

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Papers in Natural Resources

Natural Resources, School of

2017

Enhancing Dynamical Seasonal Predictions through Objective Regionalization

Saleh Satti

The Johns Hopkins University, Baltimore, ssatti1@jhu.edu

Benjamin F. Zaitchik

The Johns Hopkins University, Baltimore

Hamaba S. Babr

The Johns Hopkins University, Baltimore

Tsegaye Tadesse

University of Nebraska-Lincoln, ttadesse2@unl.edu

Follow this and additional works at: <https://digitalcommons.unl.edu/natrespapers>



Part of the [Natural Resources and Conservation Commons](#), [Natural Resources Management and Policy Commons](#), and the [Other Environmental Sciences Commons](#)

Satti, Saleh; Zaitchik, Benjamin F.; Babr, Hamaba S.; and Tadesse, Tsegaye, "Enhancing Dynamical Seasonal Predictions through Objective Regionalization" (2017). *Papers in Natural Resources*. 1156. <https://digitalcommons.unl.edu/natrespapers/1156>

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Papers in Natural Resources by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Enhancing Dynamical Seasonal Predictions through Objective Regionalization

SALEH SATTI, BENJAMIN F. ZAITCHIK, AND HAMADA S. BADR

Department of Earth and Planetary Science, The Johns Hopkins University, Baltimore, Maryland

TSEGAYE TADESSE

School of Natural Resources, University of Nebraska–Lincoln, Lincoln, Nebraska

(Manuscript received 2 June 2016, in final form 16 February 2017)

ABSTRACT

Improving seasonal forecasts in East Africa has great implications for food security and water resources planning in the region. Dynamically based seasonal forecast systems have much to contribute to this effort, as they have demonstrated ability to represent and, to some extent, predict large-scale atmospheric dynamics that drive interannual rainfall variability in East Africa. However, these global models often exhibit spatial biases in their placement of rainfall and rainfall anomalies within the region, which limits their direct applicability to forecast-based decision-making. This paper introduces a method that uses objective climate regionalization to improve the utility of dynamically based forecast-system predictions for East Africa. By breaking up the study area into regions that are homogenous in interannual precipitation variability, it is shown that models sometimes capture drivers of variability but misplace precipitation anomalies. These errors are evident in the pattern of homogenous regions in forecast systems relative to observation, indicating that forecasts can more meaningfully be applied at the scale of the analogous homogeneous climate region than as a direct forecast of the local grid cell. This regionalization approach was tested during the July–September (JAS) rain months, and results show an improvement in the predictions from version 4.5 of the Max Planck Institute for Meteorology’s atmosphere–ocean general circulation model (ECHAM4.5) for applicable areas of East Africa for the two test cases presented.

1. Introduction

East Africa (EA) is notoriously vulnerable to hydroclimatic extremes. Severe drought in the early 1980s affected large swaths of EA, resulting in crop failures that led to large migrations and widespread starvation. An estimated 16 million people were affected in Ethiopia and Sudan alone (Olsson 1993; FAO 2000). More recently, from 2011 to 2012 drought exacerbated food insecurity and left 8.8 million people in need of urgent humanitarian assistance. An estimated \$1.3 billion was requested for a humanitarian response [United Nations Office for the Coordination of Humanitarian Affairs (UN OCHR) 2011]. This drought affected multiple sectors, from agriculture and livestock to health and hygiene, and led to multiple countries declaring this drought a national disaster (UN OCHR 2011). The occurrence of multiple hydroclimatic extremes that have impacted the lives of many within EA highlights the importance

of understanding and improving seasonal forecasting in the region.

The generation of reliable forecasts at seasonal time scales has, however, proven to be a complex and elusive problem. In general, seasonal forecasts have presented a significant challenge relative to shorter-term weather forecasts. Over the past 30 yr, weather forecast skill has improved dramatically, in large part because of improved estimates of initial atmospheric conditions provided by satellite-derived observations and enhanced in situ observations (Goddard et al. 2001). Predictions on longer time scales (i.e., seasonal and interannual climate) do not benefit from these improved observations of initial atmospheric conditions, as the memory of the atmosphere is not adequate to inform forecasts beyond one or two weeks. Instead, dynamical forecasts on these longer-time-horizon forecasts rely on the initial state of climate system components that have longer memory (e.g., sea surface temperature, and soil and vegetation conditions on land) and on realistic simulation of gradually evolving atmospheric circulations and

Corresponding author: Saleh Satti, ssatti1@jhu.edu

DOI: 10.1175/JAMC-D-16-0192.1

© 2017 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](http://www.ametsoc.org/PUBSReuseLicenses) (www.ametsoc.org/PUBSReuseLicenses).

surface states. Seasonal forecasts have improved, at least in some regions across the globe, as observations of these memory components of the climate system have improved and as forecast systems have gone to higher resolution, more complete physics, and more advanced data assimilation algorithms (Goddard et al. 2001). In EA, however, the forecasting challenge is particularly acute on account of complex synoptic and meso-scale conditions, nonlinear interactions between large-scale climate modes, and subseasonal variability in teleconnections and precipitation processes (Nicholson 2000).

The evaluation of seasonal forecast skill is a challenge in its own right. The seasonal forecasting community utilizes several methods of skill scores in order to gauge the accuracy of these different methods (Goddard et al. 2001). Model skill is determined by a retrospective model evaluation, where model results are compared with observational data. Because of inherent biases and errors within the dynamical models there is a need for statistical processing of model outputs. One method for rectifying systematic model errors is to represent the forecast outputs as a percentage of ensemble forecasts that lie within an assigned category. Traditionally, these categorical forecast outputs are evaluated using categorical evaluation metrics such as rank probability skill score (RPSS), likelihood skill score, and generalized relative operating characteristics (Barnston et al. 2010).

A second important decision in forecast evaluation, but one that generally receives less attention, is the spatial basis applied when evaluating a model. Because dynamical forecasts produce gridded output, it is common practice to extract predictions at a specific location from the closest model grid cell. A potentially more forgiving approach is to evaluate models at a coarse regional scale for a box or geographical unit of interest. These grid-to-grid (GtG) and box area averaging methods are currently being used by the climate forecasting community to form seasonal forecast predictions (Jury 2014; Batté and Déqué 2011; Barnston et al. 2010).

Both the GtG and box area averaging, however, do not adequately account for spatial biases. GtG unduly penalizes the model for small spatial inaccuracies even when the overall forecast anomalies are correct. General area averaging implicitly assumes spatial matching between model and observations and also can introduce error by combining regions that have different responses to large-scale drivers. Several researchers are attempting to address this issue. Koster et al. (2008), for example, apply observed spatial correlation structures to translate model-generated forecasts' skill from locations of high skill to locations of low skill. This

transformation approach is shown to improve forecast accuracy.

Other research in precipitation prediction has illustrated the importance of isolating regions of similar variance through objective regionalization techniques in order to adequately describe the nature of large-scale influence on the area of interest (Dezfuli and Nicholson 2013). This method of regionalization divides areas into smaller homogenous regions based on the variance of a particular variable. Camberlin and Philippon (2002) use principal component analysis (PCA) in order to analyze the regional and seasonal structure of their interannual precipitation variability across EA. Performing a PCA allowed the region to be divided into two subregions with contrasting variability: Ethiopia to the northwest and Uganda–Kenya to the southeast. Other studies such as Tsidu (2012) regionalize based on self-organizing maps, separating Ethiopia into nine homogenous regions. More recent studies have also attempt to separate EA into different areas before performing seasonal predictions. Nicholson (2014) shows two relatively distinct areas within EA by delineating based on the seasonal cycle of precipitation. The first region has rainfall peaking in the July–September (JAS) months and covers Sudan and northwest Ethiopia whereas the “equatorial” region covers the horn of Africa and has its peak rainfall in March–May (MAM) and in October–December (OND).

Dynamics of the East African JAS rains

Local topography, regional winds, and large-scale drivers greatly influence precipitation variability in EA. Many studies have presented in-depth analyses of the various mechanisms that drive variability, often with the intention of improving predictability (Conway 2000; Camberlin and Philippon 2002; Gissila et al. 2004; Segele and Lamb 2005; Block and Rajagopalan 2007; Diro et al. 2011). Berhane et al. (2014) performed a broad study of the various teleconnections that influence Ethiopian highland precipitation during the JAS months. They found that teleconnection strength of various large-scale drivers varies during the June–September rainy season, with the latter months generally showing stronger associations with large-scale modes of variability in the Pacific and Indian Oceans. In the early rainy season, teleconnections are generally weaker, but there is a tendency toward associations with variability to the west, including the Atlantic Ocean, rather than the Pacific and Indian Oceans to the east.

This lack of large-scale driver consistency in precipitation throughout the rainy season presents challenges of physical process and timing of influence to

dynamical model predictions. The influence of multiple mechanisms within a similar area adds to the complexity of accurately predicting seasonal precipitation using dynamical models. Global-dynamics-related pressure systems in the Atlantic, Mediterranean Sea, and Pacific; propagating waves associated with the subseasonal Madden–Julian oscillation; and mesoscale winds responding to both remote and local variability have all been shown to influence precipitation (e.g., [Nicholson 1996](#); [Berhane and Zaitchik 2014](#)). The inability to properly capture one or more of these processes can lead to inaccurate prediction in the amount and location of seasonal precipitation.

In this study, we examine the performance of global dynamically based seasonal forecast systems in EA. In contrast to other studies, we begin with an objective regionalization of EA based on interannual precipitation variability (the primary target of seasonal forecasts) for each month of the rainy season. The regionalization is performed on observations and, independently, on each forecast system. The purposes of this study are 1) to distinguish regions that have distinct patterns of variability (presumably, differing sensitivities to large-scale climate modes), and 2) to identify systematic differences between the regionalization of observation and models, which would indicate the presence of spatial biases in the modeling systems. Once these biases are identified, it is possible to adjust for them through evaluation based on analogous region matching (ARM) in place of standard spatial match assumption (SMA) methods like grid-to-grid or box averaging. In adjusting for spatial biases, the ARM method evaluates models on the basis of their own spatial structures of variability, providing the possibility of drawing useful predictions even from a model with significant spatial biases.

2. Methods

a. Data and models

The extent of the analysis region spans from 25°N to 12°S and from 20°E to 54°E. Observed precipitation data used in this analysis were from version 2 of the Climate Hazards Infrared Precipitation with Station data (CHIRPS) at $0.05^\circ \times 0.05^\circ$ resolution, for the period from 1981 to 2010 ([Funk et al. 2015](#)). Observed SST anomalies used to identify teleconnections were extracted from the Kaplan Extended SST, version 2, dataset, which is produced at $5^\circ \times 5^\circ$ resolution ([Reynolds and Smith 1994](#); [Parker et al. 1994](#); [Kaplan et al. 1998](#)). These data were obtained from the NOAA/OAR/ESRL Physical Science Division (<https://www.esrl.noaa.gov/psd/>).

Two models were used in this analysis: the Climate Forecast System, version 2 (CFSv2), and version 4.5 of the Max Plank Institute for Meteorology's atmosphere–ocean general circulation model (ECHAM4.5). Both precipitation and SST model data were extracted from the North American Multimodel Ensemble (NMME) hindcast monthly dataset ([Kirtman et al. 2014](#)) distributed via the International Research Institute (IRI) data library. Data were available at $1^\circ \times 1^\circ$ resolution for the period from 1982 to 2010. CFSv2 was initialized 24 times to produce 24 different realizations for each separate month, while ECHAM4.5 was initialized 12 times, producing 12 realizations. These two models were selected from NMME simply as examples for the regionalization method; there was no a priori reason for choosing these models over others, though both are leading forecast systems that have been applied in previous studies of the region ([Jury 2014](#)).

b. Regionalization

Regionalization is the division of a large area into smaller regions based on the characteristics of a specific variable or set of variables. The basis of any objective regionalization is a statistical clustering algorithm that defines regions on the basis of internal homogeneity and/or metrics of difference from other clusters. Numerous algorithms are in use for climate studies ([Badr et al. 2015](#)). In this application, we apply Ward's minimum variance method because of its widespread use and its tendency to generate regions with high internal homogeneity. Ward's method develops a hierarchical approach that aims to optimize the union of similar groups while minimizing the sum of squared deviations from the group mean ([Ward 1963](#)). The method clusters data points with variance lower than an allotted threshold value and aggregates the points in order to maximize the correlation of the dataset points within a designated region. Regions that are homogenous with respect to interannual precipitation variability are expected to be relatively uniform in their response to large-scale variability and therefore serve as a good target for seasonal prediction ([Camberlin and Philippon 2002](#); [Nicholson 2014](#)). We apply Ward's method using the Hierarchical Climate Regionalization (HiClimR) package for R ([Badr et al. 2014](#)) described in [Badr et al. \(2015\)](#). HiClimR includes a range of agglomerative hierarchical clustering methods and provides pre- and postprocessing tools relevant for climate applications. Equation (1) is the Lance–Williams ([Lance and Williams 1967](#)) updated formula of Ward's method used to update the dissimilarities in agglomerative clustering. The clustering methodology, as described by [Murtagh and Legendre \(2014\)](#), measures the dissimilarity of a cluster ($i \cup j$)

relative to another cluster k , based upon coefficients (n_i , n_j , and n_k), determined via clustering rules. The HiClimR package uses the Ward1 algorithm (Murtagh 1985) where the Lance–Williams formula is written in terms of squared dissimilarities. The total error sum of squares at each merging step is used to determine a node height in the output dendrogram plot:

$$d(i \cup j, k) = (n_i + n_k)d(i, k) + (n_j + n_k)d(j, k) - n_k d(i, j). \quad (1)$$

For this application, preprocessing was performed to mask noise and to focus the analysis on areas in which JAS is the primary rainy season. Some areas within the limits of the project area do not experience a rainy season in JAS, but rather have a biannual rainy season in MAM and OND. These grid cells were masked because JAS precipitation is not of primary importance for seasonal forecasts in these areas.

Preprocessing was performed in four steps: First, in order to analyze precipitation trends in the JAS season, only data points that experience a significant increase in precipitation for those months were selected for regionalization. Points that registered a more than 7% increase in the monthly average precipitation for the months of JAS relative to all other months were used. This 7% threshold value is subjective; a 10% increase masks large parts of the EA region, while a 5% increase retains data points in locations where the MAM and OND rains dominate.

Second, any data points in the desert within the project area that receive less than 200 mm of rainfall annually were discarded. This is done to prevent an anomalous rainfall event from affecting the regionalization process. Third, the spatial resolution of the CHIRPS dataset was reduced from $0.05^\circ \times 0.05^\circ$ to $1^\circ \times 1^\circ$ to be consistent with the resolution of the seasonal forecasting models used in this analysis. This reduction also reduces the noise level within the observational dataset. Fourth, principal component analysis was applied to remove noise from the dataset. The first three principal components were retained. These four steps improved the homogeneity of the regionalization and made the regions created more statistically robust. The basic steps in preprocessing—identification of the primary rainy season, masking low precipitation areas, averaging to common spatial resolution, and analyzing lead principal components—are commonly applied in a range of climate studies. Their application to regionalization is facilitated by the fact that they are available as standard options in the HiClimR package (Badr et al. 2015). Selection of the percent threshold, the total annual rainfall used to mask the data, and the number of

principal components is subjective. Here we selected the appropriate values through visual judgment as well as multiple sensitivity runs.

Regionalization for both observations and models was performed on a relatively short 28-yr record (1982–2010), which was limited by the availability of seasonal forecast data. This short period of analysis may not be adequate to capture the different phases of the decadal forcings that affect EA precipitation. For example, MAM precipitation has been shown to decrease over this time period (Williams and Funk 2011; Lyon and DeWitt 2012; Yang et al. 2014). To ensure that regionalization results were not dominated by outliers (which could be error in the observed data) we performed regionalization 28 times, leaving one year out in each iteration. This leave-one-out repetition had little impact on forecast system regionalization, which was relatively smooth and consistent, but we did see variability in the CHIRPS-based regionalization. For the final regionalization we combine all 28 regionalizations and assign each grid cell to its most frequently assigned region. Additionally, the robustness of the regions was tested by iteratively removing 2 or 3 consecutive years from the analysis. The final regions attained from these 2- and 3-yr holdouts were consistent with the regions derived when 1 year was excluded.

c. SMA evaluation

For the first evaluation, we adopt the standard practice of evaluating model performance without adjusting for spatial model biases. This standard approach makes an SMA—that each grid cell in the model should predict the collocated grid cell in observation (GtG). Since we are interested in evaluating regional averages, we apply this SMA method at the scale of CHIRPS regions: both observed and forecast precipitation are aggregated using the CHIRPS regions, and model skill is assessed on this scale.

d. ARM evaluation

For our second method of evaluation, we relax the spatial match assumption by evaluating forecast predictions on the basis of their own regionalization rather than the CHIRPS regionalization. The motivation for this approach is the recognition that GtG differences between observed and model regions are partially due to erroneous placement of climate phenomenon captured by the model. Often, the model will capture the predictive phenomenon of interest but misplace the precipitation anomaly, in which case regionalization reveals the spatial bias of the model and can serve as a basis for making predictions based on the relevant similarities between model and observation. Equations (2a) and

(2b) show the calculated precipitation anomaly in year t for observed region obs_r , where P is the precipitation anomaly at grid cell i , and i is an index for spatial location contained in a specified region r ($i \in r$). For the ARM method, predictions are made for an analogous region mod_s that is spatially removed from obs_r . Therefore, M is not equal to N , and the spatial index j covers different grid cells from the spatial index i . For SMA it is assumed that the model's spatial regions s for which the forecast is made exactly match the observed spatial regions r . Therefore, $M = N$, and i and j are identical:

$$obs_r^t = \frac{\sum_{i \in r} P_{i \in r}^t}{N} \quad \text{and} \quad (2a)$$

$$mod_s^t = \frac{\sum_{j \in s} P_{j \in s}^t}{M}. \quad (2b)$$

e. Model skill assessment

The model's predictions for each year in the 1982–2010 hindcast archive are ranked and placed into terciles. For each year, the fraction of model realizations that falls within the lowest one-third of all realizations in the full study period is denoted as the probability of a below-normal forecast. Similarly, the fractions of forecasts that fall within the second and third terciles are placed in the normal and above-normal terciles, respectively. For example, each month's prediction in ECHAM4.5 consists of 12 realizations. Each of these realizations is ranked against the full population of 336 realizations for the study period (12 realizations for the remaining 28 yr). The fraction of the 12 realizations that fall within the first third of the ranked realizations (have values in the range of the driest 112 realizations) becomes the probability of a below-average forecast. Forecasts are demarcated into terciles of below-average, average, and above-average probability forecasts to represent the fraction of realizations that fall within the dry, middle, and wet thirds of total ranked realizations, respectively. Each year's forecast tercile probabilities are calculated using both the SMA and the ARM evaluation methods. We assess both methods by comparing their respective forecasts with observations. For this we utilize the rank probability score (RPS) for category forecasts.

f. RPS

RPS assigns a squared error based on the accuracy of the forecast. The value of the RPS depends on the value of forecast and whether the observation occurs at the category [Eq. (3); Wilks 2011]. The term F_i denotes the forecast probability, Ob_i denotes the probability of

the observation, n is the category (1, 2, or 3), and I is the total number of categories. The value of Ob_i can be either 0 or 1, and thus an event either occurs or does not occur in that category. A high RPS indicates a forecast of low accuracy:

$$RPS = \sum_{n=1}^I \left(\sum_{i=1}^n F_i - \sum_{i=1}^n Ob_i \right). \quad (3)$$

The RPS depends on proximity of the forecast probabilities to the actual observation. A forecast with a high probability two categories away from the observation will have a higher RPS than a forecast with a high probability one category away. Therefore, a forecast can perform worse than a scenario with no prior information [a climatological forecast with no prior information will assign a probability of 0.333 across all terciles (Barnston et al. 2010)].

Comparison of the forecasts with a scenario containing no prior information can be determined by the RPSS. The RPSS depends on the average RPS over all the forecasting years (RPS_{av}) and the average RPS with no prior information (RPS_{clim}):

$$RPSS = 1 - \frac{RPS_{av}}{RPS_{clim}}. \quad (4)$$

The RPSS varies from negative values to 1, with 1 being a perfect forecast and a negative value indicating that the climatological forecast RPS_{clim} outperforms RPS_{av} .

3. Results and discussion

a. Regionalization

Climate regionalization algorithms provide objective metrics that serve as a basis for dividing a large region of interest into coherent subregions. The final decision on the optimal number of regions, however, is context dependent: there is a trade-off between increasing intra-regional homogeneity (which we want to maximize) but increasing interregional correlation (which we want to minimize) as one moves from defining a few small regions to defining highly granular regions. This trade-off is evident in the dendrograms shown in Figs. 1a–c. The height of the dendrogram is a measure of the merging cost (lower is better). As one moves from top to bottom on the dendrogram, the homogeneity of regions increases but the correlation between regions also increases. For the purposes of this study we are interested in relatively large regions that have low interregional correlation and are therefore likely to represent differing sensitivities to large-scale climate variability on a

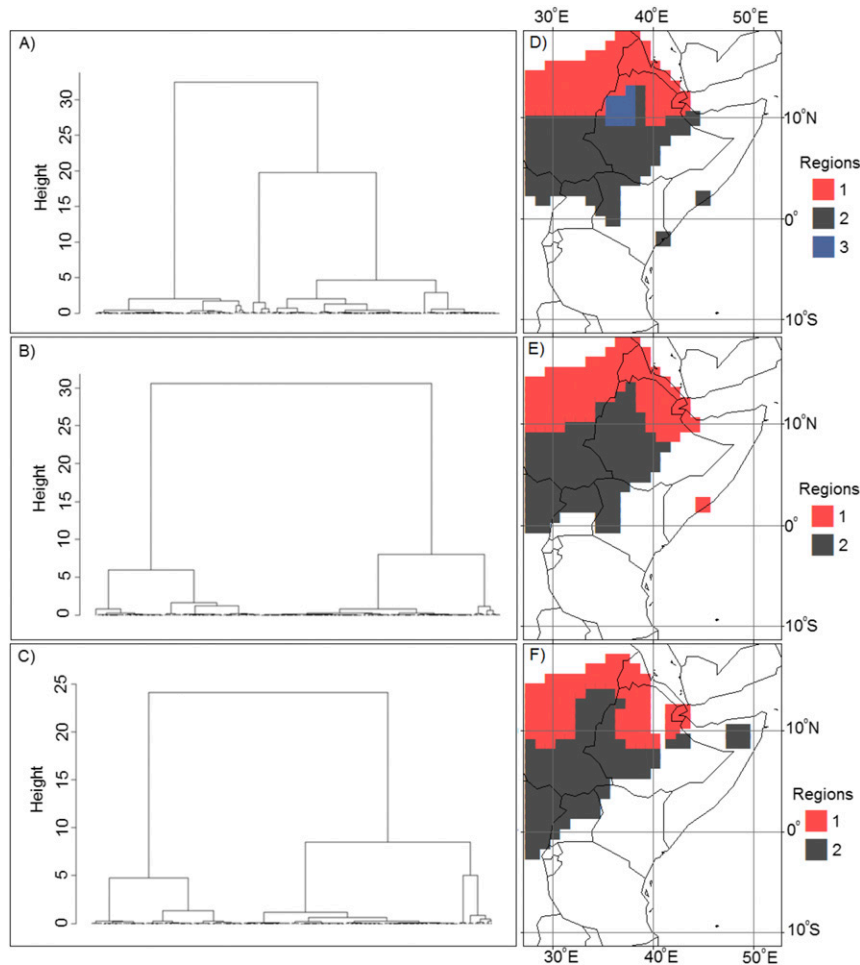


FIG. 1. CFSv2 regionalization dendrograms for (a) July, (b) August, and (c) September. Also shown are homogenous regions created using CFSv2 precipitation forecasts for (d) July, (e) August, and (f) September.

scale that GCM-based forecast systems are likely to resolve. Figure 1 shows the application of regionalization to CFSv2. Applying a threshold value that provides an acceptable balance between intra- and interregional correlation yields the maps shown in Figs. 1d–f, with three regions in July and two regions in August and September.

Figure 2 shows the same regionalization process applied to ECHAM4.5. Differences between CFSv2 and ECHAM4.5 are immediately visible: for ECHAM4.5 the regionalization statistics point to three distinct regions in July, August, and September. The spatial pattern of these regions is quite distinct from CFSv2 regions, as ECHAM4.5 tends toward an east versus west division in the southern portion of the regionalized area (ECHAM4.5 region 2 versus region 3), which is not evident in CFSv2. The extent of the ECHAM4.5 regions are also quite different from CFSv2, in large part

because ECHAM4.5 puts more rain in the eastern Horn of Africa than CFSv2 does in this season, such that ECHAM4.5 passes our precipitation threshold tests.

Regionalization based on CHIRPS precipitation observations shows more noise than the model-based regionalizations (Fig. 3). This is to be expected, since models typically smooth variability. But the magnitude of spatial heterogeneity seen in the CHIRPS regionalization is quite high (especially in August), indicative of the highly localized variability and/or challenge in measurement known to exist in the East African highlands. We choose to retain spatial discontinuities in the regionalization, in part because we use regions as a first step in a model evaluation process, rather than as an end in their own right, and in part because the heterogeneous nature of East Africa poses a challenge in distinguishing between discontinuous regions that are noise versus regions that are not.

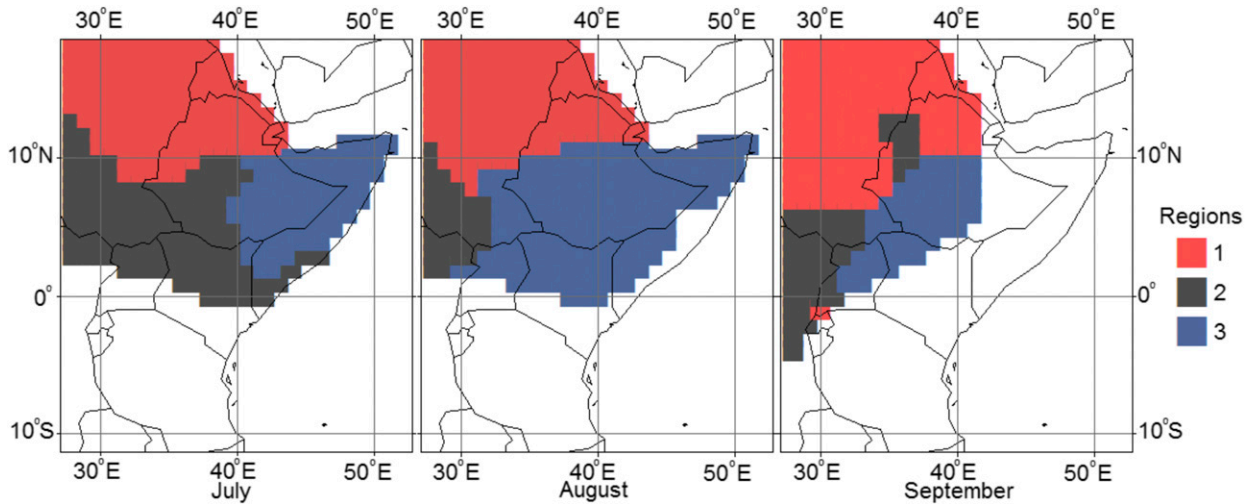


FIG. 2. Regions created using ECHAM4.5 forecasts.

Nevertheless, the regions do generally divide into a northern region (region 1) and a southeastern region (region 3), with a third region that moves between months, but lies in the southwest in August and September (region 2).

Table 1 provides a statistical summary of all regions shown in Figs. 1–3 in terms of intraregional correlation and interregional correlation. These statistics demonstrate the trade-offs inherent in picking regions. For example, CFSv2 regions 1 and 2 show high interregional correlation in all months and could potentially be combined into a single region. Doing so, however, would result in a heterogeneous region that might include areas that have differing response to the large-scale dynamics captured by the model. For all three datasets (CHIRPS,

ECHAM4.5, and CFSv2), in all months the intraregional correlation for all regions exceeds the interregional correlation between any regions.

Regionalization applied to CFSv2 and ECHAM4.5 (see Figs. 2 and 3) yields three regions for both models in the month of July. In August and September ECHAM has three separable regions while CFSv2 has only two. Correlations within regions and between regions show large intraregional correlations and low interregional correlation for ECHAM4.5, consistent with homogeneous regions (Table 1).

b. SMA and ARM

SMA model evaluation is consistent with commonly used evaluation and application techniques. It is simpler

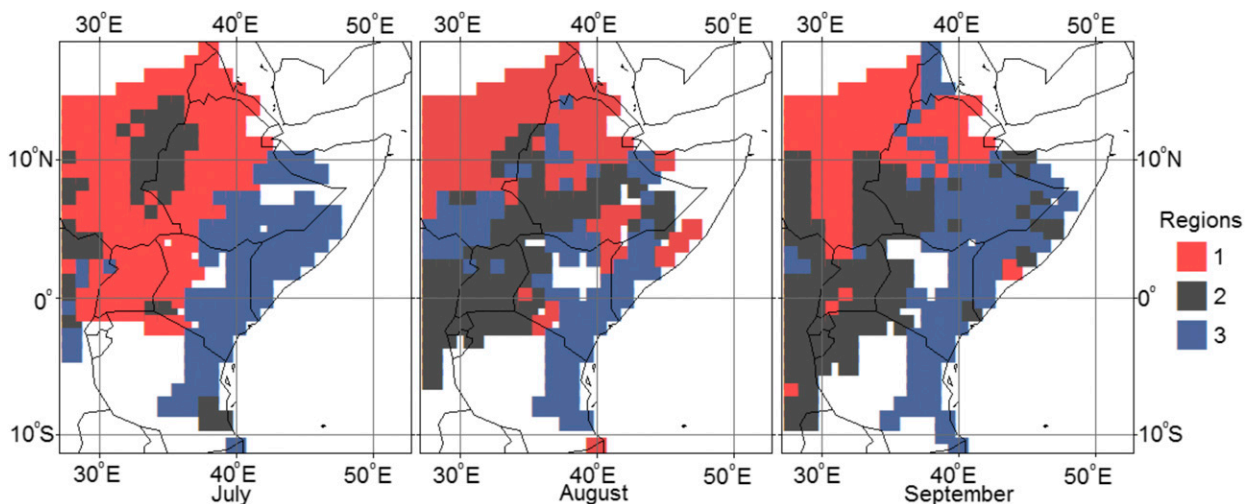


FIG. 3. Regions delineated using HiClimR and the CHIRPS dataset.

TABLE 1. Interregional and intraregional correlations for all of the regions in Figs. 1–3.

	CHIRPS	CFSv2	ECHAM4.5
July interregional			
1 and 2	0.44	0.63	0.13
1 and 3	−0.33	−0.23	0.27
2 and 3	0.12	−0.6	0.2
July intraregional			
1	0.65	0.84	0.67
2	0.5	0.9	0.87
3	0.57	0.79	0.84
August interregional			
1 and 2	0.35	0.66	−0.32
1 and 3	0.23	—	0.07
2 and 3	0.1	—	−0.16
August intraregional			
1	0.59	0.88	0.86
2	0.42	0.87	0.89
3	0.63	—	0.79
September interregional			
1 and 2	0.12	0.67	0.23
1 and 3	0.19	—	0
2 and 3	0.05	—	−0.13
September intraregional			
1	0.52	0.86	0.65
2	0.47	0.88	0.84
3	0.45	—	0.93

than ARM to implement, as it does not require that each model be regionalized, and easier to explain. For these reasons SMA is a preferable approach provided that model and observation show reasonably similar spatial patterns of variability.

To determine when this condition applies, we calculate correlations between the CHIRPS mean time series for each CHIRPS-defined region in each month and the CFSv2 and ECHAM4.5 mean time series for each region defined for those models (Table 2). These values can then be compared with the region maps presented in Figs. 1–3. Whenever there is significant correlation between regions in Table 2 that are associated with

geographically similar areas in Figs. 1–3, we conclude that SMA is a reasonable approach for evaluating model performance in that region. For example, the August CHIRPS region 1 (Fig. 3) is spatially similar to August CFSv2 region 1 (Fig. 1), and the two show statistically significant correlation, so we conclude that SMA is adequate for evaluating CFSv2 in CHIRPS region 1 for that month—CFSv2 is properly localizing the drivers of precipitation variability. Unfortunately, this approach is not satisfied in all scenarios. For example, there is extremely high correlation between July CHIRPS region 1 and July ECHAM4.5 region 3 (0.65), but the two have almost no spatial overlap (Fig. 4). Less than 5% of CHIRPS region 1 overlaps with ECHAM4.5 region 3. August is similar, with less than 26% of CHIRPS region 1 falling within ECHAM4.5 region 3.

These correlations between spatially mismatched regions suggest that ECHAM4.5 does capture a large-scale driver of precipitation variability for East Africa, but the model does not localize this phenomenon in the correct area within East Africa. For these situations, SMA is not an appropriate approach for model evaluation or application, as it fails to recognize potential value in the forecast—the correlation between CHIRPS region 1 and ECHAM4.5 region 3 would be entirely lost. To capture this phenomenon, we apply ARM for any case where there is less than a third (33.3%) overlap between the most highly correlated CHIRPS and model regions.

Table 3 shows the result of this analysis for both models in all months. There are some regions for which models fail to show significant correlation with observations regardless of whether SMA or ARM is applied. For several other combinations, however, ARM identifies significant correlations where SMA does not, suggesting that applying the ARM method could produce skillful predictions for areas where traditional SMA approaches fail to identify any significant predictive skill. Indeed, for ECHAM4.5 we find that ARM is the

TABLE 2. Correlations of the time series for each of the CHIRPS regions with the CFSv2 and ECHAM4.5 regions. Boldface values indicate positive correlations of greater than 0.35.

Month	CHIRPS	CFSv2			ECHAM4.5		
		Region 1	Region 2	Region 3	Region 1	Region 2	Region 3
July	Region 1	0.38	0.38	−0.39	0	0.38	0.65
	Region 2	0.22	0.03	0.03	0.18	0.09	0.11
	Region 3	0.07	0.13	0.05	0.01	0.13	0.2
August	Region 1	0.39	0.18	—	−0.27	−0.25	0.35
	Region 2	0.24	0.25	—	0.13	−0.01	0.33
	Region 3	0.27	0.16	—	−0.1	−0.25	0.13
September	Region 1	0.41	0.13	—	−0.07	−0.31	0.09
	Region 2	0.26	0.28	—	−0.28	−0.33	0.11
	Region 3	0.19	0.15	—	0.1	0.05	0.15

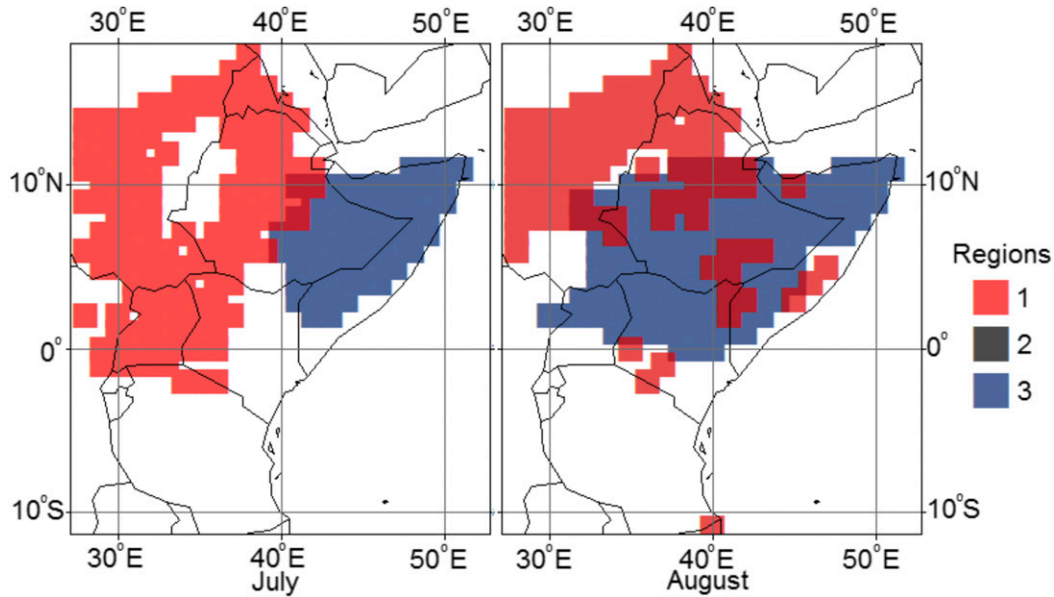


FIG. 4. Location of ECHAM4.5’s region 3 (blue) that is used to predict observed CHIRPS region 1 (red) for July and August.

only way to identify significant correlations with observations at our scale of analysis.

In comparing the RPS_{av} for ARM with that of SMA, it is seen that the ARM RPS_{av} is lower for both July and August (Table 4). A list of all the RPS for each month using both ARM and SMA is presented in the appendix. ARM outperforms SMA in 17 of the 29 yr and 23 of the 29 yr in predicting the month of July and August, respectively, when using the ECHAM4.5 forecast. The differences between yearly RPS values for ARM and SMA are marginally statistically significant for July (pairwise two-tailed t test; significance level $p = 0.08$) and highly significant for August ($p = 0.003$). RPSS values also show the value of ARM relative to SMA for these months (Table 4). Indeed, the RPSS for SMA shows negative values for both months, indicating a lower forecasting performance than having no prior information.

c. Large-scale drivers

Understanding the improved performance of the ARM requires an understanding of the dynamics at play within the region. Correlations of observed CHIRPS precipitation for August region 1 with observed SSTs (Fig. 5a) show a strong anticorrelation with the central tropical Pacific Ocean, in addition to a positive correlation in the western pacific over the Maritime Continent. In a broad sense these patterns are consistent in the maps of ECHAM4.5 SSTs correlation with ECHAM4.5 precipitation using both SMA (Fig. 5b) and

ARM (Figs. 5c). However, there is much greater similarity between observation (Fig. 5a) and ARM (Fig. 5c) correlation maps than there is between observation and SMA (Fig. 5b). This is particularly clear in the Indian Ocean, where ARM captures the positive correlation between precipitation and SSTs in the Indian ocean off the coast of southeast India, while SMA does not. Figure 5b also shows a large anticorrelation with Mediterranean SSTs, which directly opposes the relationship with observed SSTs. ARM shows no significant correlations at the 90% significance threshold, but correlations in the eastern Mediterranean are positive, matching the general tendency of observation (not shown). These correlation patterns are also consistent with CFSv2 using SMA (Fig. 5d), further illustrating the spatial bias within ECHAM4.5 and the need for spatial correction within ECHAM4.5’s precipitation outputs.

TABLE 3. Prediction method for the CHIRPS regions using CFSv2 and ECHAM4.5.

	CHIRPS regions		
	Region 1	Region 2	Region 3
CFSv2			
July	SMA	None	None
August	SMA	None	None
September	SMA	None	None
ECHAM4.5			
July	ARM	None	None
August	ARM	None	None
September	None	None	None

TABLE 4. The average RPS for ARM and SMA. Number of successes (NS) is the number of times ARM outperforms SMA over the 29-yr time span of the analysis. See the appendix for further illustration of the yearly model performance.

	Avg RPS		RPSS		NS
	SMA	ARM	SMA	ARM	
July	0.49	0.36	-0.09	0.2	18
August	0.59	0.33	-0.28	0.28	23

These similarities in SST correlations show that the ARM method in ECHAM4.5 more accurately captures large-scale dynamics that influence precipitation within the observed region 1.

The ARM result for ECHAM4.5 in August is reinforced if one looks at observed correlations between CHIRPS and SST when CHIRPS is averaged for ECHAM4.5 region 3 (i.e., the eastern Horn of Africa). These correlations are shown in Fig. 6, and it is evident that there is no significant association between rainfall in this region and the tropical Pacific. The fact that ECHAM4.5 region 3 precipitation does show correlation with SST in the tropical Pacific is further evidence that the model has shifted the true teleconnection eastward within the Horn of Africa, resulting in correlations between the eastern Horn and the Pacific that are in fact, more representative of northwest Ethiopia and Sudan—that is, observed region 1.

4. Conclusions

Multiple studies have shown the challenging nature of seasonal precipitation prediction over EA. The region contains steep precipitation gradients, is topographically complex, and is influenced by different large-scale climate dynamics in different seasons. Accurate seasonal forecast systems must capture the interplay of local, regional, and global dynamics that determine the temporal variability and spatial placement of rain within the region. The motivation for this paper is the recognition that dynamical forecast systems that capture large-scale dynamics can still fail to place precipitation variability correctly within the region. This results in low skill scores when models are evaluated or applied on the basis of traditional methods, which effectively make a spatial matching assumption of zero spatial bias. When the evaluation or application of the forecast is mediated by an objective regionalization that identifies analogous regions in model and observation, it is possible to extract meaningful information from a forecast system that would otherwise be discarded as unskillful.

This approach can be quite important for JAS EA precipitation. Objective regionalization for each of the JAS months shows that two commonly used dynamical forecast systems (ECHAM4.5 and CFSv2) regionalize quite differently from one another and also show distinct differences from regionalization based on observed

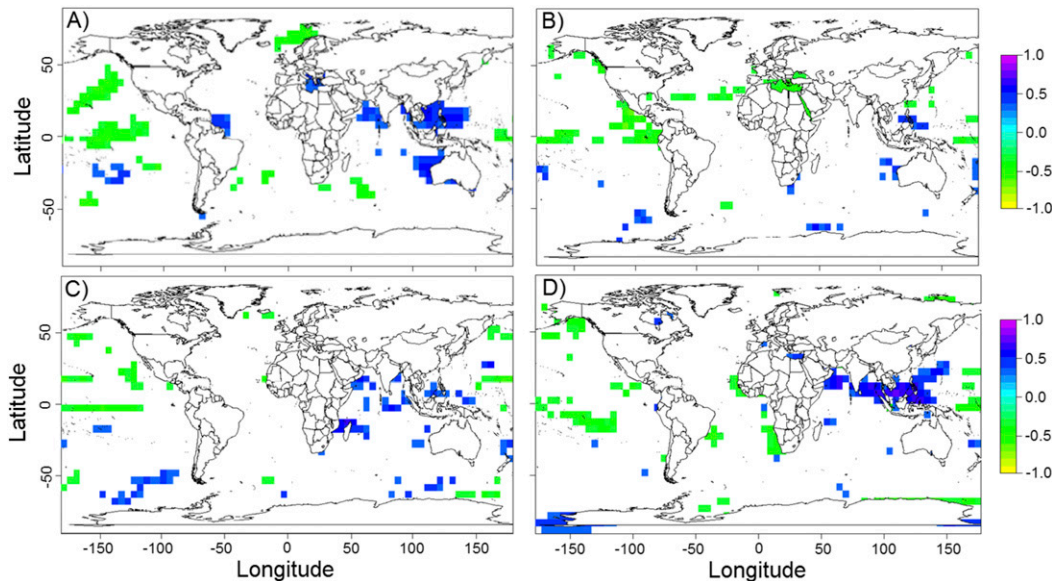


FIG. 5. Correlation of August precipitation with global gridded SST. All correlations are calculated as Spearman linear correlations and are masked at $\alpha = 0.1$. (a) Observed region 1 CHIRPS precipitation correlation with observed SST; (b) ECHAM4.5 precipitation within observed region 1 correlation with ECHAM4.5 SST (i.e., SMA method); (c) ECHAM4.5 precipitation in ECHAM4.5 region 3 with ECHAM4.5 SST (i.e., ARM approach for observed region 1); (d) CFSv2 precipitation in CFSv2 region 1 with CFSv2 SST (i.e., SMA approach).

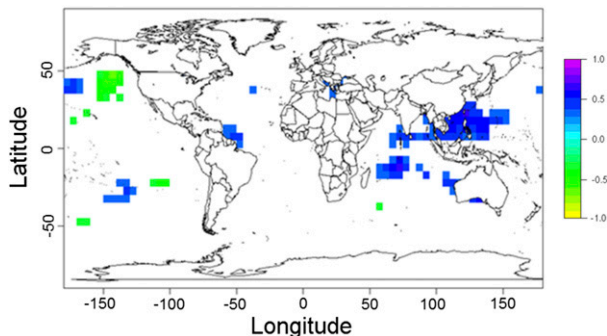


FIG. 6. Correlation of observed precipitation over Somalia with observed SSTs for August.

precipitation. Differences between observation and model regions indicate spatial biases within the models. We address this through analogous region mapping (ARM), which corrects acknowledges and adjusts for spatial bias. When compared with evaluation based on a spatial match assumption (SMA), which is similar to the traditional grid to grid approach, we find cases in which ARM allows for significant improvement. This was most clear for ECHAM4.5; at the peak of the JAS rainy season, the ARM method shows an improvement of the RPSS skill score from -0.09 to 0.20 and from -0.28 to 0.28 for the months of July and August, respectively, for a region that includes portions of the eastern Nile basin and parts of northern Ethiopia that are currently being affected by a significant El Niño-associated drought.

The RPSS results as well as the correlation maps presented in this paper show the ability of objective regionalization to improve the predictive utility of dynamic models. Future studies into this approach could vary the resolution of observational datasets in order to ascertain the impact of spatial resolution in identifying spatial biases. Similarly, use of an alternative high-resolution observation dataset could provide further insights to improve the method. Efforts are currently underway to provide even stronger satellite-gauge hybrid products for Africa, making use of nonpublic meteorological networks, such as Enhancing National Climate Services (ENACTS; Dinku et al. 2014). Unfortunately, ENACTS is not currently available for the entire study region. Ultimately, one would expect that analyses like these will contribute to continued model improvement to the point that spatial bias in dynamically based seasonal forecast systems becomes negligible. That level of model performance, however, is far from the current reality. For the foreseeable future it will be necessary to apply spatial correction methods like the regionalization approach presented in this paper in order to maximize the information contained in seasonal forecast systems.

TABLE A1. RPS for each month using both ARM and SMA. Boldface values show the average RPS scores for both methods and the total NS scores.

Run	RPS				NS	
	July	July	August	August	July	August
	SMA	ARM	SMA	ARM		
1	0.47	0.67	0.44	0.17	0	1
2	0.94	0.87	0.09	0.57	1	0
3	1.03	0.59	1.03	0.56	1	1
4	0.34	0.24	1.53	0.51	1	1
5	0.17	0.45	1.17	0.40	0	1
6	0.06	0.07	0.24	0.40	0	0
7	0.45	0.45	0.18	0.34	0	0
8	0.14	0.40	0.28	0.25	0	1
9	0.17	0.26	0.40	0.18	0	1
10	0.20	0.09	0.31	0.67	1	0
11	1.14	0.59	0.06	0.37	1	0
12	0.69	0.45	0.63	0.36	1	1
13	0.63	0.31	0.28	0.36	1	0
14	0.81	0.45	0.51	0.26	1	1
15	0.17	0.28	0.28	0.17	0	1
16	0.57	0.37	0.40	0.06	1	1
17	0.37	0.12	0.45	0.14	1	1
18	0.17	0.74	0.40	0.40	0	1
19	0.34	0.18	1.09	0.26	1	1
20	1.09	0.67	1.28	0.94	1	1
21	0.37	0.47	0.67	0.56	0	1
22	0.28	0.14	0.87	0.40	1	1
23	1.14	0.25	0.94	0.37	1	1
24	0.45	0.20	0.28	0.24	1	1
25	0.03	0.31	0.59	0.06	0	1
26	0.69	0.11	0.95	0.07	1	1
27	0.74	0.03	0.81	0.03	1	1
28	0.11	0.51	0.14	0.07	0	1
29	0.47	0.14	0.87	0.51	1	1
	0.49	0.36	0.59	0.33	18	23

Acknowledgments. This research was supported in part by the NASA Interdisciplinary Research in Earth Sciences Program, Project NNX14AL93G, as well as NASA Greater Horn of Africa Project NNX14AD30G.

APPENDIX

RPS and NS Scores

Table A1 lists all of the RPS for each month using both ARM and SMA.

REFERENCES

Badr, H. S., B. F. Zaitchik, and A. K. Dezfuli, 2014: HiClimR: Hierarchical climate regionalization. Comprehensive R Archive Network (CRAN). [Available online at <http://cran.r-project.org/package=HiClimR>.]
 —, —, and —, 2015: A tool for hierarchical climate regionalization. *Earth Sci. Inf.*, **8**, 949–958, doi:10.1007/s12145-015-0221-7.

- Barnston, A. G., S. Li, S. J. Mason, D. G. DeWitt, L. Goddard, and X. Gong, 2010: Verification of the first 11 years of IRI's seasonal climate forecasts. *J. Appl. Meteor. Climatol.*, **49**, 493–520, doi:10.1175/2009JAMC2325.1.
- Batté, L., and M. Déqué, 2011: Seasonal predictions of precipitation over Africa using coupled ocean-atmosphere general circulation models: Skill of the ENSEMBLES project multimodel ensemble forecasts. *Tellus*, **63A**, 283–299, doi:10.1111/j.1600-0870.2010.00493.x.
- Berhane, F., and B. Zaitchik, 2014: Modulation of daily precipitation over East Africa by the Madden-Julian oscillation. *J. Climate*, **27**, 6016–6034, doi:10.1175/JCLI-D-13-00693.1.
- , —, and A. Dezfuli, 2014: Subseasonal analysis of precipitation variability in the Blue Nile River basin. *J. Climate*, **27**, 325–344, doi:10.1175/JCLI-D-13-00094.1.
- Block, P., and B. Rajagopalan, 2007: Interannual variability and ensemble forecast of upper Blue Nile basin *Kiremt* season precipitation. *J. Hydrometeorol.*, **8**, 327–343, doi:10.1175/JHM580.1.
- 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.
- Conway, D., 2000: The climate and hydrology of the upper Blue Nile River. *Geogr. J.*, **166**, 49–62, doi:10.1111/j.1475-4959.2000.tb00006.x.
- Dezfuli, A., and S. Nicholson, 2013: The relationship of rainfall variability in western equatorial Africa to the tropical oceans and atmospheric circulation. Part II: The boreal autumn. *J. Climate*, **26**, 66–84, doi:10.1175/JCLI-D-11-00686.1.
- Dinku, T., P. Block, J. Sharoff, K. Hailemariam, D. Osgood, J. del Corral, R. Cousin, and M. C. Thomson, 2014: Bridging critical gaps in climate services and applications in Africa. *Earth Perspect.*, **1**, 15, doi:10.1186/2194-6434-1-15.
- Diro, G. T., D. I. F. Grimes, and E. Black, 2011: Teleconnections between Ethiopian summer rainfall and sea surface temperature: Part II. Seasonal forecasting. *Climate Dyn.*, **37**, 121–131, doi:10.1007/s00382-010-0896-x.
- FAO, 2000: The elimination of food insecurity in the Horn of Africa: A concerted government and UN agency action. Final Rep. [Available online at <http://www.fao.org/docrep/003/x8406e/x8406e00.htm>.]
- Funk, C., and Coauthors, 2015: The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data*, **2**, 150066, doi:10.1038/sdata.2015.66.
- Gissila, T., E. Black, D. I. F. Grimes, and J. M. Slingo, 2004: Seasonal forecasting of the Ethiopian summer rains. *Int. J. Climatol.*, **24**, 1345–1358, doi:10.1002/joc.1078.
- Goddard, L., S. J. Mason, S. E. Zebiak, C. F. Ropelewski, R. Basher, and M. A. Cane, 2001: Current approaches to seasonal to interannual climate predictions. *Int. J. Climatol.*, **21**, 1111–1152, doi:10.1002/joc.636.
- Jury, M. R., 2014: Evaluation of coupled model forecasts of Ethiopian highlands summer climate. *Adv. Meteorol.*, 894318, doi:10.1155/2014/894318.
- Kaplan, A., M. A. Cane, Y. Kushnir, A. C. Clement, M. B. Blumenthal, and B. Rajagopalan, 1998: Analyses of global sea surface temperature 1856–1991. *J. Geophys. Res.*, **103**, 18 567–18 589, doi:10.1029/97JC01736.
- Kirtman, B. P., and Coauthors, 2014: The North American Multimodel Ensemble: Phase-1 seasonal-to-interannual prediction; phase-2 toward developing intraseasonal prediction. *Bull. Amer. Meteor. Soc.*, **95**, 585–601, doi:10.1175/BAMS-D-12-00050.1.
- Koster, R. D., T. L. Bell, R. H. Reichle, M. J. Suarez, and S. D. Schubert, 2008: Using observed spatial correlation structures to increase the skill of subseasonal forecasts. *Mon. Wea. Rev.*, **136**, 1923–1930, doi:10.1175/2007MWR2255.1.
- Lance, G. N., and W. T. Williams, 1967: A general theory of classificatory sorting strategies. 1. Hierarchical systems. *Comput. J.*, **9**, 373–380, doi:10.1093/comjnl/9.4.373.
- Lyon, B., and D. G. DeWitt, 2012: A recent and abrupt decline in the East African long rains. *Geophys. Res. Lett.*, **39**, L02702, doi:10.1029/2011GL050337.
- Murtagh, F., 1985: *Multidimensional Clustering Algorithms*. Compstat Lectures: Lectures in Computational Statistics, J. M. Chambers, et al., Eds., Vol. 4, Physika Verlag, 134 pp.
- , and P. Legendre, 2014: Ward's hierarchical agglomerative clustering method: Which algorithms implement Ward's criterion? *J. Classif.*, **31**, 274–295, doi:10.1007/s00357-014-9161-z.
- Nicholson, S. E., 1996: A review of climate dynamics and climate variability in eastern Africa. *The Limnology, Climatology and Paleoclimatology of the East African Lakes*, T. C. Johnson and E. O. Odada, Eds., CRC Press, 25–56.
- , 2000: The nature of rainfall variability over Africa on time scales of decades to millennia. *Global Planet. Change*, **26**, 137–158, doi:10.1016/S0921-8181(00)00040-0.
- , 2014: The predictability of rainfall over the Greater Horn of Africa. Part I: Prediction of seasonal rainfall. *J. Hydrometeorol.*, **15**, 1011–1027, doi:10.1175/JHM-D-13-062.1.
- Olsson, L., 1993: On the causes of famine: Drought, desertification and market failure in the Sudan. *Ambio*, **22**, 395–403.
- Parker, D. E., P. D. Jones, C. K. Folland, and A. Bevan, 1994: Interdecadal changes of surface temperature since the late nineteenth century. *J. Geophys. Res.*, **99**, 14 373–14 399, doi:10.1029/94JD00548.
- Reynolds, R. W., and T. M. Smith, 1994: Improved global sea surface temperature analyses using optimum interpolation. *J. Climate*, **7**, 929–948, doi:10.1175/1520-0442(1994)007<0929:IGSSTA>2.0.CO;2.
- Segele, Z. T., and P. J. Lamb, 2005: Characterization and variability of Kiremt rainy season over Ethiopia. *Meteor. Atmos. Phys.*, **89**, 153–180, doi:10.1007/s00703-005-0127-x.
- Tsidu, G. M., 2012: High-resolution monthly rainfall database for Ethiopia: Homogenization, reconstruction, and gridding. *J. Climate*, **25**, 8422–8443, doi:10.1175/JCLI-D-12-00027.1.
- UN OCHR, 2011: Eastern Africa drought. Humanitarian Rep. 3, 11 pp. [Available online at <http://reliefweb.int/sites/reliefweb.int/files/resources/OCHA%20Eastern%20Africa%20Humanitarian%20Report%20No.%20-%20-%20Drought%20May%202011%20FINAL.pdf>.]
- Ward, J. H., Jr., 1963: Hierarchical grouping to optimize an objective function. *J. Amer. Stat. Assoc.*, **58**, 236–244, doi:10.1080/01621459.1963.10500845.
- Wilks, D. S., 2011: *Statistical Methods in the Atmospheric Sciences*. 3rd ed. Elsevier, 676 pp.
- Williams, A. P., and C. Funk, 2011: A westward extension of the warm pool leads to a westward extension of the Walker circulation, drying eastern Africa. *Climate Dyn.*, **37**, 2417–2435, doi:10.1007/s00382-010-0984-y.
- Yang, W., R. Seager, M. A. Cane, and B. Lyon, 2014: The East African long rains in observations and models. *J. Climate*, **27**, 7185–7202, doi:10.1175/JCLI-D-13-00447.1.