.*, Submitted.

Beraki A.F., D.G. DeWitt, W.A. Landman, and C. Olivier, (2014), Dynamical seasonal climate Prediction using an ocean–atmosphere coupled climate model developed in partnership between South Africa and the IRI. *J. Clim.*, 27, 1719-1741.

Barnston A.G., A. Leetmaa, V. Kousky, R. Livezey, E. O'Lenic, H. Van den Dool, A.J. Wagner, and D Unger, (1999), NCEP Forecasts of the El Niño of 1997-98 and Its U.S. Impacts. *Bull. Am. Meteor. Soc.*, 80: 1829-1852.

Behera, S. K., and T. Yamagata (2003), Influence of the Indian Ocean dipole on the Southern Oscillation, J. Meteorol. Soc. Japan., 81, 169‑177.

Bengtsson, L., U. Schlese, E. Roeckner, M. Latif, T. Barnett, and N. Graham, (1993), A two-tiered approach to long-range climate forecasting, *Science*, 261, 1026-1029.

Boville, B.A., and J.W. Hurrell, (1998), Comparison of the Atmospheric Circulations Simulated by the CCM3 and CSM1. *J. Clim.*, 11, 1327-1340.

Charney, J. G., and J. Shukla, (1981), Predictability of monsoons. Monsoon Dynamics, J. Lighthill and R. P. Pearce, Eds., Cambridge Univ. Press, New York, 99-109.

Chaudhari, H.S., S. Pokhrel, S. Mohanty, and S.K. Saha, (2013), Seasonal prediction of Indian summer monsoon in NCEP coupled and uncoupled model. *Theor. Appl. Climatol.*, 114, 459–477. doi:10.1007/s00704-013-0854-8

Colfescu, I., E. K. Schneider, and H. Chen, (2013), Consistency of 20th century sea level pressure trends as simulated by a coupled and uncoupled GCM. *Geophys. Res. Lett.*, 40, 3276–3280, doi:10.1002/grl.50545.

Copsey D, R.Sutton, and J. R. Knight, (2006), Recent trends in sea level pressure in the Indian Ocean region. *Geophys. Res. Lett.*, 33: L19712, DOI:10.1029/2006GL027175.

Derber, J., and A. Rosati, (1989), A global oceanic data assimilation system. *J. Phys. Oceanogr.*, 19, 1333–1347.

DeWitt, D.G., (2005), Retrospective forecasts of interannual sea surface temperature anomalies from 1982 to present using a directly coupled atmosphere-ocean general circulation model, *Mon. Wea. Rev.*, 133, 2972-2995.

Doblas-Reyes, F.J., R. Hagedorn, and T.N. Palmer, (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus* 57A, 234–252.

Fu, X.H., B. Wang, and T. Li, (2002), Impacts of air–sea coupling on the simulation of mean Asian summer monsoon in the ECHAM4 model. *Mon Wea. Rev. 130*, 2889–2904

Goddard, L., S.J. Mason, S.E., Zebiak, C.F., Ropelewski, R. Basher, and M.A. Cane, (2001), Current approaches to seasonal-to-interannual climate predictions. *Int. J. Climatol.*, 21, 1111–1152.

Graham, R. J., A. D. L. Evans, K. R. Milne, M. S. J. Harrison, and K. B. Robertson, (2000), An assessment of seasonal predictability using atmospheric general circulation models, *Q. J. R. Meteorol. Soc*., 126, 2211-2240.

Graham R.J., M. Gordon, P.J., McLean, S. Ineson, M.R., Huddleston, M.K. Davey, A. Brookshaw, and R.T.H. Barnes, (2005), A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus* 57, 320–319.

Hagedorn, R, F.J. Doblas-Reyes, and T.N. Palmer, (2005), The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. *Tellus*, 57A, 219–232.

Harris I, P. D. Jones, T. J. Osborn, and D. H. Lister, (2014), Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.*, 34 623–642.

Hartmann, H. C., T. C. Pagano, S. Sorooshian, and R. Bales (2002), Confidence builders: Evaluating seasonal climate forecasts for user perspectives, Bull. Am. Meteorol. Soc., 83, 683‑698.

Jha, B., and A. Kumar, (2009), Comparison of the atmospheric response to ENSO in coupled and uncoupled model simulations. *J. Clim.*, 137, 479-487.

Ji, M., D. W. Behringer, and A. Leetmaa, (1998), An improved coupled model for ENSO prediction and implications for ocean initialization. Part II: The coupled model. *Mon. Wea. Rev.*, 126, 1022–1034.

Kanamitsu, M., and Coauthors, (2002), NCEP dynamical seasonal forecast system 2000. *Bull. Amer. Meteor. Soc.*, 83, 1019–1037.

Kirtman, B. P., and Coauthors, (2014), The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; Phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc*., 95, 585-601, doi:10.1175/BAMS-D-12-00050.1.

Kirtman, B.P., J. Shukla, B. Huang, Z. Zhu, and E.K. Schneider, (1997), Multiseasonal Predictions with a coupled tropical ocean-global atmosphere system. *Mon. Wea. Rev.*, 125, 789-808.

Komori, N., A. Kuwano-Yoshida, T. Enomoto, H. Sasaki, and W. Ohfuchi, (2008), High-resolution simulation of the global coupled atmosphere-ocean system: Description and preliminary outcomes of CFES (CGCM for the Earth Simulator). High Resolution Numerical Modelling of the Atmosphere and Ocean, K. Hamilton and W. Ohfuchi, Eds., Springer, 241–260, Springer, New York.

Krishnamurti, T.N., C.M. Kishtawal, Z. Zang, T. LaRow, D. Bachiochi, E. Williford, S. Gadgil, and S. Surendran, (2000), Multimodel ensemble forecasts for weather and seasonal climate. *J. Clim.*, 13, 4196–4216.

Kug J.-S, I-S, Kang, and D-H., Choi, (2008), Seasonal climate predictability with Tier-one and Tier-two prediction systems. *Clim. Dyn.*, 31, 403–416. doi:10.1007/s00382-007-0264-7.

Landman, W.A., and A. Beraki, (2012), Multi-model forecast skill for mid-summer rainfall over southern Africa. *Int. J. Climatol.*, 32: 303–314. DOI: 10.1002/joc.2273.

Landman W.A., D. DeWitt, D.-E. Lee, A. Beraki, D. Lötter, (2012), Seasonal rainfall prediction skill over South Africa: 1- vs. 2-tiered forecasting systems. *Wea. Forecasting*, 27, 489-501. DOI: 10.1175/WAF-D-11-00078.1.

L'Heureux, M. L., Collins, D. C., and Hu, Z.-Z. (2012), Linear trends in sea surface temperature of the tropical Pacific Ocean and implications for the El Niño-Southern Oscillation. *Clim. Dyn.*, 1–14. doi:10.1007/s00382-012-1331-2

Mason, S. J., and N. E. Graham (2002), Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation, Q. J. R. Meteorol. Soc., 128, 2145‑2166.

Mo, K. C., and M. Ghil, (1987), Statistics and dynamics of persistent anomalies. *J. Atmos. Sci.*, 44, 877–901.

Molteni, F., L. Ferranti, M. Balmaseda, T. Stockdale, and F. Vitart, (2007), ECMWF seasonal forecast system 3. *CLIVAR Exchange*, 43, 7-9.

Murphy, A. H., (1988), Skill scores based on the mean square error and their relationships to the correlation coefficient. *Mon. Wea. Rev.*, 116, 2417–2424.

Neelin, J.D., D.S. Battisti, A.C. Hirst, F.-F. Jin, Y. Wakata, T. Yamagata, and S. Zebiak, (1998), ENSO Theory. *J. Geophys. Res*., 103, 14,261–14,290.

North, G.R., (1984), Empirical orthogonal functions and normal modes. *J. Atmos. Sci.*, 41, 879–887.

Pacanowski, R. C., and S. M. Griffes, (1998), MOM 3.0 manual. *NOAA/Geophysical Fluid Dynamics Laboratory Rep.*, 608 pp.

Palmer, T.N., and D.L.T. Anderson, (1994), The prospects for seasonal forecasting—A review paper. *Quart. J. Roy. Meteor. Soc.* 120, 755–793.

Palmer, T. N., and Coauthors, (2004), Development of a European multimodel ensemble system for seasonal-to-interannual prediction (DEMETER), *Bull. Am. Met. Soc.*, DOI: 10.1175/BAMS-85-6-853.

Philander, S.G.H., (1990), *El Niño, La-Niña and the Southern Oscillation*, 289 pp., Academic Press, San Diego, Calif.

Reynolds, C. A., P. J. Webster, and E. Kalnay (1994), Random error growth in NMC's global forecasts, Mon. Weather Rev., 122, 1281‑1305.

Roeckner, E., and Coauthors, (1996), Simulation of present-day climate with the ECHAM4 model: Impact of model physics and resolution, *Report No. 93, Max-Planck-Institut für Meteorologie, Hamburg, Germany, 171 pp*.

Ropelewski, C. F., and P. D. Jones, (1987), An extension of the Tahiti-Darwin Southern Oscillation Index, *Mon. Wea. Rev.*, 115, 2161–2165.

Shukla, R., and J. Zhu, (2014), Simulations of boreal summer intraseasonal oscillation with CFSv2 over India and western Pacific: role of air-sea coupling. *Atmosphere-Ocean*, 52, 321-330. DOI:10.1080/07055900.2014.939575.

Saha, S., S. Moorthi, X. Wu, J. Wang, S. Nadiga, P. Tripp, D. Behringer, Y.-T. Hou, H. Chuang, M. Iredell, M. Ek, J. Meng, R .Yang, M.P. Mendez, H. van den Dool, Q. Zhang, W. Wang, M. Chen, and E. Becker, (2014), The NCEP Climate Forecast System Version 2. *J. Clim.* 27, 2185-2208.

Saha, S., S. Nadiga, C. Thiaw, J. Wang, W. Wang, Q. Zhang, H.M. Van den Dool, H.-L. Pan, S. Moorthi, D. Behringer, D. Stokes, M. Peña, S. Lord, G. White, W. Ebisuzaki, P. Peng, and P. Xie, (2006), The NCEP Climate Forecast System. *J. Clim.,* 19, 3483-3517.

Saji, N. H., B. N. Goswami, P. N. Vinayachandran, and T. Yamagata, (1999), A dipole mode in the tropical Indian Ocean. *Nature*, 401, 360–363.

Stockdale, T.N., D.L.T Anderson, J.O.S. Alves, and M.A. Balmaseda, (1998), Global seasonal rainfall forecasts using a coupled ocean-atmosphere model. *Nature*, 392, 370–373.

Taylor, K. E. (2001), Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183‑7192.

Tennant, W.J., and B.C., Hewitson, (2002), Intra-seasonal rainfall characteristics and their importance to the seasonal prediction problem. *Int. J. Climatol.,* 22, 1033–1048.

Troccoli A, M. Harrison, D.L.T. Anderson, S.J. Mason, (2008), *Seasonal Climate: Forecasting and managing risk.* NATO Science Series. Earth and Environmental Sciences Vol 82. Springer: Dordrecht, The Netherlands.

Wallace, J.M., E.M. Rasmusson, T.P. Mitchell, V.E. Kousky, E.S. Sarachik, and H. von Storch, (1998), On the structure and evolution of ENSO related climate variability in the tropical Pacific: Lessons from TOGA. *J. Geophys. Res.*,103, 14,241–14,259.

Wilks, D.S., (2006), *Statistical Methods in the Atmospheric Sciences, 2nd* 320 *Edition*. Academic Press, pp. 627.

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

Yang, J., Q. Liu, and Z. Liu, (2010), Linking Observations of the Asian Monsoon to the Indian Ocean SST: Possible Roles of Indian Ocean Basin Mode and Dipole Mode. *J. Climate*, 23: 5889–5902. doi: http://dx.doi.org/10.1175/2010JCLI2962.1

Yu J.-Y., and C.R. Mechoso, (1999), A Discussion on the Errors in the Surface Heat Fluxes Simulated by a Coupled GCM. *J. Clim*. 12, 416–426.

Zhao, M., and H. H. Hendon (2009), Representation and prediction of the Indian Ocean dipole in the POAMA seasonal forecast model, Q. J. R. Meteorol. Soc., 135, 337‑352.

Zhu, J., B. Huang, L. Marx, J. L. Kinter III, M. A. Balmaseda, R-H. Zhang, Z-Z. Hu, (2012), Ensemble ENSO hindcasts initialized from multiple ocean analyses. *Geophys. Res. Lett.* **39**:L09602, DOI: 10.1029/2012GL051503.

Zhu, J., and J. Shukla, (2013), The role of air-sea coupling in seasonal prediction of Asian-Pacific summer monsoon rainfall. *J. Clim.,* 26, 5689-5697, doi: 10.1175/JCLI-D-13-00190.1.

Synopsis

This final part of the thesis has addressed the similarities and differences of the AGCM and CGCM in terms of their predictive skill in an operational environment by using a suitable model framework. In such a framework, the atmospheric model is constrained with predicted multi-model SST anomalies and its associated uncertainty envelope while the two GCMs are kept effectively similar in all other aspects. The comparative analysis has revealed that the predictive skill of the GCMs is more or less comparable in general terms and to a large extent the two models are complementing each other in one way or another. Furthermore, the study has demonstrated that the extent to which the AGCM's performance suffers in favour of the CGCM is only when the AGCM is constrained with persisted SST anomalies instead of more skilful multi-model SST forecasts. The modelling work has therefore demonstrated a circumstance under which the two GCM configurations may be able to produce similar levels of skill which may have far-reaching benefits particularly when computational resources do not support the use of CGCMs either in research or operational mode. This comparative study addresses objective 5 of the thesis through the use of a suitable modelling framework.

5. Summary and conclusions

The practice of contemporary seasonal climate prediction requires state-of-the-art GCMs. The predictive skill of seasonal forecast mainly arises from the slowly evolving boundary conditions notably SST evolution which is found to modulate significantly the mean state of weather conditions. The proper description of the interaction of the ocean and atmosphere is therefore fundamentally important for a reliable and robust representation of the coupled climate system. In fact, the coupling of the ocean and atmosphere is a minimum level of complexity required for seasonal-to-interannual climate predictions to work since this coupling is related to skilful predictions. This notion, compounded with the fast development of computational resources, stimulates the proliferation of the use of coupled models for seasonal forecasting over recent years in large modelling centres such as ECMWF and NCEP. While many numerical studies have conclusively shown the distinctive advantage of using CGCMs, AGCMs should be able to continue to provide useful forecast information and be a feasible alternative to coupled models without compromising the predictability level which may be attainable with the use of more sophisticated and memory intensive forecasting systems. This notion forms the bedrock of the thesis and is found to be highly relevant in the context of seasonal forecast practice particularly under a constrained computational resources environment. The thesis therefore strives to address the following objectives.

1. To introduce an optimally configured coupled GCM initialized with the best possible initialization strategy, in order to produce hindcasts that mimic a truly operational configuration at lead-time of several months.
2. To introduce an optimally configured atmospheric-only GCM forced with realistic initial atmospheric states and the best available description of the surface boundary conditions as reflected in projected global SST, in order to produce hindcasts that mimic a truly operational configuration at lead-time of several months.
3. To investigate the predictive potential of each forecasting system to represent key synoptic (regional) climatic systems and important diagnostic variables through the use of appropriate measures of skills

4. To identify deficiencies and sensitivities of the two systems in terms of representing climate processes in a manner that may promote further understanding of the coupled climate system and the subsequent lead to the improvement of the models.
5. To conduct performance comparison of coupled and uncoupled climate forecasting systems through standard verification procedure.

To achieve the main research objectives of the thesis, a suitable modelling framework was established which comprises of an interactive ocean-atmosphere coupled GCM (objective 1) and AGCM (objective 2). In this framework, the two GCM configurations are optimally configured in such a manner that the role of coupling on the predictive skill differences is distinguished better (objective 3). In addition, retroactive simulations (or hindcasts) are built in such a way that each ensemble set mimics an operational forecast instance (if it were issued). This approach offers not only a better insight into the relative enhancement (degradation) of forecast quality in an operational environment but also facilitates the operational implementation of the GCMs.

The use of CGCMs for seasonal forecasting in South Africa was initially not considered feasible particularly in an operational environment owing to the enormous computational needs of and complexities associated with CGCMs. Motivated by the recent advances in computing infrastructures in South Africa due to the establishment and maintenance of the CHPC and international collaboration, the thesis pioneered the emergence of a fully operational coupled ocean-atmosphere model in an effort to address objective 1. In fact, the SAWS Coupled Model (SCM) also referred to as the ECHAM4.5-MOM3-SA, is the first of its kind in Africa. The model couples the ECHAM4.5 (AGCM) and MOM3 ocean model (OGCM) using the MPMD coupler paradigm. In addition, this model employs an atmospheric initialization strategy that is different from previous versions of the model that coupled the same atmosphere and ocean models. This enhancement is a major and significant development in local numerical modelling efforts after 2009 when the South African Weather Service (SAWS) was granted Global Producing Centre (GPC) status for Long-Range Forecasting (LRF) by the World Meteorological Organization (WMO).

In the one-tiered experiment, the thesis demonstrated the robustness of the forecasting system through a thorough statistical analysis using both deterministic and probabilistic verification methods. Furthermore, the study conducts an intermodel comparison where the SCM is compared against various CGCMs administered by other international centres within the context of ENSO and IOD predictions. It is also noted that this climate drivers are commonly used in similar numerical studies as benchmarking. The comparison further contributes towards understanding the relative strength and weakness of the CGCM from an operational point of view which satisfy objectives 3 and 4 relevant to the coupled model.

In the two-tiered experiment, the primary focus is on how an optimized AGCM configuration's predictive skill may be used to test whether the AGCM can be a strong competitor for its coupled version and also establish a new baseline against which the CGCM can be tested. To achieve this objective, the thesis employs a multi-model approach in establishing an SST forcing field to constrain the AGCM. The advantage of a multi-model approach has been reported in many forecasting studies over recent years (e.g., Krishnamurti *et al.,* 2000; Palmer et al. 2004; Doblas-Reyes *et al.,* 2005; Hagedorn *et al.,* 2005; as also demonstrated here). In this process, the ECHAM4.5 AGCM is constrained by the lower boundary conditions derived from predicted SST of two CGCMs combined through equal weighting. The uncertainty amplitude of the SST (lower and upper bounds) is calculated from the combination. As in the CGCM, the AGCM is initialized with the realistic state of the atmosphere and soil moisture. Therefore, the GCMs share a great deal of resemblance in their respective configuration except for the manner in which the SST information is communicated within the GCMs. This operational configuration of the AGCM address objective 1 of the thesis, as noted earlier. Furthermore, through a pairwise sensitivity analysis, the study contributes towards the understanding of the strength and weakness of the AGCM which address objectives 3 and 4 relevant to the uncoupled model.

In the comparative experiment, as noted earlier, the framework is developed to elucidate the similarities and differences of the AGCM and CGCM in terms of their predictive skill in an operational environment in order to address the final objective (objective 5). The modelling work has therefore demonstrated a circumstance under which the two GCM configurations may be able

to produce similar levels of skill which may have far-reaching benefits particularly when computational resources do not support the use of CGCMs either in research or operational mode. This comparative study addresses objective 5 of the thesis through the use of a suitable modelling framework.

Major findings of the study are summarized as follows:

- With regard to the CGCM's experiment:
    - The CGCM is skilful in most instances in capturing the development and maturity of El-Niño and La-Niña episodes up to 8 months lead-time with reasonably low biases.
    - The evaluation of ENSO predictions reveals that the coupled model has skill levels comparable with other coupled models administered by international centres (such as NCEP, ECMWF, MF and UKMO).
    - The CGCM's ENSO skill is generally found to decay faster during the spring barrier.
    - The CGCM's fidelity in predicting upper air dynamics and surface air temperature is more pronounced as opposed to rainfall which is generally characterised by overconfidence.
    - Probabilistically, the analysis revealed that La-Niña events are more skilfully discriminated than El-Niño events by the model.
    - The CGCM is skilful up to several month lead-times in predicting the equatorial IOD during the period when IOD seasonal variation attains maturity.
    - The lower skill of IOD outside the peak season is due to the western segment of the dipole which is found to eventually contaminate the DMI.

- With regard to the optimization experiment:
    - The optimization of the AGCM leverages a large-scale consistent skill improvements for surface temperature and rainfall totals relative to the previous

forecasting system of SAWS and the AMIP2 simulations of the AGCM which is indicative of the robustness of the proposed AGCM forecast system.

- Evaluation of hindcasts reveals that the AGCM is able to forecast anomalous upper air atmospheric dynamics (circulation) over the tropics up to several months ahead.
- The contribution of the predicted sea-surface temperature, which is based on a multi-model approach, is shown to be of significant importance for best AGCM results.
- The model is able to significantly discriminate wet (warm) and dry (cold) episodes over the larger part of the globe where the skill for rainfall is more pronounced over the equatorial Pacific region.
- The study has addressed some optimization issues for atmospheric models and at the same time proposed an optimal AGCM that can serve as baseline against which more advanced models can be tested.

- With regards to the comparison experiment:
  - The analysis finds that the CGCM and AGCM are able to represent the observed spatial patterns of rainfall and they hardly differ in term of bias distributions both during the SH summer and winter seasons although both models are somewhat less skilful for winter.
  - The symmetry and position of the ITCZ and the mid-latitude storm tracks are well represented in both models with a tendency for the CGCM to overestimate the peak of the rainfall distribution in the tropics while the AGCM is slightly biased over the NH mid-latitude.
  - The study shows that the skill differences are presumably more attributed to the SST bias than the air-sea coupling process as the difference is found to mostly coincide with the neutral ENSO years or spring barrier.
  - Most differences in the skill of the GCMs arise over the tropics. Elsewhere the superiority of one model over the other is mostly indistinguishable and the skill levels are also generally lower than over the tropics.

- There is a noticeable similarity in the manner in which the GCMs vary in their skills in the prediction of ENSO characteristics and rainfall over the equatorial Indo-Pacific basins and the southern Africa sub-continent which suggests a noticeable atmospheric response to ENSO in both models. However, the equatorial IOD signature on the rainfall predictability is not clearly revealed in favor of the CGCM despite that the CGCM is found to simulate the IOD better than the AGCM.
- The AGCM's fidelity is drastically reduced when the AGCM is forced with persisting SST anomalies to the extent where the superiority of the CGCM becomes noticeable.
- The benefit of the AGCM over the CGCM due to the disproportional ensemble size in the former is found to be marginal and is incapable of shifting the superiority balance suggesting that the role of the multi-model SST forcing is crucial in leveraging a reasonable equivalency in the predictive skill of the GCMs.
- In conclusion, the comparative experiment reveals that the GCMs widely differ in their performances and the issue of superiority of one model over the other is mostly dependent on space and time (seasonality) and to the large extent they are complementing each other in one way or another.

The thesis elucidated a circumstance under which AGCMs and CGCMs may be able to produce similar levels of skill. Although only two such GCM configurations were considered in the thesis, at the very least the study has provided some guidance on how to best optimize an AGCM under circumstances of limited computational resources. Under such limitations AGCMs are often used as the only GCM for operational forecast production, a situation commonly found in developing countries such as South Africa. Despite that the CHPC computational support made it possible to conceptualize and execute this computationally intensive study, these resources are shared among a large range of users other than climate modellers, therefore placing an additional limitation on how global models to be tested can be configured and to what extent model development can happen in the region. An optimal AGCM configuration, however, depends heavily on skilful SST forecasts, here obtained through a multi-model SST forecast system. These predicted SSTs are reproduced from two CGCMs and such ocean-atmosphere interactively coupled models are therefore essential for skilful seasonal climate predictions when AGCMs are used. A significant

amount of work and investment have already gone into AGCM development, but the potential for further improvement of AGCM-based forecasts to a large extent thus depends on the improvement of CGCMs. The fact that many leading institutions are making their seasonal forecast, including SST forecasts, freely available to national and regional centres and such SST forecasts can subsequently be assimilated into AGCM operational forecast systems.

Scientific contribution of the study

The study have made a significant contribution to the state of knowledge of seasonal climate prediction which may have a far-reaching scientific and societal benefits. The key pointes that worth highlighting here may include:

- The study proposed an optimum procedure that may enhance the skill of seasonal forecast under constrained computational resources environment when the use of AGCMs remains a feasible option. The notion is successfully demonstrated here in such a way that given a suitable sea surface temperatures (SSTs) as forcing and is subject to an initialization strategy that uses realistic atmosphere and soil moisture states, AGCMs may be still relevant in the practice of contemporary seasonal climate prediction despite the proliferation of interest in the use of CGCMs globally over recent years. In addition, the study extends the notion of multi-model approach with the use of AGCMs and conclusively showing that a significant benefit of the AGCMs arises from multimodel SST forcing which is not extensively explored before.
- From a practical point of view, the study demonstrates that lack of coupling does not degrade much the predictive skill of the AGCM in favour of the CGCM which may shade an important insight to the contemporary debate of seasonal climate prediction practice.
- The study further reveals that most determinant factor in the success of seasonal prediction may be the robustness of the SST no matter whether interactively coupled or prescribed which may be an important addition to the state of knowledge.
- The study has also made a significant contribution in the development and improvement of GCMs which presumably has a far-reaching positive impact on local numerical and operational research efforts. In fact, the implementation of the forecasting systems evolved from this study for operational seasonal climate prediction has played a significant role on

the forecast skill enhancement of the SAWS' ensemble prediction system and its operational multi-model forecasting system. Furthermore, the SCM positively contributes toward sustainable development on national, regional and global domains in support of the Global Framework for Climate Services (GFCS) initiative of the WMO via the GPC platform.

Identified limitations and future perspective

The study identifies some weaknesses of the GCMs that may deserve further attention and improve the predictive skill of seasonal forecasts:

- Despite that the CGCM is found to be more skilful than the AGCM in predicting the equatorial IOD, this advantage in favour of the CGCM is not clearly manifested on the rainfall predictability of most of the regions considered in the analysis. This implies that the GCMs' atmospheric response to SST forcing is more or less dominated by ENSO. The contribution of the equatorial IOD on the predictability of rainfall particularly over the Indo-pacific region was reported in previous observational studies (e.g., Yang *et al.,* 2010). Improved physics (parameterization) or better coupling strategies may be required for better teleconnection representations or interactions among different climate components (modes).
- The AGCM appears to be weakly sensitive to soil moisture initialization and the coupling response is not clearly manifested as reported in similar modelling studies elsewhere (e.g., Seneviratne *et al.,* 2010; Koster *et al.,* 2004). This AGCM's internal weakness is presumably attributed to the framework used to couple the land surface and atmosphere or the land surface scheme itself is problematic and therefore needs additional attention.
- Although seasonal climate predictability over the extratropical atmosphere is relatively lower than over the tropics owing to strong hydrodynamical instabilities associated with baroclinicity that exits in the middle latitudes, there may be a room for further improvements in the ability of GCMs to capture those important modes of climate variability outside the tropics (e.g., the southern annular mode) which are found to be more relevant to the austral winter predictability particularly over the SH (Beraki *et al*., 2013). Using the CGCM hindcast simulations reported here, Mathole *et al*. (2014) also

indicated that the CGCM was unable to simulate the observed pole ward migration of the eddy driven southern extratropical jet stream and lower stratospheric cooling which may be attributed to the lack of proper stratospheric ozone prescription, anthropogenic forcings and the coarse vertical resolution of the model. The deficiency may be resolved with the use of Earth system models (ESMs) with better stratospheric and chemical species representation which indeed has become an emerging concept in seasonal forecasting (e.g., Smith *et al.*, 2014; Domeisen *et al.*, 2015). This new modelling challenge presents an opportunity for local model development endeavours to remain relevant and to keep pace with the global advancement of climate modelling science. South Africa, as a leading climate modelling country in Africa, should dedicate more resources toward the development of ESMs since these models are already starting to produce promising results elsewhere and has been identified as the next challenge by the local modelling community.

- Despite that the thesis has primarily been dedicated to atmospheric predictability and those climate drivers relevant to seasonal forecasting such as ENSO. More work is required to improve understanding of the processes and mechanisms behind using the tools developed in the thesis. Although ENSO is believed to be the most predictable mode of climate variability, it is not well understood why models sometime fail. For example the 2014 El Niño did not turn out as predicted. Furthermore, it is of particular interest to explore the behaviour of ENSO under the "global warming hiatus" and test the hypothesis that the oceans absorbed the heat since 2000 onward.

References

Balmaseda M and Anderson D. 2009. Impact of initialization strategies and observations on seasonal forecast skill. *Geophysical Research Letters* **36**: L01701. DOI: 10.1029/2008GL035561.

Barnston AG and Smith TM. 1996. Specification and prediction of global surface temperature and precipitation from global SST using CCA. *Journal of Climate* **9**: 2660-2697.

Barnston AG, Leetmaa A, Kousky V, Livezey R, O'Lenic E, Van den Dool H, Wagner AJ and Unger D. 1999. NCEP Forecasts of the El Niño of 1997-98 and Its U.S. Impacts. *Bulletin of the American Meteorological Society* **80**: 1829-1852.

Beraki AF and Landman WA. 2013. The role of the Southern Annular Mode in a dynamical global coupled model. *Proceedings 29th Annual conference of South African Society for Atmospheric Sciences*, 26-27 September 2013, Durban, South Africa, 84-87. ISBN 978-0-620-56626-1.

Boville BA, Hurrell JW. 1998. Comparison of the Atmospheric Circulations Simulated by the CCM3 and CSM1. *Journal of Climate,* **11**: 1327-1340.

Brinkop S and Roeckner E. 1995. Sensitivity of a general circulation model to parameterizations of cloud-turbulence interactions in the atmospheric boundary layer. *Tellus* **47**: 197-220.

Chaudhari HS, Pokhrel S. Mohanty S. Saha SK. 2013. Seasonal prediction of Indian summer monsoon in NCEP coupled and uncoupled model. *Theoretical and Applied. Climatology* **114**: 459–477. doi:10.1007/s00704-013-0854-8

Colfescu I, Schneider EK and Chen H. 2013. Consistency of 20th century sea level pressure trends as simulated by a coupled and uncoupled GCM, *Geophysical Research Letters,* **40**: 3276–3280, doi:10.1002/grl.50545.

Conil S, Douville H and Tyteca S. 2009. Contribution of realistic soil moisture initial conditions to boreal summer climate predictability. *Climate Dynamics* **32**: 75–93.

Derber J and Rosati A. 1989. A global oceanic data assimilation system. *Journal of Physical Oceanography* **19**: 1333–1347.

Dirmeyer PA, Fennessy MJ and Marx L. 2003. Low skill in dynamical prediction of boreal summer climate: Grounds for looking beyond sea surface temperature, *Journal of Climate* **16**: 995 – 1002, doi:10.1175/1520-0442(2003)016<0995:LSIDPO>2.0.CO;2.

Doblas-Reyes FJ, Hagedorn R and Palmer TN. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – II. Calibration and combination. *Tellus* **57A**: 234–252.

Doblas-Reyes FJ, Déqué M and Piedelièvre J P. 2000. Multi-model spread and probabilistic seasonal forecasts in PROVOST. *Quarterly Journal of the Royal Meteorological Society* **126**: 2069–2088.

Domeisen D, Butler A, Fröhlich K, Bittner M, Mueller WA and Baehr J. 2015. Seasonal predictability over Europe arising from El Niño and stratospheric variability in the MPI-ESM seasonal prediction system. *Journal of Climate* **28**: 256-271. doi:10.1175/JCLI-D-14-00207.1

Douville H, 2010. Relative contribution of soil moisture and snow mass to seasonal climate predictability: A pilot study. *Climate Dynamics* **34**,797–818.

Fu XH, Wang B and Li T. 2002. Impacts of air–sea coupling on the simulation of mean Asian summer monsoon in the ECHAM4 model. *Monthly Weather. Review* **130**: 2889–2904.

Goddard L, Hurrell JW, Kirtman BP, Murphy J, Stockdale T and Vera C. 2012. Two time scales for the price of one (almost). *Bulletin of the American Meteorological Society* **93:** 621–629. doi: http://dx.doi.org/10.1175/BAMS-D-11-00220.1

Goddard L, Mason SJ, Zebiak SE, Ropelewski CF, Basher R and Cane MA. 2001. Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology* **21**: 1111–1152.

Goddard L and Mason SJ. 2002. Sensitivity of seasonal climate forecasts to persisted SST anomalies. *Climate Dynamics* **19**: 619–631.

Graham RJ, Evans ADL, Milne KR, Harrison MSJ and Robertson KB. 2000. An assessment of seasonal predictability using atmospheric general circulation models, *Quarterly Journal of the Royal Meteorological Society* **126**: 2211-2240.

Graham RJ, Gordon M. McLean PJ. Ineson S. Huddleston MR. Davey MK. Brookshaw A and Barnes RTH. 2005. A performance comparison of coupled and uncoupled versions of the Met Office seasonal prediction general circulation model. *Tellus* **57**: 320–319.

Hagedorn, R, Doblas-Reyes FJ, and Palmer TN. 2005. The rationale behind the success of multi-model ensembles in seasonal forecasting – I. Basic concept. *Tellus* **57A**: 219–232.

Henderson-Seller A and McGuffeie K. 2001. Forty years of numerical climate modelling. *International Journal of Climatology* **21**:1067-1109.

Holton JR. 2004. An Introduction to Dynamic Meteorology. Elsevier Academic Press: San Diego, CA.

Jha B and Kumar A. 2009. Comparison of the atmospheric response to ENSO in coupled and uncoupled model simulations. *Journal of Climate* **137**: 479-487.

Kirtman BP, Shukla J, Huang B, Zhu Z and Schneider EK. 1997. Multiseasonal Predictions with a coupled tropical ocean-global atmosphere system. *Monthly Weather Review* **125**: 789-808.

Koster RD and Coauthors. 2004. Regions of strong coupling between soil moisture and precipitation. *Science* **305**: 1138-1140.

Krishnamurti TN, Kishtawal CM, Zang Z, LaRow T, Bachiochi D, Williford E, Gadgil S and Surendran S. 2000. Multimodel ensemble forecasts for weather and seasonal climate. *Journal of Climate* **13**: 4196–4216.

Kug J.-S, Kang I-S and Choi D-H. 2008. Seasonal climate predictability with Tier-one and Tier-two prediction systems. *Climate Dynamics* **31**: 403–416. doi:10.1007/s00382-007-0264-7.

Landman WA. 2014. How the International Research Institute for Climate and Society has contributed towards seasonal climate forecast modelling and operations in South Africa. *Earth Perspectives*, **1**, 1-13.

Landman WA and Beraki A. 2012. Multi-model forecast skill for mid-summer rainfall over southern Africa. *International Journal of Climatology* **32**: 303–314. DOI: 10.1002/joc.2273.

Landman WA, DeWitt D, Lee D-E, Beraki A and Lötter D. 2012. Seasonal rainfall prediction skill over South Africa: 1- vs. 2-tiered forecasting systems. *Weather and Forecasting* **27**: 489-501. DOI: 10.1175/WAF-D-11-00078.1.

Landman W and Mason SJ. 1999. Operational long-lead prediction of South African rainfall using canonical correlation analysis. *International Journal of Climatology* **19**: 1073-1090.

Lazenby MJ, Landman WA, Garland RM and DeWitt DG. 2014. Seasonal temperature prediction skill over southern Africa and human health. *Meteorological Applications doi:10.1002/met.1449*Mason SJ, Goddard L, Graham NE, Yulaeva E, Sun L and Arkin PA. 1999. The IRI seasonal climate prediction system and the 1997/98 El Ni˜no event. *Bulletin of the American Meteorological Society* **80**:1853–1873.

Mathole K, Ndarana T, Beraki AF and Landman WA. 2014. Assessing the importance of lower stratospheric processes on the predictability of summer rainfall over South Africa. *South African Journal of Science* **110**(3/4), Art. #2013-0161, 8 pp. http://dx.doi.org/10.1590/sajs.2014/20130161

Molteni F, Ferranti L, Balmaseda M, Stockdale T and Vitart F. 2007. ECMWF seasonal forecast system 3. *CLIVAR Exchange* **43**: 7-9.

Moore AM, and Anderson DLT. 1989. The assimilation of XBT data into a layer model of the tropical Pacific Ocean, *Dynamics of Atmospheres and Oceans* **13:** 441-464.

Palmer TN and Anderson DLT. 1994. The prospects for seasonal forecasting—A review paper. *Quarterly Journal of the Royal Meteorological Society* **120**: 755–793.

Palmer TN, Alessandri A, Anderson U, Cantelaube P, Davey M, Délécluse P, Déqué M, Díez E, Doblas-Reyes FJ, Feddersen H, Graham R, Gualdi S, Guérémy J-F, Hagedorn R, Hoshen M, Keenlyside N, Latif M, Lazar A, Maisonnave E, Marletto V, Morse AP, Orfila B, Rogel P, Terres J-M and Thomson MC. 2004. Development of a European ensemble system for seasonal to inter-annual prediction (DEMETER). *Bulletin of the American Meteorological Society* **85**: 853–872.

Saha S, Nadiga S, Thiaw C, Wang J, Wang W, Zhang Q, Van den Dool HM, Pan H-L, Moorthi S, Behringer D, Stokes D, Peña M, Lord S, White G, Ebisuzaki W, Peng P and Xie P. 2006. The NCEP Climate Forecast System. *Journal of Climate* **19**: 3483-3517.

Seneviratne SI, Corti T, Davin EL, Hirschi M, Jaeger EB, Lehner L, Orlowsky B and Teuling AJ. 2010. Investigating soil moisture-climate interactions in a changing climate: A review. *Earth-Science Reviews* **99**:125–161.

Shukla J. 1981. Dynamical predictability of monthly means. *Journal of Atmospheric Sciences* **38**: 2547–2572.

Shukla J. 1983. Comments on "Natural variability and predictability" *Monthly Weather Review* **111**: 581-585.

Smith MJ, Palmer PI, Purves DW, Vanderwel MC, Lyutsarev V, Calderhead B, Joppa LN, Bishop CM and Emmott S. 2014. Changing How Earth System Modeling is Done to Provide More Useful Information for Decision Making, Science, and Society. *Bulletin of the American Meteorological Society* **95**, 1453–1464.

Staniforth A and Wood N. 2008. Aspects of the dynamical core of a nonhydrostatic deep-atmosphere, unified weather and climate-prediction model. *Journal of Computational Physics* **227**, 3445–3464.

Stockdale TN, Anderson DLT, Alves JOS and Balmaseda MA. 1998. Global seasonal rainfall forecasts using a coupled ocean–atmosphere model. *Nature* **392**: 370-373.

Tennant WJ. 2003. An assessment of intraseasonal variability from 13-yr GCM simulations. *Monthly Weather Review* **131**: 1975–1991.

Troccoli A, Harrison M, Anderson DLT and Mason SJ. 2008. *Seasonal Climate: Forecasting and managing risk.* NATO Science Series. Earth and Environmental Sciences Vol 82. Springer: Dordrecht, The Netherlands.

Walker J and Rowntree PR. 1977. The effect of soil moisture on circulation and rainfall in a tropical model. *Quarterly Journal of the Royal Meteorological Society* **103**: 29–46. DOI: 10.1002/qj.49710343503.

Yang J, Liu Q and Liu Z. 2010. Linking Observations of the Asian Monsoon to the Indian Ocean SST: Possible Roles of Indian Ocean Basin Mode and Dipole Mode. *Journal of Climate* **23**: 5889–5902. doi: http://dx.doi.org/10.1175/2010JCLI2962.1

Yu J-Y and Mechoso CR. 1999. A Discussion on the Errors in the Surface Heat Fluxes Simulated by a Coupled GCM. *Journal of Climate* **12**: 416–426.

Zebiak SE and Cane M. 1987. El Nino-Southern Oscilation. *Monthly Weather Review* **115**: 2262-2278.