Evaluation of Empirical Statistical Downscaling Models' Skill in Predicting Tanzanian Rainfall and Their Application in Providing Future Downscaled Scenarios

HABIBA I. MTONGORI

Department of Geosciences, University of Oslo, Oslo, Norway, and Tanzania Meteorological Agency, Dar es Salaam, Tanzania

FRODE STORDAL

Department of Geosciences, University of Oslo, Oslo, Norway

RASMUS E. BENESTAD

Norwegian Meteorological Institute, Oslo, Norway

(Manuscript received 17 January 2015, in final form 14 August 2015)

ABSTRACT

Projections of three important seasonal rainfall parameters—total precipitation (p_t), wet-day mean (μ_w) and wet-day frequency (f_w)—considered to be relevant to crop agriculture were performed. Links between large-scale climate variables and local precipitation in Tanzania were investigated during the March–May (MAM), October–December (OND), and December–April (DJFMA) rainfall seasons. Variables found to have strong links were used to downscale the local precipitation for three future periods; near term, mid-century, and end of century. Downscaling models for p_t, μ_w, and f_w were calibrated using observed large-scale seasonal rainfall and projected downscaled parameters were obtained based on rainfall simulations from ensembles of GCMs. The models' skill scores were found to be sensitive to the domain size and number of leading EOFs used. The common EOF method employed in the downscaling modulated the skills depending on the GCMs used. The spread in the rainfall projections was found to be larger in OND and moderate in MAM and DJFMA. The multimodel mean projections in response to two RCPs (RCP4.5 and RCP8.5) suggest a shift toward wetter (drier) conditions (p_t) for OND (DJFMA) for all three periods. There is no uniform projection for MAM; some stations are projected to feature wetter and some drier conditions. In the midcentury and end of century, there is an increase of precipitation to about 40% for some areas getting OND rainfall and a decrease of precipitation up to about 10% for some areas getting MAM or DJFMA rainfall. Generally, the magnitude of change strongly differs across the areas.

1. Introduction

Agriculture accounts for about 43% of the annual gross domestic products (GDPs) of East African countries (Omamo et al. 2006). Yet the farming in these countries is dominated by small holders who rely on rainfall (United Republic of Tanzania 2014; Waithaka et al. 2013). Rain-fed agriculture has remained the dominant source of staple food production and the livelihood foundation of the majority of the rural poor in sub-Saharan Africa (Cooper et al. 2008). As a consequence the sector has remained susceptible to climate change and variability. Changes in climate between and within the seasons (Rowhani et al. 2011) and changes in the mean climate and the magnitudes and frequency of the extreme events (Moriondo et al. 2011) have a significant impact on crop yields and livelihood as well as food security (United Republic of Tanzania 2014). It is therefore important to have an improved understanding of the influence of climate on agricultural production

Denotes Open Access content.

Corresponding author address: Habiba I. Mtongori, Department of Geosciences, University of Oslo, P.O. Box 1022, Blindern 0315 Oslo, Norway.
E-mail: h.i.mtongori@geo.uio.no

DOI: 10.1175/JCLI-D-15-0061.1

that is needed to cope with expected changes in temperature and precipitation, and hence to minimize an increasing number of undernourished people in food-insecure regions (Rowhani et al. 2011).

Although some general conclusions about climate change and their impacts have been drawn especially at large scales, the potential damage of climate change at a particular region or farm needs to be assessed under site-specific conditions (Li et al. 2011). Moreover, there is an increasing number of decision makers from different organizations seeking climate predictions of the regional and local changes that will impact people. Since global climate models (GCMs) are at coarse resolution such that they are unable to resolve features at regional or local scales, it is therefore necessary to downscale the GCMs in order to produce scenarios of climate change at the scale required by impact assessment studies.

The downscaling of GCMs in Africa has received relatively less attention as compared to other part of the world (Hulme et al. 2005; Shongwe et al. 2011). Moreover, existing studies on regional climate modeling have mostly focused on southern Africa (e.g., Hewitson and Crane 2006; Landman and Tennant 2000) and West Africa (e.g., Paeth et al. 2005; Penlap et al. 2004), with very few in East Africa (e.g., Moore et al. 2012; Taye et al. 2011). With the exception of the statistical downscaling works done by Tumbo et al. (2010) over Same district in Tanzania and Taye et al. (2011) over Lake Victoria basin, previous downscaling work already done in East Africa/Tanzania is based on dynamical downscaling (e.g., Moore et al. 2012; Omondi et al. 2014), with analyses given on a regional or country level, which is too course to draw conclusions on impacts at local levels. While the work of Tumbo et al. (2010) was based on only one location in Tanzania, in this study we also propose the use empirical–statistical downscaling (ESD) covering different locations in the country representing homogeneous climatic zones (see, e.g., Ng'ongolo and Smyshlyaev 2010), also using several GCMs.

The current study therefore attempts to address the shortcomings of using coarse spatial resolution models for assessment of climate change for impact studies (e.g., agriculture). In particular, we aim to use ESD to explore changes in climate at local level. We rely on the hypothesis that the local rainfall is connected with larger-scale conditions and that we can use GCMs to describe the statistics of year to year variations in seasonal precipitation in Tanzania. The objectives of this work are 1) to assess and establish links between large-scale circulation patterns and local precipitation in Tanzania, 2) to develop ESD models based on the identified links, 3) to assess the skill of the ESD models, 4) to create climate change scenarios using ESD models and GCMs projections, and 5) to identify causes of the spread and differences in the multimodel downscaled results.

Regression-based ESD utilizes systematic dependency transfer functions [$y = F(x)$] between large-scale circulation (predictors) variables, x, and regional or local climate variables, y (predictand), to derive regional climate information (Leung et al. 2003). This function is typically established through training and validating the models using point observations and gridded reanalysis data (Fowler et al. 2007). The derived statistical relationships are then applied to the large-scale climate variables from GCM projections to estimate corresponding future local and regional characteristics (climate change). Practical considerations for developing an ESD scheme involve the selection of appropriate predictor variables, predictor spatial representation, domain size and location, types of transfer functions, and definition of temporal scales such as season (Leung et al. 2003).

The utility of ESD depends upon four assumptions: 1) high-quality large-scale and local data are available for an adequately long period to establish robust relationships in the current climate, 2) the link between predictands and predictors is strong in order to explain satisfactorily the local climate variability, 3) the predictor variable is well simulated by the GCM, and 4) relationships derived from recent climate remain relevant in a future climate (stationarity) (Benestad 2001a; Busuioc et al. 2001; Wilby and Wigley 2000).

The use of ESD has the advantage that 1) it is computationally very inexpensive, 2) it can provide information at smaller scales than the GCM-scale output, 3) it can be used to derive variables that are not available from regional climate models (RCMs), 4) it is able to directly incorporate observations, 5) it provides diagnostics and insights (e.g., a picture of how large and small scales are related), and 6) it can be used to assess the models. The main drawbacks of this method are that 1) ESD has underlying assumptions that can only be partly tested (Willems et al. 2012), 2) strong predictor variables do not always exist, and 3) long time series of relevant data are required to form the relationships (Jones et al. 2004; Leung et al. 2003).

2. Data and methods

a. Study area

Tanzania (Fig. 1), one of the East African countries, is located in the Southern Hemisphere. Its climate consists

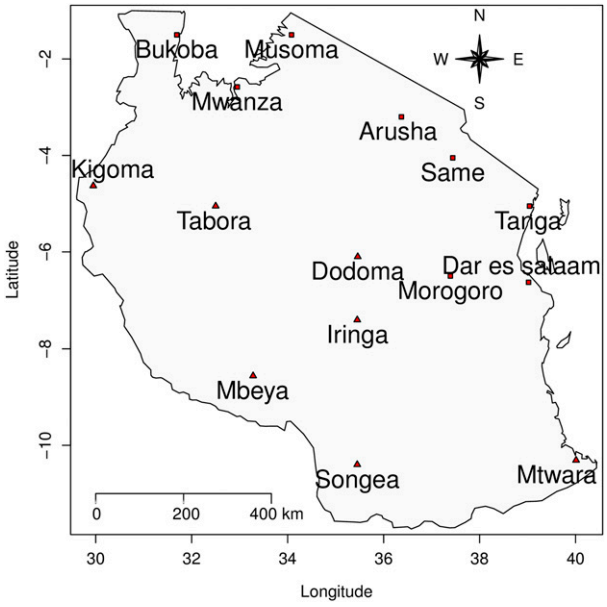

FIG. 1. Map of Tanzania showing study sites; stations marked by red squares represent areas with two rainy season (MAM and OND) and those marked by red triangles represent areas with one long rainy season (DJFMA). The station Dar es Salaam is called Dsalaam in the skill score and projection figures.

of two rainfall regimes; a bimodal one with rainfall seasons in March–May (MAM) and October–December (OND), and a unimodal one with a rainfall season in December–April (DJFMA). These rainfall regimes are controlled by the northward and southward migration of the intertropical convergence zone (ITCZ). The bimodal rainfall regime is observed over the areas located in the northern parts of the country while the unimodal rainfall regime is more typical for the southern, central, and western parts of the country, coinciding with the austral summer (December–February). The remaining months, from June to September and January to February (for bimodal areas) are relatively dry.

It is widely established that rainfall over East Africa varies in space and time (Camberlin and Philippon 2002; Indeje et al. 2000). The variability has been linked to the presence of local features (Indeje et al. 2000), variations in the dominant large atmospheric circulation (Camberlin and Okoola 2003; Mapande and Reason 2005), and global warming (Schreck and Semazzi 2004).

Some circulation oscillations have been linked to the variation of East African seasonal rainfall, ENSO (Indeje et al. 2000; Schreck and Semazzi 2004; Kijazi and Reason 2005; Ogallo 1988), the zonal circulation in the equatorial Indian Ocean (Black et al. 2003; Kabanda and Jury 1999; Marchant et al. 2007; Saji et al. 1999), the Madden–Julian oscillation (MJO) (Mpeta and Jury 2001; Mutai and Ward 2000), and the quasi-biennial oscillation (QBO) (Indeje and Semazzi 2000; Ng'ongolo and Smyshlyaev 2010). ENSO alone is found to account for almost 50% of the East African rainfall variance (Ogallo 1988).

b. *Data and choice of predictors*

1) DATA

Predictand data used for the ESD model development are daily rainfall observation records for several stations in Tanzania (Fig. 1) obtained from Tanzania Meteorological Agency (TMA). The time period available for this dataset differs from station to station with some stations going back to the 1940s and few to the 1960s. However, for this work we used data from 1979 to 2012, which matches the MERRA data period. Within this period, 99.9% of the daily data are available for analysis. However, quality control of the data including inhomogeneity that could be caused by any change to station or its operation was not performed. Predictor variables used for correlation analysis and hence ESD model development are the reanalysis of monthly mean surface and atmospheric large-scale circulation fields that were taken from the NASA's Modern-Era Retrospective Analysis for Research and Applications (MERRA), available online at http://disc.sci.gsfc.nasa.gov/daac-bin/DataHoldings.pl. The MERRA time period covers the modern era of remotely sensed data, from 1979 to 2012. A special focus of the atmospheric assimilation has been on the hydrological cycle. The reanalysis is available at a spatial resolution of 0.5° latitude × 0.67° longitude (Rienecker et al. 2011).

As predictor variables for downscaling future rainfall projections we have used GCM monthly mean rainfall simulations for the period 2006–2100. These simulations are obtained from phase 5 of the Coupled Model Intercomparison Project (CMIP5) climate change experiments (Taylor et al. 2012). This dataset is available online at http://pcmdi9.llnl.gov/esgf-web-fe/. Projection simulations forced with two representative concentration pathways, RCP8.5 and RCP4.5 (Moss et al. 2010), were used. These long-term integrations are initiated from multicentury preindustrial control integration (Taylor et al. 2012). RCP8.5 is consistent with a high emissions scenario and RCP4.5 is consistent with a midrange mitigation emissions scenario. The GCMs used in this study are listed in Table 1. The set of GCMs includes models with different spatial resolutions and degrees of complexity. Generally, the spatial resolution of CMIP5 GCMs roughly ranges from 0.5° to 4° for the atmosphere component and from 0.2° to 2° for the ocean

TABLE 1. Set of GCMs used in a multimodel ensemble projection for rainfall over Tanzania. (Expansions of acronyms are available online at http://www.ametsoc.org/PubsAcronymList.)

Model	*nx*	*ny*	Institute	Model components	Vertical layers (top)
GISS-E2-H	144	90	NASA Goddard Institute for Space Studies (Schmidt et al. 2014)	Atmosphere + ocean	40 (0.1 hPa)
HadGEM2-AO	192	145	National Institute of Meteorological Research/Korea Meteorological Administration (Baek et al. 2013)	Atmosphere + sea ice + ocean + land + atmosphere with increased resolution at the equator	38 (38 km)
CSIRO Mk3.6.0	192	96	Commonwealth Scientific and Industrial Research Organisation in collaboration with the Queensland Climate Change Centre of Excellence (Rotstayn et al. 2012)	Atmosphere + ocean	18
NorESM1-M	144	96	Norwegian Climate Centre (Bentsen et al. 2013)	Atmosphere + ocean + sea ice + land	26 (2.917 hPa)
MPI-ESM-LR	192	96	Max Planck Institute for Meteorology (MPI-M) (Giorgetta et al. 2013)	Atmosphere + ocean + sea ice + land + marine biogeochemical	47 (0.01 hPa)
CCSM4	288	192	National Center for Atmospheric Research (Meehl et al. 2012; Subramanian et al. 2011)	Atmosphere + ocean + sea ice + land	26
BCC_CSM1.1(m)	320	160	Beijing Climate Center, China Meteorological Administration (Xin et al. 2013)	Atmosphere + ocean + sea ice + land	26 (2.914 hPa)
MIROC5	256	128	Atmosphere and Ocean Research Institute (The University of Tokyo), National Institute for Environmental Studies, and Japan Agency for Marine-Earth Science and Technology (Watanabe et al. 2010)	Atmosphere + ocean + sea ice + land + aerosols	40 (3 hPa)

component (Taylor et al. 2012). In our selection, we have included GCMs of variable resolutions from very high to low.

2) PREDICTOR SELECTION

Potential predictor variables were selected by referring previous works on ESD and empirical seasonal forecasting done in Africa. These are atmospheric large-scale fields such as 850-hPa height (gpm85) (Landman et al. 2012) and sea level pressure (SLP) (Huebener and Kerschgens 2007; Landman and Tennant 2000; Zhao et al. 2005). Other fields are large-scale precipitation (Pr) (Diro et al. 2012; Greene et al. 2012) and large-scale temperature (T2m) (Ambrosino et al. 2011; Greene et al. 2012). We also included horizontal components of wind (V10m and U10m) and specific humidity (shum10m) at 10-m level.

According to Kalnay et al. (1996) and Kistler et al. (2001), the reliability of the output variables from the reanalysis can be classified in categories A to D, depending on the relative influence of the observational data input and the model forecast on the variable during data assimilation. With the exception of SLP, which is classified as a class A variable, the rest of the selected variables (i.e., gpm85, V10m/U10m, T2m, and shum10m) are considered as class B variables and Pr as a class C variable. However, because of the scarcity of data in the tropics (where this study is undertaken) the relative influence of model forecast during data assimilation in all variables is considered to be large as compared to that of observational data input. For this reason, all variables were given equal weight in the predictor selection.

Spatial correlation analysis was performed between several predictors derived from the MERRA dataset and the predictand (station observations) to identify predictors with strong links to the observed precipitation. Monthly large-scale variables (SLP, Pr, gpm85, V10m/U10m, T2m, and shum10m) were tested for their significance in describing their relationship with the local monthly precipitation values in Tanzania. This was done by considering different domains, rainy seasons, and station locations. Pr and shum10m stand out as showing significant links with the stations' observed rainfall. Since both parameters represent moisture sources, we decided to use precipitation fields with the advantages that 1) they represent the synoptic-scale processes in the atmosphere including changes in air humidity and movements of air (Linderson et al. 2004) and 2) they carry climate change signals due to anthropogenic activities (Benestad et al. 2008; Chen et al. 2006; Wilby et al. 1998) and they represent other processes not

explained by the atmospheric circulation. Finally all downscaling models were developed using the same predictor (Pr) regardless of the season, type of statistics, and location of the station.

c. Method

The ESD models used in this study are built in the following steps: First, the long-term mean for each dataset [i.e., MERRA (1979–2012) and GCMs (2006–2100)] is subtracted from that data on each grid point, and the two datasets are then combined along the time axis (concatenated) to form a single dataset covering the period 1979–2100 (see, e.g., Barnett 1999; Benestad 2001a). A bilinear spatial interpolation is then used to convert the GCM spatial grid to be in the same grid as that of MERRA. Second, a standard empirical orthogonal function (EOF) analysis (Lorenz 1956) is applied to the anomalies of the combined dataset with the common grid (which is now 0.5° × 0.67°). The eigenvectors (EOFs) produced from the EOF analysis represent patterns of variability that MERRA and GCMs data share in common. The EOF analysis applied to the combined dataset is referred to as common EOF analysis (Barnett 1999; Benestad 2001a) and its use ensures that the same large-scale spatial climatic patterns that are associated with the observed local climate anomalies are used to infer local climate changes related to the changes in the regional climate (Benestad 2002). Third, using a stepwise screening process based on the Akaike information criterion (AIC) (Akaike 1974), a multiple linear regression analyses between the predictand (i.e., observed precipitation) and the principal components (PCs) of the 20 leading EOFs of the combined dataset (MERRA + GCMs) is performed to decide the number of PCs to be included as predictors in the final models. During stepwise screening, a model that minimizes loss of information in simulating the current climate (i.e., with minimum AIC value) is retained for downscaling process. A leave-one-out cross-validation was then performed to assess the prediction skill of the models. This assessment was repeated for a test case where the station data had been replaced by random numbers for which there should be no skill. This assessment suggested that the models were not overfit for a number of stations and that the higher number of EOFs contained details related to some of the stations (see Fig. 2). The leading 20 EOFs are used because it allows more regional details in predictor fields to be used in the downscaling models especially for the MAM and DJFMA seasons that do not show a pronounced link with large-scale circulations. Before the model calibration, the best-fit linear trend is subtracted from each grid point in the observed predictor values (MERRA) and from the predictand (station observations), as the presence of a linear trend may introduce systematic biases to the model calibration (Benestad 2001b). However, the trends are included when the models make predictions. The downscaling models are calibrated with the part of the combined PCs that represents the actual observations (i.e., MERRA), and for future projections the downscaling is generated using part of the combined PCs that represents the GCMs simulations. More details of common EOF analysis and downscaling procedures can be found in Benestad (2007, 2001a, 2002).

d. ESD model implementation

Prior to statistical model development and hence downscaling, small-scale variability in the large-scale precipitation field was removed. A low-pass filter (boxcar smoother) was applied to the monthly large-scale precipitation field (reanalysis and GCMs). The width of the boxcar was 5 points (grid boxes) along the x and y axes and 3/5 points (months) along the time axis for the bimodal/unimodal climate regimes, respectively. The motive behind spatial smoothing is to reduce noise and capture the large-scale features. Temporal smoothing on the other side removes the short-term fluctuations hence emphasizes the long time fluctuations.

Predictors for model calibration are PCs of the 20 leading EOFs that represent the actual observation part of the combined dataset (i.e., MERRA). These 20 leading PCs undergo stepwise screening process in a regression method to select potential predictors that will be included in the final downscaling model. For future projections, predictors are PCs of the EOFs that were retained from the stepwise screening process but that represent the GCM simulations part of the combined dataset. Predictands are parameters that were calculated from the daily rainfall but aggregated in to seasonal basis for an individual local station and for each rainy season (i.e., MAM, OND, and DJFMA). The examined parameters are as follows: 1) frequency of wet-day occurrence (f_w), henceforth wet-day frequency [this parameter explains how frequently it rains in a season; it is calculated as the fraction of number of wet days (n_w) to the total number of days (n_a) in a season and it is measured in fraction (of counts)]; 2) seasonal total rainfall (p_t), henceforth total rainfall, the sum of rainfall amounts for all wet days in a season (measured in mm); and 3) wet-day mean (μ_w), which is calculated as the total rainfall (p_t) divided by total number of wet days (n_w) in a season; the wet-day mean is equivalent to daily rainfall intensity (measured in mm day^{-1}). Wet days are described as number of data points with rainfall ≥ 1 mm day^{-1} while all days include days with valid recordings (i.e., rainfall ≥ 0).

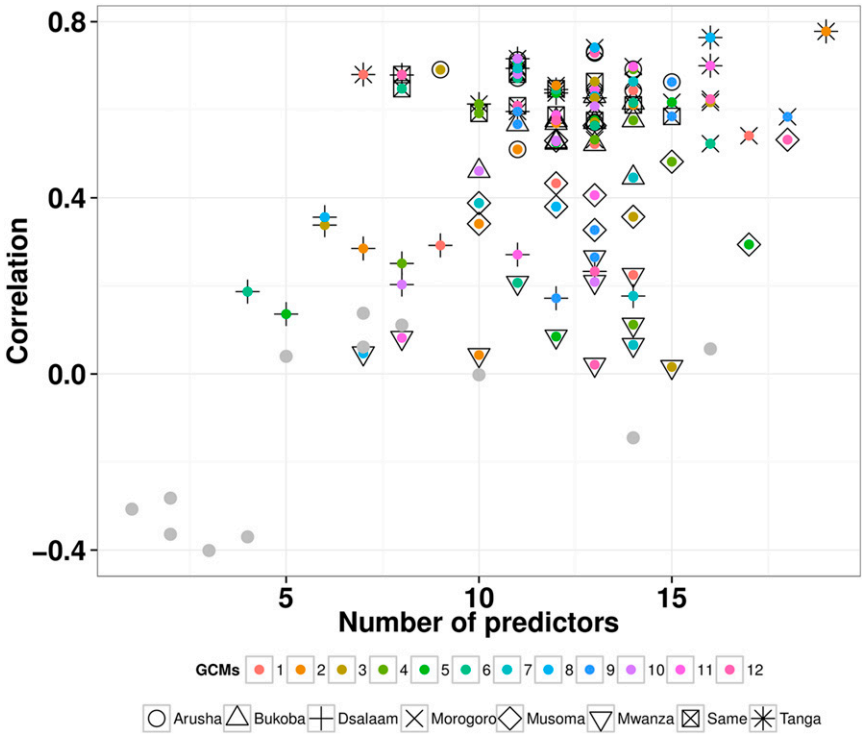

FIG. 2. Prediction skill (correlation) of the downscaling models as a function of EOFs used. Results are for the OND season. The prediction skill is established by performing a correlation test between independent (observed values not used in the model calibration) and predicted values. Each station has 12 models developed from 6 GCMs forced by either RCP4.5 or RCP8.5. The color code numbers 1–6 are models developed from the GCMs BCC_CSM1.1, CCSM4, GISS-E2-H, HadGEM2-AO, MPI-ESM-LR, and NorESM1-M that are forced with RCP4.5 and numbers 7–12 are models developed from the same GCMs but forced with RCP8.5. The solid gray circles indicate the prediction skills for the downscaling models developed from the above GCMs but when station data have been replaced by random numbers for which there should be no skill.

Changes in these parameters were calculated by comparing future 30-yr climate periods (near term = 2010–39; midcentury = 2040–69; and end of century = 2070–99) to a baseline climate period (1979–2012), both extracted from the downscaled scenario time series. Each downscaled value was obtained from individual run using a GCM that is either forced with RCP4.5 or RCP8.5. The outputs from each run were combined together (tabulated) and further statistical analysis (e.g., mean, median, percentiles) was performed. The downscaling analysis and graphics processing are performed in the R environment (http://www.R-project.org/) by using the empirical statistical downscaling packages clim.pact (Benestad 2004) and ggplot2 (Wickham 2009). These packages are open source and free.

3. Results

Since rainfall regimes across Tanzania vary significantly in space and time and since projections also vary widely between seasons and regions, results are presented for each season but for both emission scenarios combined for ease of discussion. Throughout the discussion the ESD models will be referred to simply as models and global climate models will be referred to as GCMs.

a. ESD models skill score using reanalysis data

The skill of the models discussed here is the coefficient of determination (abbreviated R^2). This tells how much of the variance of the dependent variable (predictand) is explained by the regression line (expressed as percentage). The skill of the model may be influenced by several factors such as a weak relationship between predictor(s) and predictand, poor quality of the data (station observation and reanalysis), the design of the empirical model used, the size of the predictor domain, and many others.

In this study, as in previous statistical downscaling studies (Benestad 2002), the predictor domain size was found to influence the skill of the model. When a larger

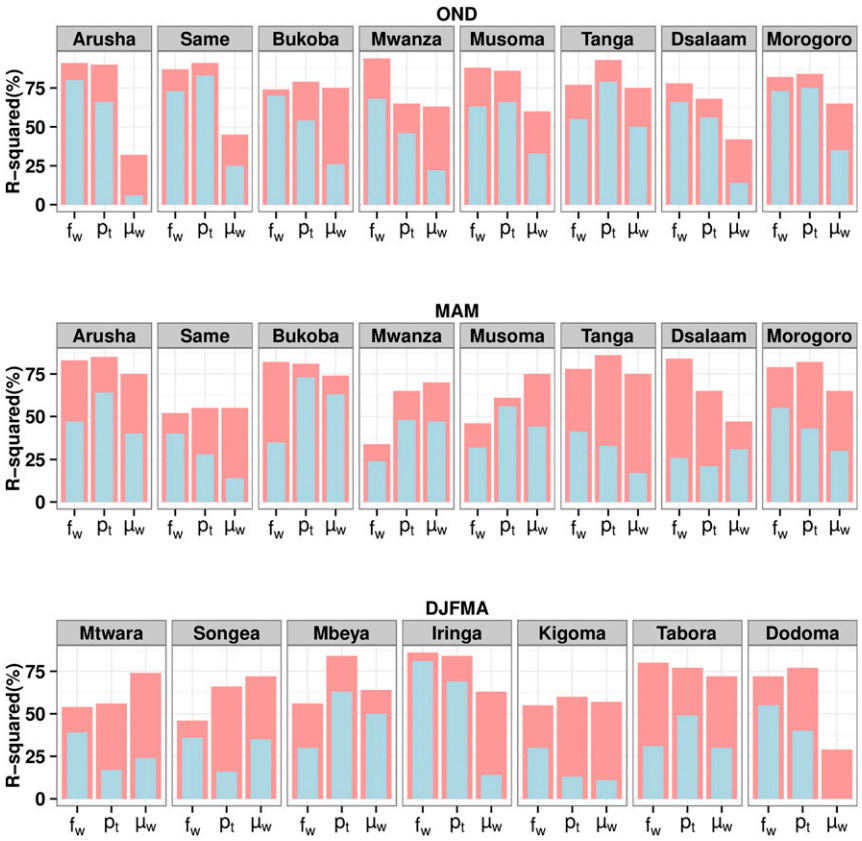

FIG. 3. Stacked bar plots for different stations and seasons showing skill scores (R^2) of the downscaled models based on MERRA data. Bars in pink were obtained when 20 leading EOFs are used in regression analysis and those in light blue were obtained when 8 leading EOFs were used. In each station category, f_w = wet-day frequency, p_t = total precipitation, and μ_w = wet-day mean.

domain was used, the model skill was generally found to be lower than when a small domain was used. Only in a few cases the opposite was detected, for some stations and particular seasons. According to Benestad (2002) it is tricky to find the optimum downscaling domain, as the domain needs not to be too large or too small. As opposed to our finding, the study of Yang et al. (2013) found that different predictor domains had less influence on the projected monthly rainfall over Taiwan. For our downscaling we have chosen the domain 6°N–18°S, 15°–60°E.

The choice of number of the leading EOFs used in the regression analysis also showed a significant impact on the final downscaling models' skill. When eight leading EOFs were passed through the stepwise-screening process in the regression analysis, the models' skill was generally weaker than when 15 or 20 leading EOFs were used. This was found for all rainfall parameters and for all seasons. It is most noticeable in the DJFMA season and for μ_w values (Fig. 3). In fact, for one station (Dodoma) there is no model selected for μ_w when the leading eight EOFs are used, implying that there was no correlation between these eight spatial fields and the station precipitation. This could be because local rather than large-scale processes, not captured by the first eight leading EOFs, are dominating in this case. Based on the above considerations and the cross-validated test results we decided to use 20 leading EOFs as an initial pool to be used by a stepwise screening.

b. The overall skill score of the models under the common EOF strategy

The algorithms in our EOF analyses are constrained to patterns (EOFs) that are common in both sets of data and their corresponding PCs that are used for model calibration. The skills of the models differ with respect to the particular GCM used (Fig. 4) as each GCM uses specific assumptions and formulations of processes in addition to having different resolution. As a result, GCMs may yield different responses to the same

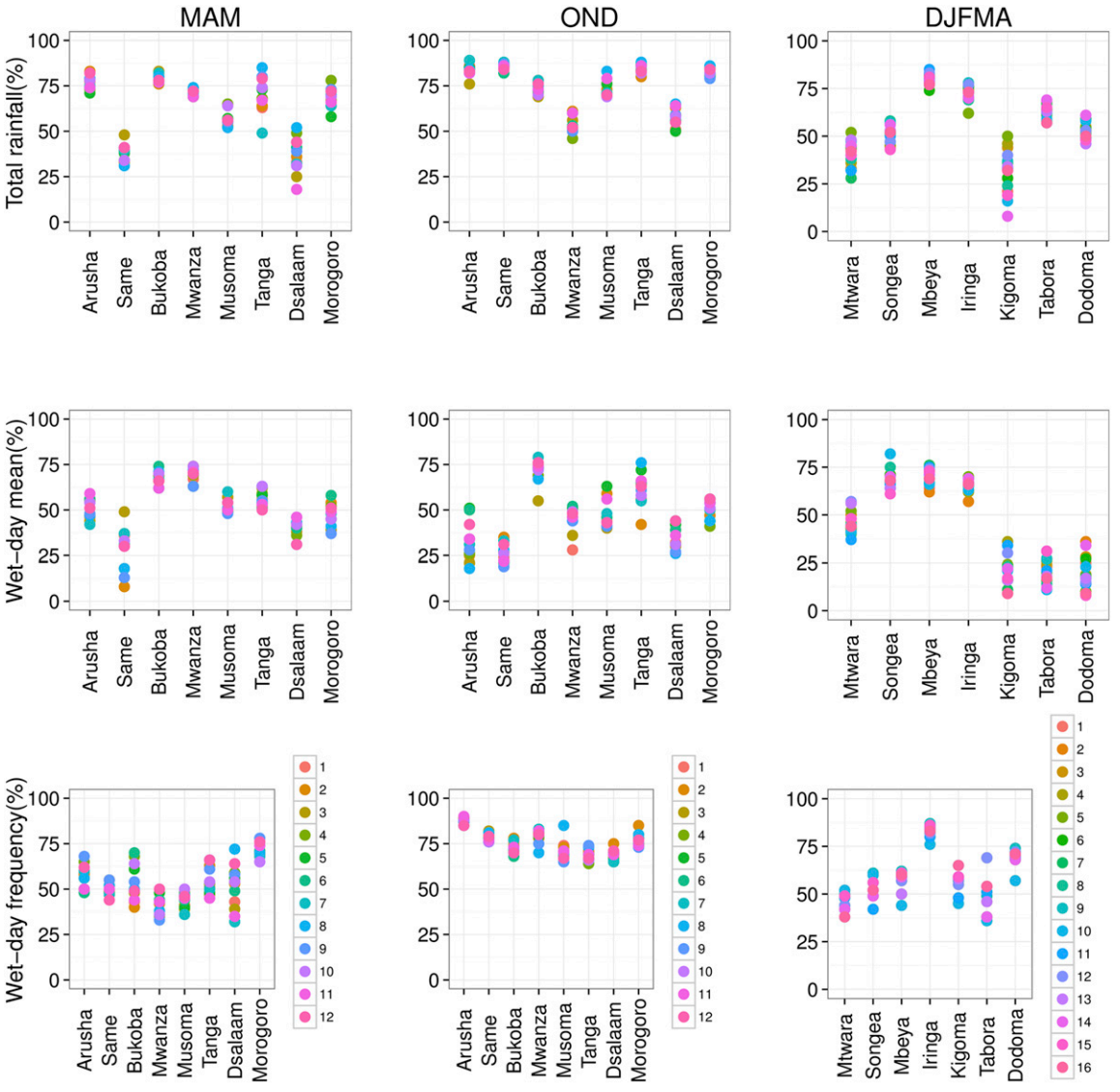

FIG. 4. Plots showing skill scores (R^2) of the downscaling models (calibrated) based on common EOFs of MERRA+GCMs simulations for the period 1979–2012. For each station colors indicate individual model skill. The skills are for three rainfall seasons, (left) MAM, (middle) OND, and (right) DJFMA, and they are presented for three rainfall parameters: (top) total rainfall, (middle) wet-day mean, and (bottom) wet-day frequency. For MAM and OND, colors numbered 1–6 are model skills derived from the common EOFs of MERRA and the GCMs BCC_CSM1.1, CCSM4, GISS-E2-H, HadGEM2-AO, MPI-ESM-LR, and NorESM1-M that were forced with RCP4.5, and numbers 7–12 are model skills obtained from common EOFs of MERRA and same GCMs but forced with RCP8.5. For DJFMA season colors numbered 1–8 are model skills derived from the common EOFs of MERRA and the GCMs BCC_CSM1.1,CCSM4, GISS-E2-H, HadGEM2-AO, MPI-ESM-LR, NorESM1-M, CSIRO Mk3.6.0, and MIROC5 that were forced with RCP4.5 and numbers 9–16 are model skills obtained from common EOFs of MERRA and same GCMs but forced with RCP8.5.

external forcings (anthropogenic and natural) and also represent internal climate variability differently. Generally, the highest skill score (between 70% and 95%) is shown for p_t and f_w for OND; exceptions are Dar es Salaam (Dsalaam in Fig. 4) and Mwanza (p_t) whose skill ranges between 50% and 75%. Skill of 50% and up to 90% is depicted for the MAM and DJFMA seasons for p_t and f_w. In particular Same and Kigoma have lower skill for p_t, and Mwanza has lower skill for f_w. Wet-day mean (μ_w) showed a wide range of skills from very low (25%) to very high (~80%) for all seasons for specific locations.

The high skill obtained in the OND season indicates that the interannual variance for p_t and f_w in the calibration period (1979–2012) is well explained by the fitted models (see, e.g., Fig. 5). The model (fit in the figure)

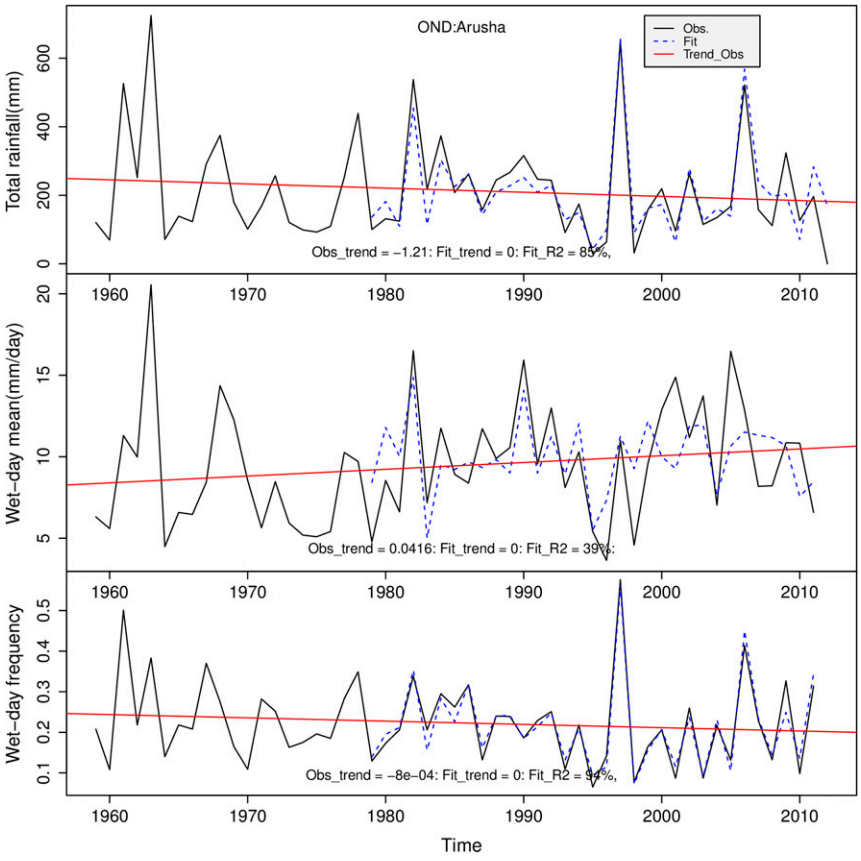

FIG. 5. Time series plot for Arusha station during OND showing simulated (dashed blue) vs observed (black) values for (top) total rainfall, (middle) wet-day mean, and (bottom) wet-day frequency. The red line represents the trend of observed values.

has well reproduced the observed total rainfall (R^2 = 85%) and the wet-day frequency (R^2 = 94%). Year-to-year variability and the maximum values recorded are well represented in the model. Analysis of residuals of regression for the period 1979–2012 (not shown) that were computed using least squares regression analysis depicted a slight linear trend (although not statistically significant) for OND. This was found for all stations and all GCMs used for μ_w, but only for three stations (Arusha, Bukoba, and Mwanza) and for some GCMs for f_w. In accordance with Hanssen-Bauer et al. (2003), such a trend may be due to the fact that models do not satisfactorily reproduce the observed fluctuations in these rainfall parameters.

In some stations, seasons and rainfall parameters models have shown to have very low prediction skills (see, e.g., Fig. 2), but still this cannot be associated with the model overfit; rather, other factors such as inability of MERRA and GCMs to simulate large-scale features are responsible. In this regard, we can generally say that the goodness of the model (skill) and its prediction skill as shown in Fig. 2 is specific to location, season, and to some extent parameter. However, even with the high skill of the downscaling models identified here, we are left with uncertainty in the downscaling of the future due to uncertainties in the GCM calculated precipitation.

c. *Downscaled scenarios*

Results presented here are downscaled multimodel ensemble scenarios for specific locations and seasons. The multimodel mean and the spread in projections from individual models are presented. The reason for this approach is that large wetting or drying may occur to some locations even if the multimodel mean has lost that signal. Moreover, the multimodel mean alone may be misleading especially when used for decision making (e.g., for adaptation purposes where the range of outcomes needs to be taken into account). The multimodel ensemble mean projection for individual stations and seasons in this study is still considered as the most likely outcome. Results are given in Figs. 6–8 for the three rainfall parameters for individual stations, seasons,

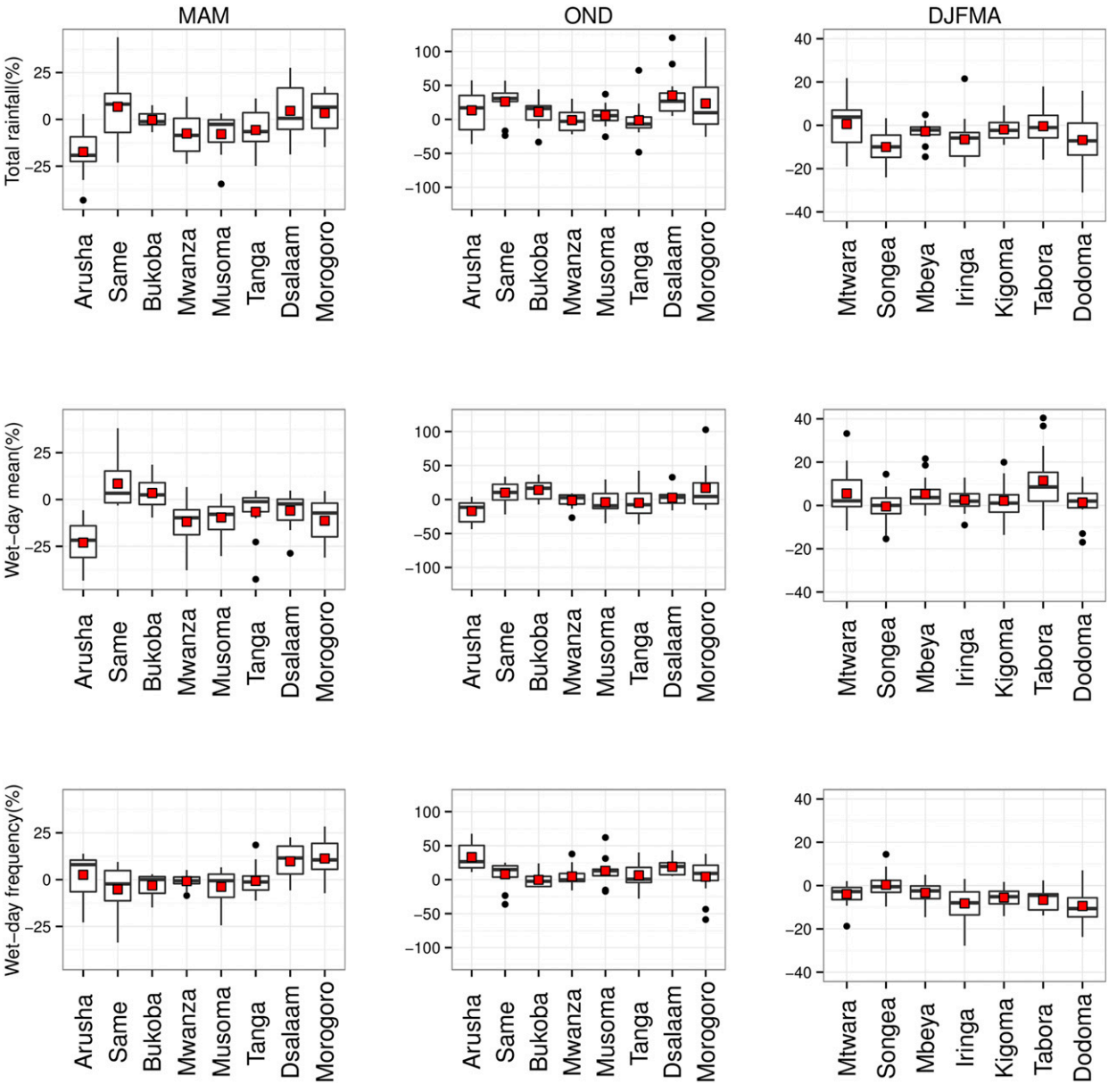

FIG. 6. Box-and-whisker plots showing percentage change of 30-yr seasonal mean of the multimodel ensemble downscaled scenarios for (left) MAM, (middle) OND, and (right) DJFMA for the period 2010–39. Changes are presented for three rainfall parameters: (top) total rainfall, (middle) wet-day mean, and (bottom) wet-day frequency. The boxes represent the interquartile range with the upper and lower hinges of the box corresponding to the first and third quartiles (the 25th and 75th percentiles), the horizontal line in the middle gives the median, and the red squares are the multimodel mean projections. The whiskers extend from the hinge to the highest value that is within 1.5 times the interquartile range of the hinge (box). Data beyond the end of the whiskers are outliers and are plotted as points. Each station presented in MAM and OND has 12 models developed from 6 GCMs (Table 1 excluding MIROC5 and CSIRO Mk3.6.0) and 2 RCPs (i.e., RCP4.5 and RCP8.5), while each station presented in DJFMA season consist of 16 models obtained from the combination of 8 GCMs (Table 1) and 2 RCPs. Note the difference in scales between the seasons in the figure.

and periods, including the variability between models (whisker plots). Their corresponding magnitudes of change are expressed in percentages. Figure 9 gives a picture of the spatial distribution of the multimodel mean projections in all three parameters across the country.

1) PROJECTION FOR 2010–39

The spread in p_t projection is depicted for all seasons and for all stations (Fig. 6) as depicted in the box-and-whisker plots. The largest spread in projection is depicted in OND (notice the difference in scales between

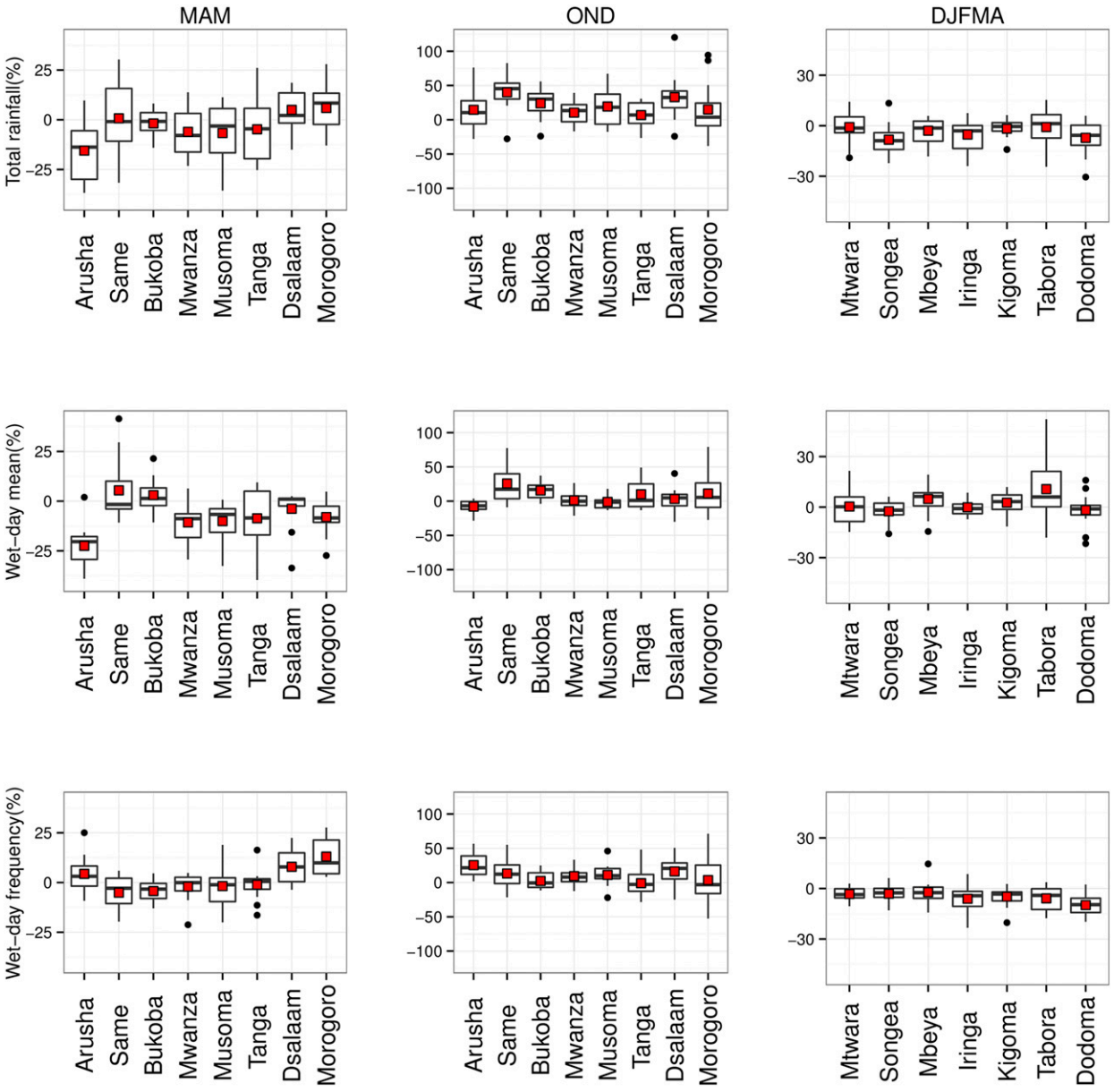

FIG. 7. As in Fig. 6, but for 2040–69.

the seasons in the figure) where individual models project more than 100% increase, namely Dar es Salam, Tanga, and Morogoro. Dar es Salaam is the only location with wetter condition as depicted by all models. This is accompanied by an increase in f_w. The multimodel mean projection shows wetter conditions (p_t) for Dar es Salaam, Same, Morogoro, Arusha, Musoma, and Bukoba during OND and they are accompanied by an increase in f_w only (Musoma and Arusha), μ_w only (Bukoba), or both (Same, Morogoro, and Dar es Salaam)

Ensemble projections of p_t are not uniform for MAM even for stations situated in the same climatic zone (e.g., the northeastern highland, represented by Arusha and Same). Almost all models project a decreasing p_t for Arusha, with a robust signal of drier conditions which is backed up by the projected decrease in μ_w. A majority of models indicate increasing rainfall for Same, Dar es Salaam, and Morogoro.

In the one rainy season (DJFMA) region there are many stations that are projected to get lower total precipitation in the ensemble mean. For most stations the wet-day mean increases, resulting in lower wet-day frequency at all stations. About 50 percentile of models indicate increasing rainfall (p_t) for Mtwara and Tabora.

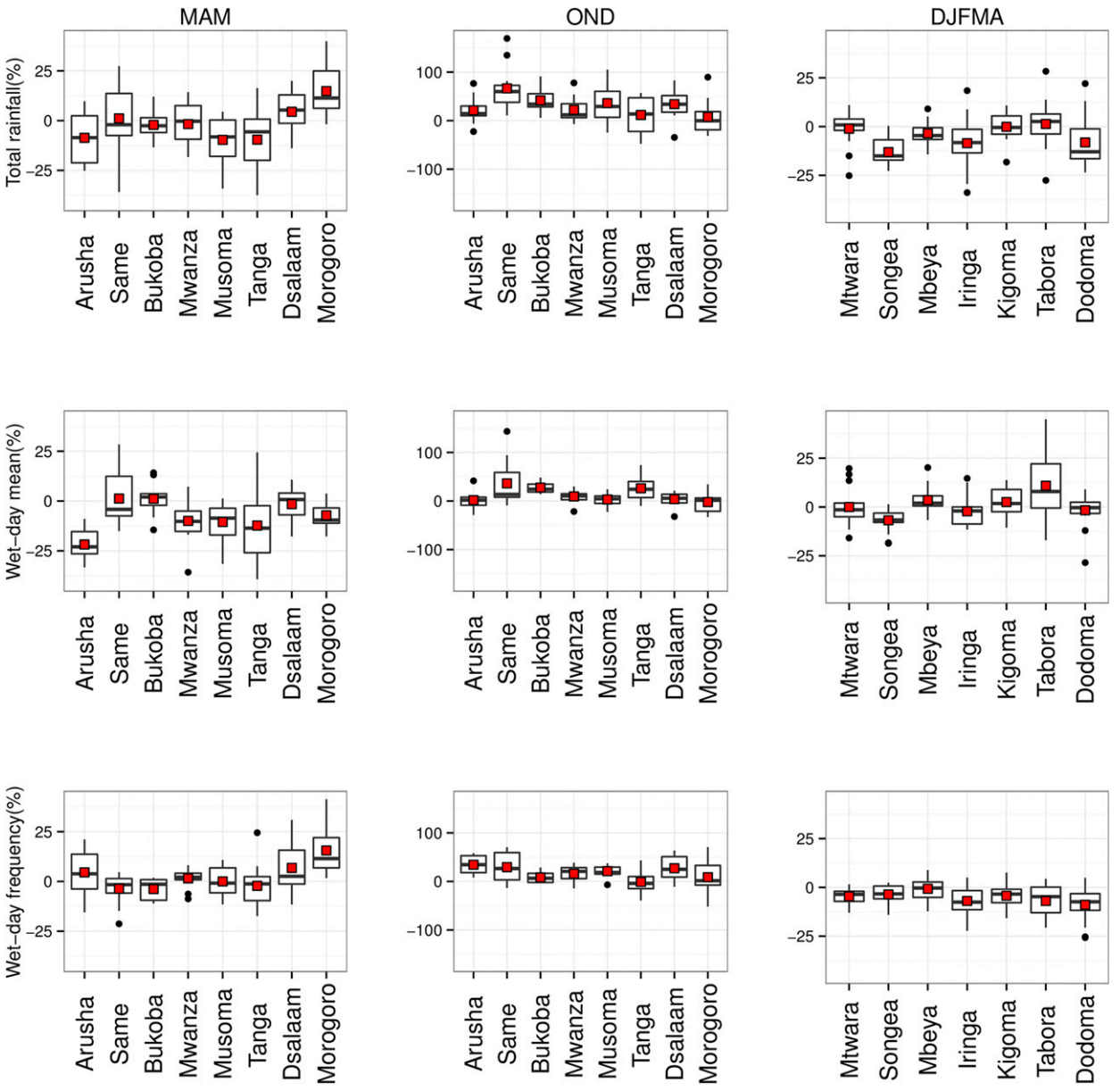

FIG. 8. As in Fig. 6, but for 2070–99.

2) PROJECTION FOR 2040–69

The change for this period (Fig. 7) also shows a large scatter in projected p_t, again in particular during OND. A majority of models point to wetter conditions as does the multimodel mean projection for all stations, both in terms of total precipitation and frequency, but still with an increase in wet-day mean (except for Arusha). In extreme cases, individual models are suggesting up to more than 100% increase of total rainfall for, for example, Dar es Salaam and Morogoro, for the near-term projection.

In the MAM projections there is less, but still substantial, spread in p_t . A majority of models point toward drier conditions for most stations (Arusha, Bukoba, Mwanza, Musoma, and Tanga), accompanied by decrease of both f_w and μ_w. Otherwise increase in p_t is depicted for Dar es Salaam, Morogoro, and Same, and it is accompanied by an increase in f_w. As for near-term projections, a robust signal toward lower total precipitation (p_t) is indicated for Arusha, along with decreasing μ_w.

A majority of models show drier conditions (p_t) for DJFMA for most stations, but the signal of change is weak as compared to other seasons (between 0% and 10%). This

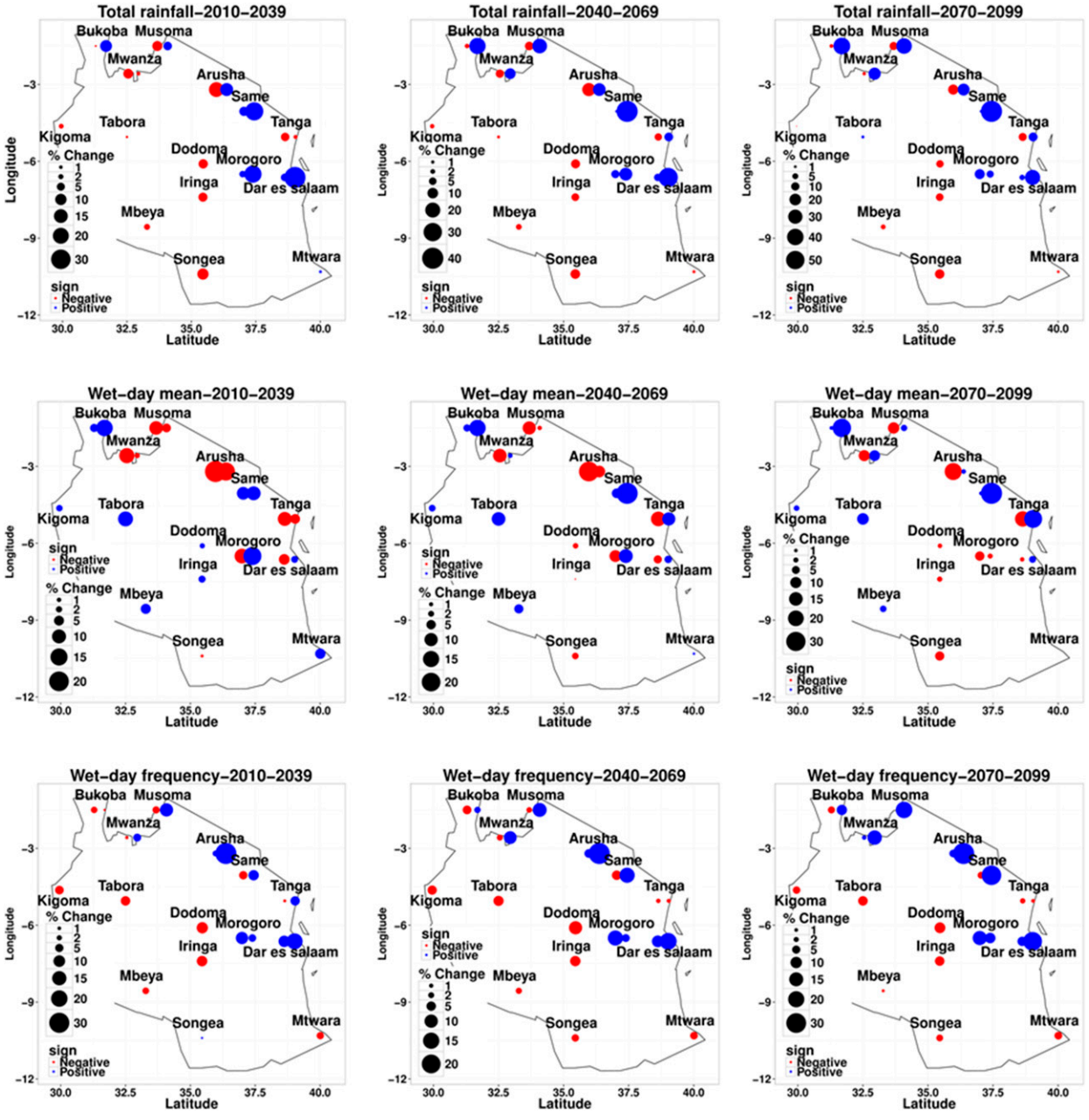

FIG. 9. Multimodel mean for downscaled local rainfall change derived from the CMIP5 GCMs experiments under RCP4.5 and RCP8.5 (combined). Changes are expressed as changes relative to the control period 1979–2012. All models are given equal weight in the calculation of the multimodel mean. Stations with double dots represent areas spanning two rainy seasons—(left) MAM and (right) OND—and stations with a single dot indicate areas with one long rainy season (DJFMA). Note the difference in scale under each map.

change is directly supported by the decrease in f_w. However, about half of the models project wetter (p_t) conditions for Tabora and Mtwara accompanied by increasing μ_w.

3) PROJECTION FOR 2070–99

The change for this period is shown in (Fig. 8). Again, the largest changes are found during OND. In this period there is a clear and robust signal toward wetter conditions (p_t) for most stations, except Morogoro and Tanga where about 50 percentile is below zero. In the ensemble mean, the increase in the total precipitation in this period is due to an increase in both f_w and μ_w. Although there is large scatter in the model projections for MAM, Dar es Salaam and Morogoro and to some extent Same show a clear signal of increasing total rainfall, which is backed up by the increase of f_w (Dar es Salaam

and Morogoro) and μ_w (Same) . While the projection of p_t over two stations (Mwanza and Bukoba) in the lake zone seems to be uncertain (i.e., boxes cross the zero line) the other station (Musoma) shows a clear signal toward a dry condition, mostly due to lower rainfall intensity (μ_w).

For DJFMA, a clear signal of decreasing p_t (represented by all models) is depicted at Songea, a station in the southern part of the country. This is due to lower frequency (f_w) and intensity (μ_w). A majority of models show drier conditions (p_t) for almost all stations, except the western part of the country (Kigoma and Tabora) and Mtwara (south) where the model projections spread through the zero line. However, the signal of change in the multimodel-mean projection is weak as compared to other seasons (between 0% and 20%). The projected dry condition (p_t) is attributed to a decrease in f_w in all stations and accompanied by either increase or decrease in μ_w at individual stations.

For all three periods, a significance test for projected changes in rainfall parameters was performed. It showed that some downscaling models provided a significant change (at 5% significant level; *t* test) especially for the periods 2040–69 and 2070–99 while some did not show any significance. It varied between the GCMs used, RCPs, stations, and downscaled parameters. Therefore, the multimodel projection presented here represents a combination of changes that are statistically significant and those that are not.

4. Discussion

We have found spatial variability in model skill for all seasons and rainfall parameters investigated. The variability is likely due to the fact that different areas have different local factors that modulate the rainfall patterns. The skill scores are in general quite high, but weak skill scores found in some cases may suggest that local processes play an important role. Ideally we would expect the skill of the models in a particular station to be equal for all GCMs used, since models are developed by using observations (reanalysis vs station observations), which are the same for all common EOF setups. However, the difference and spread depicted is probably a reason that the AIC selected variable numbers of predictors and different common EOFs for individual models whose corresponding PCs represent different fluctuations. For demonstration, Table 2 shows the coefficients of EOFs selected for two pairs of neighboring stations situated in different homogeneous climatological regions, the northeastern highlands (Arusha and Same) and northern coast (Tanga and Dar es Salaam). We show results for these stations as they showed a remarkable difference in model skill scores and projections for MAM despite being in the same homogeneous climatological zone. From Table 2 we can see that a few EOFs are selected by AIC that were common for neighboring stations, but the rest are independent (i.e., they vary between the stations). Even with similar EOFs selected, the coefficient values that describe the contribution of each EOF in explaining variability in the local precipitation are also different in magnitude and sign. Another feature depicted in the table is that some models are developed with a large number EOFs and some with a minimal number.

The number and type of EOFs picked by AIC in the regression analysis for a particular station may partly explain the differences in the depicted variable models skill scores, and the weights (regression models coefficients) may explain the remaining part of the variable skill together with the spread in GCM projections. The analysis of the time evolution (PCs) of the spatial modes (EOFs) (Fig. 10) used in the regression demonstrated that 1) the PCs of individual EOFs differs depending on the GCM used (see Fig. 10, first and third panels in both columns) and 2) models consist of PCs that do not share a common trend (see Fig. 10, second and fourth panels in both columns). The PC part of the GCMs' projections underestimates the amplitude and the periodicity of the time evolution of these spatial modes (especially for lower-order EOFs) as compared to the observed period (Fig. 10, first and third panels in the left column). However, the PCs for higher-order EOF (Fig. 10, first and third panels in the right column) are more like noisy fields and their evolution indicates comparable amplitude and the periodicity between the observed (reanalysis) and the GCM part. These noticeable behaviors in the PCs suggest that the spread in projections based on individual GCMs will even change between various projection periods.

We can therefore conclude that the spread and differences in projections can be explained by three factors that arose from the downscaling processes: 1) the number and type of EOFs picked by AIC in the regression analysis, 2) the weights of the regression coefficients, and 3) the trends in the individual PCs. In addition to these factors, the climate system noise and the uncertainty in emission scenarios are expected to cascade to the regional climate projection even with perfect GCMs used for downscaling (Mitchell and Hulme 1999). However, there is no clear evidence that the differences/spread in skill score can explain the differences/spread in projections. As an example, the OND season, which showed a remarkable spread in projections in all three periods, for p_t in particular, has very small differences in skill scores, but the opposite is true for DJFMA.

TABLE 2. Regression coefficient values for the EOFs selected in in the downscaling models using stepwise screening process. The presented values are for total precipitation in MAM and for models forced with RCP4.5. Empty boxes indicate that the EOF was not selected in that model; Dar stands for Dar es Salaam.

	Regression model coefficients							
	CCSM4				NorESM1-M			
EOFs	Tanga	Dar	Same	Arusha	Tanga	Dar	Same	Arusha
EOF1	6.10			3.73	8.73			9.33
EOF2	2.66		4.62			10.63		
EOF3	−3.25		−6.33					
EOF4	−10.32	−3.41	−4.56	−12.77	−3.61			−5.88
EOF5		6.28		5.85	−6.33			
EOF6	17.83	12.69	7.85	7.58				4.72
EOF7	8.96	11.81	9.61	16.54		−8.07	−9.94	
EOF8	17.51		5.68					11.24
EOF9		−9.81	14.31		8.46	−15.19	6.25	
EOF10					−4.70		−5.93	−3.35
EOF11	43.21	17.79	6.31					
EOF12	15.33	10.85	−8.08		38.76		5.16	12.18
EOF13	27.58	8.84		−5.45	29.00			
EOF14	−6.23	22.21		12.06	7.73			
EOF15	−13.71		19.67	7.98	−18.91	−14.15		
EOF16			8.53	13.51	−18.68			
EOF17	−29.31	−32.90		−20.25				
EOF18	34.77	29.50	11.93	12.80	−17.86	−9.77	−7.29	5.95
EOF19	6.71		−10.45	−10.52	28.30		14.42	20.44
EOF20			−19.59	−10.48		24.45	−14.71	−0.01

Further analysis of GCMs' area mean precipitation (Fig. 11 and Table 3) shows that there is no clear evidence that difference in GCMs simulations of rainfall over the downscaling domain can cause a noticeable spread in the rainfall projection, a spread that is especially large for OND for some stations. Table 3 can also be used to explain why there is no remarkable difference in projected changes between the three periods analyzed; the annual mean rainfall simulated by individual GCM over the downscaling domain is nearly the same for different periods of projection.

Regional climates are very complex due to processes that vary significantly with location and therefore respond differently to changes in global- to regional-scale influences (Christensen et al. 2013). Small-scale physical features within the country are expected to enforce a complex pattern of projected rainfall in different climatological zones even without the interaction with large-scale phenomena. Hilly areas with steep slopes, for example, enhance observed mean rainfall, especially during wet seasons (Oettli and Camberlin 2005). In Lake Victoria basin, on the other hand, the climate (rain) is induced by the lake/land breeze and the surrounding topography, which results in uneven distribution of rain over this area, with the western part of the basin receiving more rainfall than the eastern part (Anyah et al. 2006). But with the increased influx from the large-scale flow, a different story unfolds, especially for the eastern part. This phenomenon is consistent with the results from the current study during the period 2070–99 (MAM and OND seasons). There is a remarkable difference between projection of μ_w for Musoma and Mwanza (stations situated to the east and south of the lake respectively) as compared to a counterpart station in the western part of the Lake Victoria basin (Bukoba).

Anomalous low-level westerly flow (OND) and easterly flow (MAM) are projected at the end of twenty-first century (2041–60) (Cook and Vizy 2013). This suggests that even with strong influx of moisture from the Indian Ocean to the eastern part of the country (MAM), more moisture is expected to be transported to the Congo basin and leave the country dry. Moreover, more moisture influx from the Congo basin is expected during OND, when enhanced moisture convergence is expected over the country and hence more rain. Thus, changes in wind flows may partly explain projected wetter conditions for the OND season and decreased total rainfall for a majority of stations receiving rains during MAM as depicted by the majority of models in the current study, and in the period 2040–69 in particular.

A positive phase of the Indian Ocean zonal mode (IOZM) that is associated with excessive OND rains over East Africa is projected by the CMIP5 GCMs (Shongwe et al. 2011). This condition is consistent with

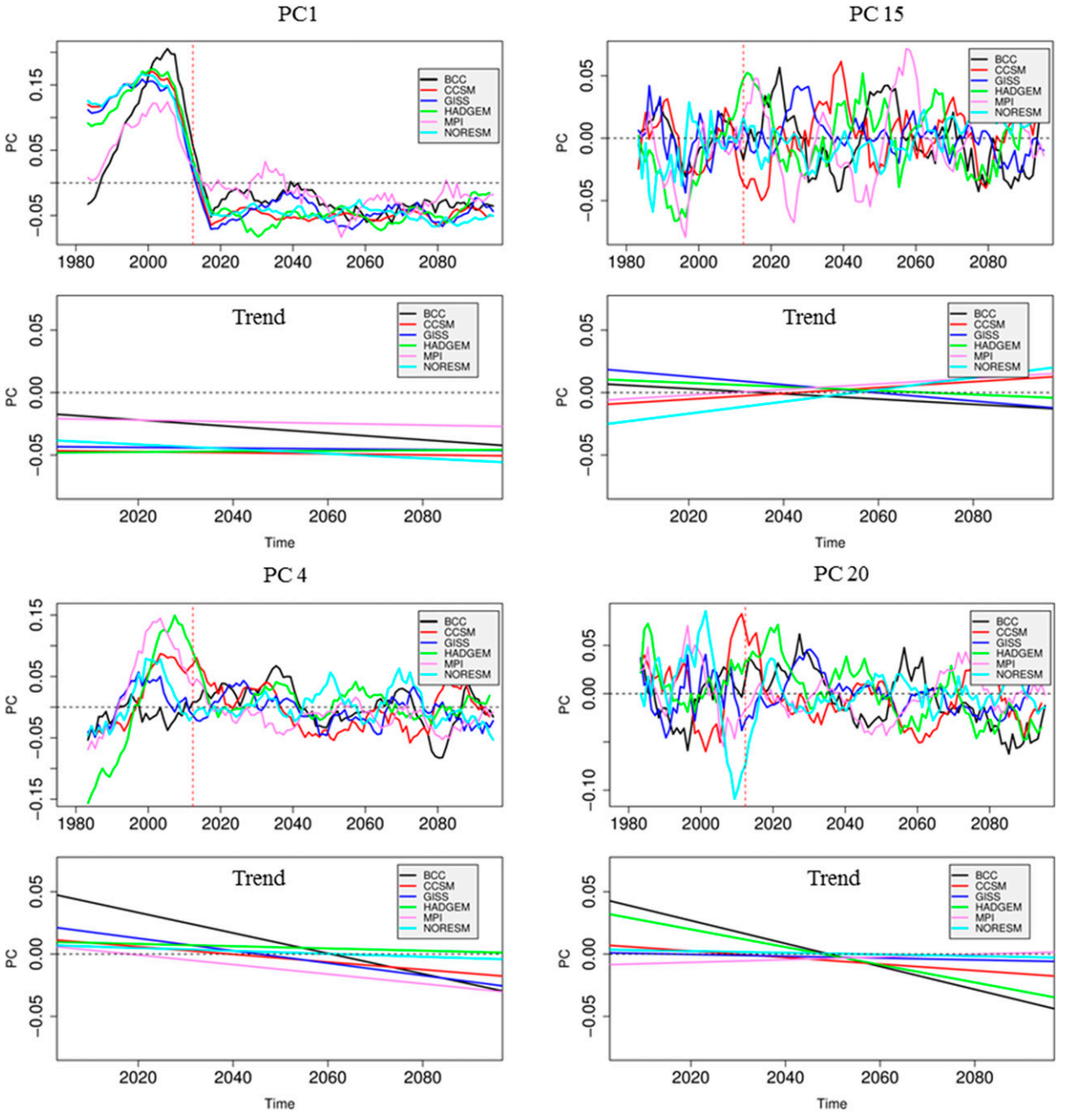

FIG. 10. Temporal evolution (PCs) and their corresponding trends for the leading EOFs used in the downscaling analysis, showing the PCs for the (left) low-order and (right) higher-order EOFs. The PCs are smoothed by using 10-yr running averages. The PCs are for MAM and for RCP4.5. The vertical dotted red line separates observation from the projection parts of the data. The downscaling domain used is 6°N–18°S, 15°–60°E.

the projected OND wetter conditions identified in this study. Global warming, on the other hand, has been directly linked to recent changes in mean rainfall (Giorgi et al. 2011) in relation to either change in f_w or μ_w (intensity) (Benestad 2013). Under global warming, the increase in precipitation intensity is driven by the greater atmospheric moisture holding capacity and the increase (decrease) in mean precipitation is driven by a correspondingly higher (lower) increase in surface evaporation (Giorgi et al. 2011). The response of these rainfall parameters to the global warming explains the thermodynamic effects of the Earth–atmosphere

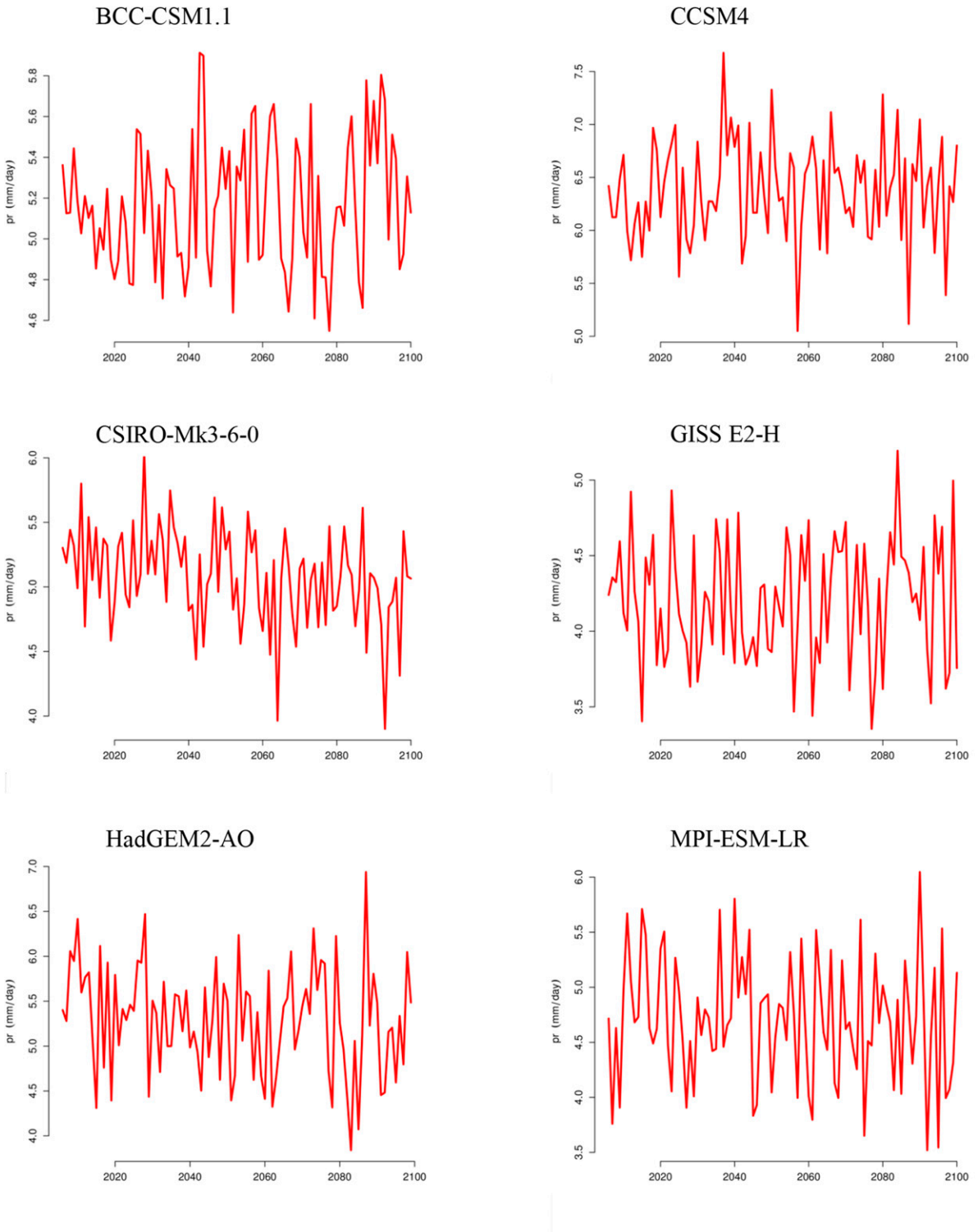

FIG. 11. OND seasonal variation of area mean rainfall (mm day^{-1}) as simulated by different GCMs. The mean is calculated as the average of precipitation values in the downscaling domain (6°N–18°S, 15°–60°E).

TABLE 3. Annual area-mean rainfall simulated by different GCMs. The mean is calculated as the average of precipitation values in the downscaling domain (6°N–18°S, 15°–60°E) and is presented for different periods of projections. Very small difference in simulated rainfall means between the periods can be identified. Some differences are statistically significant (at 5% significant level; *t* test) while some are not.

Model	Area mean total rainfall (mm day^{-1})		
	2010–39	2040–69	2070–99
BCC_CSM1.1.rcp45	3.94	3.97	4.00
BCC_CSM1.1.rcp85	4.05	4.03	4.16
CCSM4.rcp45	4.73	4.75	4.85
CCSM4.rcp85	4.65	4.91	4.83
GISS-E2-H.rcp45	4.21	4.18	4.27
GISS-E2-H.rcp85	4.18	4.16	4.16
HadGEM2-AO.rcp45	3.86	3.84	3.97
HadGEM2-AO.rcp85	3.91	3.91	4.03
MPI-ESM-LR.rcp45	3.47	3.45	3.44
MPI-ESM-LR.rcp85	3.40	3.43	3.61
NorESM1-M.rcp45	4.57	4.46	4.44
NorESM1-M.rcp85	4.54	4.54	4.44
CSIRO Mk3.6.0.rcp45	3.38	3.31	3.33
CSIRO Mk3.6.0.rcp85	3.33	3.31	3.42
MIROC5.rcp45	4.08	4.01	4.02
MIROC5.rcp85	4.14	4.12	4.19

interactions whose impact is considered to be global and to a lesser extent regional.

However, we should keep in mind that not all climatic zones will experience similar changes and thus the current study yields large spatial variability in the projected changes in the rainfall parameters. Thus there is no uniform projection for areas characterized by complex topographical features such as that of Tanzania (see also Taye et al. 2011).

This study compliments previous studies in climate analysis and projection in the region of study making use of different approaches, such as different downscaling methods (statistical or dynamical), predictors, observations (parameters and observation periods), time scale in analysis (e.g., seasonal or annual), and driving GCMs. For example, Akurut et al. (2014) found an increasing annual rainfall over the Lake Victoria basin for middle and end of this century, and these changes are attributed to an increase in μ_w and not f_w. This is not the case as depicted in this study for different areas in the basin on a seasonal basis. In Musoma as an example, increase in p_t for OND is attributed to increase in f_w and not μ_w. In the same area, rainfall is projected to increase by up to 30% for some areas in OND whereas it is expected to decrease by up to 5% in MAM. The study of Shongwe et al. (2011) over East Africa found projected wetter conditions with more intense wet seasons (OND and MAM) and less severe droughts for the period of 2050–2200. Similar results are depicted in the current study; projected wetter conditions (p_t) are predicted due to both an increase in precipitation intensity (μ_w) and an increased frequency of rainy days (f_w) at the end of the century (2070–99), for all stations receiving OND seasonal rains. For a station-specific comparison, the results of the current work are in agreement with that of Tumbo et al. (2010) for the Same station during MAM and OND seasons and the periods 2040–69 and 2070–99.

5. Summary and conclusions

We have investigated possible changes in the seasonal p_t, μ_w, and f_w using regression-based ESD. The analysis is performed for the three periods 2010–39, 2040–69, and 2070–99. Apart from the inherited uncertainty from the GCMs used in the downscaling processes, the design of the empirical models used in this study can be used to explain the differences/spread in models skills and projections in rainfall parameters. With the sample of GCMs used in this study, our analysis of the downscaled climate scenarios suggests that precipitation changes will vary across Tanzania. Spatial differences in projected climate will be experienced as it is portrayed by the observed climates (Camberlin and Philippon 2002).

Surface inhomogeneity in East Africa and Tanzania in particular makes the spatial structures of projected rainfall even more complex, especially in MAM. There is no uniform sign of change depicted by multimodel projection across different areas; therefore, drawing a conclusion of a uniform change for a region with heterogeneous surfaces is not possible. This demands the provision of location-specific climate analysis. In the near-term projection, Dar es Salaam has shown a unique wetter condition (p_t) for OND as depicted by all models. In end-of-century Morogoro (MAM), Bukoba and Same (OND), and Songea (DJFMA) also indicated a robust signal toward wetter (MAM and OND) and drier (DJFMA) conditions. In the remaining cases the likely future condition is determined by the majority of models pointing toward either wetter or drier conditions, although a few models point toward the opposite condition or do not show changes at all. The spread in projections was found to be larger in 2010–39 (e.g., in DJFMA and OND) and reduced slightly in 2040–69 and 2070–99, and it was nearly constant for MAM. A large spread in the period 2010–39 can be associated with the internal fluctuations in the climate system, which are expected to be more pronounced for the near-term projections than the effects of external forcing such as greenhouse gas increases (Collins et al. 2013).

In this study, there is no uniform conclusion regarding the magnitude of change in projected rainfall

parameters (see, e.g., Fig. 9). The projected increase in total rainfall in OND, for example, is due to the increasing frequency of wet days (f_w) (in most stations and for all projection periods) and more intense precipitation events (μ_w) (for some stations). For DJFMA, projected decreasing total precipitation is due to decreasing frequency of wet days (all stations and for all projection periods) and less intense precipitation events (only for some stations and for the periods 2040–69 and 2070–99 in particular). Although the western part of the country (Kigoma, Tabora, and Mbeya) is projected to be drier (p_t), the multimodel ensemble indicates an increase in intense precipitation events. For MAM, with the exception of Dar es Salaam, Morogoro, and Bukoba, the increase (decrease) in total precipitation is accompanied by increasing (decreasing) intense precipitation events (μ_w). In the exceptions mentioned the increase (decrease) in total precipitation is accompanied by increasing (decreasing) frequency of wet days (f_w). The magnitude of change also differs from area to area. As an example, in the northeastern highlands Same shows an increase of about 50% of OND precipitation toward the end of the century, while Arusha shows only an increase of about 20%. Likewise, Same has an increase of about 1% of MAM precipitation toward the end of the century, while Arusha shows a decrease of about 15%.

Increased total rainfall (p_t) as a result of increased frequency of wet days (f_w) alone can be important in maintaining soil moisture and hence good conditions for crops and pastures. From our results this condition is expected to be featured for some areas getting MAM rains (e.g., Dar es Salaam and Morogoro). When increased p_t is a result of increasing μ_w or both μ_w and f_w, growing conditions might be worsened due to short episodes of heavy rainfalls and related floods. Such extreme events are expected to cause a significant challenge to the agriculture sector, the economy, and the food security at the local level (United Republic of Tanzania 2014) and to other sectors sensitive to wet and dry extremes such as water resource management and construction industries (e.g., roads, dams for hydroelectric power). From our results this condition is expected to be featured largely for areas getting OND rains especially in the end-of-century time frame.

Decreasing total rainfall (p_t) as a result of decreasing frequency of wet days (f_w), either alone or in combination with the wet-day mean (μ_w), implies a higher risk of long dry spells or drought. From our results such conditions are expected to be featured in the areas spanning one rainy season (DJFMA) and in some areas getting MAM rains (e.g., Bukoba and Tanga for all projection periods and Musoma and Mwanza in the near term and at midcentury). However, a few areas getting DJFMA season rains (e.g., Kigoma, Tabora, and Mbeya) may experience some episodes of heavy and damaging rains and as a result of projected increasing heavy rainfall events (μ_w) during this season. The related impacts may include damage to crops, which in turn will affect the growth and harvests. This consequence may cause famine outbreaks and hence threaten the livelihoods of many.

Acknowledgments. This work was done under the support of the Climate Change Impacts Adaptation and Mitigation (CCIAM) project funded through the Norwegian Embassy in Tanzania. The authors wish to acknowledge the Tanzania Meteorological Agency (TMA) for providing rainfall data, the Global Modeling and Assimilation Office (GMAO), and the GES DISC for the dissemination of MERRA data. We also wish to acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP5, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. The authors also acknowledge use of the Ferret program for data processing in this paper. Ferret is a product of NOAA's Pacific Marine Environmental Laboratory. Information is available at http://ferret.pmel.noaa.gov/Ferret/.

REFERENCES

Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Autom. Control*, **19**, 716–723, doi:10.1109/TAC.1974.1100705.

Akurut, M., P. Willems, and C. Niwagaba, 2014: Potential impacts of climate change on precipitation over Lake Victoria, East Africa, in the 21st century. *Water*, **6**, 2634–2659, doi:10.3390/w6092634.

Ambrosino, C., R. E. Chandler, and M. C. Todd, 2011: Southern African monthly rainfall variability: An analysis based on generalized linear models. *J. Climate*, **24**, 4600–4617, doi:10.1175/2010JCLI3924.1.

Anyah, R. O., F. H. Semazzi, and L. Xie, 2006: Simulated physical mechanisms associated with climate variability over Lake Victoria Basin in East Africa. *Mon. Wea. Rev.*, **134**, 3588–3609, doi:10.1175/MWR3266.1.

Baek, H.-J., and Coauthors, 2013: Climate change in the 21st century simulated by HadGEM2-AO under representative concentration pathways. *Asia-Pac. J. Atmos. Sci.*, **49**, 603–618, doi:10.1007/s13143-013-0053-7.

Barnett, T. P., 1999: Comparison of near-surface air temperature variability in 11 coupled global climate models. *J. Climate*, **12**, 511–518, doi:10.1175/1520-0442(1999)012<0511:CONSAT>2.0.CO;2.

Benestad, R. E., 2001a: A comparison between two empirical downscaling strategies. *Int. J. Climatol.*, **21**, 1645–1668, doi:10.1002/joc.703.

——, 2001b: The cause of warming over Norway in the ECHAM4/OPYC3 GHG integration. *Int. J. Climatol.*, **21**, 371–387, doi:10.1002/joc.603.

——, 2002: Empirically downscaled multimodel ensemble temperature and precipitation scenarios for Norway. *J. Climate*, **15**, 3008–3027, doi:10.1175/1520-0442(2002)015<3008:EDMETA>2.0.CO;2.

——, 2004: Empirical-statistical downscaling in climate modeling. *Eos, Trans. Amer. Geophys. Union*, **85**, 417–422, doi:10.1029/2004EO420002.

——, 2007: Novel methods for inferring future changes in extreme rainfall over northern Europe. *Climate Res.*, **34**, 195–210, doi:10.3354/cr00693.

——, 2013: Are there persistent physical atmospheric responses to galactic cosmic rays? *Environ. Res. Lett.*, **8**, 035049, doi:10.1088/1748-9326/8/3/035049.

——, I. Hanssen-Bauer, and D. Chen, 2008: *Empirical-Statistical Downscaling.* World Scientific, 228 pp.

Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation of the physical climate. *Geosci. Model Dev.*, **6**, 687–720, doi:10.5194/gmd-6-687-2013.

Black, E., J. Slingo, and K. R. Sperber, 2003: An observational study of the relationship between excessively strong short rains in coastal East Africa and Indian Ocean SST. *Mon. Wea. Rev.*, **131**, 74–94, doi:10.1175/1520-0493(2003)131<0074:AOSOTR>2.0.CO;2.

Busuioc, A., D. Chen, and C. Hellström, 2001: Performance of statistical downscaling models in GCM validation and regional climate change estimates: Application for Swedish precipitation. *Int. J. Climatol.*, **21**, 557–578, doi:10.1002/joc.624.

Camberlin, P., and N. Philippon, 2002: The East African March–May rainy season: Associated atmospheric dynamics and predictability over the 1968–97 period. *J. Climate*, **15**, 1002–1019, doi:10.1175/1520-0442(2002)015<1002:TEAMMR>2.0.CO;2.

——, and R. E. Okoola, 2003: The onset and cessation of the ''long rains'' in eastern Africa and their interannual variability. *Theor. Appl. Climatol.*, **75**, 43–54, 10.1007/s00704-002-0721-5.

Chen, D. L., C. Achberger, J. Räisänen, and C. Hellström, 2006: Using statistical downscaling to quantify the GCM-related uncertainty in regional climate change scenarios: A case study of Swedish precipitation. *Adv. Atmos. Sci.*, **23**, 54–60, doi:10.1007/s00376-006-0006-5.

Christensen, J. H., and Coauthors, 2013: Climate phenomena and their relevance for future regional climate change. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1217–1308.

Collins, M., and Coauthors, 2013: Long-term climate change: Projections, commitments and irreversibility. *Climate Change 2013: The Physical Science Basis*, T. F. Stocker et al., Eds., Cambridge University Press, 1029–1136.

Cook, K. H., and E. K. Vizy, 2013: Projected changes in East African rainy seasons. *J. Climate*, **26**, 5931–5948, doi:10.1175/JCLI-D-12-00455.1.

Cooper, P. J. M., J. Dimes, K. P. C. Rao, B. Shapiro, B. Shiferaw, and S. Twomlow, 2008: Coping better with current climatic variability in the rain-fed farming systems of sub-Saharan Africa: An essential first step in adapting to future climate change? *Agric. Ecosyst. Environ.*, **126**, 24–35, doi:10.1016/j.agee.2008.01.007.

Diro, G. T., A. M. Tompkins, and X. Bi, 2012: Dynamical downscaling of ECMWF Ensemble seasonal forecasts over East Africa with RegCM3. *J. Geophys. Res.*, **117**, D16103, doi:10.1029/2011JD016997.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Review linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.*, **27**, 1547–1578, doi:10.1002/joc.1556.

Giorgetta, M. A., and Coauthors, 2013: Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *J. Adv. Model. Earth Syst.*, **5**, 572–597, doi:10.1002/jame.20038.

Giorgi, F., E. S. Im, E. Coppola, N. S. Diffenbaugh, X. J. Gao, L. Mariotti, and Y. Shi, 2011: Higher hydroclimatic intensity with global warming. *J. Climate*, **24**, 5309–5324, doi:10.1175/2011JCLI3979.1.

Greene, A. M., M. Hellmuth, and T. Lumsden, 2012: Stochastic decadal climate simulations for the Berg and Breede Water Management Areas, Western Cape province, South Africa. *Water Resour. Res.*, **48**, W06504, doi:10.1029/2011WR011152.

Hanssen-Bauer, I., E. J. Førland, J. E. Haugen, and O. E. Tveito, 2003: Temperature and precipitation scenarios for Norway: Comparison of results from dynamical and empirical downscaling. *Climate Res.*, **25**, 15–27, doi:10.3354/cr025015.

Hewitson, B. C., and R. G. Crane, 2006: Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa. *Int. J. Climatol.*, **26**, 1315–1337, doi:10.1002/joc.1314.

Huebener, H., and M. Kerschgens, 2007: Downscaling of current and future rainfall climatologies for southern Morocco. Part II: Climate change signals. *Int. J. Climatol.*, **27**, 1065–1073, doi:10.1002/joc.1457.

Hulme, M., R. Doherty, T. Ngara, and M. New, 2005: Global warming and African climate change: A reassessment. *Climate Change and Africa*, P. S. Low, Ed., Cambridge University Press, 29–40.

Indeje, M., and F. H. M. Semazzi, 2000: Relationships between QBO in the lower equatorial stratospheric zonal winds and East African seasonal rainfall. *Meteor. Atmos. Phys.*, **73**, 227–244, doi:10.1007/s007030050075.

——, ——, and L. J. Ogallo, 2000: ENSO signals in East African rainfall seasons. *Int. J. Climatol.*, **20**, 19–46, doi:10.1002/(SICI)1097-0088(200001)20:1<19::AID-JOC449>3.0.CO;2-0.

Jones, R. G., M. Noguer, D. C. Hassell, D. Hudson, S. S. Wilson, G. J. Jenkins, and J. F. B. Mitchell, 2004: Generating high resolution climate change scenarios using PRECIS. Met Office Hadley Centre, 40 pp.

Kabanda, T. A., and M. R. Jury, 1999: Inter-annual variability of short rains over northern Tanzania. *Climate Res.*, **13**, 231–241, doi:10.3354/cr013231.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.*, **77**, 437–471, doi:10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2.

Kijazi, A. L., and C. J. C. Reason, 2005: Relationships between intraseasonal rainfall variability of coastal Tanzania and ENSO. *Theor. Appl. Climatol.*, **82**, 153–176, doi:10.1007/s00704-005-0129-0.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-Year Reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247–267, doi:10.1175/1520-0477(2001)082<0247:TNNYRM>2.3.CO;2.

Landman, W. A., and W. J. Tennant, 2000: Statistical downscaling of monthly forecasts. *Int. J. Climatol.*, **20**, 1521–1532, doi:10.1002/1097-0088(20001115)20:13<1521::AID-JOC558>3.0.CO;2-N.

——, D. DeWitt, D. E. Lee, A. Beraki, and D. Lotter, 2012: Seasonal rainfall prediction skill over South Africa: One- versus two-tiered forecasting systems. *Wea. Forecasting*, **27**, 489–501, doi:10.1175/WAF-D-11-00078.1.

Leung, L. R., L. O. Mearns, F. Giorgi, and R. L. Wilby, 2003: Regional climate research: Needs and opportunities. *Bull. Amer. Meteor. Soc.*, **84**, 89–95, doi:10.1175/BAMS-84-1-89.

Li, Z., W.-Z. Liu, X.-C. Zhang, and F.-L. Zheng, 2011: Assessing the site-specific impacts of climate change on hydrology, soil erosion and crop yields in the Loess Plateau of China. *Climatic Change*, **105**, 223–242, doi:10.1007/s10584-010-9875-9.

Linderson, M. L., C. Achberger, and D. L. Chen, 2004: Statistical downscaling and scenario construction of precipitation in Scania, southern Sweden. *Nord. Hydrol.*, **35**, 261–278.

Lorenz, E. N., 1956: Empirical orthogonal functions and statistical weather prediction. Scientific report 1, Air Force Research Laboratories, 52 pp. [Available online at http://www.o3d.org/abracco/Atlantic/Lorenz1956.pdf.]

Mapande, A. T., and C. J. C. Reason, 2005: Interannual rainfall variability over western Tanzania. *Int. J. Climatol.*, **25**, 1355–1368, doi:10.1002/joc.1193.

Marchant, R., C. Mumbi, S. Behera, and T. Yamagata, 2007: The Indian Ocean dipole—The unsung driver of climatic variability in East Africa. *Afr. J. Ecol.*, **45**, 4–16, doi:10.1111/j.1365-2028.2006.00707.x.

Meehl, G. A., and Coauthors, 2012: Climate system response to external forcings and climate change projections in CCSM4. *J. Climate*, **25**, 3661–3683, doi:10.1175/JCLI-D-11-00240.1.

Mitchell, T. D., and M. Hulme, 1999: Predicting regional climate change: Living with uncertainty. *Prog. Phys. Geogr.*, **23**, 57–78, doi:10.1177/030913339902300103.

Moore, N., and Coauthors, 2012: East African food security as influenced by future climate change and land use change at local to regional scales. *Climatic Change*, **110**, 823–844, doi:10.1007/s10584-011-0116-7.

Moriondo, M., C. Giannakopoulos, and M. Bindi, 2011: Climate change impact assessment: The role of climate extremes in crop yield simulation. *Climatic Change*, **104**, 679–701, doi:10.1007/s10584-010-9871-0.

Moss, R. H., and Coauthors, 2010: The next generation of scenarios for climate change research and assessment. *Nature*, **463**, 747–756, doi:10.1038/nature08823.

Mpeta, E. J., and M. R. Jury, 2001: Intra-seasonal convective structure and evolution over tropical East Africa. *Climate Res.*, **17**, 83–92, doi:10.3354/cr017083.

Mutai, C. C., and M. N. Ward, 2000: East African rainfall and the tropical circulation/convection on intraseasonal to interannual timescales. *J. Climate*, **13**, 3915–3939, doi:10.1175/1520-0442(2000)013<3915:EARATT>2.0.CO;2.

Ng'ongolo, H. K., and S. P. Smyshlyaev, 2010: The statistical prediction of East African rainfalls using quasi-biennial oscillation phases information. *Nat. Sci.*, **2**, 1407–1416, doi:10.4236/ns.2010.212172.

Oettli, P., and P. Camberlin, 2005: Influence of topography on monthly rainfall distribution over East Africa. *Climate Res.*, **28**, 199–212, doi:10.3354/cr028199.

Ogallo, L. J., 1988: Relationships between seasonal rainfall in East Africa and the Southern Oscillation. *J. Climatol.*, **8**, 31–43, doi:10.1002/joc.3370080104.

Omamo, S. W., X. Diao, S. Wood, J. Chamberlain, L. You, S. Benin, U. Wood-Sichra, and A. Tatwangire, A. 2006: Strategic priorities for agricultural development in eastern and central Africa. International Food Policy Research Institute (IFPRI) Rep. 150, 154 pp. [Available online at http://www.ifpri.org/sites/default/files/publications/rr150.pdf.]

Omondi, P. A., and Coauthors, 2014: Changes in temperature and precipitation extremes over the Greater Horn of Africa region from 1961 to 2010. *Int. J. Climatol.*, **34**, 1262–1277, doi:10.1002/joc.3763.

Paeth, H., K. Born, R. Podzun, and D. Jacob, 2005: Regional dynamical downscaling over West Africa: Model evaluation and comparison of wet and dry years. *Meteor. Z.*, **14**, 349–367, doi:10.1127/0941-2948/2005/0038.

Penlap, E. K., C. Matulla, H. von Storch, and F. M. Kamga, 2004: Downscaling of GCM scenarios to assess precipitation changes in the little rainy season (March–June) in Cameroon. *Climate Res.*, **26**, 85–96, doi:10.3354/cr026085.

Rienecker, M. M., and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.

Rotstayn, L. D., S. J. Jeffrey, M. A. Collier, S. M. Dravitzki, A. C. Hirst, J. I. Syktus, and K. K. Wong, 2012: Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmos. Chem. Phys.*, **12**, 6377–6404, doi:10.5194/acp-12-6377-2012.

Rowhani, P., D. B. Lobell, M. Linderman, and N. Ramankutty, 2011: Climate variability and crop production in Tanzania. *Agric. For. Meteor.*, **151**, 449–460, doi:10.1016/j.agrformet.2010.12.002.

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.

Schmidt, G. A., and Coauthors, 2014: Configuration and assessment of the GISS ModelE2 contributions to the CMIP5 archive. *J. Adv. Model. Earth Syst.*, **6**, 141–184, doi:10.1002/2013MS000265.

Schreck, C. J., and F. H. M. Semazzi, 2004: Variability of the recent climate of eastern Africa. *Int. J. Climatol.*, **24**, 681–701, doi:10.1002/joc.1019.

Shongwe, M. E., G. J. van Oldenborgh, B. van den Hurk, and M. van Aalst, 2011: Projected changes in mean and extreme precipitation in Africa under global warming. Part II: East Africa. *J. Climate*, **24**, 3718–3733, doi:10.1175/2010JCLI2883.1.

Subramanian, A. C., M. Jochum, A. J. Miller, R. Murtugudde, R. B. Neale, and D. E. Waliser, 2011: The Madden–Julian oscillation in CCSM4. *J. Climate*, **24**, 6261–6282, doi:10.1175/JCLI-D-11-00031.1.

Taye, M. T., V. Ntegeka, N. P. Ogiramoi, and P. Willems, 2011: Assessment of climate change impact on hydrological extremes in two source regions of the Nile River Basin. *Hydrol. Earth Syst. Sci.*, **15**, 209–222, doi:10.5194/hess-15-209-2011.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the Experiment Design. *Bull. Amer. Meteor. Soc.*, **93**, 485–498, doi:10.1175/BAMS-D-11-00094.1.

Tumbo, S. D., E. Mpeta, M. Tadross, F. C. Kahimba, B. P. Mbillinyi, and H. F. Mahoo, 2010: Application of self-organizing maps technique in downscaling GCMs' climate change projections for Same, Tanzania. *Phys. Chem. Earth*, **35**, 608–617, doi:10.1016/j.pce.2010.07.023.

United Republic of Tanzania, 2014: Agriculture climate resilience plan, 2014–2019. Ministry of Agriculture Food Security and Cooperatives, 83 pp. [Available online at http://www.kilimo.go.tz/publications/english%20docs/ACRP_TANZANIA_ENDORSED.pdf.]

Waithaka, M., G. C. Nelson, T. S. Thomas, and M. Kyotalimye, 2013: East African agriculture and climate change: A comprehensive analysis. IFPRI Issue Brief 79, 4 pp. [Available online at http://www.ifpri.org/sites/default/files/publications/ib76.pdf.]

Watanabe, M., and Coauthors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312–6335, doi:10.1175/2010JCLI3679.1.

Wickham, H., 2009: *ggplot2: Elegant Graphics for Data Analysis.* Springer, 213 pp.

Wilby, R. L., and T. M. L. Wigley, 2000: Precipitation predictors for downscaling: Observed and general circulation model relationships. *Int. J. Climatol.*, **20**, 641–661, doi:10.1002/(SICI)1097-0088(200005)20:6<641::AID-JOC501>3.0.CO;2-1.

——, H. Hassan, and K. Hanaki, 1998: Statistical downscaling of hydrometeorological variables using general circulation model output. *J. Hydrol.*, **205**, 1–19, doi:10.1016/S0022-1694(97)00130-3.

Willems, P., K. Arnbjerg-Nielsen, J. Olsson, and V. T. V. Nguyen, 2012: Climate change impact assessment on urban rainfall extremes and urban drainage: Methods and shortcomings. *Atmos. Res.*, **103**, 106–118, doi:10.1016/j.atmosres.2011.04.003.

Xin, X.-G., T.-W. Wu, J.-L. Li, Z.-Z. Wang, W.-P. Li, and F.-H. Wu, 2013: How well does BCC_CSM1.1 reproduce the 20th century climate change over China? *Atmos. Oceanic Sci. Lett.*, **5**, 21–26.

Yang, T., C. Kuo, C. Liu, H. Tseng, and P. Yu, 2013: Effects of domain selection on singular-value-decomposition based statistical downscaling of monthly rainfall accumulation in southern Taiwan. *Terr. Atmos. Ocean. Sci.*, **24**, 369–381, doi:10.3319/TAO.2013.01.09.01(A).

Zhao, Y., P. Camberlin, and Y. Richard, 2005: Validation of a coupled GCM and projection of summer rainfall change over South Africa, using a statistical downscaling method. *Climate Res.*, **28**, 109–122, doi:10.3354/cr028109.