Dynamical downscaling of ECMWF Ensemble seasonal forecasts over East Africa with RegCM3

G. T. Diro, A. M. Tompkins, and X. Bi

Received 10 October 2011; revised 23 May 2012; accepted 2 July 2012; published 18 August 2012.

[1] Dynamical downscaling of ECMWF ERA-interim reanalysis and an ensemble of May-start ECMWF seasonal hindcasts is performed with ICTP’s regional climate model (RegCM3) over the horn of Africa for a ten year period. Using ERA-interim ‘perfect boundary’ conditions the regional model reproduced both the spatial and interannual variability of the region’s rainfall and improved the global model’s reproduction of year to year rainfall variability, capturing the teleconnection between ENSO and the region’s precipitation pattern well. The ensembles of ECMWF seasonal hindcasts and the respective downscaled RegCM3 hindcast suite were then validated in terms of the seasonal climate and deterministic and probabilistic skill scores at a one to four month lead time. Both RegCM3 and ECMWF hindcasts reproduce the spatial and temporal rainfall variability well, but overestimate the mean and variability over the Arabian peninsula and misrepresent the teleconnection between ENSO and precipitation over the western Indian Ocean. The positive bias over the Arabian peninsula in RegCM3 and the teleconnection error between ENSO and precipitation anomalies over the western Indian Ocean are due to the propagation of errors from the driving GCMs to the regional model. Nevertheless, the probabilistic assessment (ROCS and RPSS) indicated that both ECMWF and RegCM3 have significant skill suggesting the potential utility of dynamical forecasts over the region. Comparing the skill of ECMWF and RegCM3 probabilistic hindcasts, ECMWF generally performs better on grid point by grid point comparison and at homogeneous zones over Ethiopia but RegCM3 outperforms when aggregated on a country scale and when compared against high resolution rain gauge data set.


1. Introduction

[2] There is a growing demand for forecasts on local scales relevant for end-users, both of direct weather anomalies or socio-economic targeted information with strong climate drivers. Examples of the latter include crop yield [Challinor et al., 2005], health impacts [Morse et al., 2005; Thomson et al., 2006] and water resources [Nakaegawa et al., 2007] predictions and the lead times required for such applications can range from days to seasons and beyond. In order to gain climate information at the local scale which is usually a finer resolution than current generation General Circulation Models (GCMs) can provide, the GCM output must be downscaled. This can be achieved using statistical and/or dynamical downscaling techniques and both approaches have advantages and drawbacks.

[3] Statistical downscaling techniques are commonly implemented with success and are computationally efficient, but their use of historical relationships between the predictand and predictors implies that they rely on the availability of long data records of high spatial resolution with care required to avoid over fitting in their application [Wilby et al., 1998; Zorita and von Storch, 1999; Landman and Tennant, 2000]. Dynamical downscaling involves nesting a finer resolution regional model within the GCM, with the GCM providing the boundary values of the prognostic variables for the smaller model domain. Dynamical downscaling avoids the requirement of high spatial resolution of observations, since the downscaling occurs through improved resolution of surface topography and atmospheric dynamics, but has the drawback of a high computational cost. Moreover, caution is also required in the application of dynamical downscaling approaches, since poor regional model physics can compound existing errors in the driving GCM [Giorgi and Mearns, 1999]. This work is mainly concerned with the assessment of the dynamical downscaling approach.

[4] Regional downscaling has been widely employed to provide day to day short-range operational forecasts on a local scale. This application is often deterministic, and

---

1 Earth System Physics Section, Abdus Salam ICTP, Trieste, Italy.
2 Climate Change Research Center, Institute of Atmospheric Physics, Beijing, China.

Corresponding author: G. T. Diro, Earth System Physics Section, Abdus Salam ICTP, Strada Costeira 11, IT-34100 Trieste, Italy. (gtefera@ictp.it)

©2012. American Geophysical Union. All Rights Reserved.
assumes that the host GCM can adequately provide predictions of the large-scale flow over the coming days. The task of the regional model is to map the large-scale weather patterns locally in terms of variables such as temperature and precipitation, with Marsigli et al. [2001] providing evidence of the success of this approach.

[5] Dynamical downscaling has also been applied to climate model integrations, whereby the regional model is validated and tuned using an integration of the present climate for a given region before being applied to future climate integrations to give local scale information [Giorgi and Mearns, 1999]. The present-day model configuration integration is frequently performed using so-called “perfect boundary conditions” derived from reanalysis data sets such as ERA-40 [Uppala et al., 2005]. The tuned RCM is then applied to present and future climate integrations using GCM climate model boundary conditions. This approach has led to the climate application of RCMs is considered more controversial, since the quality of the GCM boundary forcing for a given region depends on its ability to represent the global teleconnections well. For example, if a GCM produces subsidence over a tropical region during a period of the year that instead should be convective and subject to low level convergence, the RCM is prevented from simulating the correct mean level of convection and it is highly unlikely that this RCM will be able to produce a correct climate sensitivity to future CO₂ increases using the lateral boundary conditions (LBC) from this GCM in this specific region in question. The degree of constraint obviously depends on the size of the domain and a number of studies have already shown how sensitivity of RCMs to LBCs is highly dependent on the choice of domain size [Seth and Giorgi, 1998; Jones et al., 1995; Giorgi and Mearns, 1999; Rauscher et al., 2006]. With very large domains the LBCs are minimal in importance and the RCM physics is able to govern the local response, but in this case the large domain implies very little gain in spatial resolution can be achieved by the RCM.

[6] Appreciation of the fact that configuration and tuning a regional model’s physics to improve regional mean climate with perfect boundary conditions does not imply that the model will provide a reliable climate sensitivity for a given region has led to more rigorous efforts to validate regional models’ ability to improve GCM statistics of intra seasonal and interannual variability [e.g., Giorgi and Bi, 2000; Christensen et al., 2007]. These efforts have been fortified by the recent interest in the development of seamless systems to provide predictions across a wide-range of temporal scales [Hurrell et al., 2009; Hazeleger et al., 2010]. Palmer et al. [2008], for example, suggest that climate model ensembles can be calibrated using the model performance in predicting seasonal anomalies. By extension, one method of assessing a regional climate model’s reliability over a specific region is to examine the ability of the model to improve the prediction of seasonal anomalies on monthly to seasonal timescales [Palmer and Anderson, 1994].

[7] Previously, Díez et al. [2005] have downscaled DEMETER seasonal forecasts and show an improvement in skill using a regional model over Northwest Europe, although the dynamical method does not outperform a statistical downscaling technique. Díez et al. [2011] conducted a downscaling experiment using ECMWF seasonal forecasting system 3 products over Europe using 5 ensemble members and found an encouraging skill in Autumn season and for the dry events. In contrast, Roads et al. [2003] claim little improvement in seasonal prediction skill over the USA despite a positive impact of higher resolution on the mean climate represented in the regional model. Lim et al. [2007] have found that the downscaled seasonal anomalies of surface air temperature matches observations better than the driving GCM simulations.

[8] Poor GCM seasonal performance over mid latitudes is likely to limit the seasonal prediction skill of the regional models. In the tropics, skillful GCM lead times are longer and the direct nature of tropical circulations implies that the ocean surface temperature contained within the regional model domain may have more potential impact on regional skill. Misra et al. [2003] and Sun et al. [2006] indicated some improvement using regional models over Brazil. In contrast, Nobre et al. [2001] indicated that, although downscaling a GCM to 80 km improved the bias and RMS error, a further nesting to refine the resolution to 20 km yielded a deteriorated skill, suggesting a need to re-tune the convection scheme and surface parameters. For Asia, Ding et al. [2006] found an improved skill using a regional climate model in a seasonal hindcast mode relative to the driving GCM, especially for the Yangtze river region.

[9] Despite the above studies, relatively little research has been conducted on the prospect of dynamical downscaling of seasonal forecasts in the tropics, and in particular concerning the African continent, where reliable seasonal predictions and climate projections have a high potential societal impact. One region that has excellent potential as a laboratory to evaluate regional dynamical seasonal forecast is the horn of Africa. The seasonal precipitation anomaly is influenced by local sea surface temperatures in the Indian Ocean, and remotely in the Pacific [Shanko and Camberlin, 1998; Segele et al., 2009a; Díez et al., 2011b]. The spatial distribution of the convective precipitation is strongly influenced by the complex topography of the Ethiopian highlands, providing the potential for regional models to excel. This study thus examines the question of whether a regional climate model can improve monthly to seasonal predictions of precipitation over Ethiopia using state-of-the-art operational seasonal forecast to provide the boundary conditions and compares this to the conventional test of simulating the regions mean precipitation climate using perfect boundary conditions provided from reanalysis. After first assessing the sensitivity of the model simulation to domain size and model configuration, the skill of RegCM3 will be presented and discussed for both perfect boundary and seasonal forecast boundary conditions. Finally, conclusions concerning the implications for the use of the regional climate model are made.

2. Data, Method and Model Setup and Sensitivity Studies

2.1. Data

[10] Various data sets were used to assess the performance of the seasonal hindcast. These are gauge data from the National Met Agency (NMA) of Ethiopia, Global Precipitation Climatology Project (GPCP) version 2.1 [Adler et al., 2003], Tropical Rainfall Measuring Mission (TRMM - 3B43) [Huffman et al., 2007] and, where necessary, ERA-interim reanalysis [Simmons et al., 2007; Dee et al., 2011] is
also used. Initially 280 rain gauge data were obtained from NMA, but after a quality control rejecting stations with a percentage of missing data greater than 25%, only 160 stations were selected for the analysis (Figure 3). Since the rain gauge stations are not uniformly distributed, the station data is converted into a gridded data set by grouping the station data into 1° latitude – 1° longitude grid boxes similar to the method used by Krishnamurthy and Shukla [2000]. GPCP is a relatively coarse resolution (2.5° × 2.5°) monthly rainfall analysis produced by merging gauge measurements and satellite estimates of rainfall [Adler et al., 2003]. TRMM 3B43 is a (0.25° × 0.25°) degrees resolution near-global data set produced by optimally merging multiple satellite estimates with monthly accumulated rain gauge analyzed data. As TRMM data is available from 1998 and thus does not significantly overlap the ENSEMBLES analysis period, it is used only for the purpose of showing a high resolution precipitation climatology for the region.

2.2. Downscaling Methods

2.2.1. Dynamical Downscaling

2.2.1.1. Regional Model and Lateral Boundary Conditions

[11] The regional model is the ICTP third generation Regional climate model RegCM3 [Pal et al., 2007]. RegCM3 is a hydrostatic sigma coordinate 3D model with the dynamical core similar to MM4 [Grell et al., 1994]. It uses the radiation scheme of CCM3 [Kiehl et al., 1996]. The land surface model is the Biosphere-Atmosphere Transfer Scheme (BATS) [Dickinson et al., 1993], while the boundary layer scheme follows Holtslag et al. [1990]. There are three choices for the convection parameterization, namely the Grell [1993], Emanuel [1991] and modified Kuo [Anthes, 1977] schemes. The two convection schemes employed in this study are the Emanuel scheme and the Grell scheme with the Fritsch and Chappell [1980] closure assumption. Large scale precipitation is parametrized by the sub grid explicit moisture scale (SUBEX) [Pal et al., 2000]. The ocean flux parametrization is described by Zeng et al. [1998].

[12] Several studies [e.g., Sylva et al., 2010; Segele et al., 2009b; Anyaha and Semazzi, 2007; Davis et al., 2009] have shown that RegCM3 has an added value compared to the driving reanalysis over various parts of Africa. Most of these studies have run RegCM3 at a resolution of 50 to 60 km. Here in this study we used a higher resolution simulation over the horn of Africa that is more suitable for impact models, using a resolution of 30 km to better resolve surface processes related to topography. Figure 1 shows the difference in the topography representation between ECMWF and RegCM2. The topography from the GCM not only underestimates the maximum height over the Ethiopian highlands but also fails to represent the location of the peak and the detailed spatial variation in height. For example the rift valley that lies between the Semien mountains in the north and Bale mountains in southern Ethiopia is not represented in the GCM. Associated with these complex topography features, it is expected that orographic forced rainfall patterns will be simulated more accurately with regional models. In the vertical, 18 σ levels are used.

[13] The lateral boundary conditions for the regional model are provided by the European Centre for Medium-range Weather Forecast (ECMWF) ERA-interim reanalysis for the perfect boundary setup) and the ECMWF atmosphere-ocean coupled model seasonal forecasting system. The seasonal forecast integrations used in this study formed ECMWF’s contribution to the down-scaling component of the multimodel ENSEMBLES project [van der Linden and Mitchell, 2009] and employed atmospheric model version cycle 29r2 with a resolution of T95L40 which corresponds to approximately 1.125 degree in the horizontal and 40 levels in the vertical. The ocean model (HOPE) has a horizontal resolution of 1 degree (0.3 degree near the equator) in the horizontal and 29 levels in the vertical. The seasonal forecasting system is very close in setup to the third generation operational system (SYSTEM 3) of which a detailed description can be found in Anderson et al. [2007], Doblas-Reyes et al. [2009] and Tompkins and Feudale [2010]. The hindcast integration ensembles consists of 9 members which are integrated for 11 years starting on May 1st and November 1st for the years of 1991 to 2001 with output stored at 6 hourly intervals. (One of the impediments to downscaling operational seasonal forecasts in general is the
archiving frequency of model or pressure level data required for the boundary conditions, which is generally lower than the minimum of 6 hours required to minimize the diurnal cycle on smaller domains. The ENSEMBLES data set represents the first significant hindcast archive with adequate temporal resolution for the downscaling task.) For the region of interest, the 6 hourly forecasts of temperature, humidity, surface pressure, geopotential heights and winds, starting on 1st May are used as the boundary conditions for the regional model.

Uncertainty in the hindcast integration can be due to initial condition uncertainty and model error. The former is accounted for by initial condition perturbations to the sea surface temperatures [Anderson et al., 2007] while the latter is addressed using a simple stochastic physics scheme [Buizza et al., 1999] in the global model. Each member of the regional model hindcast integration uses the relevant perturbed initial conditions from the driving model integration, but then downscales each member deterministically, making no attempt to address its own model uncertainty by the use of a stochastic physics scheme.

For the perfect boundary experiment, the lateral boundary conditions are derived from ERA-interim reanalysis. ERA-interim reanalysis is the latest reanalysis from ECMWF [Simmons et al., 2007; Uppala et al., 2008; Dee et al., 2011]. It uses the latest 4D var data assimilation technique and improved model physics compared to ERA-40. Although the ERA-interim reanalysis are available at every six hours interval, these are interpolated linearly in between the six hours to provide the boundary condition to the regional model. Exponential nudging was used to weight the values of boundary conditions and the model results at the buffer zones [Giorgi et al., 1993].

2.2.2. Experimental Design

For the ENSEMBLE seasonal hindcast downscaling experiment, the nine ECMWF members were downscaled by nine RegCM3 runs, all ensemble members driven separately by the corresponding nine ECMWF boundary conditions. These nine members were all run for the period May 1 to October 1 for the period between 1991 and 2000.

Regarding the land surface initialization in the RegCM3, all the RegCM3 runs are initialized with the same land surface conditions. BATS is used as the land surface model in RegCM3 and the soil moisture in BATS is initialized by specifying the fraction of soil water content as a function of land use and cover type at the first time step [Giorgi and Bates, 1989]. Therefore every year, the first integration month of May is excluded during spin up and the analysis are made for June to September; a lead time of 1–4 months.

For the perfect boundary run RegCM3 was run by ERA-interim reanalysis starting from 1 January 1991 until 1 January 2001, although the analysis is focused on the summer (JJAS) season to compare with the ENSEMBLE seasonal hindcast run. Here nesting is one way as the regional model does not impact the host global model integration. We note that the perfect boundary integration differs from the hindcast suite in one aspect since the land surface conditions are not reinitialized each May.

The surface boundary condition for the perfect boundary run is from National Ocean Atmospheric Administration’s (NOAA) Optimal Interpolated (OI)-weekly sea surface temperature (SST) [Reynolds et al., 2002], whereas for the seasonal hindcast run, the surface boundary condition of SST for the nine RegCM3 runs were obtained from the corresponding ECMWF global ensemble hindcast runs.

2.2.2. Statistical Downscaling

In addition to climatology and random forecasts, the dynamical prediction skill is also measured relative to the benchmark of a simple statistical downscaling method. There are various ways of constructing a statistical downscaling framework using a wide range of potential predictors such as model precipitation, SST, low level wind from the global model, employing methods ranging from simple regression models [e.g., Feddersen and Andersen, 2005] and analogue methods [e.g., Diez et al., 2005] to self organizing maps [e.g., Gutierrez et al., 2005]. In this paper, a very simple approach is employed using the ECMWF ensemble total precipitation over the region as a predictor in a linear regression model for the homogeneous rainfall zones over Ethiopia.

As outlined earlier, one of the main challenges in the statistical downscaling methods is access to long data records. Here in this study for instance, the 6-hourly ECMWF ensemble hindcast data set used for the dynamical downscaling is available for 10 years; too short for training and cross-validation. Instead, to train the statistical model, we are able to utilize the hindcast data set of the ECMWF system version 3 (SYS3) 11 member ensemble using the period 1981 to 1990. The period 1991–2000 is then used as an independent validation period to assess against the dynamical forecasts.

2.3. Sensitivity Experiments

Prior to downscaling multi year simulations, six simulations from May to October of 1991 were carried out to study the sensitivity of RegCM3 simulation to convection schemes and domain size. Previous RegCM3 studies over Africa have used various configurations and to produce good simulations for their respective study period and one of the major sensitivities to physics found was the convection scheme and the parameter settings for the land surface scheme. For example, Segele et al. [2009b] used MIT’s Emanuel scheme with tuned values for mass flux adjustment (α) and the fraction of condensed water that can be converted to precipitation (l), parameters and improved the simulation over the horn of Africa, whereas Sylla et al. [2010] and Mariotti et al. [2011] used Grell scheme for the convection scheme with the modified surface albedo and stomata resistance in BATS to get a good agreement with observations in many parts of Africa.

To select the model optimal configuration for this study, four simulations were carried out by using ERA-interim and ECMWF seasonal hindcasts as boundary conditions, first with modified parameters in the Emanuel scheme as in Segele et al. [2009b], and then with the modified surface albedo and stomata resistance in BATS with the Grell convection scheme following the setup of Sylla et al. [2010] and Mariotti et al. [2011]. The results (not shown) indicated that both model configurations gave rise to a similar rainfall distributions although Emanuel scheme as used in Segele et al. [2009b] gave generally less rainfall over the ITCZ region including the Ethiopian highland when forced by ERA-interim reanalysis, while the spatial distribution of rainfall simulated by the Grell scheme with modified surface parameter as in Sylla et al. [2010] and Mariotti et al. [2011] is
better captured when forced by the ECMWF ensemble hindcasts. Therefore the model configuration chosen for further study is the Grell scheme with FC closure assumption, with modified albedo and minimum stomata resistance in the BATS scheme.

[24] A further test was carried out on the sensitivity of the rainfall simulation to domain size. Two domains were set up, one covering the horn of Africa and the second smaller domain covering only Ethiopia (rectangular box in Figure 1, left) to quantify the impact of domain size in simulating the spatial pattern of summer rainfall. Figure 2 shows the precipitation simulations greatly differ for the smaller domain (compared to the larger domain) to the extent that the location of precipitation maxima is misplaced when the boundary forcing is changed from reanalysis to the GCM seasonal forecast output. Even though the small domain size reproduced the correct precipitation pattern when forced with best estimates of the state of the atmosphere from reanalysis, the outcomes are erroneous when forced with ECMWF seasonal hindcasts, since the smaller domain is more constrained to reproduce errors in the driving GCM [Seth and Giorgi, 1998; Giorgi and Mearns, 1999]. Accordingly, as a result of these sensitivity tests, the larger domain (23E to 57E, 5S to 23N) covering the horn of Africa is chosen for the subsequent simulations. As ECMWF resolution is around 125 km, this represents a ratio of 4 to 1.

3. Skill of RegCM3 Under Perfect Boundary Conditions

[25] The first test is to integrate RegCM3 using perfect boundary conditions in order to differentiate errors deriving from the driving GCM from those of the RegCM3 physics. Validation is done in terms of comparing the climatology, extreme wet and dry years and the interannual variability of the northern hemisphere summer precipitation from both
RegCM3 and the driving ECMWF ERA-interim reanalysis to other sources of precipitation data.

3.1. Climatology

[26] The 10 year JJAS climatology of RegCM3 precipitation, together with that of ERA-interim, GPCP, gauge and TRMM (Figure 3) shows that RegCM3 depicts a smaller bias compared to the short range forecasts starting from the ERA-interim reanalysis. It should be noted that the water cycle is not conserved for the amalgamated short-range ERA-interim forecasts, which are constantly reset to the reanalysis initial conditions each day. If the model climate is not the same as that of the analysis, the short range forecasts can suffer from ‘spin-up’ [e.g., Bengtsson et al., 2004]. ERA-interim in general suffers from much lower spin up in tropics than its predecessor ERA-40, though regionally it is still an issue [Dee et al., 2011]. The long-range regional model integrations could also be argued not to be in water balance, since the water vapor nudging terms at the lateral boundaries are also an artificial sources. Thus a regional model that has a climate very different from the driving global model/analysis will suffer from a similar spin up problem, in this case notable as a precipitation gradient near the model boundaries. For this reason regional model analysis disregards the adjustment zone near the lateral boundaries.

[27] The RegCM3 integrations reveal finer scale rainfall structure over the complex topography due to the higher horizontal resolution employed which can better resolve the Ethiopian highlands. Although the climatology of TRMM is calculated for a later period, it nevertheless shows similar spatial variability of rainfall associated with topography to that of RegCM3. For instance the low rainfall amounts

Figure 3. Climatology of JJAS rainfall (mm) from (top left) ERA-interim, (top right) RegCM3 (forced with ERA-interim), (middle left) TRMM, (middle right) Gauge and (bottom left) GPCP. (bottom right) The percentage error in the GPCP rainfall estimate.
within the rift valley and the secondary maxima on the eastern side of the rift valley follow the topography. These fine scale rainfall structures observed in TRMM are well reproduced by RegCM3 although there is a difference in the amount of rainfall with RegCM3 underestimating over the eastern low lands and overestimating over the highlands. RegCM3 seems to be over sensitive to the topography as it shows more detailed structures over the mountain regions than TRMM even though both RegCM3 and TRMM have more or less the same resolution. This might be due to misrepresented topographic forcing, but it also highlights the diagnostic nature of convection in the majority of atmospheric models, where diagnostic parametrization schemes prevent advection of evolving convective systems from one grid cell to another [Leung and Ghan, 1995].

[28] The observational data of GPCP on the other hand are available only at low resolution of 2.5 degrees and thus completely lacks the fine scale spatial variability across these regions. The common feature between RegCM3 and ERA-interim is that they both underestimate the rainfall over the low lands of the eastern and southern Ethiopia and overestimate over the western Indian Ocean and over the Ethiopian highlands compared to observations.

3.2. Interannual Variability

[29] To compare the temporal variability with a more regional focus, area average precipitation is taken over the five homogeneous zones used by Diro et al. [2011b] (Figure 3, middle left) and also over the entire Ethiopian domain, again using ERA-interim 1 day forecasts as a benchmark and comparing to gauge observations. All products are linearly interpolated to the gridded gauge resolution (1 x 1 degree) and only grid boxes containing at least one gauge are included in the regional average time series. Generally both ECMWF and RegCM3 reproduce the observed interannual variability of rainfall for the northwestern, central and eastern part of the country (Figure 4). RegCM3, however, captures the interannual variability better over the Northwestern and for the eastern part of the country but performs less well over the south west. ERA-interim generally underestimates the early 1990s wet anomalies compared to the late 1990s and hence misses the 1993 wet years for zone I, and zone IIa whereas RegCM3 captures these wet distinctions. When the averaging area is at the country scale the correlation between RegCM3 rainfall and observations (0.88 with GPCP and 0.83 with gauge) is greater than that of ERA-interim (Table 1). These results suggest that, when the averaging area is larger (at a country scale), RegCM3 has an added value in capturing the interannual variability of rainfall over Ethiopia better than the global short range forecasts of the driving re-analysis. This is a significant result since it implies that RegCM3 has not simply been tuned to reduce the long-term mean bias of the global model over this specific region, but rather, the configuration of the model physics combined with the improved resolution of the topography lead to a improved translation of the large-scale dynamical and thermodynamic fields into year-to-year anomalies in rainfall.

3.3. Extreme Wet (1998) and Dry (1997) Years

[30] The comparison of interannual variability of areal averaged rainfall over Ethiopia suggested that RegCM3 is able to reproduce the dry and wet years correctly including those extremes related to ENSO. Figure 4 shows that the two extreme dry and wet years from observational data (GPCP and gauge) are 1997 (dry) and 1998 (wet). These two years correspond to contrasting phases of ENSO.

[31] Observations (from GPCP) suggest that the rainfall pattern over the western Indian Ocean varies in opposition to the rainfall over the ITCZ regions over the Ethiopian highlands. For instance in 1997, the rainfall is more intense and widespread over the western Indian Ocean compared to 1998, whereas the opposite is true for the ITCZ regions over the continent. This opposite polarity in rainfall variability between the equatorial western Indian Ocean and the continent is clearly seen in Figure 5. RegCM3 reproduces well these characteristics of dipole pattern of rainfall anomalies between the equatorial western Indian Ocean and African continent associated to the opposite ENSO events.

[32] From the above results, it is fair to say that RegCM3 accurately reproduces both the spatial pattern of rainfall over the horn of Africa and also the interannual variability of Ethiopian rainfall better than the ERA-interim reanalysis. This implies that downscaling of the global ERA-interim reanalysis with RegCM3 has indeed an added value as far as rainfall is concerned. The following section will demonstrate that this is not always the case when RegCM3 is applied to downscale seasonal hindcasts of ECMWF.

4. Deterministic Verification of the Seasonal Forecast

[33] The ensemble mean and ensemble spread of JJAS total precipitation for the ten year hindcast climatology for the ECMWF and RegCM3 models indicate that the mean rainfall amount and in particular the inter-ensemble spread are larger for the regional model ensemble, especially over the Arabian peninsula, even though RegCM3 employs no additional stochastic forcing (Figure 6). Comparison of the ECMWF hindcast climatology with the observed JJAS total precipitation climatology (Figure 3, bottom) reveals that, although the ECMWF ensemble hindcast climatology slightly overestimates the summer precipitation amount over the Ethiopia, its wet bias is less than the ERA-interim reanalysis. Over the western Indian Ocean, however, the ECMWF seasonal hindcast overestimates the rainfall compared to observations and the ERA-interim reanalysis.

[34] The spatial pattern of the seasonal hindcast from RegCM3 ensemble mean on the other hand reveals fine scale structure over Ethiopia associated with the topography. For instance the maximum precipitation over the Bale mountain ridges in southern Ethiopia and over the western side of the Semien mountains is captured well whereas the global model and low resolution observational data set represent this pattern with a single precipitation maximum over the Ethiopian highlands. The hindcast seasonal total precipitation climatology of RegCM3 shows excessive rainfall over the south-west Indian Ocean and over the Arabian peninsula compared to observations. These precipitation biases did not exist or were smaller when RegCM3 was forced with a perfect boundary conditions (Figure 3, top right). This implies that the bias over the Arabian Peninsula and the western Indian Ocean in RegCM3 is due to the propagation of errors from the GCM to RegCM3. In fact, looking at the low level wind anomaly (Figure 7) it can be seen that although the East
African Low Level jet (EALLJ), and the westerly influx from Atlantic is well reproduced by the ECMWF ensemble hindcasts, the anomalous flow over the Arabian Peninsula is not. For instance during a negative ENSO event, the Arabian peninsula is dominated by the southerly anomaly (i.e., a weakening of the northerly dry flow) while the ECMWF hindcast ensemble however shows the opposite phenomenon with enhanced northerly flow and a shift in the low level convergence to over the ocean which favors convection and hence produces excess precipitation.

As noted, all the ensemble members of ECMWF exhibit similar spatial patterns of rainfall with maximum precipitation over the Ethiopian highlands and the western Indian Ocean (not shown). The precipitation over the western Indian Ocean

Table 1. Correlation of ERA-Interim and RegCM-ERA-Interim With GPCP and Gauge

<table>
<thead>
<tr>
<th></th>
<th>Zone I</th>
<th>Zone IIa</th>
<th>Zone IIIb</th>
<th>Zone I</th>
<th>Zone IIIb</th>
<th>Zone IV</th>
<th>All Zones</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERA-interim</td>
<td>0.61</td>
<td>0.80</td>
<td>-0.02</td>
<td>0.58</td>
<td>0.68</td>
<td>0.60</td>
<td>0.62</td>
</tr>
<tr>
<td>RegCM</td>
<td>0.69</td>
<td>0.51</td>
<td>0.46</td>
<td>0.72</td>
<td>0.18</td>
<td>0.45</td>
<td>0.54</td>
</tr>
</tbody>
</table>

Boldface and italics represent values that are significant at 0.05 and 0.1 levels, respectively.
compared to observations is consistently overestimated indicating that ECMWF seasonal hindcasts have a systematic wet bias over the western Indian Ocean. All ensemble members of RegCM3 also show high rainfall over the western Indian Ocean compared to observations. It should be recalled that no gauge data can supplement satellite information over the oceans, although retrieval algorithms using microwave are facilitated over a water surface [Wilheit et al., 1991; Adler et al., 2003].

The precipitation pattern of all ensemble members of RegCM3 over the African continent show a realistic fine scale structure associated mainly with the topography. This result is encouraging because JJAS is the main rainy season over the northern part of the eastern Africa continent (north of 5N) and RegCM3 has an added value in reproducing the rainfall patterns over this region accurately compared to the driving model.

Figure 8 shows the ensemble mean JJAS total rainfall pattern of the two hindcasts for the wet minus dry (1998–1997) years. The precipitation over the Arabian Peninsula and western Indian Ocean is enhanced/suppressed in both ECMWF and RegCM3 hindcasts during the La Niña/El Niño years. Over the ITCZ region of the African continent, usually north of 5N, the general conclusion is that both hindcasts show enhanced/suppressed precipitation during the La Niña/El Niño events in agreement with the observations. This anomalous positive/negative precipitation pattern over the Ethiopian highland during negative/positive phase of ENSO years is also noted in observational studies by Korecha and Barnston [2007], Diro et al. [2011b] and Segele et al. [2009a].

A comparison of Figures 5 and 8 reveals an interesting feature of the impact of boundary forcing on the solution of the simulation. In the perfect boundary forcing simulation, RegCM3 has a precipitation deficit over the western Indian Ocean during the 1998 La Niña year and excess precipitation during the 1997 El Niño year. The ECMWF seasonal hindcast forced RegCM3 simulation ensemble, however, shows the opposite signature such that the Western Indian Ocean receives excess rains during the 1998 La Niña year and deficit rains during the 1997 El Niño year. The fact that the ECMWF hindcast also shows the opposite signature suggests that the teleconnection between ENSO and rainfall over the western Indian Ocean is not captured by the ECMWF hindcast, and this error in the ECMWF hindcast is propagated to RegCM3 via the boundary forcing. In fact, a closer look at the SST teleconnection pattern between Indian Ocean and ENSO in the observed skin temperature data set reveals that during a positive/negative ENSO event the western Indian Ocean is warmer/colder. The ECMWF hindcasts ensembles however did not reproduce this direct relation between ENSO and western Indian SST with warm anomalies visible throughout the Indian Ocean (Figure 9). The western Indian Ocean appears to be one of the poorest tropical regions in the forecast, and as the RegCM3 is forced
with this predicted SST as its lower boundary condition, this error in the skin temperature can affect the convection around the western Indian Ocean and its surroundings. The poor skill of the ECMWF system 3 model in predicting Western tropical SST during the summer period in forecasts initialized in May was recently noted by Stockdale et al. [2011], although no mechanism was suggested for the short-coming in the model performance. In general, identifying the causes of such model error is not straightforward, especially in a coupled system, and requires extensive sensitivity tests to be conducted along the lines of those in Tompkins and Feudale [2010], and which are beyond the scope of the present work. Since convection has been shown to be strongly related to SST in many general observation studies [e.g., Graham and Barnett, 1987] and through idealized cloud resolving model investigations [Lau et al., 1997; Tompkins, 2001], the SST errors over the western Indian Ocean in the ECMWF forecasts may be the cause of the rainfall errors seen in the RegCM3 simulations.

[39] Examining individual ensemble members of RegCM3 and ECMWF reveals that there are a limited number of ensemble integrations that show a starkly contrasting signal compared to the other ensemble members. For instance ensemble members 1, 4 and 8 (Figure 10) of RegCM3 show contrasting signatures over Ethiopia and over the Arabian Peninsula. ECMWF ensemble members tend to show more similarity, confirming the diagnostic of lower inter-ensemble spread. An exception for ECMWF ensemble set is ensemble member 8 which exhibits suppressed precipitation over the southwestern Indian Ocean in 1998 compared to 1997 in the global model hindcasts. This feature is opposite to the other ensemble members, and this is consequently the only ensemble member that captures the ENSO-Western Indian Ocean precipitation teleconnection reasonably well although it was incorrect over northeast Africa. Further examination of the skin temperature and the low level circulation features of ensemble member 8 reveals that the ground temperature over northeast Africa is anomalously colder than the ensemble
mean and furthermore the westerly influx from the Atlantic and the southwesterly flow near the coast of the Arabian Peninsula is heavily reduced (not shown). This reduction of westerly influx over Africa is associated with the southward shift in the convergence zone which agrees with the reduction in precipitation observed in many part of eastern Africa and the Arabian peninsula.

The fact that not all ensemble members are showing the same sign of precipitation anomalies in these extreme events only serves to reiterate the well known importance of using the information of all ensemble members in a probabilistic framework rather than resorting to the ensemble mean. Figure 11 shows the time series of the ECMWF hindcast together with the dynamical (RegCM3) and statistical downscaled ensembles hindcast over the sub regions and over the Ethiopia domain in comparison to GPCP and gauge data set. The statistical downscaled ensembles has the smallest interquartile range compared to the ECMWF and RegCM3 ensembles and the correlation of the ensemble mean with observations is small except for the eastern Ethiopia. For eastern Ethiopia the correlation between the statistically downscaled ensemble mean precipitation and
observations is higher than both RegCM3 and ECMWF. This low performance of the linear regression based downscaling using ECMWF ensemble precipitation as a predictor is not surprising considering the model is trained in the 1980s where the ECMWF SYS3 hindcasts had lower skill compared to the 1990s onward due to the continuous improvements in the ocean observational network [Tompkins and Feudale, 2010]. This low performance in the model is reflected in the poor correlation between the model rainfall and observed rainfall during the training period (not shown). Additionally the rainfall over parts of Ethiopia undergoes a low frequency variability where rainfall patterns in the 1980s differ from those in the 1990s, confounding the statistical modeling approach.

RegCM3 on the other hand clearly has larger ensemble spread compared to the ECMWF hindcasts, especially over north, central and eastern Ethiopia. The Gauge data shows high rainfall values compared to GPCP for the northern region. Generally, the ECMWF ensemble hindcast reproduces the observations better than RegCM3 ensemble and the global system performs well for the northern zone (Table 2). Again, similar to the perfect boundary runs, the correlation between RegCM3 and observations is higher when aggregated over larger areas, as expected.

4.1. Subseasonal Differences in the Hindcast Skill

The assessment of the seasonal mean predictive skill is extended to examine monthly rainfall as a function of forecast lead time. Figure 12 shows the correlation of the area-averaged ensemble mean model hindcasts with observations as a function of lead time for the various sub regions over Ethiopia. Figure 12 (top) shows skill with respect to rain
Figure 11. Standardized JJAS total precipitation time series of box plot for (left) RegCM3 ensembles and (middle) ECMWF ensemble and (right) statistically downscaled ensembles from ECMWF system 3 forecast. The green circles represent the ensemble mean whereas the red stars are for GPCP and the blue cross signs represent Gauge. The standardized anomaly for each data set is created by taking out the mean and divided by the standard deviation of each respective data set.

Table 2. Correlation of ECMWF, RegCM3 and Statistical Ensemble Mean Seasonal Hindcasts Against GPCP and Gauge Data Set Over Ethiopia

<table>
<thead>
<tr>
<th>Zone I</th>
<th>Zone IIa</th>
<th>Zone IIb</th>
<th>Zone III</th>
<th>Zone IV</th>
<th>All Zones</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gauge</td>
<td>GPCP</td>
<td>Gauge</td>
<td>GPCP</td>
<td>Gauge</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>0.85</td>
<td>0.95</td>
<td>0.51</td>
<td>0.58</td>
<td>0.38</td>
</tr>
<tr>
<td>RegCM3</td>
<td>0.55</td>
<td>0.56</td>
<td>0.26</td>
<td>0.31</td>
<td>-0.03</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.43</td>
<td>0.30</td>
<td>0.05</td>
<td>-0.01</td>
<td>-0.39</td>
</tr>
</tbody>
</table>

*Boldface and italics represent values that are significant at 0.05 and 0.1 levels, respectively.*
gauges, while Figure 12 (bottom) validates using GPCP data. One would assume the skill of the seasonal hindcast would decrease monotonically with lead time with the lowest skill in September but the results show that this is not the case for either model. One strong caveat is the short period considered for the analysis which is clearly undersampling, however, the sharp drop in ECMWF hindcast skill during July over the south western/central Ethiopia and August for the north/eastern regions mimics the predictive barrier in Indian Ocean SST reported by Stockdale et al. [2011]. The skill over eastern Ethiopia is higher during the first two months in RegCM3 hindcast but when considering the whole Ethiopian domain the skill is higher toward the end of the rainy season.

5. Probabilistic Verification of the Seasonal Forecast

Probabilistic seasonal forecasting quantifies and communicates the forecast uncertainties associated with the initial conditions and the model errors. Ensemble forecasts are converted to probabilistic forecast (e.g., tercile probabilistic forecasts) by calculating the fraction of ensemble members falling into each of the tercile categories [Kumar et al., 2001; Doblas-Reyes et al., 2009]. It is noted that the limited ensemble size implies under prediction of spread which could be improved by dressing techniques [Roulston and Smith, 2003] or by parametric probabilistic density function (pdf) estimation [Tippett et al., 2007].

Since the hindcasts from ECMWF, RegCM3 and the statistical model have different tercile threshold values compared to observations due to their respective biases, both the hindcasts and the validation data sets are first converted to a standardized anomaly for intercomparison. For GPCP and gauge data, the standardized anomalies are ranked according to their tercile category. The threshold values of the GPCP normalized anomalies are used to calculate the tercile probability of the ECMWF, the statistical and dynamical downscaled ensembles. Both grid point and areal-averaged values over the Ethiopian domain were considered for the probabilistic forecast verification. Ranked Probability Skill Score (RPSS) and Relative Operational Characteristic Score (ROCS) were used to assess the skill of the probabilistic forecast.

Figure 12. Correlation of (left) ECMWF and (right) RegCM3 with observation as function of time.
RPSS measures the skill of a forecast for correctly predicting the categories in probability space [Epstein, 1969; Wilks, 1995]. RPSS tells the relative accuracy of the forecast over climatology in predicting the category that the observations fall into. For RPSS, first Ranked Probability Score (RPS\_forecast) is calculated by computing the squared difference between the forecast and the observed cumulative probabilities. Similarly, RPS\_climatology is calculated by using the climatological cumulative probability instead of the forecast from the ensembles. The Ranked Probability Skill Score (RPSS) is then defined as the ratio of the difference between RPS\_climatology and RPS\_forecast relative to the RPS\_climatology. A 100% RPSS implies a perfect probabilistic forecast while a negative value of RPSS means the skill of the forecast probability is worse than climatology. Various studies [e.g., Kumar et al., 2001; Muller et al., 2005; Tippett et al., 2007] have shown that categorical probabilities obtained from finite ensembles significantly reduce the RPSS. They suggest a modified version called ‘unbiased’ or ‘debiased’ RPSS to estimate the corresponding RPSS of the infinite ensemble size. Tippett [2008] showed that the unbiased RPSS relevant for an infinite ensemble can be obtained by adding another term in the reference forecast in effect to increase the error in the reference forecast artificially, i.e.,

$$\text{RPSS}_D = 1 - \frac{\text{RPS\_forecast}}{\text{RPS\_climatology} + \frac{1}{M}\text{RPS\_climatology}}$$

where M denote the ensemble size.

The mean debiased RPSS (compared to climatology forecast) of the ECMWF hindcast and RegCM3 hindcast from grid point by grid point computation (Figure 13) shows that for a significant part of the domain the skill is comparable to climatology at the 1–4 month lead time. There are, however, some regions where the models have positive skill. For instance, ECMWF hindcasts are skillful over the northern and southwestern part of Ethiopia and Sudan, whereas RegCM3 beats the climatological forecast over the northwestern part and southeastern Ethiopia.

When the areal average skill score is considered, ECMWF has a positive RPSS at sub-regional level for all homogeneous rainfall zones and for both validation data sets. The statistical and RegCM3 downscaled hindcasts however have a positive skill only for particular sub region. For instance the statistical method has a positive skill only over the eastern Ethiopia whereas for RegCM3 the positive skill is only over the northwestern Ethiopia (Zone I). The worst performance for both dynamical and statistical downscaling methods occurs over the southwestern part (Zone IIb) as shown in Figure 14. This region also exhibited a low skill in the statistical forecasting assessment in Diro et al. [2011c] study where they used SST as the predictor. This is due to its low correlation with SST anomalies, instead the rainfall in this region is more related with the Quasi Biennial Oscillation (QBO) [Diro et al., 2011a] which is another source of predictability for the seasonal forecasting. The fact the ECMWF resolves the stratosphere better than that of RegCM3, having more than twice the number of vertical levels, might explain the superior skill of the ECMWF model compared to RegCM3 for this part of the country.

As expected, increasing the aggregation area and analyzing the rainfall for all Ethiopia, both the ECMWF hindcast and RegCM3 have positive skill, although the ECMWF skill reduces when gauge data is used for validation purpose instead of the GPCP. In fact the skill of the RegCM3 model integrations is only significant when aggregated over the whole country.

The second verification score considered is the relative operating characteristic (ROC) score [Mason, 1982]. The ROC score tells us the ability of the forecast to discriminate
events and therefore it is a measure of resolution. The area under the ROC curve, obtained by plotting the hit rate versus false alarm rate for various probability thresholds. In this case a forecast is skillful relative to a random forecast, if the area under the ROC curve (AUC) is greater than 0.5, or equivalently if $2AUC - 1 > 0$. For the tercile forecast, hit rates and false alarm rates are computed (for the three categories i.e. dry, normal and wet) at different probability threshold levels to construct the ROC curve.

Figure 15 shows the ROC Score ($2AUC-1$) for the dry and wet tercile categories of the ECMWF and RegCM3 ensemble JJAS rainfall hindcasts when GPCP is used as a validation data set. For both models, the skill is higher for the dry category. In the case of ECMWF, the spatial coverage of the positive skill for the dry category covers most of the continental domain, while for RegCM3 on the other hand, the skill is concentrated mainly over Ethiopia and Red Sea only. For the wet category the skill is dominant over the Arabian Peninsula though this region receives less rainfall compared to northeast Africa during this season. The lowest skill in both ECMWF and RegCM3 hindcast for the wet category is over the western Indian Ocean, coinciding with the region where the ENSO -western Indian Ocean SST teleconnection is misrepresented in the ECMWF model.

Looking at the areal average, again ECMWF has better statistics for each of the homogeneous zones (not shown) than either dynamical or statistical downsampling methods. For the larger aggregated area (averaged over all zones), all the three models have positive skill for the above normal and below normal categories. It is interesting to note that RegCM3 and the statistical model have better skill scores compared to ECMWF when gauge data is used as validation data set whereas the reverse is true, i.e. ECMWF is the better model when GPCP is used for validation. This is the case for all categories as shown in Table 3. While this could be due to the fact that GPCP is produced on a spatial scale more similar to the global model, another potential reason for this is that gridded rainfall products, including (and especially) GPCP, are systematically used at ECMWF for evaluation of new physics parameterizations and their subsequent tuning during the development of the prototype system of each new model cycle of the operational system at ECMWF [e.g., Tompkins et al., 2007]. Model upgrades must improve, or at least not degrade, the precipitation climate errors compared to GPCP data (one of a large array of climate metrics and predictability skill scores). Thus the model is effectively “tuned” to these data sets over time, where the terminology refers to decisions concerning parametrization scheme upgrades and developments, in addition to the choice of individual parameter settings. This constraint is weaker for data sets not used in the evaluation suite, such as the raingauge data set presented here, where the model error will random-walk toward reduced errors over time. Regional climate models undergo many fewer upgrades, and therefore have fewer model generations to perfect the tuning process. In summary, the rain gauge data in this respect can be considered a
more independent data set with which to validate the two modeling systems.

6. Discussion

The experiments conducted in this paper have confirmed many previous studies concerning the importance of lateral boundary condition forcing for regional model integrations. For the East Africa domain, the regional model physics combined with the improved representation of the Ethiopian highland topography, resulting in precipitation forecasts with small biases and better year-to-year variability than the short range forecasts driven with ERA-interim.

Repeating the experiment using the ECMWF seasonal forecast model, the global model was found to suffer from some systematically biases in the region, particularly with respect to the seas surface temperature anomalies in the Indian Ocean in ENSO years. The teleconnections associated with these biases led to an incorrect lower level wind direction over the target region. These dynamical wind errors reduced the ability of the global model to predict seasonal anomalies, although the model still had significant skill at 1–4 month lead times if rainfall is aggregated to climatically homogeneous zones or at the Ethiopian national level. The propagation of these errors in the boundary conditions to the regional model has negatively impacted the regional model ability to even represent the precipitation climatology of the region well, with anomalous precipitation features appearing over the Arabian peninsula. As a result,

![ROC Score for the dry category of ECMWF hindcast](image1)

![ROC score for the dry category of RegCM3 hindcast](image2)

![ROC Score for the Wet category of ECMWF hindcast](image3)

![ROC score for the Wet category of RegCM3 hindcast](image4)

Figure 15. ROC score for the (top) dry category and (bottom) wet category, (left) from ECMWF hindcast and (right) from RegCM hindcast.

Table 3. ROC Area (Area Under the ROC Curve) for ECMWF, RegCM3 and Statistical Ensemble Seasonal Hindcasts Against GPCP and Gauge Data Aggregated Over the Five Rainfall Zones Over Ethiopia

<table>
<thead>
<tr>
<th>ROCs</th>
<th>Dry</th>
<th>Normal</th>
<th>Wet</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Gauge</td>
<td>GPCP</td>
<td>Gauge</td>
</tr>
<tr>
<td>ENSEMBLE</td>
<td>0.67</td>
<td>0.86</td>
<td>0.54</td>
</tr>
<tr>
<td>RegCM3</td>
<td>0.86</td>
<td>0.64</td>
<td>0.69</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.69</td>
<td>0.69</td>
<td>0.62</td>
</tr>
</tbody>
</table>
the regional model was only skillful in a subset of zones in terms of month 1–4 lead time predictions. Point wise skill measures also failed to improve relative to the global model, despite the significantly higher resolution.

This result highlights a potential pitfall of regional climate modeling studies in which regional models are configured for a partial region using reanalysis and then applied to downscale present day and future climate integrations. While careful consideration is often given as to which GCMs give the best performance and should be selected for dynamical downscaling they may still be unable to reproduce correct statistics (mean and variability) of precipitation over the target region of focus, with convective zones shifted in location, and the regional model will then struggle to recuperate the regional climate. The use of a numerical weather prediction-like, seasonal forecasting framework, repeatedly reinitializing the global model from observed initial conditions for the atmosphere and ocean could offer a method of increasing confidence in global models selected for downsampling, as similarly suggested previously by Phillips et al. [2004], Klein et al. [2006] and Rodwell and Palmer [2007].

Discovering a systematic bias for a given region in the host model, that is likely to propagate to the regional model, how can this best be corrected? Presently, most approaches apply bias correction to the regional model output [e.g., Piani et al., 2009] before application to end-user models. However this approach does not allow the regional dynamical model to gain any benefit from its higher resolution when the global model LBCs are biased and of poor quality. A modified regional modeling paradigm could instead be to focus on the interface between the global and regional climate models, that is, to attempt to correct the SST fields used from the global models before their use to drive the regional models at its surface boundaries. This modified approach would allow to remove the systematic SST bias in the driving GCM. It will not, however, improve the wind errors at the LBCs associated with the poor teleconnections - these would have to be addressed using a regime-dependent bias correction of the LBCs themselves, although implementing such a scheme is far from straightforward.

7. Summary

In this study, both ERA-interim reanalysis and ECMWF seasonal hindcast ensembles are dynamically downscaled to a higher spatial resolution for the northern hemisphere summer season using RegCM3 for the Eastern Africa domain.

Sensitivity experiments were first conducted concerning the experimental setup and these confirmed previous findings of Seth and Giorgi [1998] in that with accurate boundary forcing from reanalysis the domain size is less important; sensitivity to domain size is much more relevant with the use of seasonal forecasts (and by extension global climate model) integrations. Larger domains obviously reduce the impact of the boundary forcing but global models systematic errors still propagate to the regional model.

The perfect boundary condition (reanalysis) integrations show that RegCM3 reproduced the spatial and the interannual variability well compared to observations. The result assures the utility of RegCM3 under a near correct boundary conditions. The model improved forecast skill scores for all regions in Ethiopia relative to the global model coarser resolution short range forecasts initialized every day from the ERA-interim reanalysis. The use of seasonal forecasts, instead resulted in a significant systematic errors in the precipitation over some part of the domain in the regional model. The global model poorly represented the Indian Ocean SST anomalies. This error in the lower boundary condition resulted in incorrect lateral boundary dynamical forcing and led to a strong deterioration in the regional model precipitation forecast, not least being the excess precipitation bias over the Arabian peninsula, a region that climatologically receives little rain during summer and the rains are usually limited to the highlands over the western coast of the Arabian peninsula.

For the interannual variability it is observed that grid point comparisons of correlations, RPSS, and ROCS using GPCP as a validation data set suggest that RegCM3 is unable to improve on the global forecasting system; in fact in many areas the point by point validation was inferior. Even averaging over climatologically homogeneous areas of Ethiopia showed that RegCM3 had positive skill only in a subset of regions, while the global model was skillful in all areas. It is notable that when using a higher resolution gridded gauge data set (with 160 stations) for an independent validation, the RegCM3 performance is improved with respect to the global model, both with the reanalysis and seasonal forecast integrations. Nevertheless the deterministic verification shows that the regional model is inferior when used in seasonal forecast mode, irrespective of the validation data set used.

The probabilistic validation revealed that both hindcast ensembles have skill in discriminating wet and dry events compared to random forecasts. Similarly, both model ensembles predicted the probability of each category better than climatology when aggregated over a large areas and at the country scale. This is an encouraging result for a one to four month lead time forecast, and shows that both the global and regional modeling systems have potential to help the community of the horn of Africa in early warning system development for impacts modeling.

The probabilistic verification skill scores did show that at a country-wide scale, RegCM3 does improve the skill of the global forecast, but only when using rain gauges for validation. The deterministic assessment though, on the other hand suggested the improved resolution of the regional model does not benefit seasonal forecast skill significantly, even over a region of highly varying and complicated topography such as Ethiopia where such models would be expected to excel. While this seems a disappointing result, the fact that the model can give a positive impact when given reliable boundary forcing gives reason for optimism, since as global models improve, it appears that further benefit can be gained from resolution. In the meantime, an extension to the present dynamical downscaling paradigm could be to apply a bias correction technique to the global model SST fields prior to their use in dynamical regional downscaling, and possibly also the LBCs, although the best approach to achieve the latter remains unclear.

Future work will examine the potential benefit of such an approach in the seasonal forecasting framework, that allows for explicit validation of the modeling system in a hindcast mode as used here. These futures studies will employ larger domains covering the whole of Africa and the Indian Ocean to include the large-scale rainfall controls over Africa to be
References


Epstein, E. N. (1969), Stochastic dynamic prediction, Tellus, 21, 739–759.


Muller, W. A., C. Appenzeller, F. J. Doblas-Reyes, and M. A. Liniger (2005), A debiased ranked probability skill score to evaluate probabilistic ensemble forecasts with small ensemble sizes, J. Clim., 18, 1513–1523.


